Compiled Comments - Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research

Guide Notice Number: NOT-OD-23-091

February 21, 2023 – April 24, 2023

# **Table of Contents**

- **1.** Aaron W Dobbs
- 2. <u>Natalie Moffett</u>
- 3. University of Oregon Libraries
- 4. Mayank Verma
- 5. Gail Johnson
- **6.** Daniel Gorelick
- **7.** Peter L. Elkin, MD
- **8.** Anonymous
- **9.** Kenneth Pawlak
- **10.** Bobby Hollingsworth
- **11.** Anonymous
- 12. Center for emerging and neglected diseases uc berkeley
- **13.** Sandra Poulson
- **14.** Adam Armstrong
- **15.** Federico Leva
- **16.** Catherine Christian-Hinman
- **17.** Anonymous
- **18.** Sonal Sathe
- 19. Anonymous
- **20.** Tiffany Atkins
- **21.** Damien Camany
- 22. Peggy Lentz
- **23.** Rebecca Braddock
- **24.** Thurman McGarian
- 25. Stephen J. Kron
- **26.** Peggy Lentz
- **27.** Christopher Marcum
- 28. Aligning Science Across Parkinson's
- **29.** Ana Chicas-Mosier
- **30.** Andrea Bertke
- **31.** Federation of American Societies for Experimental Biology
- **32.** NC State University
- **33.** Jessica Moise
- 34. Luella Allen-Waller
- **35.** Elizabeth McNally
- **36.** cOAlition S
- **37.** Robert Weinberg
- 38. Nils Walter
- 39. Clifford B. Saper, MD, PhD
- 40. Amy Wright
- 41. Anonymous
- **42.** Lynn Brielmaier

- 43. Gail A. Bishop
- 44. Mr. Stephen P. Panossian
- **45.** Phil Hurvitz
- **46.** Andriy Fedorov
- 47. Robin Ely MD
- 48. Bill & Melinda Gates Foundation
- 49. American Academy of Neurology
- **50.** American Society of Colon & Rectal Surgeons/ Journal Diseases of the Colon & Rectum
- **51.** Sonal Sathe
- **52.** Steve Pieper
- **53.** Peer Community In
- **54.** <u>Libraries and Sponsored Projects Administration</u>
- **55.** Carl Tuttle
- **56.** Open Research Funders Group
- **57.** Public Knowledge Project
- **58.** Steven D. Smith
- 59. Mark Peifer
- **60.** NC State University Libraries
- **61.** John Vaughen
- **62.** Duke University
- **63.** American Psychological Association
- 64. American Society of Clinical Oncology, American College of Physicians, NEJM Group,
  American Heart Association, American Diabetes Association, American Society of
  Anesthesiologists, American Urological Association, American Thoracic Society, American
  Gastr
- **65.** University of Illinois Chicago
- **66.** European Bioinformatics Institute EMBL-EBI
- **67.** The American Society for Cell Biology
- 68. Massachusetts Medical Society New England Journal of Medicine (NEJM)
- 69. ORCID
- 70. Association of American Medical Colleges
- 71. <u>American Chemical Society</u>
- **72.** AcademyHealth
- 73. Oregon State University Libraries and Press
- **74.** American Association for the Advancement of Science
- **75.** Public Library of Science (PLOS)
- **76.** Frontiers Media Inc
- 77. The American Association of Immunologists (AAI)
- **78.** Congress of Neurological Surgeons
- **79.** Northern Illinois University
- 80. ASAPbio
- **81.** American Thoracic Society
- **82.** American Society for Microbiology
- 83. Elsevier

- **84.** American Society for Investigative Pathology
- 85. STM (The International Association of Scientific, Technical and Medical Publishers)
- **86.** Council of Medical Specialty Societies
- **87.** Alliance for Nursing Informatics
- 88. Association of American Universities
- **89.** The Alliance for Aging Research
- 90. American Society for Pharmacology and Experimental Therapeutics
- **91.** The Ohio State University
- **92.** Association of American Publishers
- 93. Medical Library Association & Association of Academic Health Sciences Libraries
- **94.** Taylor and Francis Group
- 95. Huntsman Cancer Institute
- 96. American Urological
- **97.** Springer Nature
- 98. Copyright Clearance Center
- 99. American Society for Radiation Oncology
- 100. Copyright Clearance Center
- 101. Research Data Access and Preservation Association
- 102. American Gastroenterological Association
- **103.** American Society for Nutrition
- **104.** Crossref
- 105. Simon Bacon
- **106.** Data Discovery Collaboration
- 107. American Society of Hematology
- **108.** Center for Open Science
- 109. Case Western Reserve University Libraries
- 110. American Medical Informatics Association
- 111. Epilepsy Foundation
- 112. Massachusetts Institute of Technology MIT Libraries
- **113.** OASPA
- **114.** Wolters Kluwer
- 115. Research Organization Registry (ROR)
- **116.** Association of Research Libraries
- **117.** Wilev
- 118. Seventeen Science Societies
- **119.** The Allen Institute
- 120. Society for Industrial and Applied Mathematics (SIAM)
- 121. Women's Health Issues
- **122.** U.S. Repository Network
- 123. American Society of Mechanical Engineers
- 124. Association of Public and Land-grant Universities
- **125.** COGR
- **126.** American Association for Cancer Research
- 127. University of California Office of the President

- 128. Carrie Nelson and Cameron Cook
- 129. Arizona State University -ASU Library
- **130.** American Association for Dental, Oral, and Craniofacial Research
- **131.** Figshare (Digital Science)
- **132.** American Society of Clinical Oncology
- 133. Jonathan Saunders
- **134.** Creative Commons
- 135. The Gerontological Society of America
- **136.** American College of Radiology
- **137. SPARC**
- 138. Federation of Associations in Behavioral and Brain Sciences
- 139. Booz Allen Hamilton, Inc.
- 140. Robin Puett
- **141.** Gary McDowell
- 142. American College of Rheumatology
- **143.** Fully OA Publishers

**Submit date: 2/21/2023** 

I am responding to this RFI: On behalf of myself

Name: Aaron W Dobbs

Name of Organization: US Citizen and Member of the Public

Type of Organization: Other

Type of Organization-Other: Member of the Public

Role: Member of the public

# 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Upon publication, \*all\* US government funded (especially NIH) research results and publications derived therefrom shall be submitted by the author indexed in PMC (or successor resources). Researchers are encouraged to publish wherever they wish, with the stipulation that the results of the research and any publications derived therefrom shall be publicly indexed in PMC (or successor resources).

# 2. Steps for improving equity in access and accessibility of publications.

Upon publication, \*all\* US government funded (especially NIH) research results and publications derived therefrom shall be made publicly available in the Public Domain. Researchers are encouraged to publish wherever they wish, with the stipulation that the results of the research and any publications derived therefrom shall be shall be in the Public Domain.

# 3. Methods for monitoring evolving costs and impacts on affected communities.

Acknowledging current profitability levels of major research publishers, \*all\* US government funded (especially NIH) research results and publications derived therefrom should not be subject to publication fees. If this position is too extreme, then a one-time payment for publication should be funded by NIH, not to exceed the 10-year average of the journal issue subscription price divided by the 10-year average page count of each article in an issue.

4. Early input on considerations to increase findability and transparency of research.

**Submit date: 2/22/2023** 

I am responding to this RFI: On behalf of myself

Name: Natalie Moffett

Name of Organization: Student - Washington State University

Type of Organization: Not applicable

Role: Member of the public

1. How to best ensure equity in publication opportunities for NIH-supported investigators.

2. Steps for improving equity in access and accessibility of publications.

The websites fonts sizes can vary wildly, making them larger and more consistent would be better. Prefer bolding and underlining to differentiate titles/authors/section headings. Avoid italics when possible.

3. Methods for monitoring evolving costs and impacts on affected communities.

Publications should announce where your application fees go -what percentage pays for reading and fact checking, how much is spent on formatting or printing, and how much is simply getting kicked into a publishers coffers.

4. Early input on considerations to increase findability and transparency of research.

Washington vs Wichita State Universities can be hard to tell apart, I imagine similar problems occur at Other universities and can make it difficult to track down research and scientists for follow-up questions.

**Submit date: 2/22/2023** 

I am responding to this RFI: On behalf of an organization

Name: Alicia Salaz

Name of Organization: University of Oregon Libraries

Type of Organization: University

Role: Institutional official

# 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Allowing federal funds to be used to pay individual author publication charges (APCs) to support 'gold' open access publishing is helpful to the funded author in the short-term, but fuels inflationary price increases to APCs that negatively impact all authors across the research publication system in the US and around the world and entrench growing inequities in access to publishing. As a professional information steward, I have serious concerns about continuing this policy, and would encourage NIH and Other federal agencies to consider expanding and encouraging "green" open access deposits for compliance, for instance, depositing an accepted version of a manuscript into PMC, or a local institutional or disciplinary repository. NIH dollars should go directly towards supporting these government, non-profit, or academy-owned infrastructures, and not to individual researchers to take to the non-competitive publishing marketplace.

A very large percentage of federally funded APC fees go directly to for-profit publishers, whose profit margins on scientific publishing have regularly exceeded 30-40%. Meanwhile, subscription read fees for scholarly journal databases as public and publicly-funded academic libraries across the nation continue to climb. At the system level, this is not a good or effective use of public money or good stewardship of taxpayer funds.

- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.

Is information gleaned from monitoring actionable? APC fees from for-profit publishers have already inflated to excessive levels. For-profit publishers say they offer discounts and waivers to provide equitable access to publishing. This is highly misleading. The availability, hoops, hurdles, and administrative overhead required to know about and secure these waivers is prohibitive. Our Library strongly advocates for NIH to pursue policies that move towards low-cost, open public infrastructures for sharing the products of publicly funded research (such as PMC); thereby driving down market demand both for for-profit journal subscriptions and open access publishing charges, and hopefully lowering costs for consumers and producers of publicly funded research across the board.

4. Early input on considerations to increase findability and transparency of research.

**Submit date:** 2/23/2023

I am responding to this RFI: On behalf of myself

Name: Mayank Verma

Name of Organization: UTSW

Type of Organization: University

Role: Scientific researcher

- 1. How to best ensure equity in publication opportunities for NIH-supported investigators.
- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.

Journals fees should be capped for NIH dollar expenses.

4. Early input on considerations to increase findability and transparency of research.

**Submit date: 2/24/2023** 

I am responding to this RFI: On behalf of myself

Name: Gail Johnson

Name of Organization: University of Rochester

Type of Organization: University

Role: Scientific researcher

1. How to best ensure equity in publication opportunities for NIH-supported investigators.

2. Steps for improving equity in access and accessibility of publications.

In principal I agree that having published works immediately available is good. However, depending on the journal, the cost of open access publishing can be prohibitively expensive. Unfortunately NIH cannot put caps on the amount journals can charge, which would solve the problem (point 3 below). Alternatively a possible solution would be that if the paper has been uploaded to a preprint server such as bioRxiv then publishing using the subscription model with an embargo period should be allowable.

- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

**Submit date: 2/24/2023** 

I am responding to this RFI: On behalf of myself

Name: Daniel Gorelick

Name of Organization: Baylor College of Medicine

Type of Organization: University

Role: Scientific researcher

# 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

One option is to modify the NIH Public Access Plan to allow grantees to post manuscripts as preprints on a server like bioRxiv or medRxiv. This would immediately make results from NIH-funded research freely accessible. Preprints are free to post and free to read but are not

peer-reviewed. Following posting of a preprint, scientists are free to submit their manuscript for peer-review and publication

in any journal, whether subscription or open access. This would minimize the compliance burden on NIH-supported researchers and also maintains the flexibility of NIH-supported researchers to publish their final, peer-reviewed manuscript in any journal. For details on the advantages and limitations of this idea, see the attached PDF.

- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.

In the proposed policy, NIH-funded publications would be deposited in PubMed Central immediately following publication. People could read peer-reviewed manuscripts for free immediately following publication without waiting for a 12 month embargo to elapse.

Advantages: All NIH-funded research would be free to read immediately to anybody in the world.

Disadvantages: This policy does not address article publication costs (APC) to scientists. Journal publishers would be free to set APCs as they wish, with no maximums, as is the case currently. Scientists would need to find ways of paying the APCs, using either grant money, institutional funds or personal funds.

The attached PDF discusses ways to reduce these APCs and still maintain open access, either by having NIH pay APCs directly to journals, by having NIH cap the amount of grant money to be used to pay APCs (eg no more than \$3000 per manuscript), by the NIH publishing more journals that would be free to read and free to publish (similar to the existing Environmental Health Perspectives, published by NIEHS), or some combination of these policies.

4. Early input on considerations to increase findability and transparency of research.

# **Uploaded File:**

Gorelick-OpenAccess-Proof-Final.pdf

Email: gorelick@bcm.edu

# Reducing open access publication costs for biomedical researchers in the U.S.A.

Daniel A. Gorelick<sup>1,\*</sup> and Ye Li<sup>2,\*\*</sup>

Edited by Lisa D. Cervia and Grant A. Knappe

# **HIGHLIGHTS**

- Biomedical publishing is adopting an open access model, where peer-reviewed manuscripts are free to read, but authors pay a fee to the journal to publish their manuscript
- Funding agencies often pay publication fees on behalf of scientists
- Publication fees are rising much faster than inflation, putting a burden on scientists and funding agencies to adopt open access policies that reduce costs to authors

Open Access (OA) publishing is a critical route for biomedical researchers to broadly disseminate their research results and comply with policies from funding agencies. A popular business model for OA publishing requires scientists to pay an article processing charge (APC). In the last two decades, APCs have risen well beyond inflation, posing a burden to scientists and funding agencies that often pay APCs on behalf of scientists. There are no policies in place that address the rising costs of APCs. Here, we examined the history of OA in biomedical research and analyzed the benefits and limitations of different OA policies and their effects on APCs.

Biomedical researchers disseminate the results of their research by publishing results in peer-reviewed journals. Historically, publication operated under a subscription model. Scientists would submit manuscripts to journals, journal publishers would organize peer-review and decide whether or not to publish the manuscript. Once the manuscript was published, readers would pay publishers to read the published articles. Journal subscription fees were typically paid for by universities or other research organizations on behalf of their faculty. The subscription model restricts access of scientific results to those that can afford to pay to read them. A single university cannot afford to subscribe to every journal.

Additionally, the ability to distribute and read journal articles digitally, on the World Wide Web, made it easier for readers to access journal articles. In response, biomedical research, among other disciplines, is adopting an open access (OA) model [1], where journal articles are free to read online.

When the peer-reviewed journal articles are free to read, innovative business models are needed to pay for the cost of publishing, in place of the subscription model. One popular model is to require that authors pay an article processing charge (APC) to the journal's publisher upon acceptance of a scientific manuscript for publication. Currently, the APCs are rising and outpacing inflation. Between 2012 and 2016, a study of 10 leading universities in the United Kingdom found that the average APC paid by or on behalf of scientists increased 16%, while the consumer price index, a marker of inflation, increased 5% [2]. A meta-analysis of variations in APC trends from 2011 to 2021 also demonstrated the increases of average APCs are well above the inflation rate for most of the journal categories examined [3]. For example, the average APC change is 186% in comparison with inflation for the journals with high Impact Factors in the Journal Citation Report (JCR) [4] and indexed by the Directory of Open Access Journals (DOAJ) [5]. As the number of scientific papers increases annually, and as more scientific papers are published with open access rather than on a subscription basis, this creates a cycle of positive reinforcement that could drive APCs higher. In this policy paper, we discuss ways to reduce open access publication costs for biomedical researchers in the U.S.A. We recognize that issues of open access and APCs affect scientists of all disciplines all over the world, and there may not be a single policy optimal for all disciplines in all locations. In the United States, the government spends more money funding biomedical research than funding physics, chemistry and other disciplines [6]. Thus, we restrict our analysis to biomedical research, though some of the policy options explored here may be relevant for other disciplines.

# A brief history of open access in biomedical research

In 1996, the Journal of Clinical Investigation became the first prominent biomedical journal to be completely free to read online. Instead of libraries paying a subscription fee, authors paid an article processing charge to have their manuscripts published in the Journal of Clinical Investigation [7]. In

The authors declare no conflict of interest.

© 2021 The Author(s)

<sup>&</sup>lt;sup>1</sup>Department of Molecular and Cellular Biology, Center for Precision Environmental Health. Baylor College of Medicine. Houston. TX

<sup>&</sup>lt;sup>2</sup>MIT Libraries, Massachusetts Institute of Technology, Cambridge, MA

 $<sup>^*</sup>$ Email: Daniel.Gorelick@bcm.edu

<sup>\*\*</sup>Email: yel@mit.edu

the 2000s, more journals began publishing peer-reviewed articles online, making it easier for open access to thrive. In 2000, the National Institutes of Health (NIH) established PubMed Central (PMC), a free, online, full-text archive of biomedical and life sciences journal literature [8]. Since 2008, the United States Government required that peer-reviewed manuscripts produced by investigators funded by the NIH be deposited in PubMed Central no later than 12 months after the official date of publication in a peer-reviewed journal [9]. The 12 month embargo was a compromise between publishers, concerned about losing revenue if their subscription material was available to read immediately, and scientists, funders, and the general public, who wanted results free to read immediately [10]. The 12 month embargo only partially solved the open access problem. Scientists who can't pay journal subscription costs are still at a competitive disadvantage compared to colleagues that can read published results immediately and begin follow-up studies. When we discuss Open Access publications in this article, we focus on peer-reviewed manuscripts or published version of records that are free to read immediately upon acceptance to the journal, with no embargo period following publication.

In 2018, a consortium of European funding agencies, cOAlition S, created Plan S, an initiative to promote full and immediate Open Access to research publications [11]. Funding agencies that are cOAlition S members, such as Wellcome Trust, require their grantees to publish their results in "Open Access Journals, on Open Access Platforms, or made immediately available through Open Access Repositories without embargo" and may provide funds for their grantees to publish OA. Plan S, however, does not directly address open access publication costs. Plan S does not determine who will pay APCs and encourages, but does not require, publishers to be transparent about their costs to define a fair market value for APCs [12]. Moreover, none of the federal funding agencies in the U.S. participates in cOAlition S, in part because the White House Office of Science and Technology expressed concerns about telling researchers where to publish (cOAlition S grantees cannot publish in journals that lack an open access option as of January 1, 2021) [13].

# Article processing charges (APC)

As a popular model for OA, APC can cover the publishing costs and solve issues regarding access. But there are no regulations that limit the APC amount. For example, publishing in an influential journal is seen as a professional benefit, one that scientists think can lead to promotions and increase their chances of receiving grant funding (money provided by the government or private foundations to pay for biomedical research). The demand to publish in influential journals is high, therefore such journals can charge higher APCs, as has occurred in the last decade [3].

How do scientists pay APCs? Some universities provide scientists with funds for this purpose, although they often only partially cover the full APCs for biomedical journal articles [14]. More commonly in Biomedical fields, funding agencies, such

as the National Institutes of Health, enable authors to use grant money to pay APCs. However, funding agencies and universities often do not have an official limit on the amount of money that can be used to pay publication charges. When the funder is a government organization, such as the U.S. National Institutes of Health (NIH), then this invites scrutiny of whether paying high APCs supports the funding agency's mission.

# **Current OA publishing policies**

Open access publication policies for biomedical journals are generally driven by research funding agencies. Papers containing research funded by the U.S. National Institutes of Health must be free to read no later than one year following publication [9]. Papers containing research funded by the Howard Hughes Medical Institute (effective January 1, 2022), the Bill & Melinda Gates Foundation, Wellcome Trust, and the World Health Organization must be free to read immediately upon publication [15, 16]. Many funding agencies in the European Union and United Kingdom require that researchers make their published manuscripts free to read immediately [17].

Journals offer a mix of open access and subscription options, depending on the journal. Some subscription-based journals offer an open access option for individual manuscripts, which are often referred to as hybrid journals. Other journals are entirely open access, or entirely subscription-based. The hybrid journals can result in confusion around whether a publisher is "double dipping" — charging authors APCs while charging institutions for subscriptions fees. Some medical journals do not offer authors OA options when their research is funded by agencies that do not mandate OA publishing.

Journal publishers typically follow government regulations or funder mandates regarding open access. For example, for research funded by the NIH, subscription-based journals will make manuscripts free to read one year after publication, by depositing the peer-reviewed manuscripts in PubMed Central. In some cases, subscription-based journals make published articles freely available after six months rather than the one year embargo period mandated by the NIH, for example journals published by the American Medical Association and the U.S. National Academy of Sciences [18].

Journals can be published by for-profit companies (e.g., Elsevier, SpringerNature) or non-profit (or not-for-profit) organizations including scientific societies (e.g., Company of Biologists, American Association for the Advancement of Science). Sometimes a non-profit scientific society will contract a for-profit publishing company to publish their journal (e.g., Developmental Biology, the official journal of the non-profit Society for Developmental Biology, is published by Elsevier). In every case, journals have an incentive to maximize revenue, to either increase shareholder profits in the case of for-profit publishing companies, or to increase the ability to provide services to members and benefit the public good, in the case of scientific societies and other

non-profit organizations (two examples of exceptions to this rule of maximizing revenue, the journals Environmental Health Perspectives and Quantitative Science Studies, are discussed later). For profit publishers may choose to use profits to benefit the scientific community and the public good (at the expense of returning profits to shareholders), but they are under no legal obligation to do so. During the process of transitioning from the traditional subscription model to the OA publishing model, maximizing revenue means charging a higher APC, publishing more articles, or both.

The more prestigious or influential the journal, the higher the APC [19, 20]. For example, SpringerNature charges \$11,500 (€9500) to publish OA in Nature (Article Influence Score 22), \$5,380 to publish in Nature Communications (Article Influence Score 5.5) and \$1,870 to publish in Scientific Reports (Article Influence Score 1.9) [21]. No matter how influence is measured, whether using Impact Factor, article influence score, or other metric such as citability, there is a positive correlation between APC and journal influence [22]. The article influence score (AIS) measures the average influence of each of a journal's articles over the first five years after publication [20]. AIS is similar to the journal impact factor but corrects for self-citations. AIS is normalized so that the mean article in the Clarivate Journal Citation Reports (JCR) database has an AIS of 1.00. In 2019, the average article in Nature Communications had 5.5 times the influence of the average article in the JCR. Scientists use journal impact factors and to assign a relative measure of influence to biomedical journals (the higher the impact factor, the more influential and prestigious the journal), but impact factor and AIS are not absolute measures and should be used with caution [23]. APCs positively correlate with journal influence among journals from the same publisher (Fig. 1).

There are no regulations that limit APCs, although some funding agencies have set a maximum on the APCs they are willing to pay (see Policy Type 3, below). Scientists value prestige as professional currency and pay publication or subscription costs using someone else's money, typically money provided by the funding agency supporting the research project or the research institution where the scientists work. Scientists from lower income countries or underprivileged institutions are left with little or no options to pay APCs. Some publishers may selectively provide APC waivers or discount for low- and middle-income countries [24] but they may not be as consistent and transparent as needed [25]. In addition, a recent study found that global health researchers in low- and middle-income countries (LMIC) cite non-OA articles less than researchers from upper-income countries [26]. This suggests that researchers from LMIC are not reading and citing non-OA manuscripts as frequently as OA manuscripts, likely because of the cost to read non-OA manuscripts.

The publication polices of funding agencies, research institutions, governments, scientific societies and publishers are related and influence each other. Here, we focus on policies of funding agencies, since they have the broadest

and most fundamental impact. If funding agencies require open access, then most scientists must comply. Individual research institutions may also have their own OA policies, but these policies often provide opt-out options for researchers and thus are not enforced 100% in practice [27,28]. In addition, although we are not focusing on open data policies in this article, they are important components of open science policies along with the OA publishing policies.

# Policy Type 0: Maintain the status quo but eliminate the open access embargo

In this model, NIH-funded publications would be deposited in PubMed Central immediately following publication. People could read peer-reviewed manuscripts for free immediately following publication without waiting for a 12 month embargo to elapse.

Advantages: This policy would make the United States compliant with Plan S. All NIH-funded research would be free to read immediately to anybody in the world.

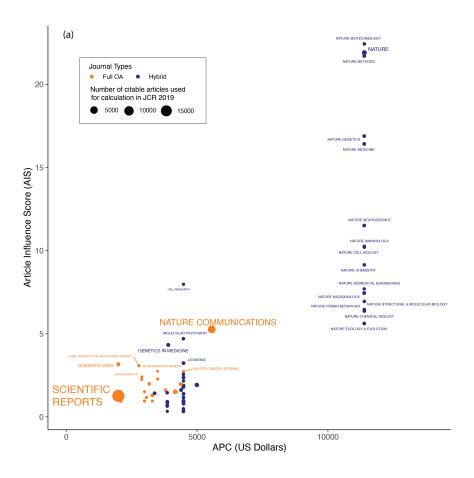
Disadvantages: This policy does not address the costs of APCs to scientists. Journal publishers would be free to set APCs as they wish, with no maximums, as is the case currently. Scientists would need to find ways of paying the APCs, using either grant money, institutional funds or personal funds. Additionally, changing the NIH public access policy would likely require action by the U.S. Congress.

# Policy Type 1: Funding agency pays open access costs directly

In this model, open access journals are directly supported by funding agencies such that there is no charge to funded authors to publish and no charge to read published articles. Funding agencies might publish a journal themselves or pay a publishing company so that grantees are charged nothing to publish. Several existing journals demonstrate the feasibility of this model. Environmental Health Perspectives is the most influential journal in the field of toxicology and environmental health sciences. Published by the National Institute of Environmental Health Sciences at the NIH, it charges authors no publication fees and is free to read [29, 30]. Environmental Health Perspectives has been published using this model since 2004 and has no plans to change their publishing model.

Wellcome Trust and the Bill & Melinda Gates Foundation pay a for-profit publisher, F1000 Research Ltd., to publish Wellcome Open Research and Gates Open Research journals, respectively [31,32]. These journals are free to read and charge authors no article publication costs (CC BY license) provided the author's research is funded by the Wellcome Trust or the Gates Foundation. In contrast to Environmental Health Perspectives, these journals are newer and so their reputation and influence in the biomedical sciences is untested.

Advantages: This is the least expensive business model for scientists because scientists would not pay APCs. This model currently works at a small scale for funders with steady financial resources, as evidenced by the journals discussed



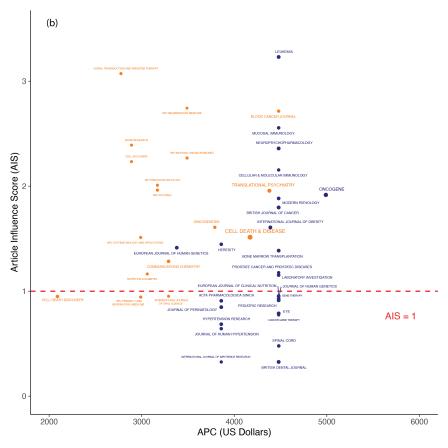


Figure 1: Article Processing Charges (APC) and Article Influence Scores (AIS) of biomedical journals published by SpringerNature under their Nature Portfolio. (a) shows all biomedical journals, (b) focuses on journals with APC between \$2000 and \$6000 and AIS less than 3.5. AIS measures the average influence of each of a journal's articles over the first five years after publication, where AIS > 1 has above average influence and AIS < 1 has below average influence. A journal with an AIS of three means the average article in that journal had three times the influence of the average article in the Clarivate InCites Journal Citation Reports (JCR). APC data were retrieved from https://www.springernature.com/gp/open-research/journals-books/journals on March 4, 2021. The journal impact data including AIS for 2019 were retrieved from JCR 2019 on March 4, 2021.

above. This model would be straightforward to execute, requiring changes to internal NIH policy but not approval by the U.S. Congress.

Disadvantages: Can this model be scaled up to accommodate the millions of peer-reviewed biomedical research articles published annually [33]? If scaled up, will it limit the publishing options for those researchers who are not funded by these funding agencies? If a funding agency pays for research and supervises publication of the same work, then there is the potential for conflicts of interest to arise. It would be important to keep journals scientifically and editorially independent from funding agencies. Note that Environmental Health Perspectives demonstrates that it is possible for the U.S. government to publish a peer-reviewed scientific journal that is editorially and scientifically independent from the National Institutes of Health.

# Policy Type 2: Funding agencies mandate preprint publications

In this model, funding agencies would require grantees to post manuscripts as preprints on a server like *bioRxiv* or *medRxiv* [34]. Preprints are free to post and free to read but are not peer-reviewed. Following posting of a preprint, scientists are free to submit their manuscript for peer-review and publication in any journal, whether subscription or open access.

Advantages: This policy could be implemented rapidly, at little or no cost to funding agencies or research scientists. Results would be free to read immediately.

Disadvantages: The sudden and increased emphasis on preprints would require biomedical researchers to read and comment on preprints, a new ecosystem for biomedical research. It is an empirical question whether this preprint ecosystem will flourish for biomedical research as it has for physics. While preprints are standard in physics, physicists still publish in peer-reviewed journals following preprint submission. Comparing publication approaches in physics and biomedical research is confounded by the differences in the size of each field. The physics preprint server arXiv receives on average 16,000 manuscripts per month [35]. In contrast, PubMed listed approximately 2.5 million peer-reviewed papers per month in 2020 (for a total of 30.4 million) [36], while the biomedical preprint server bioRxiv received about 3300 manuscripts per month (for a total of 40,022).

Additionally, this plan does not address article publication charges for open access journals. Once scientists deposit manuscripts on preprint servers, they may still publish manuscripts in peer-reviewed journals. The desire to publish in

high impact journals, and to pay the associated APCs, may not be diminished by the number of preprints. If scientists rely on preprint servers only, an effective peer-review mechanism of preprints will need to be adopted broadly and also allow simple differentiation between peer-reviewed and non peer-reviewed manuscripts. This differentiation is particularly important for biomedical research due to public health implications. It would be disastrous to confuse the public with potential misinformation that would have been corrected by peer-review.

# Policy Type 3: Funding agencies cap payments of APCs for grantees

In the current system, many grantees use funder's money to pay APCs. Funding agencies could set a standard or universal maximum APC in collaboration with publishers. If enough funding agencies and publishers work together to negotiate a maximum APC, then this negotiated APC could drive the adoption of a standard cost of publication in most journals. To achieve such collaborations, certain levels of transparencies in the true cost of publishing and willingness to build a sustainable ecosystem together are essential. There is precedent for funding agencies setting a universal cost standard on scientific goods and services. The NIH set a minimum salary for postdoctoral fellows funded by F32 grants [37]. This minimum is widely followed, even at universities in the U.S. where the head of the laboratory pays postdocs using non-NIH funds. Similarly, the NIH sets a maximum cap on annual salaries for principle investigators, the professors that lead research grants, which makes it expensive for universities to pay professors an annual salary higher than this cap [37]. Thus, the NIH maximum allowable APC could become a universal maximum APC.

An appropriate monetary cap for APCs is complicated to determine. Since 2014, the Austrian Science Fund (FWF) caps payments for its grantees at €2,500 per manuscript [38]. It is not known whether APCs have stayed the same or gone down as a result. Globally, open access costs increased since 2014 [3], suggesting that the FWF is too small to impact APCs. What factors should be used to determine an appropriate monetary cap is beyond the scope of this article and will take collaborations among the stakeholders to investigate.

Effective January 2021, the Wellcome Trust will "cover fair and reasonable APCs for articles published in fully OA journals" but does not list a specific dollar amount [39]. Will the Wellcome Trust allow grantees to pay \$2500 to publish in the *Journal of Biological Chemistry*, when it costs \$600 less to publish in *Scientific Reports*? Both journals have similar subject matter scopes, article influence scores and impact factors.

Converting an existing subscription-based or hybrid journal to a completely open access journal could be expensive [40]. Some argue that these costs should be subsidized by funding agencies or governments. Others argue that the market should be allowed to find a solution to a funding agency cap on APCs. In the market approach, for-profit publishers would implement a business model where they are profitable despite a cap on the APC. This approach discourages scientists from advising publishers how to run their business.

There is precedent for the market approach, where an open access journal found a way to remain viable despite a low APC. The *Journal of Infometrics* is a hybrid OA journal, founded in 2006, published by Elsevier. The editorial board was concerned about high APCs and unhappy with the lack of autonomy afforded by working with Elsevier [41]. In 2019, the editorial board of the *Journal of Infometrics* resigned and started a new journal, *Quantitative Science Studies*. The *QSS* editorial board negotiated with several publishers before choosing MIT Press. *QSS* has the same scope as the *Journal of Infometrics*, but the APC is \$800 compared to \$2000 [41]. If funding agencies capped APCs, existing journals might find a way to remain profitable, or new journals might arise, a la *QSS*, to meet the requirement for low APCs.

Advantages: Compliance would be close to 100%, as it would be difficult or impossible for scientists to spend funder's money contrary to the funder's wishes. Lower APCs would benefit all researchers, not just those directly funded by influential funding agencies. We note that a maximum APC could be determined by funding agencies and publishers working together. However, it is also possible that their conflicting interests could necessitate funding agencies acting unilaterally to set a maximum APC. Funding agencies have powerful leverage over their grantees.

Disadvantages: Funding agencies such as the NIH have extensive bureaucracies that will hamper implementation of an APC cap. Additionally, it will be difficult to determine an appropriate dollar amount for an APC cap. Should the NIH determine an APC cap based on what they are willing to pay, or based on publication costs? The latter requires transparency in the cost of publishing, which does not exist. In 2018, European research institutions formed Plan S, a plan to have all research funded by public grants published in open access journals. Plan S attempted to set APC caps but received criticism on how the cap should be negotiated and maintained. Publishers argued in favor of the highest possible APC, while researchers and university administrators argued that publishers are powerful and will negotiate with funding agencies to increase the APC regularly, as occurs today with subscription costs [42]. Now, instead of an APC cap, Plan S advocates for a Price Transparency Framework where publishers provide voluntary data on how much it costs to publish articles [12]. If the publishing industry can be more transparent regarding the true cost of publishing, it could help funding agencies set a reasonable APC cap. On the other hand, Plan S may not have sufficient power to enforce publisher participation because there are no direct incentives for publishers to reduce the APCs. Funders may need to spend additional resources on enforcement and help publishers, especially smaller society publishers, with the transition

A funder's cap on APCs could restrict authors' intellectual freedom in where they can publish. For example, if Nature refuses to lower their APC, then funding agencies could prohibit their funded researchers from publishing in Nature. One can argue that since the funding agencies are paying for the research, they have a say in how and where the results are disseminated. On the other hand, this sets up a potential conflict of interest. For example, the NIH has restrictions on the use of human embryonic stem cells in research projects. It could be damaging for the scientific enterprise if these research restrictions extended to publication restrictions for stem cell research. As discussed in Policy Type 1, it would be important to keep journals scientifically and editorially independent from funding agencies.

Some have argued that the open market encourages scientists to publish in journals that deliver a higher quality of service, such as more rapid peer review. However, a recent study found no significant correlation between average review time and APC [22], suggesting that scientists are not paying higher APCs to receive better service.

# Policy Type 4: Funding agencies require both preprint posting and cap payments for APCs

Manuscripts would be immediately free to read as preprints. Once a preprint is published, scientists could submit the manuscript to any peer-reviewed, open access journal that charges an APC within the funder's limits. A possible outcome is that scientists publish preprints and then submit manuscripts to peer-reviewed subscription-based journals, as this is the cheapest path for authors (subscription-based journals charge low or no fees for authors). Therefore, for this policy to be effective, funding agencies would also need to prohibit non-OA publication and cap APCs.

Advantages: Research results would be free to read immediately, due to preprint requirement. During the conventional, peer-reviewed publication process, funder's money is protected and spent efficiently. Biomedical scientists might rely more heavily on discussing findings as preprints (similar to physics) rather than always waiting for peer review, which could take months.

Disadvantages: Requires substantial changes to funder policy and research community culture, making implementation difficult. As discussed in policy type 2, the emphasis on preprints would require biomedical researchers to read and comment on preprints, a new ecosystem for biomedical research. It's an empirical question whether this preprint ecosystem will flourish for biomedical research.

# Policy Type 5: Universities negotiate payments to journals so faculty can publish and read for free

Subscription-based journals charge a fee to read but low or no additional publication fees. Universities currently pay

journal subscription fees on behalf of their research scientists. Universities could also pay APC fees on behalf of scientists. As the number of subscription-based journals decline and OA journals become more prevalent, some of a university's budget for subscription fees could transform into covering open access publishing fees. Universities could negotiate APCs with publishers. Universities could apply to funding agencies for money to pay publication charges on behalf of entire faculty [43]. The University of California system and PLOS negotiated an agreement whereby the university will pay some or all of the APC on behalf of its faculty if they publish in a PLOS journal [44]. The University of California more recently enacted a similar but complex deal with Elsevier, where the university will pay \$1000 towards grant-funded authors' APC (authors are expected to use research funds to cover the balance) [45]. The University of Cambridge signed agreements with many publishers to enable researchers at the university to publish their primary research and review articles in open access journals. If the manuscript's corresponding author is affiliated with the University of Cambridge, then authors will not be charged publication fees. This agreement covers more than 5000 journals published by Springer, Wiley, PLOS and the Company of Biologists, among other publishers [27].

Advantages: An inexpensive business model for biomedical researchers because research scientists would pay reduced or no publication fees. The approach works with the existing publication and biomedical journal ecosystem and few changes are required for implementation.

Disadvantages: This policy will be difficult to scale and could lead to different APCs for different universities. Initially, this policy might be feasible for large universities such as Cambridge, or multiple large universities that operate under a single board of directors, such as the University of California system, but smaller universities and colleges would either be left to fend for themselves and negotiate at a disadvantage compared to large universities, or they would need to form coalitions with other institutions, which would take time. Additionally, negotiations between universities and publishers could favor large and established publishers, who control a larger share of the market. A large publisher, like Elsevier, publishes hundreds of journals and would have substantial leverage in a negotiation to determine APC. In contrast, a smaller publisher, such as a scientific society that publishes less than 10 journals, would have less leverage in a negotiation. This could create publishing disparities where scientists at some universities are unable to publish in some journals. To explore new pricing models, Massachusetts Institute of Technology is negotiating with publishers to have price based on value-added services provided by the publisher, and not based on the number of articles published by affiliated authors multiplied by the APCs [46].

# Conclusions

Long term, the least expensive solution for biomedical researchers is to have funding agencies pay the costs of OA publication by publishing their own journals and/or by directly funding journals that are published by a third party. This way, funding agencies could use their power to negotiate lower APCs, while scientists can publish wherever they like for free or at a lower cost if they are not directly funded for paying APCs. Biomedical researchers, funding agencies and publishers could work together to ensure that open access journals are reputable and not predatory. Predatory journals are those self-serving publications that accept publication fees and disseminate manuscripts without any quality check [47]. In this scenario, funding agencies would pay the cost of publication and create a list of prohibited, predatory journals, so the incentive to publish in predatory for-profit journals would plummet.

One concern is that by lowering APCs, journals will decline in quality. Less revenue or lower profit margin for publishers could increase the number of non-rigorous journals, in which the quality of peer review is low and/or the ability to detect fraud, such as image manipulations, is poor. The existance of prestigious and influential journals that charge relatively low or no APCs argues against this outcome. Journals such as *eLife, Environmental Health Perspectives* and *PLOS Biology*, whose APCs are thousands of dollars less than similarly influential journals published by Elsevier and SpringerNature, suggests that journals can charge lower APCs without sacrificing quality.

In the short term, a graded implementation might cushion the blow of a rapid transition. One idea is to start by having funding agencies mandate preprint publications, and subsequently have funding agencies cap APCs. Over time, this would lead to agencies directly funding journals that are free to read and free to publish. If the transition takes several years, this will give universities, funding agencies, scientists and publishers time to prepare and adjust.

Moving from a subscription-based model to a new model where all journals are OA raises difficult questions. How will universities access articles previously published in subscription-based journals? Many of these papers were published before funder requirements that manuscripts be free to read one year following publication. Will universities continue to pay access fees to archived papers in journals that no longer publish using a subscription model? Some existing license agreements may include terms for perpetual access to subscribed content but many do not. If implementing some of the options above, funders and universities may need to use their leverage with publishers to ensure that older publications remain accessible.

What happens to professional journal editors, those scientists that work for publishers and edit manuscripts full time? In the face of reduced APC and reduced revenues, journals will need to decide whether the cost of paying professional editors is worthwhile. In the current publishing ecosystem, most journals do not employ professional editors. The fact that so many journals eschew professional editors suggests that professional editors are not required to publish a successful, influential journal. Currently, at most biomedical

journals peer-review and editing are performed by scientists for free. This need not change if any of the policies above are implemented. If the quality of peer-review remains high and scientists continue to perform peer review for free, then publishers may find that they can reduce costs by paying editors less (or not paying editors at all).

Finally, these policies will have substantial impact on the publishing industry over all. If funders cap APCs or publish OA journals themselves, profit margins for many publishers will likely decrease. Some journals and/or publishers could go out of business. Other publishers might find a way to reduce costs, make publishing more efficient and remain profitable. Funders could consider providing additional support on a case-by-case basis for vulnerable publishers, such as small non-profit scientific societies, during the transition. Scientists have been contributing their expertise as authors, editors, and peer reviewers, mostly for free, to the publishing ecosystem for more than a century. Publishers, funders and universities should make that ecosystem more equitable and sustainable. Journals like Environmental Health Perspectives and Wellcome Open Research demonstrate that peer-reviewed results can be published at reasonable cost to scientists and funding agencies. Scientists and funding agencies should ask how sustainable it is in the long-term to spend the extra tens of thousands of dollars per article to publish in highly prestigious journals like Nature. Communicating results as inexpensively as possible, while still maintaining rigorous criticism and discussion, will benefit scientists, universities, funding agencies and ultimately every person whose life is improved by discoveries made in biomedical research labs.

#### **Acknowledgements**

We thank Drs. Prachee Avasthi, Peter Rogers and Richard Sever for helpful discussions during the preparation of this manuscript, and Drs. Rachel Arey, Courtney Hodges and Margot Williams for feedback on initial drafts of the manuscript. The idea for this manuscript arose following a spontaneous discussion between D.A.G and Dr. Margaret A Goodell at a department holiday party in 2019. We thank Dr. Goodell for her insights on publication practices at scientific societies. We thank Laura Hanscom for her feedback and insights from the academic library perspective.

# Citation

Gorelick, D. A. & Li, Y. Reducing open access publication costs for biomedical researchers in the U.S.A. *MIT Science Policy Review* **2**, 90-98 (2021). https://doi.org/10.38105/spr.4nulqfjf3t.

# **Open Access**



This MIT Science Policy Review article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to

the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

#### References

- [1] Suber, P. Open Access. The MIT Press Essential Knowledge series (MIT Press, Cambridge, MA, USA, 2012). Online: https://openaccesseks.mitpress.mit.edu/.
- [2] Monitoring the transition to open access: December 2017. Tech. Rep., Universities UK (2017). https: //www.universitiesuk.ac.uk/policy-and-analysis/ reports/Pages/monitoring-transition-openaccess-2017.aspx. Accessed: July, 2021.
- [3] Morrison, H., Borges, L., Zhao, X., Kakou, T. L. & Shanbhoug, A. N. Open access article processing charges 2011 -2021. Preprint. (2021). Online: http://ruor.uottawa.ca/ handle/10393/42327.
- [4] Journal Impact Factor Journal Citation Reports (2021). Online: https://clarivate.com/webofsciencegroup/solutions/journal-citation-reports/. Clarivate, Web of Science Group, Accessed: July, 2021.
- [5] Directory of Open Access Journals (2021). Online: https://doaj.org/. Directory of Open Access Journals, Accessed: July, 2021.
- [6] Federal Funds for Research and Development: Fiscal Years 2019–20. NSF 21-329, National Center for Science and Engineering Statistics (NCSES), National Science Foundation, Alexandria, VA (2021). Online: https://ncses.nsf.gov/ pubs/nsf21329/. Accessed: July, 2021.
- [7] Jackson, S. Free access to scientific publications: contrasting the JCI approach to Plan S. The Journal of Clinical Investigation 129, 440–441 (2020). https://doi.org/10.1172/JCI126932.
- [8] PMC Overview (2021). Online: https://www.ncbi.nlm. nih.gov/pmc/about/intro/. US National Library of Medicine, National Institutes of Health, Accessed: July, 2021.
- [9] NIH Public Access Policy Details (2009). Online: https://publicaccess.nih.gov/policy.htm. National Institute of Health, Accessed: February, 2021.
- [10] The Grand Compromise of U.S. Public Access Programs: Going Green (2016). Online: https://www.newswise.com/articles/the-grand-compromise-of-u-s-public-access-programs-going-green. DOE Science News Source, Newswise, Accessed: July, 2021.
- [11] About | Plan S and cOAlition S (2021). Online: https: //www.coalition-s.org/about/. cOAliationS, Accessed: July, 2021.
- [12] Price and Service Transparency Frameworks (2021). Online: https://www.coalition-s.org/price-and-service-transparency-frameworks/. Plan S: Making full & immediate Open Access a reality, Accessed: April, 2021.
- [13] An Interview with OSTP Director Kelvin Droegemeier (2019). Online: https://www.aip.org/fyi/2019/interview-ostp-director-kelvin-droegemeier. American Institute of Physics, Accessed: July, 2021.
- [14] Tananbaum, P. G. North American Campus-Based Open Access Funds: A Five-Year Progress Report. Tech. Rep., SPARC, the Scholarly Publishing and Academic Resources Coalition (2014). https://sparcopen.org/wp-content/uploads/ 2016/01/OA-Fund-5-Year-Review.pdf.

- [15] Sherpa Juliet Browse Funders (2021). Online: https://v2.sherpa.ac.uk/view/funder\_list/1.html. JISC Digital Resources, Open Access, Accessed: April, 2021.
- [16] Keeley, J. & Rosen, M. HHMI Announces Open Access Publishing Policy (2020). Online: https://www.hhmi.org/ news/hhmi-announces-open-access-publishingpolicy. Howard Hughes Medical Institute, Accessed: April, 2021.
- [17] Organisations endorsing Plan S and working jointly on its implementation (2021). Online: https://www.coalitions.org/organisations/. Plan S: Making full & immediate Open Access a reality, Accessed: April, 2021.
- [18] PMC Journals (2021). Online: https://www.ncbi.nlm.nih. gov/pmc/journals/. PubMed Central, US National Library of Medicine, National Institutes of Health, Accessed: April, 2021.
- [19] West, J. D., Bergstrom, T. & Bergstrom, C. T. Cost Effectiveness of Open Access Publications. *Economic Inquiry* 52, 1315–1321 (2014). https://doi.org/10.1111/ecin.12117.
- [20] Bergstrom, C. T., West, J. D. & Wiseman, M. A. The Eigenfactor™ Metrics. *Journal of Neuroscience* 28, 11433–11434 (2008). https://doi.org/10.1523/JNEUROSCI.0003-08.2008.
- [21] Sakellaropoulou, R. Introducing open access options for Nature and the Nature research journals (2021). Online: https://www.springernature.com/gp/researchers/the-source/blog/blogposts-open-research/open-access-options-for-nature-and-the-nature-research-journals/18782320. Springer Nature, For Researchers, Accessed: April, 2021.
- [22] Gray, R. J. Sorry, we're open: Golden open-access and inequality in non-human biological sciences. *Scientometrics* 124, 1663–1675 (2020). https://doi.org/10.1007/s11192-020-03540-3.
- [23] Walters, W. H. Information Sources and Indicators for the Assessment of Journal Reputation and Impact. *The Reference Librarian* 57, 13–22 (2016). https://doi.org/10.1080/02763877.2015.1088426.
- [24] Research4Life. Open access publishing: options for reduced or waived publication charges. Online: https://extranet.who.int/hinari/en/oaapc\_info.php.
  World Health Organization, Accessed: July, 2021.
- [25] Powell, Andrea. Guest Post APC Waiver Policies; A Job Half-done? (2021). Online: https://scholarlykitchen. sspnet.org/2021/04/19/guest-post-apc-waiverpolicies-a-job-half-done/. The Scholarly Kitchen, Accessed: July, 2021.
- [26] Smith, E. et al. Knowledge sharing in global health research the impact, uptake and cost of open access to scholarly literature. Health Research Policy and Systems 15, 73 (2017). https://doi.org/10.1186/s12961-017-0235-3.
- [27] Open access agreements (2021). Online: https: //www.openaccess.cam.ac.uk/publishing-openaccess/open-access-agreements. Publishing Open Access, University of Cambridge, Accessed: April, 2021.
- [28] Open access policies at MIT (2021). Online: https: //libraries.mit.edu/scholarly/mit-open-access/. Scholarly Publishing, MIT Libraries, Accessed: April, 2021.
- [29] Open Access | Environmental Health Perspectives (2021). Online: https://ehp.niehs.nih.gov/about-ehp/open-access. Environmental Health Perspectives, Accessed: April, 2021.
- [30] Olden, K. EHP Moves to Open Access. Environmental Health Perspectives 112, A13-A14 (2004). https://doi.org/10. 1289/ehp.112-a13.
- [31] About Wellcome Open Research | How It Works | Beyond A Research Journal (2021). Online: https://wellcomeopenresearch.org/about. Wellcome Open Research, Accessed: July, 2021.
- [32] About Gates Open Research | How It Works | Beyond A Research Journal (2021). Online: https:

- //gatesopenresearch.org/about. Gates Open Research, Accessed: July, 2021.
- [33] MEDLINE PubMed Production Statistics (2020). Online: https://www.nlm.nih.gov/bsd/medline\_pubmed\_ production\_stats.html. U.S. National Library of Medicine, Accessed: July, 2021.
- [34] Sever, R., Eisen, M. & Inglis, J. Plan U: Universal access to scientific and medical research via funder preprint mandates. *PLoS Biology* 17, e3000273 (2019). https://doi.org/10. 1371/journal.pbio.3000273.
- [35] arXiv Monthly Submissions (2021). Online: https://arxiv.org/stats/monthly\_submissions. arXiv.org | stats | montly submission rates, Accessed: April, 2021.
- [36] 2020 MEDLINE/PubMed Baseline (2020). Online: https://www.nlm.nih.gov/bsd/licensee/2020\_stats/2020\_LO.html. National Library of Medicine, National Institutes of Health, Accessed: April, 2021.
- [37] Salary Cap, Stipends, & Training Funds (2021). Online: https://www.niaid.nih.gov/grants-contracts/ salary-cap-stipends. NIH: National Institute of Allergy and Infectious Diseases, Accessed: April, 2021.
- [38] New Policy for Open Access and Publication Costs (2014).
  Online: https://www.fwf.ac.at/en/news-and-media-relations/news/detail/nid/20141219-2097/.
  FWF Der Wissenschaftsfonds, Accessed: April, 2021.
- [39] Open Access Policy (2021). Online: https://wellcome.org/grant-funding/guidance/open-access-guidance/open-access-policy. Grant Funding | Wellcome, Accessed: April. 2021.
- [40] McNutt, M. Opinion: "Plan S" falls short for society publishers-and for the researchers they serve. Proceedings of the National Academy of Sciences of the United States of America 116, 2400–2403 (2019). https://doi.org/10. 1073/pnas.1900359116.
- [41] Waltman, L., Larivière, V., Milojević, S. & Sugimoto, C. R. Opening science: The rebirth of a scholarly journal. *Quantitative Science Studies* 1, 1–3 (2020). https://doi.org/10.1162/qss\_e\_00025.
- [42] Plan S[how me the money]: why academic-led initiatives represent a more equitable, less costly publishing future (2018). Online: https://blogs.lse.ac.uk/impactofsocialsciences/2018/10/24/plan-show-me-the-money-why-academic-led-initiatives-represent-a-more-equitable-less-costly-publishing-future/. Impact of Social Sciences, The London School of Economics and Polictal Science, Accessed: April, 2021.
- [43] R\(\textit{\textit{Mtingen}}\), J.-A. & Sweeney, D. Financing open-access publication after 2024. Nature 572, 586 (2019). https://doi.org/10.1038/d41586-019-02547-y.
- [44] PLOS and the University of California announce open access publishing agreement (2020). Online: https://theplosblog.plos.org/2020/02/plos-and-the-university-of-california-announce-open-access-publishing-agreement/. The Official PLOS Blog, Public Library of Science, Accessed: February, 2021.
- [45] Anderson, R. Six Questions (with Answers!) about UC's and Elsevier's New Transformative Deal (2021). Online: https://scholarlykitchen.sspnet.org/2021/03/ 25/six-questions-with-answers-about-ucs-andelseviers-new-transformative-deal/. The Scholarly Kitchen, Accessed: April, 2021.
- [46] Publishing, S. MIT Framework for Publisher Contracts. Online: https://libraries.mit.edu/scholarly/publishing/ framework/. MIT Libraries, Accessed: July, 2021.
- [47] Grudniewicz, A. et al. Predatory journals: no definition, no defence. Nature 576, 210–212 (2019). Bandiera\_abtest: a Cg\_type: Comment Number: 7786 Publisher: Nature Publishing Group Subject\_term: Research management, Publishing, https://doi.org/10.1038/d41586-019-03759-y.

**Submit date: 2/24/2023** 

I am responding to this RFI: On behalf of myself

Name: Peter L. Elkin, MD

Name of Organization: University at Buffalo

Type of Organization: University

Role: Scientific researcher

# 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Encourage authors to site Data sets that they utilized in their research in addition to articles.

In order to best understand datasets made available encourage researchers when possible to collaborate the data providers.

2. Steps for improving equity in access and accessibility of publications.

once accepted for publication encourage journals and authors to post the pre-print to a pre-print server.

3. Methods for monitoring evolving costs and impacts on affected communities.

Have a monitoring program that collects ongoing data for data driven decision making.

4. Early input on considerations to increase findability and transparency of research.

This is a good step forward. It falls short in a few areas.

- 1. By giving a HIPAA exception you guarantee poor compliance with the aim of this regulation. Instead what you should have done is to establish required methods for sharing of de-identified and separately for limited datasets with proper authentication and security and privacy. This would exponentially increase compliance with the intent of this policy.
- 2. You did not go far enough to specify interoperability. APIs for the most part provide only syntactic interoperability. There are considerable issues with administrative code sets. The federal government should require the use of ontology to move toward semantic interoperability. This would require SNOMED CT, LOINC and RxNorm be used to code clinical data and the sequence ontology and UniProt for molecular data. Additionally clinical data should be provided in one of the three most common observational database formats (OMOP, i2b2 or PCORNet).

Email: elkinp@buffalo.edu

**Submit date: 2/25/2023** 

- 1. How to best ensure equity in publication opportunities for NIH-supported investigators.
- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

Email: lancejr38@outlook.com

**Submit date: 2/26/2023** 

I am responding to this RFI: On behalf of myself

Name: Kenneth Pawlak

Type of Organization: Other

**Role:** Member of the public

1. How to best ensure equity in publication opportunities for NIH-supported investigators.

When will the public reading room be reopened?

- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

Email: k.pawlak@comcast.net

**Submit date: 2/26/2023** 

I am responding to this RFI: On behalf of myself

Name: Bobby Hollingsworth

Name of Organization: Harvard Medical School

Type of Organization: University

Role: Scientific researcher

# 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

I agree that authors should have some semblance of choice in publication medium; however, the existence of prestige signalling through journal title has allowed bloated for-profit journals to increase APCs at the expense of the American taxpayer. This must stop. NIH funds should cap the allowable APC payed through NIH grant funds and mandate non-profit publishing in order to maximize research dollars and create better incentives for publication, resource sharing, and hiring. In a separate request for comments, the NIH asks for opinions on the postdoc experience--funds payed through APCs could instead be payed out to researchers doing the work, encouraging innovation and talent recruitment. Additionally, current publishers are extremely biased through editors that reach out to particular labs to fast-track papers, particularly when there is perceived competition with Other journals. Create a watchdog system for monitoring and reporting such unethical behavior.

# 2. Steps for improving equity in access and accessibility of publications.

The NIH could help create searchable databases that ease data parsing. Take for example, this paper: PMC8380731. The screening data are deposited as raw counts data with barcodes, rather than processed gene names and statistics relevant to the figure itself. Clearly, follow-up hits are being obscured, which is unacceptable research practice when funds come from the taxpayer. In addition to data management plans, such case examples should be subject to reporting, and folks at the NIH should insure compliance with data deposition standards.

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

Mandate publishing in nonprofit groups only, particularly ones that register with the NIH. Encourage consistent research release through alternative platforms such as pubpub, and index these alternatives in pubmed.

4. Early input on considerations to increase findability and transparency of research.

**Submit date: 2/28/2023** 

I am responding to this RFI: On behalf of myself

Role: Scientific researcher

1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Because publication in high-quality open access journals is often associated with significant costs, the option to deposit author-accepted manuscripts rather than final published manuscripts in PMC is a critical one. The goal should be to reduce inequalities and barriers to researchers disseminating their work to scholarly communities while also making the results broadly accessible to the public. The plan as written seems workable in these respects.

- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

Submit date: 3/2/2023

I am responding to this RFI: On behalf of an organization

Name: Julia schaletzky

Name of Organization: Center for emerging and neglected diseases uc berkeley

**Type of Organization:** University

**Role:** Bioethicist

# 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Whatever is decided evaluate administrative impact - very few profs have admin support and it has become very hard to attract funding, publish, hire and motivate students and fulfill ever increasing compliance and training requirements. Nih systems are onerous already. Pushing the most disadvantaged investigators to spend hours on new policy compliance creates undue burden and inequities

#### 2. Steps for improving equity in access and accessibility of publications.

Most important for public is not access to raw data for basic research but clinical trial data ...this is required to be released but isn't and nih complains but does not enforce. This would be transformational for patients and should be prioritized

# 3. Methods for monitoring evolving costs and impacts on affected communities.

This is a waste of money since it's been well documented that fees are not justified and inherently inequitable. Instead of "documenting" NIH should change the situation and put pressure on publishing houses. This data exists already

#### 4. Early input on considerations to increase findability and transparency of research.

Clinical data needs to be released - enforcement needed

Email: jschaletzky@berkeley.edu

Submit date: 3/2/2023

I am responding to this RFI: On behalf of myself

Name: Sandra Poulson

Role: Scientific researcher

### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

It makes a great deal of sense to install more transparency at the level of review to ensure that competing interests (considering that reviewers often perform research themselves in adjacent areas of expertise) do not prevent the publication of publicly funded research. Although the NIH does not want to disrupt the broad discresion for authors to choose where to publish, it does make sense to have some sort of oversight board to give a rating to journals for how fair the reviews are, for how reasonable the demand for additional evidence is, how frequently feedback targets authors for being female or not from an elite University, and frankly how rude or helpful the feedback is. It might be interesting to anonymize the reviews and rate on a journal level to try to push for accountability at the journal level for managing its reviewer pool.

#### 2. Steps for improving equity in access and accessibility of publications.

I feel that effort to make older articles, e.g. 1950-1995, accessible is worthwhile. These articles may not currently be accessible online without access beyond a paywall, and making them accessible would greatly benefit the public, including students and educators. The older research has valuable information but is often inaccessible.

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

The fact that it costs more to publish with an open access option is ridiculous. It costs researchers more and therefore costs more grant money even though the research was already funded by the public. There should be no additional cost to publish open source, and the publishers should be thankful that the public funded the research that the publishing company now uses in its publications. No company should be charging more for researchers to publish open access. And institutions that house researchers that produce publicly funded research should not be charged exorbitant fees (library fees) to access publisher material. This type of fee increases overhead demanded by the institution on each researcher which takes away from funds that could go toward research. This issue is made more difficult in that the editors of the journals are professors who do not wish to no longer have income from the publisher, but this is public money and it should be going toward buying supplies and funding the postdoctoral and graduate labor that makes research happen.

# 4. Early input on considerations to increase findability and transparency of research.

After making older articles (e.g. 1960-1995) available, it would be useful to add keywords to help make the research searchable like today's articles are.

It would be a fantastic idea to add searchability of methods used, as opposed to just keyword topics. E.g. "RNAscope" as a searchable term if the article used the RNAscope kit. It would be useful to build a catalog for each articles that used mouse brain tissue, coronal slices, immunohistochemistry. It would be so amazing to search for techniques and find several papers that successfully used the method to

determine whether the method was worthwhile to spend the time and resources to do it. I feel this would eliminate wasted money. It would also be a useful tool to use to review the usefulness of a method.

Email: <a href="mailto:sandrapoulson@fastmail.com">sandrapoulson@fastmail.com</a>

**Submit date: 3/2/2023** 

I am responding to this RFI: On behalf of myself

Name: Adam Armstrong

Name of Organization: Saint-Gobain

Type of Organization: Biotech pharmaceutical company

Role: Scientific researcher

- 1. How to best ensure equity in publication opportunities for NIH-supported investigators.
- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

Transparency of research - communicating critical geometries associated with functions of research.

Lots of the great NIH research I have encountered is based on critical geometries being developed or improved upon to enable specific device functions of research. Having research aligned to a standard 3D file format (maybe stp) and 2D dimensional file format (maybe dxf) would make that communication easier. Further, having those files shared openly would vastly increase research transparency and speed further supporting developments.

Email: adam.armstrong@saint-gobain.com

Submit date: 3/2/2023

I am responding to this RFI: On behalf of myself

Name: Federico Leva

Name of Organization: Dissem.in/CAPSH

Type of Organization: Professional org association

Role: Member of the public

1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Expeditious archival on PMC should remain the primary avenue to ensure public access. NIH could invest more in enriching metadata with URLs declaring the copyright status of manuscripts, especially for works not deposited by the publishers.

- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.

To reduce the costs for researchers and research entities, while increasing compliance, NIH should further invest in reducing the friction involved in getting works deposited in PMC. There's a need for increased participation in preprint deposit and more support for researchers whose employer doesn't have an established institutional repository or open access office to help.

In particular, NIH could expand its cooperation with trusted LinkOut repositories so that they can proactively archive works which are subject to the policy, and if necessary contact authors to facilitate any further required step. Some multidisciplinary repositories may also be able to provide such a service for Other agencies affected by the new OSTP policy, thereby increasing economies of scale and decreasing costs for authors and NIH. NIH could for example contract a repository to develop software and processes to actively solicit, or collect from openly available collections on the web, manuscripts to add to the repository's collection; the NIH could share metadata about grants and authors, as well as HR information for authors of works which may be considered US government works; NIH could further provide a copyright license to the repository, authorising it to host and distribute works which the US federal government has the right to use. This would allow the repository or repositories to preserve millions of works which are already known to be in the public domain, openly licensed or Otherwise licensed to the US federal government, but remain paywalled.

4. Early input on considerations to increase findability and transparency of research.

Email: federicoleva@tiscali.it

**Submit date:** 3/3/2023

I am responding to this RFI: On behalf of myself

Name: Catherine Christian-Hinman

Name of Organization: University of Illinois Urbana-Champaign

Type of Organization: University

Role: Scientific researcher

1. How to best ensure equity in publication opportunities for NIH-supported investigators.

2. Steps for improving equity in access and accessibility of publications.

3. Methods for monitoring evolving costs and impacts on affected communities.

It would be more equitable for all NIH-funded investigators if open access publication costs for research on NIH grants were directly paid by the NIH, rather than as direct costs on the grants. The grant budgets are already spread very thin (especially for modular R01s or smaller grants), and the costs of OA publishing, now often running into multiple thousands of dollars, can be the equivalent to a month or more of a trainee stipend. In addition, if work is funded by an NIH grant but published after the grant funding period is completed, the grant funds are no longer available to cover these costs.

Create a portal where PIs can input information on the journal, publisher, manuscript, and associated grant(s), and instruct publishers that if a manuscript is citing an NIH grant as support, NIH will pay for the OA fees. This will also further incentivize submission to OA journals, as they will not require an outlay of increased costs on the part of researchers.

4. Early input on considerations to increase findability and transparency of research.

**Submit date:** 3/8/2023

I am responding to this RFI: On behalf of myself

**Type of Organization:** University

Role: Scientific researcher

# 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

The current proposal is resulting in the publishing landscape radically changing. Just this week two of the scientific journals I have traditionally published in have gone to open-access. While they advertise this in the way of 'diversity' and 'equity' it is anything but. Now, those without sufficient funding, will be unable to publish in high quality journals.

In essence, work not supported by the federal government is becoming unpublishable due to these strategies. The federal government should either prevent the use of funding for publications in order to drive down he cost of open-access publishing or consider setting up additional internal publishing opportunities.

As it stands, the federal government, and NIH, are now fully subsidizing the scientific publishing industry.

#### 2. Steps for improving equity in access and accessibility of publications.

NIH should limit the amount of funding they will allow to be used per publication. If they do not, journals will continue to charge fees that are not in line with the reality of the marketplace and scientific advances will be slowed.

I strongly recommend the NIH to develop a journal associated with each of its centers.

# 3. Methods for monitoring evolving costs and impacts on affected communities.

The NIH should absolutely do this. They should also monitor the new number of journals developed from publishers and the number of journals that go open access as a result of this.

# 4. Early input on considerations to increase findability and transparency of research.

The NIH should adopt the ORCID as the standard method to report on scholarly activity. It should replace the existing structure and should be used on biosketches.

**Submit date: 3/10/2023** 

I am responding to this RFI: On behalf of myself

Name: Sonal Sathe

Type of Organization: University

1. How to best ensure equity in publication opportunities for NIH-supported investigators.

2. Steps for improving equity in access and accessibility of publications.

Equity in access and accessibility of publications certainly should include the human and machine-readable forms of journal articles. As a person and academic in training with a visual impairment, though, I must say that not all figures and text are easily readable or accessible with use of a screen reader or machines. Nor are searches in NCBI or PubMed easily understandable with speech-to-text functions. Perhaps a priority for the NIH should include an accessibility audit with, and by, disability allies in order to make the rich body of literature available for all to (literally or figuratively) see.

I also strongly feel PI's training PhD students should strive to be inclusive of visual impairments beyond just the letter on an accommodation form. It is one thing to abide by the law (and if you do not, you do not deserve to be a mentor at all.) It is anOther to be truly supportive.

- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

**Submit date:** 3/15/2023

I am responding to this RFI: On behalf of myself

Type of Organization: Not applicable

1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Researchers may appreciate a description of how NIH plans to monitor compliance of DMS Plans.

- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

**Submit date:** 3/19/2023

I am responding to this RFI: On behalf of myself

Name: Tiffany Atkins

Name of Organization: Alphastar

Type of Organization: Other

Type of Organization-Other: Working on building business now

Role: Scientific researcher

1. How to best ensure equity in publication opportunities for NIH-supported investigators.

I would like to be published in public access if it was my work or a part of a team work

2. Steps for improving equity in access and accessibility of publications.

Making sure researchers are accurate and sources as well as letting people involved be involved with publication

3. Methods for monitoring evolving costs and impacts on affected communities.

I think so far NIH has done good in being reasonable

4. Early input on considerations to increase findability and transparency of research.

**Submit date:** 3/22/2023

I am responding to this RFI: On behalf of myself

Name: Damien Camany

Name of Organization: Self

Type of Organization: Other

Type of Organization-Other: Self motivation

Role: Member of the public

- 1. How to best ensure equity in publication opportunities for NIH-supported investigators.
- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

Email: d.camany@yahoo.com

**Submit date: 3/23/2023** 

I am responding to this RFI: On behalf of myself

Name: Peggy Lentz

Name of Organization: Henry Ford Health System

Type of Organization: Nonprofit research organization

Role: Member of the public

1. How to best ensure equity in publication opportunities for NIH-supported investigators.

All Journals/publishers should utilize Method A only for depositing the manuscripts for Public Access Compliance. Having Methods A-D is confusing to PI's - all should be done by the Journal/Publisher.

2. Steps for improving equity in access and accessibility of publications.

I think the publishers are greedy - the 12 month embargo goes against making valuable information available to the public. Do away with the embargo period.

3. Methods for monitoring evolving costs and impacts on affected communities.

The fees are exorbitant - \$3,000 to publish is too much...

4. Early input on considerations to increase findability and transparency of research.

**Submit date: 3/23/2023** 

I am responding to this RFI: On behalf of myself

Name: Rebecca Braddock

Type of Organization: Not applicable

- 1. How to best ensure equity in publication opportunities for NIH-supported investigators.
- 2. Steps for improving equity in access and accessibility of publications.

My daughter died of a very rare cancer with an average 11-month survival period after diagnosis. A 12-month embargo on publication of research findings makes the information unreachable and unusable for the patients and caregivers.

I checked, and I could access the studies at the clinic or hospital, or at a University library. However, poor internet connection and time restraints made that impossible.

Since her death, I've been searching the internet and saving many medical journal articles pertaining in some way to her disease. The cause is unknown, and the treatment hasn't changed for 40 years.

She was 24 when diagnosed, and died at 26. Researchers need to be gathering all the information they can from patients and caregivers. If access is denied to very recent research, some of our hope for progress is lost. Clinical trials.gov OFTEN doesn't publish results, even though the trial may be finished. The whole logiam of communication needs to be removed.

3. Methods for monitoring evolving costs and impacts on affected communities.

I wish there were a way to look up data at NIH, NCI, and CCDI to see if information is being shared between researchers in the U.S. about my daughter's very rare extrapulmonary neuroendocrine carcinoma.

4. Early input on considerations to increase findability and transparency of research.

Email: rbraddock 55@yahoo.com

**Submit date: 3/23/2023** 

I am responding to this RFI: On behalf of myself

Name: Thurman McGarian

Name of Organization: Private individual

Type of Organization: Not applicable

Role: Member of the public

1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Access through the NIH website and associated publications offer free and extensive amounts of information. A short cut to available printed information would be welcome.

- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.

Registration or Other identification standards would ensure that equal access to appropriate amounts of data.

4. Early input on considerations to increase findability and transparency of research.

A less cumbersome index of information would shorten time spent navigating the website and getting information on the way to the consumer.

**Submit date: 3/25/2023** 

I am responding to this RFI: On behalf of myself

Name: Stephen J. Kron

Name of Organization: University of Chicago

Type of Organization: University

Role: Scientific researcher

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

I really don't understand how any of this addresses equity. Bias is baked into the system at lots of levels. Of course, some people think they are being treated unfairly and perhaps they are, but the current system is built around unfair advantages at every level. Clearly, there is some underlying message here, but if you are not from a favored institution, not doing favored style of work, don't up with the favored answer, etc. then you are going to be affected by bias at the level of funding, publication and so on. Is that going to be addressed?

#### 2. Steps for improving equity in access and accessibility of publications.

Sure, NIH paid for it, they get to decide when it becomes public.

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

NIH needs to shift publication costs to the institutions. They should not be allowable costs beyond a standard fee that the publisher may charge at their discretion. That fee could be \$2500, assuming several factors like having provided peer review with at least three reviewers, and Other services that would be valuable to NIH.

#### 4. Early input on considerations to increase findability and transparency of research.

Sure, go ahead. It seems unlikely that this will really help anything, but it might. The burden is not going to be that significant on researchers.

**Submit date: 3/28/2023** 

I am responding to this RFI: On behalf of myself

Name: Peggy Lentz

Name of Organization: Henry Ford Health System

Type of Organization: Health care delivery organization

**Role:** Member of the public

1. How to best ensure equity in publication opportunities for NIH-supported investigators.

All Journals and Publishers should accommodate Method A for NIH funded research publication. The process is very confusing.

2. Steps for improving equity in access and accessibility of publications.

All Journals/Publishers should take care of making the manuscript available i.e. Method A. The embargo period should be eliminated - they are being greedy.

3. Methods for monitoring evolving costs and impacts on affected communities.

Open access is to much i.e. \$3,000 is significant.

4. Early input on considerations to increase findability and transparency of research.

**Description:** public access

**Submit date: 3/30/2023** 

I am responding to this RFI: On behalf of myself

Name: Christopher Marcum

Type of Organization: Not applicable

Role: Member of the public

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

One of the easiest ways to support equity in scholarly publication opportunities is to expand the incentive and reward structure by giving credit (i.e. for grant consideration, hiring decisions, etc) to researchers for: participating in peer-review and editorial activities; depositing pre-prints in agency designated repositories; making source code open source and publicly accessible; supporting trainees; publishing datasets in agency designated repositories; and Other activities beyond just publishing.

#### 2. Steps for improving equity in access and accessibility of publications.

I strongly encourage NIH to direct their intramural and extramural funded researchers to use green open access model to submit their author-accepted manuscripts in PMC with no embargo or delay. No Other model is more equitable - its free and easy. I also strongly recommend that NIH clearly articulate this position to the scholarly publishing community

- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

I strongly recommend that NIH require all researchers, intramural and extramural, to acquire ORCIDs, that their home institutions acquire RORs, and that all publicly accessible research products associated with NIH support acquire DOIs. It would be beneficial for NIH to join the DOE-sponsored ORCID Consortium.

**Description:** Writing in my personal capacity as a scientist and advocate for open science.

**Submit date: 3/30/2023** 

I am responding to this RFI: On behalf of an organization

Name: Sonya Dumanis

Name of Organization: Aligning Science Across Parkinson's

Type of Organization: Other

Type of Organization-Other: Research Initiative

Role: Scientific researcher

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Encourage the use of preprints or the posting of author accepted manuscripts in public repositories to allow for early sharing of research. Research without paywalls is a necessary, rather than optional, step in research communication. In our experience, preprints are fast, equitable, and flexible and can be used to describe many types of research outputs and findings including data papers, null results, and incremental progress. Preprints can also be used as a training step to guide researchers on what else needs to be linked in later versions to ensure all outputs are deposited. The cost of open access fees from publishers can be prohibitive for researchers to share these articles Otherwise.

#### 2. Steps for improving equity in access and accessibility of publications.

Require open licensing to ensure ongoing global access to research and embrace CC-BY or an equivalent license as the minimal license required for all research outputs generated by its funding. Only through open licensing can research be truly reusable.

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

Survey costs of publishing and corresponding author demographics for those already utilizing the NIH Manuscript Submission System to deposit author-accepted manuscripts in PubMedCentral (PMC) versus those being deposited in PMC through journals directly during the same time period.

#### 4. Early input on considerations to increase findability and transparency of research.

Require the utilization of persistent identifiers such as ORCIDs for all researchers, DOIs for publications, grants, datasets, and Other research outputs, RRIDs for tangible resources, and RORs for institutions to increase reusability and findability of outputs.

Align on how credit for outputs is given such as utilization of the emerging CRediT taxonomy standards to acknowledge the emerging team science approach being utilized for studying complex conditions and changing the current incentive structure towards recognizing collaborations.

Coordinate with all the US Federal Agencies and the open research community to collectively establish best practices and standards so that open access and data sharing, for example, are tracked and reported on in a consistent manner. There are already community-developed best practices such as standards for data/software sharing and FAIR principles. If each funding body counts data sharing in a different way, compliance and meta analysis will be all but meaningless. The US Government could take

on a transformative role by developing a shared agenda in policy implementation, best practices, and by investing in a common set of standards, tools, and centralized support structures.

Coordinate with the global community when developing open science standards. Science is international. Many emerging best practices are coming from around the world and being reported on by UNESCO, the European Commission, and country-specific entities. Any shared agenda developed in the US would be incomplete without attempting to build convergence globally.

Invest in training support and education materials to ensure that the next generation of researchers are aware of best practices, improve the user experience of sharing outputs, and reduce the barriers to complying with emerging requirements.

#### **Uploaded File:**

2023-03-30-ASAP-Comments-to-NIH-RFI\_v2.docx

**Description:** We have uploaded a more in depth overview of our interest in these policies and point by point considerations to each of the sections of the RFI.

Email: sdumanis@parkinsonsroadmap.org



# ASAP's Response to NIH RFI on the Plan to Enhance Public Access to the Results of NIH-Supported Research

Aligning Science Across Parkinson's (ASAP) is a research funding initiative that aims to accelerate the pace of discovery and inform the path to a cure for Parkinson's disease through collaboration, research-enabling resources, and data sharing. It was intentionally created to be open by design - incorporating open science principles into grantee policies from the start. ASAP's policies require that grantees share all relevant datasets, code, resources, and protocols underlying a published manuscript and that the manuscript be shared first as a preprint at the time of, or before, journal submission as an open-access article. ASAP believes that open research is a necessary step towards collaborative research, which, in turn, accelerates outcomes.

ASAP is keenly interested in the rapid progress made at the federal level in the US to catalyze a move towards open research broadly. Alignment of policies, best practices, standards, and metrics is crucial at this stage to ensure that implementation improves equity and instills a new incentive structure that shifts the research culture away from being predominantly individualized and competitive. The current culture rewards publishing of findings in high-impact research journals as the primary means of communication, which leads to closed licensing and outputs, slow science, and a lack of equity in terms of access to research opportunities and funding. Moreover, <u>studies</u> suggest that add-on open access fees in high-impact journals impact researchers with limited resources for selecting this as an option.

Along with other members of the open research community, ASAP hopes that NIH will build upon the current practices and infrastructures, including sharing early discoveries through preprints, utilization of the persistent identifier networks, and leveraging the emerging compliance monitoring tool chain to improve discovery of outputs and assessment of open science compliance. Reinvention of these early best practices and tools would subvert the progress being made and reported on by pioneers in open science. ASAP has published a Blueprint on its own best practices with the goal of creating a shared set of standards, tools, and best practices.

Below are ASAP's thoughts related to the four sections of the NIH public access plan.

- 1. How to best ensure equity in publication opportunities for NIH-supported investigators.
  - Encourage the use of preprints or the posting of author accepted manuscripts in
    public repositories to allow for the early sharing of research. Research without
    paywalls is a necessary, rather than optional, step in research communication. In our
    experience, preprints are fast, equitable, and flexible and can be used to describe many
    types of research outputs and findings, including data papers, null results, and
    incremental progress. Preprints can also be used as a training step to guide researchers
    on what else needs to be linked in later versions to ensure all outputs are deposited. The
    cost of open access fees from publishers can be prohibitive for researchers to share
    these articles otherwise.

- 2. Steps for improving equity in access and accessibility of publications.
  - **Require open licensing** to ensure ongoing global access to research and embrace CC-BY or an equivalent license as the minimal license required for all research outputs generated by its funding. Only through open licensing can research be truly reusable.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
  - Survey costs of publishing and corresponding author demographics for those already utilizing the NIH Manuscript Submission System to deposit author-accepted manuscripts in PubMedCentral (PMC) versus those being deposited in PMC through journals directly during the same time period.
- 4. Early input on considerations to increase findability and transparency of research.
  - Require the utilization of persistent identifiers such as ORCIDs for all researchers, DOIs for publications, grants, datasets, and other research outputs, RRIDs for tangible resources, and RORs for institutions to increase reusability and findability of outputs
  - Align on how credit for outputs is given such as utilization of the emerging <u>CRediT taxonomy standards</u> to acknowledge the emerging team science approach being utilized for studying complex conditions and changing the current incentive structure towards recognizing collaborations.
  - Coordinate with all the US Federal Agencies and the open research community to collectively establish best practices and standards so that open access and data sharing, for example, are tracked and reported on in a consistent manner. There are already community-developed best practices such as standards for data/software sharing and FAIR principles. If each funding body counts data sharing in a different way, compliance and meta analysis will be all but meaningless. The US Government could take on a transformative role by developing a shared agenda in policy implementation and best practices and investing in a common set of standards, tools, and centralized support structures.
  - Coordinate with the global community when developing open science standards.
     Science is international. Many emerging best practices are coming from around the world and being reported on by UNESCO, the European Commission, and country-specific entities. Any shared agenda developed in the US would be incomplete without attempting to build convergence globally.
  - Invest in training support and education materials to ensure that the next generation
    of researchers are aware of best practices, improve the user experience of sharing
    outputs, and reduce the barriers to complying with emerging requirements.

We look forward to learning more about NIH's emerging policies to increase findability and transparency of research.



Submit date: 4/3/2023

I am responding to this RFI: On behalf of myself

Name: Ana Chicas-Mosier

**Type of Organization:** Not applicable

Role: Scientific researcher

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Publication is expensive and the price increases with prestige of the journal and to ensure open access. To respond to these sources of inequity, the NIH could require publications fees be included in budgetary requests or as readily available supplemental awards along with the proposed open-access requirements. The NIH has weight that can be used to push journal publishers to reduce the cost of publication so that PIs at smaller institutions or without a formal affiliation can still afford to publish in journals with the highest impact factors. By only forcing open-access requirements, the NIH does not adequately respond to the cost paid by scientists to publish their studies.

#### 2. Steps for improving equity in access and accessibility of publications.

The core of data in manuscripts is often presented in graphical formats. Graphs, even in machine-readable document formats, are not typically accessible to people with vision impairments. To increase equity on this front, the NIH can establish a single format for graph printing that can be deciphered by machine-readers, require detailed text descriptions of the graph, and push journal publishers to use this format.

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

The additional costs that are associated with impact factor disproportionately weed-out smaller institutions and those with smaller endowments. In monitoring these trends, the NIH could investigate the number of published articles from universities/PIs with >\$500 mill in assets vs <\$500 mill and <\$100 mill.

4. Early input on considerations to increase findability and transparency of research.

Submit date: 4/6/2023

I am responding to this RFI: On behalf of myself

Name: Andrea Bertke

Name of Organization: Virginia Tech

Type of Organization: University

Role: Scientific researcher

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

NIH should consider negotiating with publishers to reduce public access fees. \$8,000-13,000 for public access fees, in addition to regular publication fees, is excessive and limits the ability of researchers to choose where to publish. These fees are also wasteful and would be more appropriately used for the research, rather than the publisher for simply releasing an article to the public immediately upon publishing. Why is \$13,000 needed to release an article?

#### 2. Steps for improving equity in access and accessibility of publications.

Again, negotiate with or regulate the publishers that charge excessive fees for open access and those that charge excessive fees for public access, even if the open access fee is paid by the researcher.

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

I disagree with the statement "ensure that they REMAIN reasonable and equitable." Fees are not reasonable and equitable now, so they cannot REMAIN so. Many researchers would like to publish in higher impact journals but are unable to do so because of the excessive publishing and open access fees. MDPI has expanded exponentially because they charge lower fees and make research available quickly, even though MDPI was once considered a predatory and questionable publisher. Nature Neuroscience charges a \$13,000 open access fee in addition to publication fees, after a 6-month to 2-year review process, and then charges the public \$35-65 for access to an article. That is not reasonable nor equitable, for either the researchers or the public. These types of fees amplify the perception that higher impact journals are only for the prestigious researchers who have excess funding to pay these fees. Since the modular R01 budget has not increased with increasing costs of research, many researchers must make the choice of using that \$13,000 for research or for excessive open access fees.

#### 4. Early input on considerations to increase findability and transparency of research.

Use a single identifier, instead of multiple identifiers that must be used for different purposes.

**Submit date:** 4/6/2023

I am responding to this RFI: On behalf of an organization

Name: Kevin C. Kregel, PhD

Name of Organization: Federation of American Societies for Experimental Biology

Type of Organization: Professional orgassociation

Role: Institutional official

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Researchers interpret NIH's response to the August 2022 Office of Science and Technology Policy Memo as a preference for charging reasonable publishing costs to the direct portion of grants. This misunderstanding could lead grantees down a path that increases overall costs to NIH and slows progress of research activities. Item III.D.1 notes "NIH intends to develop supplemental information that elaborates on and clarifies allowable costs for publication, consistent with these conditions." We encourage that such supplemental information covers all allowable paths for charging publishing costs, including from indirect costs and Other University general or restricted funds.

Reasons many researchers do not want to add publishing costs to the direct portion of their grants include:

- On the University side, publications are primarily supported through the library purchasing subscriptions, and increasingly open access. University general or restricted funds are the source of the support for library purchases, and while varying, include diverse streams: federal and private research and development grants (indirect costs), but also federal and state library funds, state and local tax allocations, direct fundraising by librarians, tuition and fees, and endowments. Without a new injection of funds into the direct portion of grants, or a commitment to move existing general funds now supporting publications to directs, an added cost to the direct portion of grants would result in reduced funds to support postdocs, graduate students, and research support staff; to purchase equipment and supplies; to support travel to conferences and Other career development opportunities.
- The additional administrative burden would further distract researchers from research activities. In today's approach, a team of societies, publishers, librarians, and institutional grant managers work together to arrange payment, support compliance, ensure proper metadata, and deposition to PubMed Central. If the costs are added to the direct portion of grants, the researcher is likely to be expected to handle some of these activities, such as the payment of fees, or deposition of manuscripts, processes which take time.

Researchers at larger institutions are better positioned to adapt, with libraries already implementing new arrangements (e.g., transformative agreements, subscribe to open) that do not impact the direct portion of the grant. Therefore, FASEB encourages NIH to allow flexibility for institutions to use indirect funds for a variety of publishing models, and to encourage institutions to continue to use the diverse revenue streams beyond indirect costs available in the general and restricted funds to support the costs of publication and make the transition to the realities of the new policy easier and more achievable for researchers.

Researchers from underserved populations, including early career researchers, those from historically excluded backgrounds, and those at less research-intensive institutions, do not have assured access to the aforementioned arrangements. Likewise, some societies are too small to handle detailed negotiations to make such accommodations. These disparities are already a reality. Many societies provide waivers, which the author may find an inconvenience and a barrier, with potential required actions such as requesting a waiver, and submitting a manuscript without assurance that a waiver will be provided until the manuscript is accepted. Waivers are provided at societies' expense, and we recognize this as a stop-gap solution that does not fully support equity. NIH could alleviate these issues by dedicating publishing resources for underserved researchers and societies and by providing guidance to program officers on addressing equity in publication opportunities.

#### 2. Steps for improving equity in access and accessibility of publications.

By virtue of their broad membership and core missions, scholarly societies are well-positioned to improve equity in access to publications for many stakeholders. However, financial support for these effortsis lacking. With proper funding, scholarly societies would be ideal partners to improve equity in access and accessibility. Examples of practical steps that could be taken more broadly, and are being experimented with at societies, include plain language summaries, alt text for images, creating more videos, working with media on news stories, and engaging through social media. Societies are also well-situated to develop educational materials and facilitate training to support researchersand the broader diverse community on improving communication around the scientific process and a specific field of science. To facilitate this, resources from NIH could be specifically allocated to address the financial need for domain-specific experts, including scholarly societies.

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

We were unable to identify a comparable approach taken by NIH to monitor fees for Other research services or outputs. FASEB recommends that NIH not monitorpublication fees lest the impact result in a system that favors quantity over quality. Any reference to a specific cost or price could have the unintended consequence of driving the system towards a 'one size fits all' pricing structure that negatively affects quality of resulting publications. The building blocks of scientific integrity - best practices and standards, ethical behavior, and the principles of honesty and objectivity - can lead to improved rigor, and reproducibility and must be included with public access approaches. While peer reviewers are not paid, peer review is far from free of cost. Upholding scientific integrity during peer review and publication increases costs through additional human time and adoption of innovative technologies. FASEB appreciates NIH's recognition of the value of peer-reviewed research publications and the services provided by scholarly societies to further scientific understanding and improve human health. Therefore, we encourage continued valuation of scientific integrity.

While there might be an interest in monitoring whether funded researchers are requesting more total resources in the direct versus indirect portion of the grant and resultant changes in awarded amounts over time, this would be challenging to monitor without an effective baseline. Similarly, there are many variables that must be considered; a few examples include the growing costs of ensuring quality against papermills or image manipulation, the number of articles published (output) which may grow if public

access achieves the goals of open science and drives rapid advances in science, and the changing demographics or preferences for services provided by different societies.

Monitoring equity in funded grants will be important, as is understanding where and how the system is developing and evolving. To obtain a snapshot of the current environment and assess impact of policy changes, NIH could compare the total, median, and mean number of publication fees in the direct portion of grants for different stakeholder groups over time and as a percentage of total published articles funded by the agency.

#### 4. Early input on considerations to increase findability and transparency of research.

FASEB isencouraged by NIH's commitment to engage withexisting identifier infrastructure and standards already in use across many scholarly societies. Requiring ORCID (Open Researcher and Contributor ID) for the corresponding and/or submitting author has been seamless for integration into societies' manuscript submission, peer review, and publication systems; requiring ORCID for all co-authors has posed more challenging but is improving with time. FASEB supportsNIH adoption of a DOI (Digital Object Identifier) overlay on existing grants; this activity could foster a more connected ecosystem of grants, publications, and data.

#### **Uploaded File:**

FASEB\_NIH\_RFI\_Final\_Letter\_Submitted\_04-06-2023.pdf

**Description:** The PDF attached is a full letter from the FASEB President that includes answers to the above four specific questions, as well as additional relevant information and some clarification sought from NIH with regards to manuscript types and guidance planned

Email: <a href="mailto:dhenderson@faseb.org">dhenderson@faseb.org</a>



### Representing Over 115,000 Researchers

6120 Executive Blvd., Suite 230, Rockville, MD 20852 | faseb.org

April 6, 2023

Lyric Jorgenson, PhD Acting Director, Office of Science Policy and Acting NIH Associate Director for Science Policy The National Institutes of Health 6705 Rockledge Drive, Suite 630 Bethesda, MD 20892

RE: FASEB Comments in Response to NOT-OD-23-091, Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research

Comments transmitted electronically via RFI Web form on April 6, 2023

Dear Dr. Jorgenson,

The Federation of American Societies for Experimental Biology (FASEB) appreciates the opportunity to provide comments in response to NOT-OD-23-091, Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research. FASEB is comprised of 27 scientific societies, collectively representing over 115,000 biological and biomedical researchers. As nonprofit scholarly scientific societies, we have missions that are well-aligned with the NIH mission.

Scholarly scientific societies were founded to convene researchers in a field and advance a particular branch of science. FASEB and our member societies have long accomplished this goal through various means, including establishing best practices and standards, policy feedback, workforce and career development, awards and recognition, advocacy, education, and communicating advances in science through publications, conferences, and other means. Over the last decade specifically, we have committed to improving diversity, equity, accessibility, and inclusion in the sciences, and are implementing major investments and activities to drive this change. We are led by and represent many of the same scientists who conduct research funded by NIH. As nonprofits, revenues we collect are reinvested in advancing science and supporting the research community.

FASEB recognizes the value of a refined framework to advance public access and the potential benefits of the taxpayers having access to trusted scientific information. We commend NIH's commitment to broad engagement and to iterative work on this plan. Our specific responses to the questions within the Request for Information are noted below.

How to best ensure equity in publication opportunities for NIH-supported investigators. NIH policy already allows supported researchers to charge reasonable publishing costs - NIH seeks information on additional steps it might consider taking to ensure that proposed changes to implementation of the

### NIH Public Access Policy do not create new inequities in publishing opportunities or reinforce existing ones.

Researchers interpret NIH's response to the <u>August 2022 Office of Science and Technology Policy Memo</u> as a *preference* for charging reasonable publishing costs to the direct portion of grants. This misunderstanding could lead grantees down a path that increases overall costs to NIH and slows progress of research activities. Item III.D.1 notes "NIH intends to develop supplemental information that elaborates on and clarifies allowable costs for publication, consistent with these conditions." We encourage that such supplemental information covers all allowable paths for charging publishing costs, including from indirect costs and other university general or restricted funds.

Reasons many researchers do not want to add publishing costs to the *direct* portion of their grants include:

- On the university side, publications are primarily supported through the library purchasing subscriptions, and increasingly open access. University general or restricted funds are the source of the support for library purchases, and while varying, include *diverse* streams: federal and private research and development grants (indirect costs), but also federal and state library funds, state and local tax allocations, direct fundraising by librarians, tuition and fees, and endowments. Without a new injection of funds into the direct portion of grants, or a commitment to move existing general funds now supporting publications to directs, an added cost to the direct portion of grants would result in reduced funds to support postdocs, graduate students, and research support staff; to purchase equipment and supplies; to support travel to conferences and other career development opportunities.
- The additional administrative burden would further distract researchers from research activities. In today's approach, a team of societies, publishers, librarians, and institutional grant managers work together to arrange payment, support compliance, ensure proper metadata, and deposition to PubMed Central. If the costs are added to the direct portion of grants, the researcher is likely to be expected to handle some of these activities, such as the payment of fees, or deposition of manuscripts, processes which take time.

Researchers at larger institutions are better positioned to adapt, with libraries already implementing new arrangements (e.g., transformative agreements, subscribe to open) that do not impact the direct portion of the grant. Therefore, FASEB encourages NIH to allow flexibility for institutions to use indirect funds for a variety of publishing models, and to encourage institutions to continue to use the diverse revenue streams beyond indirect costs available in the general and restricted funds to support the costs of publication and make the transition to the realities of the new policy easier and more achievable for researchers.

Researchers from underserved populations, including early career researchers, those from historically excluded backgrounds, and those at less research-intensive institutions, do not have assured access to the aforementioned arrangements. Likewise, some societies are too small to handle detailed negotiations to make such accommodations. These disparities are already a reality. Many societies provide waivers, which the author may find an inconvenience and a barrier, with potential required actions such as requesting a waiver, and submitting a manuscript without assurance that a waiver will be provided until the manuscript is accepted. Waivers are provided at societies' expense, and we recognize this as a stop-gap solution that does not fully support equity. NIH could alleviate these issues by dedicating publishing resources for underserved researchers and societies and by providing guidance to program officers on addressing equity in publication opportunities.

Steps for improving equity in access and accessibility of publications. NIH welcomes input on other steps that could be taken to improve equity in access to publications by diverse communities of users, including researchers, clinicians and public health officials, students and educators, and other members of the public.

By virtue of their broad membership and core missions, scholarly societies are well-positioned to improve equity in access to publications for many stakeholders. However, financial support for these efforts is lacking. With proper funding, scholarly societies would be ideal partners to improve equity in access and accessibility. Examples of practical steps that could be taken more broadly, and are being experimented with at societies, include plain language summaries, alt text for images, creating more videos, working with media on news stories, and engaging through social media. Societies are also well-situated to develop educational materials and facilitate training to support researchers and the broader diverse community on improving communication around the scientific process and a specific field of science. To facilitate this, resources from NIH could be specifically allocated to address the financial need for domain-specific experts, including scholarly societies.

Methods for monitoring evolving costs and impacts on affected communities. NIH seeks information on effective approaches for monitoring trends in publication fees and equity in publication opportunities.

We were unable to identify a comparable approach taken by NIH to monitor fees for other research services or outputs. FASEB recommends that NIH not monitor publication fees lest the impact result in a system that favors quantity over quality. Any reference to a specific cost or price could have the unintended consequence of driving the system towards a 'one size fits all' pricing structure that negatively affects quality of resulting publications. The building blocks of scientific integrity – best practices and standards, ethical behavior, and the principles of honesty and objectivity – can lead to improved rigor, and reproducibility and must be included with public access approaches. While peer reviewers are not paid, peer review is far from free of cost. Upholding scientific integrity during peer review and publication increases costs through additional human time and adoption of innovative technologies. FASEB appreciates NIH's recognition of the value of peer-reviewed research publications and the services provided by scholarly societies to further scientific understanding and improve human health. Therefore, we encourage continued valuation of scientific integrity.

While there might be an interest in monitoring whether funded researchers are requesting more total resources in the direct versus indirect portion of the grant and resultant changes in awarded amounts over time, this would be challenging to monitor without an effective baseline. Similarly, there are many variables that must be considered; a few examples include the growing costs of ensuring quality against papermills or image manipulation, the number of articles published (output) which may grow if public access achieves the goals of open science and drives rapid advances in science, and the changing demographics or preferences for services provided by different societies.

Monitoring equity in funded grants will be important, as is understanding where and how the system is developing and evolving. To obtain a snapshot of the current environment and assess impact of policy changes, NIH could compare the total, median, and mean number of publication fees in the direct portion of grants for different stakeholder groups over time and as a percentage of total published articles funded by the agency.

Early input on considerations to increase findability and transparency of research. NIH seeks suggestions on any specific issues that be considered in efforts to improve use of PIDs and metadata, including information about experiences institutions and researchers have had with adoption of different identifiers.

FASEB is encouraged by NIH's commitment to engage with existing identifier infrastructure and standards already in use across many scholarly societies. Requiring ORCID (Open Researcher and Contributor ID) for the corresponding and/or submitting author has been seamless for integration into societies' manuscript submission, peer review, and publication systems; requiring ORCID for *all* coauthors has posed more challenging but is improving with time. FASEB supports NIH adoption of a DOI (Digital Object Identifier) overlay on existing grants; this activity could foster a more connected ecosystem of grants, publications, and data.

#### **Additional Clarification**

FASEB requests more clarification about the types of manuscripts subject to NIH's public access plan, specifically, whether review articles, perspectives, commentaries, or editorials would be included. In our experience, this type of content is most often developed outside of research grants, with content development supported by society staff.

Regarding point III.C.1 (proposes to clarify how NIH-supported investigators may retain sufficient rights to NIH-supported peer-reviewed manuscript), we recommend that NIH work closely with the community on the development of any planned guidance. The 1940 Statement of Principles on Academic Freedom and Tenure is the professional standard of academic freedom widely endorsed and included in handbooks at most colleges and universities. It entitles faculty to 'full freedom in research and in the publication of the results.' Positive partnerships between government agencies, institutions, researchers, and other stakeholders – including scholarly societies - form the foundation for the success of the economic enterprise. FASEB supports researchers having the academic freedom to choose where they communicate and share their research findings, including their preferred choice of journal and their preferred license for any reuse.

#### Conclusion

FASEB commends NIH for its commitment to engaging and iterating to improve the plan for public access and to develop a policy that allows researchers to comply more readily. As the largest coalition of biological and biomedical researchers in the United States, we hope to continue the discussion, and offer to work with NIH to host dedicated events convening the variety of stakeholders impacted.

Sincerely,

Kevin C. Kregel, PhD FASEB President

**Submit date:** 4/6/2023

I am responding to this RFI: On behalf of an organization

Name: Heather Patisaul

Name of Organization: NC State University

Type of Organization: University

Role: Institutional official

1. How to best ensure equity in publication opportunities for NIH-supported investigators.

2. Steps for improving equity in access and accessibility of publications.

One of the biggest problems we foresee with NIH's plan is that all the data must be shared before the end of the project. That does not mesh with reality. Lots of publications occur after the project is over, sometimes long after. Researchers are going to need more time to get their data out. Publication is slow, particularly now given the increased demand on reviewer time. Giving people at least a year after the project officially ends would help ensure people have the time they need to be compliant.

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

As both a researcher and an administrator, compliance cost is our biggest concern. It's not just publication fees, which will likely be a big problem since we will now have to become more dependent on open access journals, but lots of Other "invisible" compliance costs that will fall on universities. Storing the data, curating the data, hiring data managers who can put all of this data in repositories or similar are all going to be expensive problems for our University. Data management and sharing isn't logistically simple. The data storage costs alone are easily going to cost us millions of dollars and we do not have the budget for that. And while there are great databases for some kinds of data, such as GWAS and Other genetic data, there are not obvious places for data like animal behavior data, imaging data, or computer code for analyses done in R or similar. Also, getting data into those databases can be very challenging. Some require at least some level of basic coding skills, which a lot of researchers do not have. The administrative burden (both time and cost) this could create for universities is potentially enormous and will disproportionately impact institutions like ours that doesn't have a big NIH portfolio and/or a medical school. Wealthier schools will already have a lot of infrastructure in place for managing patient data and that kind of thing. Others, like ours, will struggle. Researchers are going to need a lot of data management help and hiring those people is expensive. Whole industries will likely arise just to manage all the data NIH now expects we manage and share. It is strongly recommended NIH help shoulder this burden and have people in place to help researchers with data management.

#### 4. Early input on considerations to increase findability and transparency of research.

Not all data is easily put in a database that generates PIDs. So for some kinds of data this is going to be a challenge. It is also unclear how the metadata is supposed to be formatted or what it should include. Ideally it is organized for easy curation and/or systematic review or Other processes but accomplishing that is going to take professional data managers. Most researchers are not going to be able to do this on their own. Who is going to pay for that? If NIH wants things prepared and shared in a specific way, they should have the staff and people in place to help facilitate that. Part of the issue for us is understanding

the intent of NIH's data sharing plan. How is NIH envisioning the data will be used? How is the public supposed to access it? Is that even a reasonable expectation (particularly given the enormous financial and time burden this is going to place on researchers and their institutions)? If there are "higher priority" data sets NIH could start with and pilot to create guidance for Other kinds of data, that would be helpful.

**Submit date:** 4/9/2023

I am responding to this RFI: On behalf of an organization

Name: Jessica Moise

Type of Organization: Not applicable

1. How to best ensure equity in publication opportunities for NIH-supported investigators.

An additional 2% should be added to the A component of the F&A specifically restricted to entering into Institution-wide Institutional publishing accounts or to support staff who handle assisting Investigators with the new requirements.

- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

**Submit date: 4/10/2023** 

I am responding to this RFI: On behalf of myself

Name: Luella Allen-Waller

Name of Organization: The University of Pennsylvania

Type of Organization: University

Role: Scientific researcher

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

The NIH should provide funding specifically to subvent the costs to investigators to publish in the open access form of journals, which are often much more expensive than subscriber-access versions. This will reduce disproportionate funding burdens on early-career researchers and researchers in less well-resourced institutions who hope to make their findings publicly available. In addition, the NIH should seek to enter into agreements with all life science publishers to submit final published articles to PMC so that the general public can access quality articles of interest without the undue burden of expensive subscriptions.

#### 2. Steps for improving equity in access and accessibility of publications.

I support removal of the embargo, and believe that all peer-reviewed papers should be made immediately accessible to the general public.

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

I support monitoring publication fees, and reducing them, especially for open-access publications. NIH should create a data analyst position to track costs associated with publication and dissemination of results for all relevant positions, and to liaise with major publishers to understand cost increase trends and where that funding is going.

#### 4. Early input on considerations to increase findability and transparency of research.

I have not had experience with this.

**Submit date:** 4/10/2023

I am responding to this RFI: On behalf of myself

Name: Elizabeth McNally

Name of Organization: Northwestern University Feinberg School of Medicine, The Journal of Clinical

Investigation

Type of Organization: Other

Type of Organization-Other: Medical School and Journal Editor-In-Chief

Role: Scientific researcher

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Scientific publishing relies on peer review, since peer review provides an independent assessment of scientific discoveries. However, the peer review system is inherently a biased process. The peer review system favors investigators with experience, which has some merit given the role of experience in experimental design and interpretation. Editors should strive for balance when weighing differences of opinions between authors and reviewers, especially because authors typically remain blinded to the reviewers' identity. Opportunities to publish non peer-reviewed work in the form of preprint servers is highly valuable for multiple reasons. Preprint servers allows authors to make information accessible in a timely manner, and authors may cite manuscripts submitted to preprint servers on NIH biosketches and grant proposals. Not all manuscripts posted to a "preprint" server will ultimately appear in a peer-reviewed journal. There is likely value in having a long term archive for non-peer reviewed work since it helps disseminate work, albeit in the absence of peer review.

#### 2. Steps for improving equity in access and accessibility of publications.

With immediate access of published work, journals will lose subscription revenues and the increased cost will be passed on to authors, and hence the NIH. Publication fees have been rising, and it is not evident that all increase in publication fees derive from increasing costs. Larger publishers have economy of scale, while smaller, society-led journals may not have this advantage. In comparison to the large for-profit publishing houses, society-led journals may be nonprofit entities, and society-led journals hold an historically valuable role in the dissemination of science and opinion, which can influence science and science policy in critical ways. While it is tempting to establish flat fee recommendations for publishing, flat fees might endanger the smaller, nonprofit, society-led journals.

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

The NIH should keep in mind the wide range of publishers from the larger, for-profit publishing houses to the nonprofit society-led journals. The financial models supporting these different publishers are strikingly different, and the costs associated with publishing may also differ. Equity in publishing should not favor one financial publishing model. However, equity in publishing might consider opportunities to highlight predatory journals that have inadequate peer review and purely profit-seeking motives since this would be of value to the scientific community and public.

#### 4. Early input on considerations to increase findability and transparency of research.

Persistent identifiers for manuscripts should be an internationally agreed upon format given the international conduct of science. Persistent identifiers for authors should not replace names or identities since knowledge of who and where science was produced is relevant to the role of experience as an investigator.

Email: elizabeth.mcnally@northwestern.edu

**Submit date:** 4/11/2023

I am responding to this RFI: On behalf of an organization

Name: Robert Kiley

Name of Organization: cOAlition S

Type of Organization: Nonprofit research organization

Role: Institutional official

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

cOAlition S has long championed the view that their funded researchers should have the freedom to submit their manuscripts to any journal of their choice, irrespective of any open access (or public access) mandate they may be subject to. As such, funded researchers should have the freedom to submit manuscripts to both fully OA journals and subscription/hybrid journals, whilst also being able to honour the conditions of any public access mandate.

Publication costs should not be borne by the author

When a manuscript is accepted for publication in a fully OA journal, any publication costs charged by the publisher - like an APC - should be met by the funder. This is the approach cOAlition S has long endorsed.

Avoiding double payments in hybrid journals

However, when publishing in a subscription journal/hybrid journal, we do not believe a funder should pay an APC (or similar open access publishing fee), as the costs incurred by the publisher in publishing that article have already been met by the journals' subscribers.

Retaining author rights

To ensure that NIH funded researchers can always seek to publish in their journal of choice while at the same time making their papers available in public access via a repository, we strongly support the NIH proposal, outlined in section III. C. 1, to "develop language that NIH-supported investigators may use for submission with their peer-reviewed manuscripts to journals to retain rights to make the peer-reviewed manuscript available post-publication in PMC as soon as processing is complete, without an embargo period".

Many funders within cOAlition S - including the Bill and Melinda Gates Foundation, Howard Hughes Medical Institute, Wellcome and UK Research and Innovation (UKRI) - have adopted a similar approach, providing templated language which researchers must include in the manuscripts they submit to publishers.

By way of example, the Wellcome grant conditions include the following clauses:

7.4. You hereby grant a CC-BY Public Copyright Licence to all future Author Accepted Manuscripts (AAMs). If you allow Others to own copyright in AAMs, you must ensure they grant such a licence.

7.5. All submissions of original research to peer-reviewed journals must contain the following statement:

"This research was funded in whole or in part by the Wellcome Trust [Grant number]. For the purpose of open access, the author has applied a CC-BY public copyright licence to any author accepted manuscript version arising from this submission."

By requiring researchers to include the language (in clause 7.5) in their submissions, they are giving notice to the publisher of a prior licence. Publishers must either respect this - and allow the author to make the AAM available at the time of publication under the specified licence - or reject the submission.

In the two years or so since this approach was introduced by many cOAlition S funders, we are only aware of one example where a publisher rejected a manuscript due to the existence of a prior licence. In contrast there are many examples where an AAM has been made freely available at the time of publication (with a CC BY licence), but where the publisher version (the so-called Version of Record (VoR)), is paywalled.

See below three examples of articles published in 2023 where the AAM is freely available and licensed CC BY, but the VoR is paywalled with a more restrictive licence.

1. Article published in Nature Cell Biology, January 2023.

AAM, published under CC BY licence, freely available at: https://europepmc.org/article/MED/36650381#free-full-text;

VoR, paywalled and published under an exclusive licence to Springer Nature Limited, available at: <a href="https://dx.doi.org/10.1038%2Fs41556-022-01053-0">https://dx.doi.org/10.1038%2Fs41556-022-01053-0</a>

2. Article published in Journal of Virology, February 2023

AAM, published under a CC BY licence available at: <a href="https://europepmc.org/article/MED/36749077#free-full-text">https://europepmc.org/article/MED/36749077#free-full-text</a>;

VoR, paywalled, copyright of the American Society for Microbiology, All Rights Reserved, available at: https://dx.doi.org/10.1128%2Fjvi.00039-23

3. Article published in Journal of Immunology, March 2023

AAM, published under a CC BY licence, available at: <a href="https://europepmc.org/article/MED/36695776#free-full-text">https://europepmc.org/article/MED/36695776#free-full-text</a>;

VoR, paywalled, copyright of the The American Association of Immunologists, available at: <a href="https://dx.doi.org/10.4049%2Fjimmunol.2200211">https://dx.doi.org/10.4049%2Fjimmunol.2200211</a>

2. Steps for improving equity in access and accessibility of publications.

We are delighted that the NIH will remove the 12-month embargo period for NIH-supported publications.

Using licenses that allow sharing and reuse

However, to ensure that this research can be used by a large and diverse community of users, it is imperative that this work is properly licensed in ways which facilitate this.

For example, it may be desirable to translate an article from English to anOther language, such that it can be read by communities where English is not their first language. Equally, there may be value in creating a lay-person summary of a research article, such that it could be made accessible to non-experts. In both examples cited here, this would only be possible if third parties had the right to create derivative works, which is only possible under specific licences.

Beyond the need to create derivatives, some third parties may wish to re-use NIH-funded work which could have commercial implications, such as re-using a figure from an article for inclusion in a commercially published textbook. To ensure this is possible, the ability to reuse NIH funded research for commercial purposes must be made explicit in the licence which accompanies the research article.

It is also worth stressing that re-using images/figures from an article to create or enhance a page on a resource like Wikipedia, is only possible if the images are free of copyright or in the public domain. See: <a href="https://en.wikipedia.org/wiki/Wikipedia:Image\_use\_policy">https://en.wikipedia.org/wiki/Wikipedia:Image\_use\_policy</a>

All the use cases described here can be enabled if the NIH makes it a requirement that research findings which arise from its funding are made available under a Creative Commons Attribution licence (CC BY) or similar licence.

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

We agree with the NIH that it is important to monitor trends in publication fees, to ensure they are reasonable and equitable.

Price and services transparency.

One way cOAlition S is seeking to do this is through the free, online Journal Comparison Service (JCS), which we have developed.

The primary purpose of the JCS is to provide those who procure publishing services (typically libraries, library consortia, and funders) with the ability to quickly compare journal publishing services and fees. As of March 2023, 28 publishers have shared their data through this service.

Although the JCS holds data on journal APCs and subscriptions - and will retain such data to enable longitudinal analyses to be conducted over time - the service also provides information on the services publishers provide (copy editing, managing peer review, marketing etc) and the proportion of the total price which is allocated to each service. As such we believe it will be possible for users to determine whether the fees levied are commensurate with the services provided.

Consequently, one way the NIH could operationalise its ambition to monitor trends in publication fees is by strongly encouraging publishers who publish NIH-funded research to make their price and service data available through the JCS.

#### 4. Early input on considerations to increase findability and transparency of research.

The widespread adoption of PIDs will both reduce the burden on researchers (as information required for publisher and funder systems can be pre-populated) and provide all users with richer and more accurate data. For example, a funder reporting system, which requires grantees to disclose a list of publications arising from their grant, will get more accurate metadata if the publication data is pulled from services like Crossref or SCOPUS, using the researchers ORCID id as its match point.

Regarding specific actions NIH could consider to further encourage the adoption of PIDs, we would make two recommendations:

1. Require researchers to have an ORCID iD.

Although NIH already makes good use of ORCID - for example by allowing researchers to populate their SciENcv and eRA Commons records using their ORCID iD - having an ORCID iD is not yet a requirement for NIH applicants and grant holders.

However, we would like to suggest that, as part of the grant application process, all applicants are required to have an ORCID iD, and for that PID to be validated as part of the application process. A number of funders - such as Wellcome and HHMI - already require their researchers to have an ORCID iD.

By implementing this change, NIH can be assured that every funded researcher has a valid ORCID iD, which will make downstream reporting far simpler, as all published papers carrying the ORCID iD can be automatically added to the researcher's ORCID record.

Although some may argue that mandating the use of ORCID will discourage Other researcher identification systems to be developed, there is no need for multiple systems in this space, especially given the fact that ORCID is run as a community initiative, governed by a Board of Directors representative of its membership with wide stakeholder representation.

2. Assign a DOI to every grant awarded by the NIH

A number of funders within cOAlition S - including Wellcome and the Austrian Science Fund (FWF) - mint a DOI for every grant they award.

This approach has two distinct benefits.

Firstly, it enables the funder to make a trusted assertion in the researchers' ORCID record that they are in receipt of a Wellcome (or FWF) grant. Consequently, when anyone else looks at this ORCID record - maybe a funder considering a new award, or an institution determining a promotion or tenure decision - they can be assured that the applicant does hold the award they claim.

Secondly, it enables third party systems - such as publisher submission systems - to query Other sources (such as the Crossref registry) to prepopulate the submission system with the correct name of the funder and the specific grant ID. And, if the article is eventually published, then the Grant DOI will become part of its public metadata, enabling the funder (or the researcher) to unambiguously identify all the articles which have arisen from that grant.

#### **Uploaded File:**

Response-to-NIH\_cOAlition\_S.pdf

**Description:** Formatted version of responses provided via the form.

## Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research: response from cOAlition S

cOAlition S is an international consortium of research funding and performing organisations committed to accelerating the transition to open access. See <a href="https://www.coalition-s.org/">https://www.coalition-s.org/</a> for further details.

1. How to best ensure equity in publication opportunities for NIH-supported investigators.

The NIH Public Access Plan aims to maintain the existing broad discretion for researchers and authors to choose how and where to publish their results. Consistent with current practice, the NIH Public Access Plan allows the submission of final published articles to PMC (in cases where a formal agreement is in place) to minimize the compliance burden on NIH-supported researchers and also maintains the flexibility of NIH-supported researchers to submit the final peer-reviewed manuscript. These submission routes are allowed regardless of whether or not the journal uses an open access model, a subscription model of publishing, or other publication model. This flexibility aims to protect against concerns that have been raised about certain publishing models potentially disadvantaging early career researchers and researchers from limited-resourced institutions or under-represented groups. NIH policy already allows supported researchers to charge reasonable publishing costs against their awards. NIH seeks information on additional steps it might consider taking to ensure that proposed changes to implementation of the NIH Public Access Policy do not create new inequities in publishing opportunities or reinforce existing ones

#### Response from cOAlition S

cOAlition S has long championed the view that their funded researchers should have the freedom to submit their manuscripts to any journal of their choice, irrespective of any open access (or public access) mandate they may be subject to. As such, funded researchers should have the freedom to submit manuscripts to both fully OA journals and subscription/hybrid journals, whilst also being able to honour the conditions of any public access mandate.

#### Publication costs should not be borne by the author

When a manuscript is accepted for publication in a fully OA journal, any publication costs charged by the publisher – like an APC – should be met by the funder. This is the approach cOAlition S has long endorsed.

#### Avoiding double payments in hybrid journals

However, when publishing in a subscription journal/hybrid journal, we do not believe a funder should pay an APC (or similar open access publishing fee), as the costs incurred by the publisher in publishing that article have already been met by the journals' subscribers.

#### **Retaining author rights**

To ensure that NIH funded researchers can <u>always</u> seek to publish in their journal of choice while at the same time making their papers available in public access via a repository, we strongly support the NIH proposal, outlined in section III. C. 1, to "develop language that NIH-supported investigators may use for submission with their peer-reviewed manuscripts to journals to retain rights to make the peer-reviewed manuscript available post-publication in PMC as soon as processing is complete, without an embargo period".

**Commented [JR1]:** I found the original sentence hard to parse

Many funders within cOAlition S – including the Bill and Melinda Gates Foundation, Howard Hughes Medical Institute, Wellcome and UK Research and Innovation (UKRI) – have adopted a similar approach, providing templated language which researchers must include in the manuscripts they submit to publishers.

By way of example, the Wellcome grant conditions include the following clauses:

- 7.4. You hereby grant a CC-BY Public Copyright Licence to all future Author Accepted Manuscripts (AAMs). If you allow others to own copyright in AAMs, you must ensure they grant such a licence.
- 7.5. All submissions of original research to peer-reviewed journals must contain the following statement:

"This research was funded in whole or in part by the Wellcome Trust [Grant number]. For the purpose of open access, the author has applied a CC-BY public copyright licence to any author accepted manuscript version arising from this submission."

By requiring researchers to include the language (in clause 7.5) in their submissions, they are giving notice to the publisher of a prior licence. Publishers must either respect this – and allow the author to make the AAM available at the time of publication under the specified licence – **or** reject the submission.

In the two years or so since this approach was introduced by many cOAlition S funders, we are only aware of one example where a publisher rejected a manuscript due to the existence of a prior licence. In contrast there are many examples where an AAM has been made freely available at the time of publication (with a CC BY licence), but where the publisher version (the so-called Version of Record (VoR)), is paywalled.

See below three examples of articles published in 2023 where the AAM is freely available and licensed CC BY, but the VoR is paywalled with a more restrictive licence.

1. Article published in Nature Cell Biology, January 2023.

**AAM**, published under CC BY licence, freely available at:

 $\underline{\text{https://europepmc.org/article/MED/36650381\#free-full-text}};$ 

**VoR,** paywalled and published under an exclusive licence to Springer Nature Limited, available at: <a href="https://dx.doi.org/10.1038%2Fs41556-022-01053-0">https://dx.doi.org/10.1038%2Fs41556-022-01053-0</a>

2. Article published in Journal of Virology, February 2023

**AAM**, published under a CC BY licence available at:

https://europepmc.org/article/MED/36749077#free-full-text;

**VoR**, paywalled, copyright of the American Society for Microbiology, All Rights Reserved, available at: <a href="https://dx.doi.org/10.1128%2Fjvi.00039-23">https://dx.doi.org/10.1128%2Fjvi.00039-23</a>

3. Article published in *Journal of Immunology*, March 2023

**AAM**, published under a CC BY licence, available at: https://europepmc.org/article/MED/36695776#free-full-text;



2. Steps for improving equity in access and accessibility of publications.

Removal of the currently allowable 12-month embargo period for NIH-supported publications will improve access to these research products for all. As noted in the NIH Public Access Plan, NIH also plans to continue making articles available in human and machine-readable forms to support automated text processing. NIH will also seek ways to improve the accessibility of publications via assistive devices. NIH welcomes input on other steps that could be taken to improve equity in access to publications by diverse communities of users, including researchers, clinicians and public health officials, students and educators, and other members of the public.

#### Response from cOAlition S

We are delighted that the NIH will remove the 12-month embargo period for NIH-supported publications.

#### Using licenses that allow sharing and reuse

However, to ensure that this research can be used by a large and diverse community of users, it is imperative that this work is properly licensed in ways which facilitate this.

For example, it may be desirable to translate an article from English to another language, such that it can be read by communities where English is not their first language. Equally, there may be value in creating a lay-person summary of a research article, such that it could be made accessible to non-experts. In both examples cited here, this would only be possible if third parties had the right to create derivative works, which is only possible under specific licences.

Beyond the need to create derivatives, some third parties may wish to re-use NIH-funded work which could have commercial implications, such as re-using a figure from an article for inclusion in a commercially published textbook. To ensure this is possible, the ability to reuse NIH funded research for commercial purposes must be made explicit in the licence which accompanies the research article.

It is also worth stressing that re-using images/figures from an article to create or enhance a page on a resource like Wikipedia, is only possible if the images are free of copyright or in the public domain. See: <a href="https://en.wikipedia.org/wiki/Wikipedia:Image">https://en.wikipedia.org/wiki/Wikipedia:Image</a> use policy

All the use cases described here can be enabled if the NIH makes it a requirement that research findings which arise from its funding are made available under a Creative Commons Attribution licence (CC BY) or similar licence.

3. Methods for monitoring evolving costs and impacts on affected communities. NIH proposes to actively monitor trends in publication fees and policies to ensure that they remain reasonable and equitable. NIH seeks information on effective approaches for monitoring trends in publication fees and equity in publication opportunities.

#### Response from cOAlition S

We agree with the NIH that it is important to monitor trends in publication fees, to ensure they are reasonable and equitable.

#### Price and services transparency

One way cOAlition S is seeking to do this is through the free, online <u>Journal Comparison Service</u> (JCS), which we have developed.

Commented [JR2]: I really like these examples, very pedagogical!

The primary purpose of the JCS is to provide those who procure publishing services (typically libraries, library consortia, and funders) with the ability to quickly compare journal publishing services and fees. As of March 2023, 28 publishers have shared their data through this service.

Although the JCS holds data on journal APCs and subscriptions – and will retain such data to enable longitudinal analyses to be conducted over time – the service also provides information on the services publishers provide (copy editing, managing peer review, marketing etc) and the proportion of the total price which is allocated to each service. As such we believe it will be possible for users to determine whether the fees levied are commensurate with the services provided.

Consequently, one way the NIH could operationalise its ambition to monitor trends in publication fees is by strongly encouraging publishers who publish NIH-funded research to make their price and service data available through the JCS.

4. Early input on considerations to increase findability and transparency of research. Section IV of the NIH Public Access Plan is a first step in developing the NIH's updated plan for PIDs and metadata, which will be submitted to OSTP by December 31, 2024. NIH seeks suggestions on any specific issues that should be considered in efforts to improve use of PIDs and metadata, including information about experiences institutions and researchers have had with adoption of different identifiers.

#### Response from cOAlition S

The widespread adoption of PIDs will both reduce the burden on researchers (as information required for publisher and funder systems can be pre-populated) and provide all users with richer and more accurate data. For example, a funder reporting system, which requires grantees to disclose a list of publications arising from their grant, will get more accurate metadata if the publication data is pulled from services like Crossref or SCOPUS, using the researchers ORCID id as its match point.

Regarding specific actions NIH could consider to further encourage the adoption of PIDs, we would make two recommendations:

#### 1. Require researchers to have an ORCID iD.

Although NIH already makes good use of ORCID – for example by allowing researchers to populate their SciENcv and eRA Commons records using their ORCID iD – having an ORCID iD is not yet a requirement for NIH applicants and grant holders.

However, we would like to suggest that, as part of the grant application process, all applicants are **required** to have an ORCID iD, and for that PID to be validated as part of the application process. A number of funders – such as Wellcome and HHMI – already require their researchers to have an ORCID iD.

By implementing this change, NIH can be assured that every funded researcher has a valid ORCID iD, which will make downstream reporting far simpler, as all published papers carrying the ORCID iD can be automatically added to the researcher's ORCID record.

Although some may argue that mandating the use of ORCID will discourage other researcher identification systems to be developed, there is no need for multiple systems in this space, especially given the fact that ORCID is run as a community initiative, governed by a Board of Directors representative of its membership with wide stakeholder representation.

#### 2. Assign a DOI to every grant awarded by the NIH

Commented [JR3]: official ORCI spelling is "ORCID iD".

A number of funders within cOAlition S – including Wellcome and the Austrian Science Fund (FWF) – mint a DOI for every grant they award.

This approach has two distinct benefits.

Firstly, it enables the funder to make a trusted assertion in the researchers' ORCID record that they are in receipt of a Wellcome (or FWF) grant. Consequently, when anyone else looks at this ORCID record – maybe a funder considering a new award, or an institution determining a promotion or tenure decision – they can be assured that the applicant does hold the award they claim.

Secondly, it enables third party systems – such as publisher submission systems – to query other sources (such as the Crossref registry) to prepopulate the submission system with the correct name of the funder and the specific grant ID. And, if the article is eventually published, then the Grant DOI will become part of its public metadata, enabling the funder (or the researcher) to unambiguously identify all the articles which have arisen from that grant.

**Submit date: 4/11/2023** 

I am responding to this RFI: On behalf of myself

Name: Robert Weinberg

Name of Organization: Whitehead Institute for Biomedical Research/MIT

Type of Organization: Nonprofit research organization

Role: Scientific researcher

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

This continues the trend to complicate receiving and maintaining administratively an NIH grant, making it even more cumbersome and bureaucratic in order to secure and preserve grant funding. Those of you who applaud all these new steps continue to make it less and less attractive to apply for and maintain an NIH grant, making it more and more laborious administratively to keep the grant and administer it properly. (The amount of time spent by PIs and AAs in negotiating the labyrinthine rules grows progressively with each year.) I suppose it's the job of administrators to add more and more layers of detailed regulations. These new requirements continue the onward march of oppressive bureaucratization of NIH grant applications and reporting of awarded grants at a time when NIH grants are increasingly unable to support many of the experiments that are being proposed. I suppose you will only be happy when you totally smOther the program with more and more layers of bureaucratic regulation. You will have secured a pyrrhic victory, having proudly added all of these new bells and whistles to grant programs that increasingly no one is interested in applying for any more. Robert Weinberg, Whitehead Institute/MIT

- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

I am responding to this RFI: On behalf of myself

Name: Nils Walter

Name of Organization: University of Michigan

Type of Organization: University

Role: Scientific researcher

# 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Europe has found an equitable solution in the form of Plan S (<a href="https://www.coalition-s.org/">https://www.coalition-s.org/</a>) that supports open-access publication without burdening shrinking (in terms of inflation-corrected dollars) grant budgets. Such a program will a) remove open-access inequities among both researchers/authors and the public; b) reduce the extra work and cost involved in the duplication publication on the NIH MS system; c) make a single, consistent, peer-reviewed version of a publication available immediately upon publication (and often at time of acceptance) that publishers already offer; and d) allow the NIH/government to negotiate discounted open-access pricing from all publishers, rather than the current "Wild West" where highly reputed journals can charge large sums biased toward improving the publisher's bottom line.

### 2. Steps for improving equity in access and accessibility of publications.

Europe has found an equitable solution in the form of Plan S (<a href="https://www.coalition-s.org/">https://www.coalition-s.org/</a>) that supports open-access publication without burdening shrinking (in terms of inflation-corrected dollars) grant budgets. Such a program will a) remove open-access inequities among both researchers/authors and the public; b) reduce the extra work and cost involved in the duplication publication on the NIH MS system; c) make a single, consistent, peer-reviewed version of a publication available immediately upon publication (and often at time of acceptance) that publishers already offer; and d) allow the NIH/government to negotiate discounted open-access pricing from all publishers, rather than the current "Wild West" where highly reputed journals can charge large sums biased toward improving the publisher's bottom line.

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

Europe has found an equitable solution in the form of Plan S (<a href="https://www.coalition-s.org/">https://www.coalition-s.org/</a>) that supports open-access publication without burdening shrinking (in terms of inflation-corrected dollars) grant budgets. Such a program will a) remove open-access inequities among both researchers/authors and the public; b) reduce the extra work and cost involved in the duplication publication on the NIH MS system; c) make a single, consistent, peer-reviewed version of a publication available immediately upon publication (and often at time of acceptance) that publishers already offer; and d) allow the NIH/government to negotiate discounted open-access pricing from all publishers, rather than the current "Wild West" where highly reputed journals can charge large sums biased toward improving the publisher's bottom line.

#### 4. Early input on considerations to increase findability and transparency of research.

On top of a "Plan S-like" solution, publishers could be encouraged to work with NIHMS to develop faster, more versatile access options, most likely through AI tools.

I am responding to this RFI: On behalf of myself

Name: Clifford B. Saper, MD, PhD

Name of Organization: Beth Israel Deaconess Medical Center, Harvard Medical School

Type of Organization: University

Role: Scientific researcher

# 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

The costs of publication must be borne by someone. In the subscription model this was largely users, and most of the burden fell on University libraries and industry. While non-academic members of the public would have had to pay high fees for access, in fact very few private individuals read the primary scientific literature, and under the current system this is available within one year anyway. In the open access model, the costs fall on the scientists and clinicians who publish. University libraries are happy about this because it takes pressure off their budget, but industry is ecstatic because they get a free ride. Scientists have to pay these fees out of their grant budgets, if they have grants, but there has been no increase in funds available for this purpose, so they come out of the scientific budget. Worse, many clinical research studies are done without benefit of NIH grants, and there is often no way to pay for the publication fees in the open access model. This problem falls disproportionately on individuals who work at institutions that do not have large discretionary budgets, i.e., clinicians and scientists who serve underprivileged segments of society. A fair and equitable system would be similar to countries in Europe where there is a single fee shared by government and universities and industry and paid to publishers, which gives their scientists the ability to publish in open access journals without additional fees, and gives the entire country access to publications.

# 2. Steps for improving equity in access and accessibility of publications.

Mandating immediate access essentially will undermine and destroy the subscription model, without coming up with an adequate replacement. As noted above, this will unfairly punish investigators who do not have NIH grants, investigators who work at institutions that have low budgets because they serve poor people, and will unfairly benefit rich universities and industry (who hire the staff to promote this model).

# 3. Methods for monitoring evolving costs and impacts on affected communities.

The way to do this fairly is for the US to establish a consortium of industry, universities, and the federal government, to provide funds, proportionate to their use of published scientific material, to publishers, who would then eliminate both manuscript processing and access fees. This is similar to what has evolved, for example, in Germany.

# 4. Early input on considerations to increase findability and transparency of research.

While it is laudable to have work done by NIH-funded investigators available to Other investigators and the public, realistically it costs about 10% of the cost of the actual research to establish and maintain such databases. This is a very time-consuming and expensive proposition. Investigators cannot do it in

a few minutes of their spare time. To provide a robust and searchable archive will, realistically, require the NIH to devote about \$4B a year of its budget just to do this. It would also help if there were a national infrastructure in place that investigators could use to deposit data. We are now relying upon institutional databases, with no funds provide for establishing them, putting the data into a searchable format, monitoring the deposits to make sure that they actually occur, and providing public access to those databases. It is extremely unlikely that the NIH will be able to succeed in its goals with the current plans. More likely than not, we will have a system like the ClinicalTrials.gov database, where more than half of those who should be contributing are scofflaws.

The moral is: without adequate funding and infrastructure, these plans are burdensome, reduce research efficiency, and are likely to provide nothing of value in the long run.

Email: csaper@bidmc.harvard.edu

I am responding to this RFI: On behalf of myself

Name: Amy Wright

Type of Organization: Not applicable

Role: Scientific researcher

# 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

I come from a University with limited resources but which is not a PUI. I am a chemist and would typically publish in ACS published journals. The ACS open access costs are very high (in excess of \$4000 for an ACS member). ACS is selling reduced cost open access to large Universities/University systems (Transformative program) and provides free open access publication to PUIs. Almost all of the ACS journals are moving to full open access because of the removal of the currently allowed 12 month embargo. Since I work at a low resource, high research (but not R1) University I am caught with having to pay the full open access fee to publish. I can publish in MDPI journals for a reduced cost, but I think overall I will be publishing less under these new rules rather than more - so in the long term my program will suffer (fewer pubs=fewer grants). Eventually I will simply not be able to continue in my field. I envision that the day when research conducted at smaller Universities simply stops (and our students will not have access to that experience reducing workforce readiness overall). Just the top Universities will get most of the grants and diversity will be greatly reduced.

I have had a number of R21 grants which have had the same cap in total direct costs (\$275k over 2 years) for probably 25 years. Colleagues with R15s are equally stuck with budget caps from the dark ages. These types of projects (and even some continuing RO1s) don't have a lot of extra \$\$ for publishing OA. Perhaps it is time for these budget caps to be changed.

- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.

See my comments above- there are a lot of us stuck in the worst category for publishing OA and our Universities don't have the resources to buy reduced cost OA publishing.

4. Early input on considerations to increase findability and transparency of research.

I am responding to this RFI: On behalf of myself

Type of Organization: Not applicable

# 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

You have created the problem that you seek the solution to! Publication via subscription journals, whether published by professional societies or commercial publishers, had no bias or barrier Other than quality. Now, by requiring open access, you put a financial burden on people with low levels of funding. Already, I ask my students, do you want an Other paper published or do you want a month of your stipend. This is the tradeoff that we now face due to the belief that there was a problem with lack of access to research results.

# 2. Steps for improving equity in access and accessibility of publications.

There are very few people who will read and use the results of research who do not have ready access to a library. Libraries routinely use interlibrary loan to get access to the journals they do not have subscription for. The access problem existed only in the minds of political activists.

Now that the problem has been created, the best solution would be for NIH to pay directly for publication, in addition to, not part of, the awarded funding in the grant. This is a little like the American Chemical Society (ACS) Petroleum Research Fund did (I do not know the current policy). They would pay the so-called "page charges" if you published in an ACS journal. Page charges went way when commercial publishers began competing with professional society publishers.

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

NIH can ask for accounting of publication costs in annual reports. NIH can also ask if people combined results into conglomerate papers to avoid paying for multiple papers.

### 4. Early input on considerations to increase findability and transparency of research.

I have no idea what problem this attempts to solve. Each area of science has abstracting and search tools. Chemical Abstracts Service is the oldest and best in the US. Scientists are already asked to use numerical identifiers (ORCID) to overcome problems of inconsistent use of names or changes in names. An ORCID could be required to get a grant or to publish a paper.

I am responding to this RFI: On behalf of myself

Name: Lynn Brielmaier

Name of Organization: ALS Problem Solvers

Type of Organization: Patient advocacy organization

Role: Patient advocate

1. How to best ensure equity in publication opportunities for NIH-supported investigators.

2. Steps for improving equity in access and accessibility of publications.

Open up the NIH lecture series to the public. It is a simple matter to mirror the video seminars onto a cyber secure website outside of NIH or HHS.

eg. For this Nirenberg Lecture, Patapoutian will speak on his latest research on the structure-function relationships of Piezo proteins and their roles in somatosensation and interoception.

Note: This is a special Monday, Monday, Monday WALS.

\* Join and you will learn such fascinating trivia as this: Ardem Patapoutian is the only WALS speaker to be featured on an Armenian postage stamp. And Marshall Nirenberg is the only NIH scientist to be featured on a Palau Islands postage stamp.

Most of these science seminars ARE NOT available to the public.

3. Methods for monitoring evolving costs and impacts on affected communities.

https://www.manuscriptedit.com/scholar-hangout/market-trends-open-access-publishing/

https://authorservices.taylorandfrancis.com/choose-open/publishing-open-access/open-access-costfinder/

https://theplosblog.plos.org/2023/04/open-science-indicators/ (datasets published every six months)

4. Early input on considerations to increase findability and transparency of research.

Make data compatible with OMAP.

Please contact Danielle Boyce at dboyce3@jhu.edu, ok to say I sent you.

**Uploaded File:** 

NIH-Seminar-series-fail\_Screenshot-2023.jpg

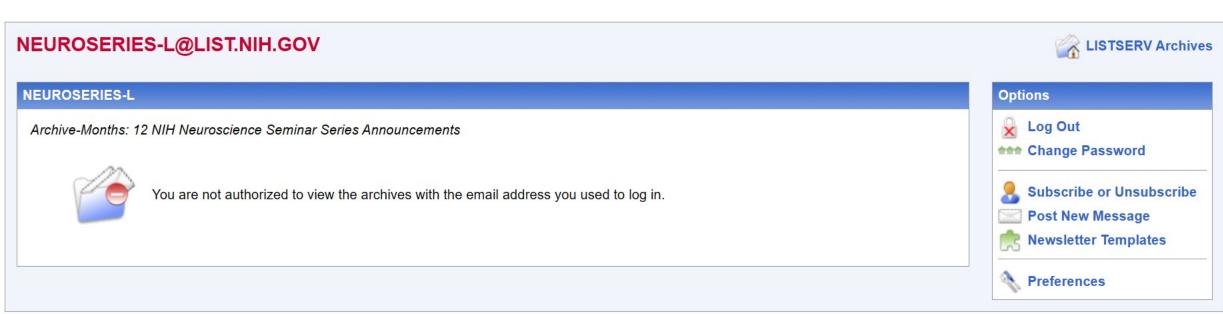
**Description:** NIH seminar list access denial.

Email: lynnbr2@att.net

Subscriber's Corner Email Lists



# **NEUROSERIES-L Home Page**



I am responding to this RFI: On behalf of myself

Name: Gail A. Bishop

Name of Organization: The University of Iowa

Type of Organization: University

Role: Scientific researcher

# 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

The major challenge of all plans that mandate early or immediate Open Access for all scientific publications from NIH-supported researchers is that there are significant costs associated with scientific publishing. If publishers can no longer use subscription revenues to support such costs, they will continue and accelerate the trend already underway of passing on more and more costs to authors. The costs of publishing a paper have risen considerably over the past 10-15 years, but the NIH modular budget has gone unchanged. At the same time, costs of research personnel, particularly the cost of benefits, have also increased substantially. Thus, mandating more and more open access without providing any cost relief steadily decreases the funds available to NIH-funded PIs to spend on the actual research project. Additionally, scientists who receive funding from major foundations, such as Howard Hughes or Wellcome Trust, receive funds from these entities for open access publishing, but not the majority of researchers, further increasing the advantages that scientists with additional resources such as these, or large endowments from some institutions, have, and risks narrowing the field of those who can contribute to scientific discovery.

- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.

HOW will NIH 'ensure that publication fees remain reasonable and equitable'. What will NIH do if, as seems likely in a mandated 'immediate open access for all publications', such fees for its grantee authors rise substantially?

4. Early input on considerations to increase findability and transparency of research.

I am responding to this RFI: On behalf of myself

Name: Mr. Stephen P. Panossian

Name of Organization: Unemployed

Type of Organization: Not applicable

Role: Scientific researcher

# 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Allowing the submission of final, published articles to PMC satisfies the centralized distribution aspect of published scientific research. However, many reputable journals are becoming open access, which may diminish the need for PMC. Furthermore, it does not address the initial hurdle of publishing research results in a reputable scientific journal, that of the publishing fees. These fees can range from several hundred to several thousand of US dollars, which can be a sizable obstacle for researchers from developing nations and small colleges in the Western world. These publishing fees siphon away valuable funding that principal investigators could use for resources and/or temporary labor (who would gain valuable experience during the research process). Ensuring equity in publication opportunities requires financing the publication fees and monitoring the expenditures. This is an additional responsibility that the PI will need to undertake to prove responsible spending of research funds. Thank you for considering my perspective on this topic.

### 2. Steps for improving equity in access and accessibility of publications.

Ensuring Internet access to researchers in developing nations improves their ability to submit research articles in human and machine-readable forms for publication. It also improves download access for the diverse communities of users, who can then read research results online and/or print them as they need. Furthermore, it will increase the accessibility of publications via assistive-technology, such as language translation and text-to-speech applications. While NIH cannot endorse any specific software or hardware, it could initiate the development of those applications and/or devices. Thank you for considering my perspective on this topic.

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

NIH's proposal to monitor trends in publication fees and policies to ensure that they remain reasonable and equitable is welcome. One challenge will be defining the landscape of reputable US and foreign scientific journals. NIH needs to consider whether the journals they monitor actively or passively constitutes an endorsement of those publications. Thank you for considering my perspective on this topic.

# 4. Early input on considerations to increase findability and transparency of research.

One aspect NIH may consider regarding the effort to increase findability and transparency of research is leading the release of standards for metadata governing content and format, and for datasets, the standards for datasheet contents. AnOther aspect to consider is, should PMC provide an easy user interface for researchers to document their publications with additional PIDs and metadata?

Furthermore, would NIH support financially the efforts to "retrofit" older publications with the latest metadata and PIDs? This would incentivize the standardization process. Thank you for considering my perspective on this topic.

Email: <a href="mailto:stephenpanossian@gmail.com">stephenpanossian@gmail.com</a>

I am responding to this RFI: On behalf of myself

Name: Phil Hurvitz

Name of Organization: University of Washington

**Type of Organization:** University

Role: Scientific researcher

- 1. How to best ensure equity in publication opportunities for NIH-supported investigators.
- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

It would be great to develop a standard for PDF metadata so that publications brought in to reference managers would populate fields correctly. I have had to spend considerable effort editing records in EndNote and Mendeley because fields are not populated correctly. At the same time, it would be good to encourage publishers to enter metadata correctly--I have had to edit many records because the PDF metadata are simply incorrect (wrong journal or author names, page numbers, etc.). Thanks for your consideration.

I am responding to this RFI: On behalf of myself

Name: Andriy Fedorov

Name of Organization: Brigham And Women's Hospital

Type of Organization: Not applicable

Role: Scientific researcher

# 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

The lack of consideration in this RFI survey for the requirements and strategies aimed to support sharing of research data - in addition to sharing of the publications - is unfortunate.

This is particularly important in consideration of ensuring equity in publication opportunities. Researchers have highly uneven access to the data needed both to conduct innovative research and to validate findings that have already been published. As a result, investigators from large institutions that have access to data have exceedingly stronger opportunities to receive even more funding. At the same time, at NIH, there appears to be lack of clear strategy and infrastructure investment to encourage and enable data sharing, and discourage and (as much as possible) mitigate hoarding of the data in the individual NIH-funded labs and institutions. Infrastructure currently being established by NIH to support collection and sharing of data does not appear to have plans or commitments in place to ensure continuing funding of the repository and guarantee longevity of the deposited data for \*\*any\*\* period of time, which is in direct contradiction of the principles set forth by the National Science and Technology Council! The new data sharing policy introduced by NIH does not affect the existing peer review process, which means what is shared and how will be decided by administrators and lawyers - not scientists! It is very likely that the outcome of the new policy will be datasets of limited utility, due to lack of scientific oversight of the approaches used to share those datasets.

To sum up, there is urgent need to 1) develop strategy for sharing datasets produced by NIH-funded projects, 2) establish plans for the development of the technology to enable such strategy; 3) establish framework and policies to support longevity of the repositories. It is very important that items 1 and 2 in the above are developed with participation of the domain experts representing research community, and that the process is transparent.

- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

Email: andrey.fedorov@gmail.com

Name: Robin Ely MD

Name of Organization: Integrative and Regenerative Medicine

Type of Organization: Health care delivery organization

**Role:** Medical provider

1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Totally agree!

2. Steps for improving equity in access and accessibility of publications.

Totally agree- motivated "patients" are often more informed than their doctors-

3. Methods for monitoring evolving costs and impacts on affected communities.

Can't speak to this

4. Early input on considerations to increase findability and transparency of research.

Recommend an immediate update process to clinicaltrials.gov and pubmed-

No embargoes, no blocks-

Information that can save or improve a life should be FREE-

This obviously is not a capitalistic view which puts a price on information sharing- but there should be a way around it -

I am responding to this RFI: On behalf of an organization

Name: Ashley Farley

Name of Organization: Bill & Melinda Gates Foundation

Type of Organization: Nonprofit research organization

Role: Institutional official

# 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

This response to the "NIH Plan to Enhance Public Access to the Results of NIH-Supported Research" request for public input is submitted on behalf of the Bill & Melinda Gates Foundation. Guided by the belief that every life has equal value, the Bill & Melinda Gates Foundation works to help all people lead healthy, productive lives. In developing countries, it focuses on improving people's health and giving them the chance to lift themselves out of hunger and extreme poverty. In the United States, it seeks to ensure that all people—especially those with the fewest resources—have access to the opportunities they need to succeed in school and life. Based in Seattle, Washington, the foundation is led by CEO Mark Suzman, under the direction of co-chairs Bill Gates and Melinda French Gates and the board of trustees. In 2022, the foundation's annual giving exceeded \$7 billion USD and continues to grow.

Since 2015, the Bill & Melinda Gates Foundation has had a strong Open Access (OA) Policy that is included in all grant agreements with no exemption. The scope of the OA policy enables the unrestricted access and reuse of all peer-reviewed published research funded, in whole or in part, by the foundation, including any underlying data sets. The policy implementation changed beginning January 1st 2021 to align with Plan S as the foundation joined cOAlition S in 2018. The OSTP Nelson Memo is a welcomed and much needed policy change to further advance OA publishing and establish it as the norm for research communication. However, implementation specifics will be important for adoption to ensure that OA publishing becomes more equitable and a real option for any researcher anywhere. With seven years of policy implementation, the foundation shares its learnings below.

Opportunity to expand publishing equity beyond NIH-supported investigators. While most funders focus primarily on equitable options for its own grantee author community - it is equally important that funders understand the influence their policies have on the wider community. We have learned this at the foundation: while we may be able to afford a range of OA fees on the behalf of our grantee authors - this promotes a business model that impedes equity for non-funded authors.

Recognizing outputs that do not require the publisher's version of record (VoR) - such as preprints, author accepted manuscripts, archived versions - is critical for equity within the entire publication ecosystem. Along with providing multiple routes to compliance, authors must retain their rights at the point of submission - safeguarding author choice from being stripped away by publisher policies and practices. As much as possible, the NIH should use its voice and influence to push back on publisher bad practices and decouple research dissemination from business models. Publishing will become more equitable as the academic career incentives shift from focusing on faulty metrics stymied in prestige publications. Strong signaling of the validity of open sharing of funded research by the NIH is a vital step.

# 2. Steps for improving equity in access and accessibility of publications.

Open licenses improve equity in access and accessibility of publications. Open and flexible licensing (particularly CC-BY) allow for increased innovation through discoverability and curation of published research. As more and more research is accessed for translation and further research, there will be less reliance on standard article formats and publisher versions of record. Funders and institutions should value different forms of outputs, such as plain language summaries, preprints, micropublications, protocols, case studies, and data notes. Supporting such a variety will expand the burgeoning AI and ML technologies to parse, summarize, and further disseminate research information. Foreign language translations of language in the author's native language will further expand reach and accessibility. Support and prioritize knowledge translation for the general public to be able to read, understand, and implement knowledge into their communities.

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

Trends in costs and community impact through robust compliance tracking. With improved tracking of data through the foundation's partnership with OA.Works the trend appears that the foundation policy is resulting in an increase of OA publications yearly, however the foundation is paying more year on year for fewer publications. In Other words, APC pricing continues to rise year over year for major publishers without notice, topping out in some cases at over \$12,000. Recent research projects show that funders are ineffectual in affecting APC pricing and it has been near impossible to define a "reasonable fee". While we presume that it does not cost \$12k to publish an article in a highly-selective journal, we do recognize that costs are incurred to support the publishing process. Funders have been advocating for more cost and price transparency but these initiatives have been slow to achieve impact with little publisher participation. The APC model is not sustainable or equitable and Other models have been slow to be tested or implemented. Readers and authors currently pay for cumbersome and antiquated publisher technology, various revenue streams (paywalled content, subscriptions, etc), and print-based legacy processes. Understanding actual publication costs and system improvements will help us realize a sustainable publishing model that is free to both readers and authors. More exploration and financial support for alternative models, such as Diamond and S20 is needed.

#### 4. Early input on considerations to increase findability and transparency of research.

Consistent metadata increases findability and transparency of research. Steps must be taken to increase funder metadata in the publication record to allow for proper attribution and discovery of funded research across multiple platforms and indexing services. Ideally consistent metadata travels with the research from inception to early sharing to data management to publication. There are various Persistent Identifiers (PIDs) projects to help improve this issue, however uptake is slow and requires adoption from a lot of actors. .

The Bill & Melinda Gates Foundation wishes to again express our gratitude and support for the work of the NIH, the OSTP, and Other federal agencies to advance a more open, equitable, and inclusive research ecosystem. We appreciate the opportunity to comment on this draft plan, and we are eager to assist in its rollout.

#### **Uploaded File:**

NIH\_PublicAccessRFI\_Gates.docx

Email: ashley.farley@gatesfoundation.org

This response to the "NIH Plan to Enhance Public Access to the Results of NIH-Supported Research" request for public input is submitted on behalf of the Bill & Melinda Gates Foundation. Guided by the belief that every life has equal value, the Bill & Melinda Gates Foundation works to help all people lead healthy, productive lives. In developing countries, it focuses on improving people's health and giving them the chance to lift themselves out of hunger and extreme poverty. In the United States, it seeks to ensure that all people—especially those with the fewest resources—have access to the opportunities they need to succeed in school and life. Based in Seattle, Washington, the foundation is led by CEO Mark Suzman, under the direction of co-chairs Bill Gates and Melinda French Gates and the board of trustees. In 2022, the foundation's annual giving exceeded \$7 billion USD and continues to grow.

Since 2015, the Bill & Melinda Gates Foundation has had a strong Open Access (OA) Policy that is included in all grant agreements with no exemption. The scope of the OA policy enables the unrestricted access and reuse of all peer-reviewed published research funded, in whole or in part, by the foundation, including any underlying data sets. The policy implementation changed beginning January 1st 2021 to align with Plan S as the foundation joined coalition S in 2018. The OSTP Nelson Memo is a welcomed and much needed policy change to further advance OA publishing and establish it as the norm for research communication. However, implementation specifics will be important for adoption to ensure that OA publishing becomes more equitable and a real option for any researcher anywhere. With seven years of policy implementation, the foundation shares its learnings below.

- Opportunity to expand publishing equity beyond NIH-supported investigators. While most funders focus primarily on equitable options for its own grantee author community it is equally important that funders understand the influence their policies have on the wider community. We have learned this at the foundation: while we may be able to afford a range of OA fees on the behalf of our grantee authors this promotes a business model that impedes equity for nonfunded authors. Recognizing outputs that do not require the publisher's version of record (VoR) such as preprints, author accepted manuscripts, archived versions is critical for equity within the entire publication ecosystem. Along with providing multiple routes to compliance, authors must retain their rights at the point of submission safeguarding author choice from being stripped away by publisher policies and practices. As much as possible, the NIH should use its voice and influence to push back on publisher bad practices and decouple research dissemination from business models. Publishing will become more equitable as the academic career incentives shift from focusing on faulty metrics stymied in prestige publications. Strong signaling of the validity of open sharing of funded research by the NIH is a vital step.
- Open licenses improve equity in access and accessibility of publications. Open and flexible
  licensing (particularly CC-BY) allow for increased innovation through discoverability and curation
  of published research. As more and more research is accessed for translation and further
  research, there will be less reliance on standard article formats and publisher versions of record.
  Funders and institutions should value different forms of outputs, such as plain language
  summaries, preprints, micropublications, protocols, case studies, and data notes. Supporting
  such a variety will expand the burgeoning AI and ML technologies to parse, summarize, and

further disseminate research information. Foreign language translations of language in the author's native language will further expand reach and accessibility. Support and prioritize knowledge translation for the general public to be able to read, understand, and implement knowledge into their communities.

- Trends in costs and community impact through robust compliance tracking. With improved tracking of data through the foundation's partnership with OA.Works the trend appears that the foundation policy is resulting in an increase of OA publications yearly, however the foundation is paying more year on year for fewer publications. In other words, APC pricing continues to rise year over year for major publishers without notice, topping out in some cases at over \$12,000. Recent research projects show that funders are ineffectual in affecting APC pricing and it has been near impossible to define a "reasonable fee". While we presume that it does not cost \$12k to publish an article in a highly-selective journal, we do recognize that costs are incurred to support the publishing process. Funders have been advocating for more cost and price transparency but these initiatives have been slow to achieve impact with little publisher participation. The APC model is not sustainable or equitable and other models have been slow to be tested or implemented. Readers and authors currently pay for cumbersome and antiquated publisher technology, various revenue streams (paywalled content, subscriptions, etc), and printbased legacy processes. Understanding actual publication costs and system improvements will help us realize a sustainable publishing model that is free to both readers and authors. More exploration and financial support for alternative models, such as Diamond and S20 is needed.
- Consistent metadata increases findability and transparency of research. Steps must be taken to increase funder metadata in the publication record to allow for proper attribution and discovery of funded research across multiple platforms and indexing services. Ideally consistent metadata travels with the research from inception to early sharing to data management to publication. There are various Persistent Identifiers (PIDs) projects to help improve this issue, however uptake is slow and requires adoption from a lot of actors.

The Bill & Melinda Gates Foundation wishes to again express our gratitude and support for the work of the NIH, the OSTP, and other federal agencies to advance a more open, equitable, and inclusive research ecosystem. We appreciate the opportunity to comment on this draft plan, and we are eager to assist in its rollout.

I am responding to this RFI: On behalf of an organization

Name of Organization: American Academy of Neurology

Type of Organization: Professional org association

Role: Medical provider

# 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Dear Dr. Tabak,

The American Academy of Neurology (AAN) is the world's largest neurology specialty society representing more than 40,000 neurologists and clinical neuroscience professionals. The AAN is dedicated to promoting the highest quality patient-centered neurologic care. A neurologist is a physician with specialized training in diagnosing, treating, and managing disorders of the brain and nervous system. These disorders affect one in six people and include conditions such as multiple sclerosis (MS), Alzheimer's disease (AD), Parkinson's disease, stroke, migraine, epilepsy, traumatic brain injury, ALS, and spinal muscular atrophy.

The AAN greatly appreciates the opportunity to provide feedback in response to the "Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research" from the National Institutes of Health (NIH). While the AAN is supportive of the goal of enhancing public access to the results to NIH-supported research, the AAN is deeply concerned that the NIH Public Access Plan as described in NOT-OD-23-091 will be highly disruptive to the ongoing operations and article quality of Neurology® and Neurology Clinical Practice®.

The AAN is deeply concerned that the NIH Public Access Plan will result in numerous unintended consequences, resulting from the need for journals like Neurology® and Neurology Clinical Practice® to substantially modify their revenue models. The AAN believes that changes to the underlying business model stemming from implementation of the NIH Public Access Plan will likely necessitate a shift of financial responsibility from subscribers to the researchers seeking to have their research published, creating substantial additional barriers for those seeking publication. The rapid implementation of the NIH plan, specifically the elimination of the 12-month embargo, is extremely disruptive and may negatively impact the financial underpinnings of scholarly publishing and dissemination. The AAN is alarmed by the potential for the NIH Public Access Plan to create substantial inequity in those able to contribute to the body of peer-reviewed published scientific research. The AAN is a long-standing partner in ensuring the rapid dissemination of critical discoveries and improvements stemming from NIH-supported research and is eager to collaborate with the NIH in support of policies that enhance public access, while ensuring that policy changes do not detrimentally impact the research pipeline and the ability of the AAN's journals to continue to provide critical value to researchers and the broader community impacted by neurologic disease.

AAN Publications Impacted by the NIH Public Access Plan

As the leading clinical neurology journal worldwide, Neurology® is directed to physicians concerned with diseases and conditions of the nervous system. The journal's purpose is to advance the field by

presenting new basic and clinical research with emphasis on knowledge that will influence the way neurology is practiced. The journal is at the forefront in disseminating cutting-edge, peer-reviewed information to the neurology community worldwide. Editorial content includes Research, Clinical/Scientific Notes, Views & Reviews (including Medical Hypothesis papers), Issues of Neurological Practice, Historical Neurology, Neurolmages, Humanities, Disputes & Debates: Editors' Choice, and position papers from the American Academy of Neurology. Contents appearing solely online include the Patient Page, CME Quizzes, Podcasts, and play-in-place video.

Neurology Clinical Practice® focuses mainly on two aspects of neurologic care: 1) Clinical research on patient-reported outcomes and quality, including original research articles and meta-analyses/systematic reviews; and 2) Commentaries, reviews, and research articles on general practice, billing and coding, wellness and burnout, diversity and inclusion in the workplace, telehealth, health care policy, and financial management.

1. How to best ensure equity in publication opportunities for NIH-supported investigators.

The request for information (RFI) states that the "NIH seeks information on additional steps it might consider taking to ensure that proposed changes to implementation of the NIH Public Access Plan do not create new inequities in publishing opportunities or reinforce existing ones." As stated previously, the AAN is deeply concerned that the NIH Public Access Plan is likely to contribute to substantial inequity in relation to who has the resources to contribute to the body of peer-reviewed, published research. The AAN believes that the NIH Public Access Plan is predicated on a belief that implementation is unlikely to have a substantial impact on journal sustainability under the existing business model. The AAN believes that the current subscription model used for Neurology® and Neurology Clinical Practice® is equitably accessible to researchers submitting their work as there are no fees for submitting a paper to either publication. Upon submission, authors are able to receive valuable feedback on the paper, prior to the paper being published in a journal within the AAN's family of journals.

The AAN is concerned that the NIH Public Access Plan will result in changes to the underlying publication business model resulting in AAN journals at least partially needing to be funded through article processing charges (APCs) and Other fees borne by authors. While this policy may result in greater immediate access to published literature for individuals who do not subscribe to the AAN's journals, the AAN believes that this policy significantly disadvantages researchers who are either unfunded or have limited funding to allocate towards the APCs and Other fees that are necessitated by the NIH Public Access Plan.

In order to make the peer-reviewed content accessible without an embargo, and in recognition of the AAN's continued support in aiding researcher compliance with NIH requirements, the AAN asks that the NIH policy refrain from requiring reuse rights under licenses that restrict our ability to establish copyright and preserve the downstream revenue associated with the final version of record. The value we provide to our research community is at risk when content is under licenses that allow broad re-use of content, particularly for commercial purposes.

While the NIH Public Access Plan states that "NIH currently allows funding to be used to cover costs of publication, consistent with the NIH Grants Policy Statement, 7.9 Allowability of Costs/Activities. Under the NIH Public Access Policy, NIH has clarified that publication costs, including article processing charges often associated with open access publishing, may be charged to NIH grants and contracts" provided

that certain conditions are met. The AAN is concerned regarding the lack of clarity surrounding the amount of funding that will be available and the length of time for which it may be available. Additionally, it would be helpful for the NIH to precisely define the conditions under which a submitted paper may claim NIH funding and/or under which conditions the public access mandate will apply. It is currently unclear how the NIH Public Access Plan applies to a number of potential cases that a journal may encounter. The AAN requests clarification on each of the scenarios below:

- NIH-supported researchers submit for publication after grant funding has elapsed and they no longer have funding to cover APCs or Other fees.
- Cases in which researchers are not funded by the NIH but cite long-running NIH studies or analyses that use data from NIH-studies.
- Instances in which an author is receiving NIH-funding for a subject Other than the topic of the work that is seeking publication.

The AAN also notes that one of the conditions is that "costs are charged consistently regardless of the source of support." The AAN requests clarification regarding whether the same fees must be charged to all researchers, even those without adequate funding to cover APCs or Other fees that may be covered under an NIH grant or contract. Additionally, the AAN requests clarification regarding the impact of the NIH Public Access Plan on researchers that publish research using data from NIH-funded studies after the relevant study has been concluded and as such do not have NIH support to compensate for APCs and Other necessary fees. Does this requirement preclude journals from accounting for a lack of author resources in determining an appropriate fee? The AAN is concerned that in both of these cases, there will be inequitable access for researchers to access the AAN's robust peer-review and publication process.

Furthermore, the AAN is concerned that access to funding to account for APCs and Other necessary fees may systematically bias the types of research able to receive peer-review and publication within Neurology® and Neurology Clinical Practice®. This is in part due to variability in funding levels for long-standing NIH priorities and is also attributable to long-standing biases, that the NIH is currently working to address, relating to the link between funding gaps and the identities of researchers, as well as the topics chosen by those researchers. Additionally, a recent study found that publishing open access drops significantly for researchers from middle-income countries whose waivers for fees are either smaller discounts or non-existent. The AAN is concerned that the NIH Public Access Plan will exacerbate inequities for the global scientific community by forcing a large-scale shift to open access models that will price underfunded researchers out of the market.

To avoid these adverse consequences, the NIH could consider financial agreements with journals and publishers to directly cover the fees that will be required to support the NIH Public Access Plan, including compensating for changes to the underlying business model. Such agreements could then allow all authors, regardless of funding source, to continue to publish in the publication venue of their choice without directly incurring fees. Alternatively, the NIH could consider including a minimum threshold of funding on which to apply this proposed mandate.

While it may appear that the NIH Public Access Plan is the most equitable for readers, the NIH should be aware that institutional subscribers have tools to assess metrics relating to the value of a subscription,

including the relative volume of journal content that is freely available to the public. Institutions then use this information to determine whether to continue subscribing to a particular journal. The AAN is concerned, absent substantial modification and clarification, that many institutions may decide not to renew existing subscriptions once this plan is implemented, necessitating a substantial modification of the existing business model for Neurology® and Neurology Clinical Practice®. Historically the value of subscribing to individual journals has been evaluated by subscribers at the point of purchase. Journals were therefore required to demonstrate their value to the end user or institution who purchased the subscription. To do so, the AAN provides metrics to illustrate the value of Neurology® and Neurology Clinical Practice®, including both usage data and impact factors. By removing the need to demonstrate quality to the end user or institutional subscriber, and instead shifting towards a business model that is at least partially predicated on fees generated based on the volume of submissions, publications will be incentivized to maximize revenues by accepting as many manuscripts as necessary without regard for quality of science or impact.

In addition to supporting the dissemination of the highest-quality and most impactful research in neurology and neuroscience, the AAN's current subscription-based publishing model supports author equity by providing equal opportunity for all authors to submit for review and publication by the journal, and benefit from the peer review process, as well as the journal's editorial oversight, production, and dissemination without charge. The AAN's peer review and publication process adds substantial value to authors as they refine their submission throughout the peer-review process and to the broader neurology and neuroscience community through the development of supplemental content aimed at enhancing reader understanding of published articles. These substantial additions in value are reflected in the subscription price for AAN journals and the costs borne by the journal to engage in these activities may not be able to be recouped under the NIH Public Access Plan. As such, the AAN is concerned that our ability to continue to develop supplemental content and provide timely and robust peer-review at the same scale will be negatively impacted by this policy. Editorial operations that produce credible, validated, accessible and timely scientific papers may be weakened under the NIH Public Access Plan, due to budget shortfalls. This may result in slower peer review and/or a less rigorous review overall. Editorial offices and publishers are also addressing numerous Other issues, including equity, diversity, and inclusion, scientific and editorial misconduct such as plagiarism, data and image manipulation, conflict of interests, author impersonation or fabrication, papermill output and ethical violations, all of which may be hindered under this plan.

The AAN takes its role in preserving the scientific integrity of research published in our journals very seriously. The reputations of the AAN's journals and the AAN itself relies on being a provider of trusted content. The AAN is committed to expedient but thorough review and publication of research that affects patient care. Maintaining this trusted role in society, at a time when disinformation is rampant, requires a significant amount of investment. Vigilance in publication research integrity and conflict of interest management not only aligns with the AAN's mission but, more importantly, gives confidence to clinicians and researchers that the information we publish has been verified and is reliable. Diligent peer review, management and public disclosures of conflicts, and data and figure integrity checks are vital parts of the process. These services are critical to production of a final product our members can rely upon as they conduct vital research and deliver evidence-based care, but they also require direct and substantial expense. Significant staff training and resources could be endangered if the AAN loses revenue in the form of cancelled subscriptions, insufficient total APC income, and lost licensing fees for

approved reuse of content. The AAN believes it is critical that NIH account for the impact of decreased revenues on our ability to continue to offer the full range of services now protecting the scientific integrity of research published in our journals.

The AAN requests additional clarification regarding author self-deposit of the accepted manuscript on PMC as an acceptable method of compliance with the NIH Public Access Plan. While the NIH requirement is for authors to deposit, many publishers facilitate this submission to ensure the version published by the journal is the one deposited. The AAN is concerned that this policy may necessitate that journals charge additional fees to researchers for deposit to PMC to ensure compliance and consistency in cases in which the author fails to submit directly.

The AAN also requests clarification regarding NIH's statement that it will limit "inappropriate uses" of NIH-supported articles, "such as redistribution of PMC content for sale." Would this include a publisher's reuse of material from their own publications for a derivative commercial product, if that material is also hosted in PMC?

# 2. Steps for improving equity in access and accessibility of publications.

The RFI notes "removal of the currently allowable 12-month embargo period for NIH-supported publications will improve access to these research products for all. As noted in the NIH Public Access Plan, NIH also plans to continue making articles available in human and machine-readable forms to support automated text processing. NIH will also seek ways to improve the accessibility of publications via assistive devices. NIH welcomes input on Other steps that could be taken to improve equity in access to publications by diverse communities of users, including researchers, clinicians and public health officials, students and educators, and Other members of the public."

Although the NIH is not promoting one specific publishing business model, the AAN believes that the NIH Public Access Plan will likely result in a substantial weakening of the current subscription-based model for the AAN's journals, which may require a substantial modification of the existing model to more closely resemble a Gold OA model. The existing Green OA model with a 12-month embargo is currently underwritten by subscription, licensing, and advertising revenue. Removal of the 12-month embargo undermines the AAN's ability to recoup investment in content-related and infrastructure costs including, stipends for editors, validation of publication research integrity, content recruitment, development and enrichment through production of ancillary material, submission and peer review systems, editorial tools such as plagiarism detection, digital platforms, and dissemination. The AAN also invests in the development of capabilities for ensuring that content is tagged and presented in a way that is useful to adaptive devices needed by users with visual and auditory disabilities.

With regard to improving access for individuals outside of the typical subscriber or society, the AAN routinely produces and/or publishes infographics, short form article summaries, and patient pages. All of the AAN's guidelines are also published for free public consumption. Absent a direct link to the hosted page on the Neurology.org website, users on PMC have no chance to discover this content. With zero-embargo, it is possible that usage and visits to Neurology.org will fall substantially and along with it, usage of this added value content. This will discourage the AAN from continuing to invest in this content. Additionally, advertising revenue is a substantial component of the business model supporting Neurology® and Neurology Clinical Practice®. By demanding that all papers that report on NIH funded

research appear in PMC with zero embargo, the NIH is restricting usage of content on the publisher sites and thereby significantly threatening advertising revenue.

As noted above, all of these activities add substantial value for researchers and readers and are reflected in the subscription cost. The AAN is concerned that we will need to investigate new means of supporting content and infrastructure costs by directly charging authors APCs and Other necessary service-based charges. Furthermore, the NIH Public Access Plan erodes the longer-term value of subscriptions for journals with significant amounts of federally funded content, creating a perverse incentive related to publishing NIH-supported research in AAN journals.

In addition to clinical practice guidelines which are immediately made free to the public, the AAN routinely makes Other content that is less than 12 months old freely available to any reader. With a zero-embargo policy, the AAN will be forced to decide whether they can continue to make this content available for free. This will be true regardless of whether the AAN ultimately decides to maintain a subscription model under the NIH Public Access Plan or if the AAN shifts to a Gold OA model.

The work of converting Word files into machine readable, highly tagged extensible markup language (XML) is important, particularly for readers in need of assistive devices. Doing so also aids in search and discovery. The AAN believes there is a duplication in effort in creating XML and metadata for content reporting on NIH funded projects. To support equity in access to publications and to support automated text processing, the NIH could compensate journals or publishers for depositing high quality XML machine readable content instead of processing XML a second time via a licensing agreement. Only 11% of publishers depositing content to PMC have agreements with the National Library of Medicine (NLM) whereby they deposit already parsed and tagged XML. Whether these agreements will continue without an embargo remains to be seen. A licensing arrangement would boost compliance of deposits into PMC. While not every journal or publisher will have the ability to enter into such an arrangement, the majority, including the AAN, are already investing in XML processing.

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

The RFI notes that "NIH proposes to actively monitor trends in publication fees and policies to ensure that they remain reasonable and equitable. NIH seeks information on effective approaches for monitoring trends in publication fees and equity in publication opportunities."

If the NIH moves forward with implementing the Public Access Plan, the AAN recommends that the NIH monitor each publication separately and avoid using average calculations. Costs involved in publishing vary across the industry and are highly dependent on specialty, as well as the nature of the publication. Even within the field of medicine, clinical journal costs vary across practitioner type and specialty. Furthermore, the costs involved in publishing in a highly selective journal, like Neurology® and Neurology Clinical Practice®, both of which receive a high volume of submissions, are vastly different from the costs of publishing in less selective journals or those with lower volumes of submissions. Costs and revenue streams vary drastically depending on many factors such as audience, circulation, ranking, article quality, supplemental materials, number of articles published, field/specialty, and distribution method.

The AAN requests clarification regarding how NIH will operationalize its approach to monitoring costs and impacts on affected communities. Specifically, the AAN requests clarification regarding how the NIH

will determine the affected communities and whether it will include the entire medical publishing ecosystem and the broader neuroscience and neurologic community. The AAN also requests clarification regarding how the NIH will determine whether publication fees and policies are "reasonable and equitable." We note that reasonable and equitable fees may vary greatly across the industry and that there is no one-size-fits-all approach. The AAN also believes that this determination may vary under different models including read and publish models and multi-payer models. Further, services rendered to authors vary by journal, which can affect the overall fee structure. The AAN firmly believes that authors need to be given the freedom to choose the journal most appropriate for their research.

The AAN is concerned that the NIH Public Access Plan may impose substantial additional reporting burden on publishers and urges the NIH to engage in a transparent process to determine and evaluate the most appropriate monitoring method(s). There are several complications in tracking publication fees for the NIH in this scenario. We ask that any method the NIH chooses to take the following into consideration:

- It is not uncommon for authors to report NIH funding on manuscripts related to funded projects many years after the grant is officially closed. These papers will be subject to the mandate and yet no further reports (or direct grant expenses) will be accrued. However, these are direct expenses and therefore should be tracked to fully assess the impact of the proposed plan.
- Open access fees should be specifically tracked as separate from any Other publication fees to truly assess the cost of the mandate. In Other words, non-OA fees (such as page fees, submission fees, and color charges) that may already exist should not be included in the tracking associated with this mandate.
- Discounts given (whether by author request or as a result of society membership) and waivers should be noted in the expense tracking so as to avoid skewing the averages. The NIH should account for whether societies and journals are subsidizing author fees to understand the full impact of the mandate.
- The mandate will likely force institutions and industry to pay for publication fees on papers that report on NIH funded research. If an NIH funded author is on a paper but does not have any grant money left to pay publication fees, another author on the paper or the author's institution will have to pay. To truly understand the impact that this proposed policy is having, the NIH should be tracking exactly who is paying the fees.
- Many NIH funded authors will be able to take advantage of Read and Publish agreements that their institutions have made with publishers. As such, the grant money may not be used to pay publication fees. Still, this is an expense to the authors that ties directly to their grant funding and should be tracked by the NIH to gather a complete picture of the impact of the policy.

#### 4. Early input on considerations to increase findability and transparency of research.

The NIH is requesting "suggestions on any specific issues that should be considered in efforts to improve use of PIDs and metadata, including information about experiences institutions and researchers have had with adoption of different identifiers."

The AAN believes it would be beneficial for PMC to only include metadata for papers that report on NIH funded research, with the accepted content, either the peer reviewed accepted manuscript or the

version of record, being accessible only on the publisher site. The NIH Public Access Plan states that "NIH will continue to enable broad reuse of NIH-supported articles through services that allow for automated retrieval and downloading of full text and metadata, consistent with available license terms. NIH-supported peer-reviewed manuscripts, Other article files as license terms allow, and article metadata are made available by PMC in formats and through channels that enable text mining, large-scale machine-analysis, and computation. These machine-readable article datasets also include retractions, corrections, and expressions of concern." Although the AAN appreciates this commitment, the AAN believes that our recommendation could achieve several important outcomes including:

- Elimination of the administrative burden on authors to make deposits in PMC. NLM could use the already existing CrossRef metadata APIs to fuel PMC as a metadata repository. Compliance would be promoted as publishers would have an incentive to invest in the metadata deposits to CrossRef as the NLM would be a strong driver of traffic to journal sites.
- Elimination of the NLM expense of "processing" content for display in PMC.
- Serving the users by aiding in discoverability of value-add related content on the publisher site, ensuring that addenda are displayed, and mitigating some of the financial damage the NIH Public Access Plan will have on societies and society journals.

If the NIH declines to implement the above recommendation, to minimize implementation burden, the AAN believes that NIH should utilize existing infrastructure already widely adopted across the industry to support findability and transparency of research. The current persistent identifier and metadata structure is supported by publishers through sponsorship and membership in organizations such as CrossRef and ORCID. Publishers also work with National Information Standards Organization to ensure metadata remains current, accessible, and included in the cost to prepare for content dissemination. By adopting persistent identifiers already in use in scholarly publishing, journals can include persistent links to critical pieces of research for the users to access.

Publishers are very interested in and have been early adopters of persistent identifiers in the scholarly communication life cycle. Digital Object Identifiers (DOIs) have been the backbone of online journal publishing since the 1990s. Much like the NIH requirement for grantees to have ORCIDs, many journals, including those within the AAN's family of journals, require or encourage authors to use ORCID to assist in author disambiguation. The AAN's family of journals recently updated our tracking system and authors are now required to use ORCID and FUNDREF as persistent IDs to disambiguate authors and credit funders. Further, publishers make use of FunderID and ROAR identification to again disambiguate human input data received by authors. We encourage the NIH to engage with the AAN, publishers, and the PID community of partners to use or adapt what has already been created. We highly encourage the NIH to employ DOIs for grants as well as require DOIs for datasets published.

Lastly, a commitment from the NIH to adopt persistent identifiers already in use should end the NLM practice of replacing publisher DOIs in the references of papers in PubMed. The NLM does not have permission from publishers or authors to make material changes to the deposited manuscripts. By stripping DOIs from reference links or choosing to include links to the PMC versions instead of the Version of Record, the NLM is unnecessarily restricting the user's access to associated editorials, letters to the editor, podcasts, infographics, and Other added value content hosted by Neurology® and/or Neurology Clinical Practice®.

# **Uploaded File:**

Final-AAN-Comments-on-NIH-Public-Access-Plan.pdf

**Description:** Please see the attached for the full comments from the AAN

Email: mkerschner@aan.com



201 Chicago Avenue Minneapolis, Minnesota 55415

> Tel: (612) 928-6100 Fax: (612) 454-2744

> > AAN.com

#### President

Orly Avitzur, MD, MBA, FAAN Tarrytown, New York

#### **President Elect**

Carlayne E. Jackson, MD, FAAN San Antonio, Texas

#### **Vice President**

Janis M. Miyasaki, MD, MEd, FRCPC, FAAN Edmonton, Alberta, Canada

#### Secretary

Sarah M. Benish, MD, FAAN Minneapolis, Minnesota

#### **Treasurer**

Charles C. Flippen II, MD, FAAN Los Angeles, California

# **Immediate Past President**

James C. Stevens, MD, FAAN Fort Wayne, Indiana

#### **Directors**

Wayne E. Anderson, DO, FAHS, FAAN San Francisco, California

Brenda Banwell, MD, FAAN Philadelphia, Pennsylvania

Bruce H. Cohen, MD, FAAN Chair, Advocacy Committee Akron, Ohio

Charlene E. Gamaldo, MD, FAASM, FAAN Baltimore, Maryland

James N. Goldenberg, MD, FAAN Lake Worth, Florida

Larry B. Goldstein, MD, FAHA, FAAN Lexington, Kentucky

Lily Jung Henson, MD, MMM, FAAN Stockbridge, Georgia

Shannon M. Kilgore, MD, FAAN Palo Alto, California

Brett M. Kissela, MD, MS, FAHA, FAAN Cincinnati, Ohio

Brad C. Klein, MD, MBA, FAAN

Chair, Medical Economics

and Practice Committee

Willow Grove, Pennsylvania

José G. Merino, MD, M Phil, FAHA, FAAN Editor-in-Chief, Neurology® Washington. DC

> Bruce Ovbiagele, MD, MSc, MAS, MBA, FAAN San Francisco, California

Maisha T. Robinson, MD, MSHPM, FAAN Chair, Member Engagement Committee Jacksonville, Florida

> Non-voting Board Member Mary E. Post, MBA, CAE Chief Executive Officer Minneapolis, Minnesota

April 14, 2023

Lawrence A. Tabak, DDS, PhD Acting Director National Institutes of Health 9000 Rockville Pike Bethesda, MD 20892

# **RE:** Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research [NOT-OD-23-091]

Dear Dr. Tabak,

The American Academy of Neurology (AAN) is the world's largest neurology specialty society representing more than 40,000 neurologists and clinical neuroscience professionals. The AAN is dedicated to promoting the highest quality patient-centered neurologic care. A neurologist is a physician with specialized training in diagnosing, treating, and managing disorders of the brain and nervous system. These disorders affect one in six people and include conditions such as multiple sclerosis (MS), Alzheimer's disease (AD), Parkinson's disease, stroke, migraine, epilepsy, traumatic brain injury, ALS, and spinal muscular atrophy.

The AAN greatly appreciates the opportunity to provide feedback in response to the "Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research" from the National Institutes of Health (NIH). While the AAN is supportive of the goal of enhancing public access to the results to NIH-supported research, the AAN is deeply concerned that the NIH Public Access Plan as described in NOT-OD-23-091 will be highly disruptive to the ongoing operations and article quality of *Neurology*® and *Neurology Clinical Practice*®.

The AAN is deeply concerned that the NIH Public Access Plan will result in numerous unintended consequences, resulting from the need for journals like *Neurology® and Neurology Clinical Practice®* to substantially modify their revenue models. The AAN believes that changes to the underlying business model stemming from implementation of the NIH Public Access Plan will likely necessitate a shift of financial responsibility from subscribers to the researchers seeking to have their research published, creating substantial additional barriers for those seeking publication. The rapid implementation of the NIH plan, specifically the elimination of the 12-month embargo, is extremely disruptive and may negatively impact the financial underpinnings of scholarly publishing and dissemination. The AAN is alarmed by the potential for the NIH Public Access Plan to create substantial inequity in those able to contribute to the body of peer-reviewed

published scientific research. The AAN is a long-standing partner in ensuring the rapid dissemination of critical discoveries and improvements stemming from NIH-supported research and is eager to collaborate with the NIH in support of policies that enhance public access, while ensuring that policy changes do not detrimentally impact the research pipeline and the ability of the AAN's journals to continue to provide critical value to researchers and the broader community impacted by neurologic disease.

# **AAN Publications Impacted by the NIH Public Access Plan**

As the leading clinical neurology journal worldwide, *Neurology*® is directed to physicians concerned with diseases and conditions of the nervous system. The journal's purpose is to advance the field by presenting new basic and clinical research with emphasis on knowledge that will influence the way neurology is practiced. The journal is at the forefront in disseminating cutting-edge, peer-reviewed information to the neurology community worldwide. Editorial content includes Research, Clinical/Scientific Notes, Views & Reviews (including Medical Hypothesis papers), Issues of Neurological Practice, Historical Neurology, Neurolmages, Humanities, Disputes & Debates: Editors' Choice, and position papers from the American Academy of Neurology. Contents appearing solely online include the Patient Page, CME Quizzes, Podcasts, and play-in-place video.

Neurology Clinical Practice® focuses mainly on two aspects of neurologic care: 1) Clinical research on patient-reported outcomes and quality, including original research articles and meta-analyses/systematic reviews; and 2) Commentaries, reviews, and research articles on general practice, billing and coding, wellness and burnout, diversity and inclusion in the workplace, telehealth, health care policy, and financial management.

# 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

The request for information (RFI) states that the "NIH seeks information on additional steps it might consider taking to ensure that proposed changes to implementation of the NIH Public Access Plan do not create new inequities in publishing opportunities or reinforce existing ones." As stated previously, the AAN is deeply concerned that the NIH Public Access Plan is likely to contribute to substantial inequity in relation to who has the resources to contribute to the body of peer-reviewed, published research. The AAN believes that the NIH Public Access Plan is predicated on a belief that implementation is unlikely to have a substantial impact on journal sustainability under the existing business model. The AAN believes that the current subscription model used for *Neurology*® and *Neurology Clinical Practice*® is equitably accessible to researchers submitting their work as there are no fees for submitting a paper to either publication. Upon submission, authors are able to receive valuable feedback on the paper, prior to the paper being published in a journal within the AAN's family of journals.

The AAN is concerned that the NIH Public Access Plan will result in changes to the underlying publication business model resulting in AAN journals at least partially needing to be funded through article processing charges (APCs) and other fees borne by authors. While

this policy may result in greater immediate access to published literature for individuals who do not subscribe to the AAN's journals, the AAN believes that this policy significantly disadvantages researchers who are either unfunded or have limited funding to allocate towards the APCs and other fees that are necessitated by the NIH Public Access Plan.

In order to make the peer-reviewed content accessible without an embargo, and in recognition of the AAN's continued support in aiding researcher compliance with NIH requirements, the AAN asks that the NIH policy refrain from requiring reuse rights under licenses that restrict our ability to establish copyright and preserve the downstream revenue associated with the final version of record. The value we provide to our research community is at risk when content is under licenses that allow broad re-use of content, particularly for commercial purposes.

While the NIH Public Access Plan states that "NIH currently allows funding to be used to cover costs of publication, consistent with the NIH Grants Policy Statement, 7.9 Allowability of Costs/Activities. Under the NIH Public Access Policy, NIH has clarified that publication costs, including article processing charges often associated with open access publishing, may be charged to NIH grants and contracts" provided that certain conditions are met. The AAN is concerned regarding the lack of clarity surrounding the amount of funding that will be available and the length of time for which it may be available. Additionally, it would be helpful for the NIH to precisely define the conditions under which a submitted paper may claim NIH funding and/or under which conditions the public access mandate will apply. It is currently unclear how the NIH Public Access Plan applies to a number of potential cases that a journal may encounter. The AAN requests clarification on each of the scenarios below:

- NIH-supported researchers submit for publication after grant funding has elapsed and they no longer have funding to cover APCs or other fees.
- Cases in which researchers are not funded by the NIH but cite long-running NIH studies or analyses that use data from NIH-studies.
- Instances in which an author is receiving NIH-funding for a subject other than the topic of the work that is seeking publication.

The AAN also notes that one of the conditions is that "costs are charged consistently regardless of the source of support." The AAN requests clarification regarding whether the same fees must be charged to all researchers, even those without adequate funding to cover APCs or other fees that may be covered under an NIH grant or contract. Additionally, the AAN requests clarification regarding the impact of the NIH Public Access Plan on researchers that publish research using data from NIH-funded studies after the relevant study has been concluded and as such do not have NIH support to compensate for APCs and other necessary fees. Does this requirement preclude journals from accounting for a lack of author resources in determining an appropriate fee? The AAN is concerned that in both of these cases, there will be inequitable access for researchers to access the AAN's robust peer-review and publication process.

Furthermore, the AAN is concerned that access to funding to account for APCs and other necessary fees may systematically bias the types of research able to receive peer-review and publication within *Neurology® and Neurology Clinical Practice®*. This is in part due to

variability in funding levels for long-standing NIH priorities and is also attributable to long-standing biases, that the NIH is currently working to address, relating to the link between funding gaps and the identities of researchers, as well as the topics chosen by those researchers. Additionally, a recent study found that publishing open access drops significantly for researchers from middle-income countries whose waivers for fees are either smaller discounts or non-existent. The AAN is concerned that the NIH Public Access Plan will exacerbate inequities for the global scientific community by forcing a large-scale shift to open access models that will price underfunded researchers out of the market.

To avoid these adverse consequences, the NIH could consider financial agreements with journals and publishers to directly cover the fees that will be required to support the NIH Public Access Plan, including compensating for changes to the underlying business model. Such agreements could then allow all authors, regardless of funding source, to continue to publish in the publication venue of their choice without directly incurring fees. Alternatively, the NIH could consider including a minimum threshold of funding on which to apply this proposed mandate.

While it may appear that the NIH Public Access Plan is the most equitable for readers, the NIH should be aware that institutional subscribers have tools to assess metrics relating to the value of a subscription, including the relative volume of journal content that is freely available to the public. Institutions then use this information to determine whether to continue subscribing to a particular journal. The AAN is concerned, absent substantial modification and clarification, that many institutions may decide not to renew existing subscriptions once this plan is implemented, necessitating a substantial modification of the existing business model for Neurology® and Neurology Clinical Practice®. Historically the value of subscribing to individual journals has been evaluated by subscribers at the point of purchase. Journals were therefore required to demonstrate their value to the end user or institution who purchased the subscription. To do so, the AAN provides metrics to illustrate the value of Neurology® and Neurology Clinical Practice®, including both usage data and impact factors. By removing the need to demonstrate quality to the end user or institutional subscriber, and instead shifting towards a business model that is at least partially predicated on fees generated based on the volume of submissions, publications will be incentivized to maximize revenues by accepting as many manuscripts as necessary without regard for quality of science or impact.

In addition to supporting the dissemination of the highest-quality and most impactful research in neurology and neuroscience, the AAN's current subscription-based publishing model supports author equity by providing equal opportunity for all authors to submit for review and publication by the journal, and benefit from the peer review process, as well as the journal's editorial oversight, production, and dissemination without charge. The AAN's peer review and publication process adds substantial value to authors as they refine their

\_

<sup>&</sup>lt;sup>1</sup> Hoppe, Travis A, et al. "Topic Choice Contributes to the Lower Rate of NIH Awards to African-American/Black Scientists." Science Advances, American Association for the Advancement of Science, 9 Oct. 2019, https://www.science.org/doi/10.1126/sciadv.aaw7238.

<sup>&</sup>lt;sup>2</sup> Powell, Andrea, et al. "Achieving an Equitable Transition to Open Access for Researchers in Lower and Middle-Income Countries [ICSR Perspectives]." SSRN, International Center for the Study of Research, 12 June 2020, https://papers.ssrn.com/sol3/papers.cfm?abstract\_id=3624782.

submission throughout the peer-review process and to the broader neurology and neuroscience community through the development of supplemental content aimed at enhancing reader understanding of published articles. These substantial additions in value are reflected in the subscription price for AAN journals and the costs borne by the journal to engage in these activities may not be able to be recouped under the NIH Public Access Plan. As such, the AAN is concerned that our ability to continue to develop supplemental content and provide timely and robust peer-review at the same scale will be negatively impacted by this policy. Editorial operations that produce credible, validated, accessible and timely scientific papers may be weakened under the NIH Public Access Plan, due to budget shortfalls. This may result in slower peer review and/or a less rigorous review overall. Editorial offices and publishers are also addressing numerous other issues, including equity, diversity, and inclusion, scientific and editorial misconduct such as plagiarism, data and image manipulation, conflict of interests, author impersonation or fabrication, papermill output and ethical violations, all of which may be hindered under this plan.

The AAN takes its role in preserving the scientific integrity of research published in our journals very seriously. The reputations of the AAN's journals and the AAN itself relies on being a provider of trusted content. The AAN is committed to expedient but thorough review and publication of research that affects patient care. Maintaining this trusted role in society, at a time when disinformation is rampant, requires a significant amount of investment. Vigilance in publication research integrity and conflict of interest management not only aligns with the AAN's mission but, more importantly, gives confidence to clinicians and researchers that the information we publish has been verified and is reliable. Diligent peer review, management and public disclosures of conflicts, and data and figure integrity checks are vital parts of the process. These services are critical to production of a final product our members can rely upon as they conduct vital research and deliver evidence-based care, but they also require direct and substantial expense. Significant staff training and resources could be endangered if the AAN loses revenue in the form of cancelled subscriptions, insufficient total APC income, and lost licensing fees for approved reuse of content. The AAN believes it is critical that NIH account for the impact of decreased revenues on our ability to continue to offer the full range of services now protecting the scientific integrity of research published in our journals.

The AAN requests additional clarification regarding author self-deposit of the accepted manuscript on PMC as an acceptable method of compliance with the NIH Public Access Plan. While the NIH requirement is for authors to deposit, many publishers facilitate this submission to ensure the version published by the journal is the one deposited. The AAN is concerned that this policy may necessitate that journals charge additional fees to researchers for deposit to PMC to ensure compliance and consistency in cases in which the author fails to submit directly.

The AAN also requests clarification regarding NIH's statement that it will limit "inappropriate uses" of NIH-supported articles, "such as redistribution of PMC content for sale." Would this include a publisher's reuse of material from their own publications for a derivative commercial product, if that material is also hosted in PMC?

# 2. Steps for improving equity in access and accessibility of publications.

The RFI notes "removal of the currently allowable 12-month embargo period for NIH-supported publications will improve access to these research products for all. As noted in the NIH Public Access Plan, NIH also plans to continue making articles available in human and machine-readable forms to support automated text processing. NIH will also seek ways to improve the accessibility of publications via assistive devices. NIH welcomes input on other steps that could be taken to improve equity in access to publications by diverse communities of users, including researchers, clinicians and public health officials, students and educators, and other members of the public."

Although the NIH is not promoting one specific publishing business model, the AAN believes that the NIH Public Access Plan will likely result in a substantial weakening of the current subscription-based model for the AAN's journals, which may require a substantial modification of the existing model to more closely resemble a Gold OA model. The existing Green OA model with a 12-month embargo is currently underwritten by subscription, licensing, and advertising revenue. Removal of the 12-month embargo undermines the AAN's ability to recoup investment in content-related and infrastructure costs including, stipends for editors, validation of publication research integrity, content recruitment, development and enrichment through production of ancillary material, submission and peer review systems, editorial tools such as plagiarism detection, digital platforms, and dissemination. The AAN also invests in the development of capabilities for ensuring that content is tagged and presented in a way that is useful to adaptive devices needed by users with visual and auditory disabilities.

With regard to improving access for individuals outside of the typical subscriber or society, the AAN routinely produces and/or publishes infographics, short form article summaries, and patient pages. All of the AAN's guidelines are also published for free public consumption. Absent a direct link to the hosted page on the Neurology.org website, users on PMC have no chance to discover this content. With zero-embargo, it is possible that usage and visits to Neurology.org will fall substantially and along with it, usage of this added value content. This will discourage the AAN from continuing to invest in this content. Additionally, advertising revenue is a substantial component of the business model supporting *Neurology*® and *Neurology Clinical Practice*®. By demanding that all papers that report on NIH funded research appear in PMC with zero embargo, the NIH is restricting usage of content on the publisher sites and thereby significantly threatening advertising revenue.

As noted above, all of these activities add substantial value for researchers and readers and are reflected in the subscription cost. The AAN is concerned that we will need to investigate new means of supporting content and infrastructure costs by directly charging authors APCs and other necessary service-based charges. Furthermore, the NIH Public Access Plan erodes the longer-term value of subscriptions for journals with significant amounts of federally funded content, creating a perverse incentive related to publishing NIH-supported research in AAN journals.

In addition to clinical practice guidelines which are immediately made free to the public, the AAN routinely makes other content that is less than 12 months old freely available to any

reader. With a zero-embargo policy, the AAN will be forced to decide whether they can continue to make this content available for free. This will be true regardless of whether the AAN ultimately decides to maintain a subscription model under the NIH Public Access Plan or if the AAN shifts to a Gold OA model.

The work of converting Word files into machine readable, highly tagged extensible markup language (XML) is important, particularly for readers in need of assistive devices. Doing so also aids in search and discovery. The AAN believes there is a duplication in effort in creating XML and metadata for content reporting on NIH funded projects. To support equity in access to publications and to support automated text processing, the NIH could compensate journals or publishers for depositing high quality XML machine readable content instead of processing XML a second time via a licensing agreement. Only 11% of publishers depositing content to PMC have agreements with the National Library of Medicine (NLM) whereby they deposit already parsed and tagged XML. Whether these agreements will continue without an embargo remains to be seen. A licensing arrangement would boost compliance of deposits into PMC. While not every journal or publisher will have the ability to enter into such an arrangement, the majority, including the AAN, are already investing in XML processing.

# 3. Methods for monitoring evolving costs and impacts on affected communities.

The RFI notes that "NIH proposes to actively monitor trends in publication fees and policies to ensure that they remain reasonable and equitable. NIH seeks information on effective approaches for monitoring trends in publication fees and equity in publication opportunities."

If the NIH moves forward with implementing the Public Access Plan, the AAN recommends that the NIH monitor each publication separately and avoid using average calculations. Costs involved in publishing vary across the industry and are highly dependent on specialty, as well as the nature of the publication. Even within the field of medicine, clinical journal costs vary across practitioner type and specialty. Furthermore, the costs involved in publishing in a highly selective journal, like *Neurology*® and *Neurology Clinical Practice*®, both of which receive a high volume of submissions, are vastly different from the costs of publishing in less selective journals or those with lower volumes of submissions. Costs and revenue streams vary drastically depending on many factors such as audience, circulation, ranking, article quality, supplemental materials, number of articles published, field/specialty, and distribution method.

The AAN requests clarification regarding how NIH will operationalize its approach to monitoring costs and impacts on affected communities. Specifically, the AAN requests clarification regarding how the NIH will determine the affected communities and whether it will include the entire medical publishing ecosystem and the broader neuroscience and neurologic community. The AAN also requests clarification regarding how the NIH will determine whether publication fees and policies are "reasonable and equitable." We note that reasonable and equitable fees may vary greatly across the industry and that there is no one-size-fits-all approach. The AAN also believes that this determination may vary under different models including read and publish models and multi-payer models. Further, services rendered to authors vary by journal, which can affect the overall fee structure. The

AAN firmly believes that authors need to be given the freedom to choose the journal most appropriate for their research.

The AAN is concerned that the NIH Public Access Plan may impose substantial additional reporting burden on publishers and urges the NIH to engage in a transparent process to determine and evaluate the most appropriate monitoring method(s). There are several complications in tracking publication fees for the NIH in this scenario. We ask that any method the NIH chooses to take the following into consideration:

- It is not uncommon for authors to report NIH funding on manuscripts related to funded projects many years after the grant is officially closed. These papers will be subject to the mandate and yet no further reports (or direct grant expenses) will be accrued. However, these are direct expenses and therefore should be tracked to fully assess the impact of the proposed plan.
- Open access fees should be specifically tracked as separate from any other publication fees to truly assess the cost of the mandate. In other words, non-OA fees (such as page fees, submission fees, and color charges) that may already exist should not be included in the tracking associated with this mandate.
- Discounts given (whether by author request or as a result of society membership) and waivers should be noted in the expense tracking so as to avoid skewing the averages.
   The NIH should account for whether societies and journals are subsidizing author fees to understand the full impact of the mandate.
- The mandate will likely force institutions and industry to pay for publication fees on papers that report on NIH funded research. If an NIH funded author is on a paper but does not have any grant money left to pay publication fees, another author on the paper or the author's institution will have to pay. To truly understand the impact that this proposed policy is having, the NIH should be tracking exactly who is paying the fees.
- Many NIH funded authors will be able to take advantage of Read and Publish agreements that their institutions have made with publishers. As such, the grant money may not be used to pay publication fees. Still, this is an expense to the authors that ties directly to their grant funding and should be tracked by the NIH to gather a complete picture of the impact of the policy.

# **4.** Early input on considerations to increase findability and transparency of research

The NIH is requesting "suggestions on any specific issues that should be considered in efforts to improve use of PIDs and metadata, including information about experiences institutions and researchers have had with adoption of different identifiers."

The AAN believes it would be beneficial for PMC to only include metadata for papers that report on NIH funded research, with the accepted content, either the peer reviewed accepted manuscript or the version of record, being accessible only on the publisher site. The NIH Public Access Plan states that "NIH will continue to enable broad reuse of NIH-supported articles through services that allow for automated retrieval and downloading of full text and metadata, consistent with available license terms. NIH-supported peer-reviewed manuscripts,

other article files as license terms allow, and article metadata are made available by PMC in formats and through channels that enable text mining, large-scale machine-analysis, and computation. These machine-readable article datasets also include retractions, corrections, and expressions of concern." Although the AAN appreciates this commitment, the AAN believes that our recommendation could achieve several important outcomes including:

- Elimination of the administrative burden on authors to make deposits in PMC. NLM
  could use the already existing CrossRef metadata APIs to fuel PMC as a metadata
  repository. Compliance would be promoted as publishers would have an incentive to
  invest in the metadata deposits to CrossRef as the NLM would be a strong driver of
  traffic to journal sites.
- Elimination of the NLM expense of "processing" content for display in PMC.
- Serving the users by aiding in discoverability of value-add related content on the
  publisher site, ensuring that addenda are displayed, and mitigating some of the
  financial damage the NIH Public Access Plan will have on societies and society
  journals.

If the NIH declines to implement the above recommendation, to minimize implementation burden, the AAN believes that NIH should utilize existing infrastructure already widely adopted across the industry to support findability and transparency of research. The current persistent identifier and metadata structure is supported by publishers through sponsorship and membership in organizations such as CrossRef and ORCID. Publishers also work with National Information Standards Organization to ensure metadata remains current, accessible, and included in the cost to prepare for content dissemination. By adopting persistent identifiers already in use in scholarly publishing, journals can include persistent links to critical pieces of research for the users to access.

Publishers are very interested in and have been early adopters of persistent identifiers in the scholarly communication life cycle. Digital Object Identifiers (DOIs) have been the backbone of online journal publishing since the 1990s. Much like the NIH requirement for grantees to have ORCIDs, many journals, including those within the AAN's family of journals, require or encourage authors to use ORCID to assist in author disambiguation. The AAN's family of journals recently updated our tracking system and authors are now required to use ORCID and FUNDREF as persistent IDs to disambiguate authors and credit funders. Further, publishers make use of FunderID and ROAR identification to again disambiguate human input data received by authors. We encourage the NIH to engage with the AAN, publishers, and the PID community of partners to use or adapt what has already been created. We highly encourage the NIH to employ DOIs for grants as well as require DOIs for datasets published.

Lastly, a commitment from the NIH to adopt persistent identifiers already in use should end the NLM practice of replacing publisher DOIs in the references of papers in PubMed. The NLM does not have permission from publishers or authors to make material changes to the deposited manuscripts. By stripping DOIs from reference links or choosing to include links to the PMC versions instead of the Version of Record, the NLM is unnecessarily restricting the user's access to associated editorials, letters to the editor, podcasts, infographics, and other added value content hosted by *Neurology*® and/or *Neurology Clinical Practice*®.

# Conclusion

As the world's largest neurology specialty society, the AAN is deeply committed to ensuring that equitable access to the most current and impactful clinical neurology research is widely available. The AAN welcomes the opportunity to continue our longstanding collaborative relationship with the NIH to ensure that any plan that may disrupt the existing business model for the AAN family of journals is implemented in a way that minimizes adverse consequences and achieves the administration's aim of promoting broad access to NIH-funded research. The AAN urges the NIH to heed our recommendations in response to this RFI to ensure continued equitable access to clinical neurology research. Please contact Patty Baskin, the Executive Editor of the AAN's family of journals at <a href="mailto:pbaskin@aan.com">pbaskin@aan.com</a> or Matt Kerschner, the AAN's Director, Regulatory Affairs and Policy at <a href="mailto:mkerschner@aan.com">mkerschner@aan.com</a> with any questions or requests for additional information.

Sincerely,

Orly Avitzur, MD, MBA, FAAN

Orly Chippon MD

President, American Academy of Neurology

I am responding to this RFI: On behalf of an organization

Name: Susan Galandiuk, MD

Name of Organization: American Society of Colon & Rectal Surgeons/ Journal Diseases of the Colon &

Rectum

Type of Organization: Professional org association

Role: Scientific researcher

- 1. How to best ensure equity in publication opportunities for NIH-supported investigators.
- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

# **Uploaded File:**

Galandiuk-DCR-Letter-NIH-Request-for-Comments-4.13.23-.pdf

 $\textbf{Description:} \ \textbf{American Society of Colon \& Rectal Surgeons' Journal "Diseases of the Colon \& Rectum"}$ 

Editor-In-Chief" response to request for comments

Email: s0gala01@louisville.edu



April 13, 2023

Dr. Alondra Nelson

Deputy Assistant to the President and Deputy Director for Science and Society

Performing the Duties of Director, Office of Science & Technology Policy (OSTP)

Subject: Ensuring Free, Immediate, and Equitable Access to Federally Funded Research

Dear Dr. Nelson,

I am the Editor-in-Chief of the Society owned journal *Diseases of the Colon & Rectum*. The journal is the official journal of the American Society of Colon & Rectal Surgeons and is the leading international publication for content relating to diseases affecting the colon and rectum. The society has more than 4,000 members from the United States and across the globe. Our journal has been published for 65 years and is seen as the premier scholarly journal in colorectal surgery. During most of this time, the journal has been published by Wolters Kluwer (formerly as JB Lippincott). I am an academic surgeon, holding an endowed Professorship at a public state University, at which I have worked for more than thirty years. During this time, I have served on many different medical journal editorial boards, as well as on the executive committees of many national and international surgical societies. My comments below are made in this context.

Removal of the 12-month embargo period will adversely affect scholarly publishing in several ways. We strongly urge this not be enacted. Specifically...

- 1. Removal of the 12-month embargo will have a reciprocal effect upon the level of support that publishers provide to academic society partners. This support enables academic societies to publish peer-review journals, to fund educational programs for healthcare professionals, to fund scholarships of fellows and surgeons in training, as well as numerous other educational activities. Taken together, all these programs are dedicated to helping members improve their skills and clinical management of their colorectal surgery patients. The impact of the outreach we provide is global.
  - A) Currently, publishers help fund the introduction of innovative new journal content and methods of education, production of clinical practice guidelines, education of Society members regarding peer review, ethical issues in publication, technical skills in performing

- surgical procedures, information regarding advocacy and many others, all of which improves scholarship in the field of medicine.
- B) Representing the expert consensus of multiple peer-review research studies, our clinical practice guidelines are made "freely" available so that healthcare professionals and all patients may benefit.
- C) Creation of a video library for helping patients deal with stoma problems is an example of how a journal can provide valuable patient resources especially when there is limited access to specialty nursing services.
- D) The fact is all the aforementioned activities lose their funding in an OA pay-to-publish model.
- 2. Removal of the 12-month embargo will also eventually lead to significant reduction or elimination of the current Green-OA model as is desired by Plan S advocates. This would pose a great burden on many authors, researchers and institutions.
  - A) Non-federally funded authors (which in 2021 according to the Web of Science accounted for 69% of total US manuscript output) or those with limited resources would be especially disadvantaged in this scheme as they would need to publish through author supported models (Article Processing Charges [APC]). This would create significant author inequity.
  - B) Researchers with current federal funding will need to allocate more of their limited grant budgets toward APC, leaving less available for performance of research. This is unwise use of federal research funds and will quickly deplete already constrained budgets from performing the valuable research itself. Two examples: (a) funding for an NIH K08 covers salary but reserves only a minimal amount of about \$25,000 per year for the actual research itself, and (b) an NIH R03 has a small budget of \$50,000 per year, which is barely enough to cover the expertise for the personnel and the ever-increasing laboratory and equipment expenses.
  - C) Institutions and academic departments, faced with decreasing clinical reimbursement, will not finance the "pay to publish" model, which would apply to the majority of currently published academic research output. This would stifle academic development and in time, reduce the pipeline of researchers applying for federal funding.

- D) Emphasis on this type of publishing model will encourage the proliferation of so-called "predatory" publishers and journals who seek to profit from APC without an increase in overall quality standards.
- E) In an era, where the rapid speed of publication and online review systems facilitate the proliferation of unethical practices on the behalf of authors, reviewers, and publishers, changes that will lead to financial harm to a publishing model that has supported quality peerreview, and ethical standards in publishing is unwise.

A final note. At a recent medical multidisciplinary conference, held prior to the publication of your call for public comments, I had been asked to speak about the future of medical publishing. During my talk, I asked Department Chairs in the audience how they would fund their unit's manuscript output if the one-year embargo period was ever to be removed and there was a move toward "gold OA." Unanimously... they simply did not have the funding to do so!

I thank you for the opportunity to respond and for your consideration,

Sincerely,

Susan Galandiuk, MD

Editor-in-Chief, Diseases of the Colon & Rectum Price Family Endowed Professor of Surgery

Guliel

Director, Division of Colon & Rectal Surgery University of Louisville

I am responding to this RFI: On behalf of myself

Name: Sonal Sathe

Name of Organization: Virginia Tech

Type of Organization: University

Role: Scientific researcher

# 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Increasing visibility and engagement of those with documented disabilities as defined under the Americans with Disabilities Act is one critical part of advancing research---especially because visibility of NIH-supported investigators with ADA-documented disabilities is often ignored.

#### 2. Steps for improving equity in access and accessibility of publications.

a)Assistive devices are one part of the equation, but examining and determining the user experience of these devices is also critical to advance the NIH agenda for both improving access to publications and as an actual specific aim for research itself. A person with a vision impairment, for instance, needs accessible websites to complete their literature review in order to set up the dissertation---and so much more.

b) In addition, removal of paywalls for certain articles would be most helpful to support open-access initiatives and to remove cost as a barrier. Not all institutions cover all databases or articles, and sometimes those articles are needed to form the basis for new and innovative research.

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

As mentioned above, removal of paywall is critical, but also monitoring the actual incidence and prevalence of said paywalls might be helpful when surveying NIH-supported investigators' efforts in a literature search.

Tracking and monitoring publication fees for journals for open access will be critical for this purpose. PLoS journals (One, Digital Health, Water, etc.) and Frontiers (Public Health, Nutrition, Digital Health) are two examples of journals that tend to have steep fees to publish; a challenge when a student is seeking to get an article published in those venues.

# 4. Early input on considerations to increase findability and transparency of research.

Clear and specific verbiage associated with metadata, and guides for visual and hearing impaired users, are most welcome.

Email: sss20a@vt.edu

I am responding to this RFI: On behalf of myself

Name: Steve Pieper

Name of Organization: Isomics, Inc.

Type of Organization: Other

Type of Organization-Other: Technology consulting and independent research small business.

Role: Scientific researcher

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

The scientific publication process has become a primarily yardstick for determining academic promotions and this should be recognized and separated from the purely scientific role. NIH should encourage academic leaders to recognize contributions like tool development and data curation as critical scientific output. The current system generates too many junk publications that only exist to puff up resumes. The NIH itself is also stuck with this paper-counting bias, and even the Public Access Plan is guilty of assuming that more papers means more progress.

This focus on papers also makes it hard for investigators to invest money in supporting their peers by developing open source software and training people to use it. I believe that as a general rule resources invested to support an open source softwares tools maintained by communities of skilled users and developers result in a much better value to science than the corresponding investment in scholarly publications.

# 2. Steps for improving equity in access and accessibility of publications.

This is all great. Definitely requiring the data and articles to be freely available is a really good thing.

The NIH should also consider how to ensure the quality of the articles, since there is already a lot of dubious literature and machines are learning to generate even more of it.

# 3. Methods for monitoring evolving costs and impacts on affected communities.

NIH should really fully support alternatives to the current for-profit scientific publication model. Researchers provide free labor to these companies in the form of publications and reviews and then they are charged for the publications and blocked by paywalls from reading Others.

#### 4. Early input on considerations to increase findability and transparency of research.

No particular suggestions.

I am responding to this RFI: On behalf of an organization

Name: Thomas Guillemaud

Name of Organization: Peer Community In

Type of Organization: Other

**Type of Organization-Other:** Nonprofit publisher and preprint peer-review service

Role: Scientific researcher

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

The NIH should index preprints that have been peer-reviewed, and in particular those that have been validated by academic preprint peer-review and validation services such as Peer Community In (<a href="https://peercommunityin.org">https://peercommunityin.org</a>), regardless of the source of funding for the studies that form the basis of these publications.

The NIH should publicly state, as Other international research institutions have done (e.g. Coalition S, <a href="https://www.coalition-s.org/statement-on-peer-reviewed-publications/">https://www.coalition-s.org/statement-on-peer-reviewed-publications/</a>), that peer-reviewed and validated articles, including preprints, are considered by the NIH in all its evaluation works to have the same a priori value as articles published in journals after peer-review.

A public statement such as that of Coalition S ("'peer reviewed publications' - defined here as scholarly papers that have been subject to a journal-independent standard peer review process with an implicit or explicit validation- are considered to be of equivalent merit and status as peer-reviewed publications that are published in a recognized journal or on a platform.") would be useful.

- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

Email: <a href="mailto:contact@peercommunityin.org">contact@peercommunityin.org</a>

I am responding to this RFI: On behalf of an organization

Name: Libraries and Sponsored Projects Administration

Name of Organization: University of Minnesota

Type of Organization: University

# 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

We support this intention to remove existing burdens and avoid creating new ones for NIH funded researchers. However, we urge NIH to consider the potential for these policies to ripple and cause inequities for non-funded projects and researchers. Submitting to PMC has been a requirement for NIHfunded research for over a decade, and removing the allowed embargo period will not introduce a compliance burden. However, NIH must be aware of the limitations on institutional capacity to help authors comply with this requirement. Currently, this responsibility falls on the PI or the journal, and care should be taken to not move that responsibility to the institution, which would create inequities for authors at less resourced institutions. The current policy requires only that the text of the accepted (final, peer-reviewed) version of the manuscript be shared. Continuing this will avoid authors being required to pay an article processing charge (APC) for each article that results from their grant. We have been carefully monitoring the development and implementation of Plan S in Europe, which has focused much of its efforts on read-and-publish agreements and transformative journals, which are based on the APC model of open access (OA). We are not alone in our concerns that publishers are taking advantage of the requirements for immediate open access for research funded by cOAlition S members, as we have seen a large increase in the number of publishers who are focusing their efforts on APC-based OA. One path Plan S supports is transformative journals, in which journals avow that they will achieve annual growth of OA content and "flip" to full OA when they reach a certain percentage of OA content published per year. cOAlition S and publishers who have registered their journals as "transformative" have not adequately defined what model the fully OA journals will use. If they all move to APC-based models, significant portions of the world will be prohibited from sharing their research. This will introduce new inequities for all researchers. Those who do not have funding, or do not have sufficient funding, will be unable to share their work. Researchers from the Global South may be affected more dramatically, but in the United States, many researchers do not have grant funding, and many institutions would be unable to pay for all articles from their institution to be made OA. The University of Minnesota publishes more than 8,000 journal articles per year. APCs vary widely in price, but at \$2500 per article, the University would need to find an additional \$20 million to fund publishing—an amount that is insurmountable. Currently, many publishers have aligned their policies with this and allow for sharing of the author-accepted manuscript (AAM) in any non-profit repository after 12 months. It is possible that publishers may be unwilling to alter their policies to allow for immediate deposit of articles to PMC (although if they were to decide not to accept articles from NIH-funded researchers, they would miss out on significant high quality research). To help ensure authors are able to publish in the journals that are most appropriate for their audience, NIH could increase support for alternative methods. Plan S includes a path for compliance that is based on "green" OA, in which the AAM is deposited into a repository and no APC is required (provided the journal is not fully OA). This path affirms longstanding strategies for green open access that predate widespread adoption of APCs, such

as institutional open access policies - while also providing new tools to researchers and Other advocates. The "Rights Retention Strategy" approach has the potential to address the inequities that will arise from continuing or increased reliance on APC-based publishing models. In addition to the members of cOAlition S, the Ligue des Bibliothà ques Europà ennes de Recherche - Association of European Research Libraries (LIBER) support the rights retention strategy for enabling access to publicly funded research. (https://libereurope.eu/article/liber-supports-coalitions-rights-retention-strategy-to-ensureopen-access-to-publicly-funded-research/). If authors will need to negotiate rights to share their articles to comply with NIH's policy, we would like for NIH to provide very specific guidance and templates for authors to use. Many publishers use "click-through" copyright transfer systems that are opaque to the researchers, so there needs to be very clear instructions for how to ensure they do not accidentally agree to something that is counter to NIH policy. COAlition S provided an analysis of an example publishing agreement from Taylor and Francis (T&F copyright advice. Author, beware. February 9, 2023. https://www.coalition-s.org/blog/tf-copyright-advice-author-beware/), which outlines the many ways publishers can use obscure language to conceal from authors what they are committing to when agreeing to publish in a particular journal. The burden of understanding and negotiating this legal agreement should not be solely on the researcher. It should also not be a new burden that is placed on their institution to manage on behalf of their researcher. One solution would be to require publishers accepting NIH-funded manuscripts to indicate clearly in their copyright assignment materials either whether the journal is or is not compliant with NIH publishing requirements, or a statement embedded in their copyright assignment processes that in the event of a conflict between the NIH requirements and that of the journal, the NIH requirements will take precedence. Although NIH will allow for publication fees to make their work publicly accessible to be paid from grant funding, an APC-based publishing system would prevent the many researchers who do not have funding from sharing their research. This would have negative effects on all researchers, including those funded by the NIH. A common theme at the 2023 United Nations Open Science Conference (https://www.un.org/en/library/OS23) was that open science, and open sharing of publications, is necessary for the world to achieve the United Nations' Sustainable Development Goals. For example, researchers from the Global South conduct important research on climate change, which is essential for all, including those in the Global North. NIH should establish policies that proactively avoids predictable adverse outcomes. NIH should also consider increasing support for more equitable publishing models. "Diamond" open access publishing is free for all readers and free for all authors to publish. Support for diamond OA is growing, as demonstrated by investments from Science Europe and statements from Deans at some of the most prestigious universities in the US (https://libraries.mit.edu/news/librariessupport-3/34036/) and researchers in the United Kingdom (https://docs.google.com/document/d/1ZAIPDvECb5Zm1pqAf0I1f0sjcBqPbkPGMvGlhaCz6IM/edit#). Science Europe, cOAlition S, OPERAS, and the French National Research Agency also jointly developed an Action Plan for Diamond Open Access with steps that NIH could consider undertaking to support this open access model (https://zenodo.org/record/6282403#.ZDhEvXbMI2w). Examples of options for NIH support in this space could include direct grants to Diamond OA publishers, support for meetings among these publishers, and educating NIH-funded researchers about Diamond OA journal options. Additionally, because of pressures to optimize "impact" of publications, researchers often prefer "big name" journals over less well known ones—NIH could support Diamond OA by promoting specific Diamond OA journals relevant to NIH areas of focus or by building processes into future grant application assessments that reward diamond OA publication in ways that adjust for lower "impact".

#### 2. Steps for improving equity in access and accessibility of publications.

We support NIH's goals of making full text articles and related metadata available and accessible to the public. We strongly encourage continuing to make the full text and metadata of articles available via API, which enables text-based and text-mining research that is not possible with many closed-access and restrictively licensed articles. We also strongly support NIH's goals of making articles accessible via screen reader and encourage guidance for researchers to make tables and figures more accessible, including providing alternative text as well as descriptive captions. We applaud NIH's desires to make public articles more understandable to a broader audience. NSF already requires PIs to submit brief project outcome reports written for a public audience. We would encourage NIH to adapt a similar policy to increase accessibility of the research to a broader audience. Additionally, we want to encourage as much clarity as possible in the scientific articles to encourage interdisciplinary collaboration; for example, including less jargon, using active voice, and clearly defining abbreviations.

# 3. Methods for monitoring evolving costs and impacts on affected communities.

We appreciate NIH's commitment to ensure that publication fees do not increase due to the new public access policy. However, publication fees for many journals are already unreasonable and inequitable. Based on data from Web of Science, the average APC for the top 10 journals in which NIH-funded articles were published had an average APC of \$3,434, and APCs can reach as high as \$11,690 per article. These costs are already consuming significant portions of NIH grants, reducing the amount of funding available for conducting research. It is important to monitor publisher fees, but NIH must be willing and able to act if publishers increase fees to ensure researchers do not face ever-increasing burdens for publication. NIH must define what they consider to be unreasonable, and must take into account that, based on past experience, publishers will continually increase article processing charges (APCs) and are likely to set APCs at the maximum that NIH allows. The current public access policies and ones that will result from the 2022 OSTP memo are based on providing access to federally funded research for taxpayers. These policies are motivated by ensuring the public has timely access to the results of federally funded research. It remains important to find the proper balance between ready access to results and ensuring that federal research dollars are primarily devoted to conducting the actual research, rather than paying publishers.

#### 4. Early input on considerations to increase findability and transparency of research.

In order for all aspects of NIH funded research to be available and findable, we strongly encourage NIH to urge researchers to adopt a standard data citation method to link the articles with the associated datasets. We would also like to see guidance from NIH and Other scientific communities on how best to apply PIDs to various parts of a larger study in order to make sure the components are clearly linked, identified, and findable. For example, some repositories assign DOIs for each file within a study, while Others assign a global DOI for the set of files within the project. Unregulated proliferation of PIDs likely will make findability MORE difficult as individual datasets or articles may be associated with multiple identifiers and cited inconsistently. Linkages between components and PIDS associated with the research study should be both human readable and machine actionable, and ideally in a central metadata aggregator. AnOther consideration for PIDs is the cost associated with minting them - DOIs are costly for repositories or entities who are creating them. However, less costly PIDs (such as ARKs and handles) lack the central metadata infrastructure for discoverability that DOI agencies like DataCite and CrossRef provide.

**Uploaded File:** 

 $RFI-NIH-public-access-response\_UMN\_UL-SPA\_2023.pdf$ 

**Description:** PDF of comments

Email: hunt0081@umn.edu

# Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research

Comments from University of Minnesota Libraries (UL) & Sponsored Projects Administration (SPA)

#### Authors:

Allison Langham-Putrow, Scholarly Communications and Engineering Liaison Librarian (UL) Alicia Hofelich Mohr, PhD, Research Support Coordinator (LATIS, an integral UL partner) Jenny McBurney, Interim Government Publications and Social Sciences Librarian (UL) Pamela Webb, Associate Vice President for Research Administration (SPA) Shanda Hunt, Public Health Librarian and Data Curation Specialist (UL) Shannon Farrell, Interim Research Data Services Librarian (UL)

1. How to best ensure equity in publication opportunities for NIH-supported investigators. The NIH Public Access Plan aims to maintain the existing broad discretion for researchers and authors to choose how and where to publish their results. Consistent with current practice, the NIH Public Access Plan allows the submission of final published articles to PubMed Central (PMC) (in cases where a formal agreement is in place) to minimize the compliance burden on NIH-supported researchers and also maintains the flexibility of NIH-supported researchers to submit the final peer-reviewed manuscript. NIH seeks information on additional steps it might consider taking to ensure that proposed changes to implementation of the NIH Public Access Policy do not create new inequities in publishing opportunities or reinforce existing ones.

We support this intention to remove existing burdens and avoid creating new ones for NIH funded researchers. However, we urge NIH to consider the potential for these policies to ripple and cause inequities for non-funded projects and researchers. Submitting to PMC has been a requirement for NIH-funded research for over a decade, and removing the allowed embargo period will not introduce a compliance burden. However, NIH must be aware of the limitations on institutional capacity to help authors comply with this requirement. Currently, this responsibility falls on the PI or the journal, and care should be taken to not move that responsibility to the institution, which would create inequities for authors at less resourced institutions. The current policy requires only that the text of the accepted (final, peer-reviewed) version of the manuscript be shared. Continuing this will avoid authors being required to pay an article processing charge (APC) for each article that results from their grant. We have been carefully monitoring the development and implementation of Plan S in Europe, which has focused much of its efforts on read-and-publish agreements and transformative journals, which are based on the APC model of open access (OA). We are not alone in our concerns that publishers are taking advantage of the requirements for immediate open access for research funded by cOAlition S members, as we have seen a large increase in the number of publishers who are focusing their

efforts on APC-based OA. One path Plan S supports is transformative journals, in which journals avow that they will achieve annual growth of OA content and "flip" to full OA when they reach a certain percentage of OA content published per year. cOAlition S and publishers who have registered their journals as "transformative" have not adequately defined what model the fully OA journals will use. If they all move to APC-based models, significant portions of the world will be prohibited from sharing their research. This will introduce new inequities for all researchers. Those who do not have funding, or do not have sufficient funding, will be unable to share their work. Researchers from the Global South may be affected more dramatically, but in the United States, many researchers do not have grant funding, and many institutions would be unable to pay for all articles from their institution to be made OA. The University of Minnesota publishes more than 8,000 journal articles per year. APCs vary widely in price, but at \$2500 per article, the University would need to find an additional \$20 million to fund publishing—an amount that is insurmountable. Currently, many publishers have aligned their policies with this and allow for sharing of the author-accepted manuscript (AAM) in any non-profit repository after 12 months. It is possible that publishers may be unwilling to alter their policies to allow for immediate deposit of articles to PMC (although if they were to decide not to accept articles from NIH-funded researchers, they would miss out on significant high quality research). To help ensure authors are able to publish in the journals that are most appropriate for their audience, NIH could increase support for alternative methods. Plan S includes a path for compliance that is based on "green" OA, in which the AAM is deposited into a repository and no APC is required (provided the journal is not fully OA). This path affirms longstanding strategies for green open access that predate widespread adoption of APCs, such as institutional open access policies - while also providing new tools to researchers and other advocates. The "Rights Retention Strategy" approach has the potential to address the inequities that will arise from continuing or increased reliance on APC-based publishing models. In addition to the members of cOAlition S, the Ligue des Bibliothèques Européennes de Recherche – Association of European Research Libraries (LIBER) support the rights retention strategy for enabling access to publicly funded research.

(https://libereurope.eu/article/liber-supports-coalitions-rights-retention-strategy-to-ensure-open-access -to-publicly-funded-research/). If authors will need to negotiate rights to share their articles to comply with NIH's policy, we would like for NIH to provide very specific guidance and templates for authors to use. Many publishers use "click-through" copyright transfer systems that are opaque to the researchers, so there needs to be very clear instructions for how to ensure they do not accidentally agree to something that is counter to NIH policy. COAlition S provided an analysis of an example publishing agreement from Taylor and Francis (T&F copyright advice. Author, beware. February 9, 2023. https://www.coalition-s.org/blog/tf-copyright-advice-author-beware/), which outlines the many ways publishers can use obscure language to conceal from authors what they are committing to when agreeing to publish in a particular journal. The burden of understanding and negotiating this legal agreement should not be solely on the researcher. It should also not be a new burden that is placed on their institution to manage on behalf of their researcher. One solution would be to require publishers accepting NIH-funded manuscripts to indicate clearly in their copyright assignment materials either whether the journal is or is not compliant with NIH publishing requirements, or a statement embedded in their copyright assignment processes that in the event of a conflict between the NIH requirements and that of the journal, the NIH requirements will take precedence. Although NIH will allow for publication fees to make their work publicly accessible to be paid from grant funding, an APC-based publishing system would prevent the many researchers who do not have funding from

sharing their research. This would have negative effects on all researchers, including those funded by the NIH. A common theme at the 2023 United Nations Open Science Conference (https://www.un.org/en/library/OS23) was that open science, and open sharing of publications, is necessary for the world to achieve the United Nations' Sustainable Development Goals. For example, researchers from the Global South conduct important research on climate change, which is essential for all, including those in the Global North. NIH should establish policies that proactively avoids predictable adverse outcomes. NIH should also consider increasing support for more equitable publishing models. "Diamond" open access publishing is free for all readers and free for all authors to publish. Support for diamond OA is growing, as demonstrated by investments from Science Europe and statements from Deans at some of the most prestigious universities in the US (https://libraries.mit.edu/news/libraries-support-3/34036/) and researchers in the United Kingdom (https://docs.google.com/document/d/1ZAIPDvECb5Zm1pqAf0I1f0sjcBqPbkPGMvGIhaCz6IM/edit#). Science Europe, cOAlition S, OPERAS, and the French National Research Agency also jointly developed an Action Plan for Diamond Open Access with steps that NIH could consider undertaking to support this open access model (https://zenodo.org/record/6282403#.ZDhEvXbMI2w). Examples of options for NIH support in this space could include direct grants to Diamond OA publishers, support for meetings among these publishers, and educating NIH-funded researchers about Diamond OA journal options. Additionally, because of pressures to optimize "impact" of publications, researchers often prefer "big name" journals over less well known ones-NIH could support Diamond OA by promoting specific Diamond OA journals relevant to NIH areas of focus or by building processes into future grant application assessments that reward diamond OA publication in ways that adjust for lower "impact".

2. Steps for improving equity in access and accessibility of publications. Removal of the currently allowable 12-month embargo period for NIH-supported publications will improve access to these research products for all. As noted in the NIH Public Access Plan, NIH also plans to continue making articles available in human and machine-readable forms to support automated text processing. NIH will also seek ways to improve the accessibility of publications via assistive devices. NIH welcomes input on other steps that could be taken to improve equity in access to publications by diverse communities of users, including researchers, clinicians and public health officials, students and educators, and other members of the public.

We support NIH's goals of making full text articles and related metadata available and accessible to the public. We strongly encourage continuing to make the full text and metadata of articles available via API, which enables text-based and text-mining research that is not possible with many closed-access and restrictively licensed articles. We also strongly support NIH's goals of making articles accessible via screen reader and encourage guidance for researchers to make tables and figures more accessible, including providing alternative text as well as descriptive captions. We applaud NIH's desires to make public articles more understandable to a broader audience. NSF already requires PIs to submit brief project outcome reports written for a public audience. We would encourage NIH to adapt a similar policy to increase accessibility of the research to a broader audience. Additionally, we want to encourage as much clarity as possible in the scientific articles to

encourage interdisciplinary collaboration; for example, including less jargon, using active voice, and clearly defining abbreviations.

3. Methods for monitoring evolving costs and impacts on affected communities. NIH proposes to actively monitor trends in publication fees and policies to ensure that they remain reasonable and equitable. NIH seeks information on effective approaches for monitoring trends in publication fees and equity in publication opportunities.

We appreciate NIH's commitment to ensure that publication fees do not increase due to the new public access policy. However, publication fees for many journals are already unreasonable and inequitable. Based on data from Web of Science, the average APC for the top 10 journals in which NIH-funded articles were published had an average APC of \$3,434, and APCs can reach as high as \$11,690 per article. These costs are already consuming significant portions of NIH grants, reducing the amount of funding available for conducting research. It is important to monitor publisher fees, but NIH must be willing and able to act if publishers increase fees to ensure researchers do not face ever-increasing burdens for publication. NIH must define what they consider to be unreasonable, and must take into account that, based on past experience, publishers will continually increase article processing charges (APCs) and are likely to set APCs at the maximum that NIH allows. The current public access policies and ones that will result from the 2022 OSTP memo are based on providing access to federally funded research for taxpayers. These policies are motivated by ensuring the public has timely access to the results of federally funded research. It remains important to find the proper balance between ready access to results and ensuring that federal research dollars are primarily devoted to conducting the actual research, rather than paying publishers.

4. Early input on considerations to increase findability and transparency of research. Section IV of the NIH Public Access Plan is a first step in developing the NIH's updated plan for persistent identifiers (PIDs) and metadata, which will be submitted to OSTP by December 31, 2024. NIH seeks suggestions on any specific issues that should be considered in efforts to improve use of PIDs and metadata, including information about experiences institutions and researchers have had with adoption of different identifiers.

In order for all aspects of NIH funded research to be available and findable, we strongly encourage NIH to urge researchers to adopt a standard data citation method to link the articles with the associated datasets. We would also like to see guidance from NIH and other scientific communities on how best to apply PIDs to various parts of a larger study in order to make sure the components are clearly linked, identified, and findable. For example, some repositories assign DOIs for each file within a study, while others assign a global DOI for the set of files within the project. Unregulated proliferation of PIDs likely will make findability MORE difficult as individual datasets or articles may be associated with multiple identifiers and cited inconsistently. Linkages between components and PIDS associated with the research study should be both human readable and machine actionable, and ideally in a central metadata aggregator. Another consideration for PIDs is the cost associated with minting them - DOIs are costly for repositories or entities who are creating them. However, less costly

PIDs (such as ARKs and handles) lack the central metadata infrastructure for discoverability that DOI agencies like DataCite and CrossRef provide.

I am responding to this RFI: On behalf of myself

Name: Carl Tuttle

Type of Organization: Not applicable

Role: Patient advocate

# 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

STOP propagating the false Lyme disease narrative on NIH funded research through omission of the truth, facts, and opposing scientific references. End the debate and find a cure for persistent/chronic infection.

The Lyme disease debacle stems from the NIH funded "Klempner Antibiotic Trials" which set the stage for treatment denial leaving hundreds of thousands (if not millions worldwide) in a debilitated state. Patient testimony across America is describing a disease that is destroying lives, ending careers while leaving its victim in financial ruin.

These "antibiotic trials" in the early 2000's were stopped short at 90 days whereas there are many known infections requiring months to years of antibiotics to clear the infection.

Leprosy for example (Now known as Hanson's disease) is curable with long term antibiotics. In some cases, it may take two years to clear the infection.

CDC: Hansen's Disease (Leprosy)

# http://www.cdc.gov/leprosy/treatment/index.html

On a personal note, it took two years to clear a chronic prostatitis in my early twenties and when symptoms returned no one questioned the need to prescribe additional antibiotics or a different combination. It was the advent of Bactrim that finally cleared the infection.

These so-called "antibiotic trials" were stopped at 90 days and prove nothing. In 2017 scientists at the Tulane National Primate Research Center reported evidence of persistent and metabolically active B. burgdorferi after antibiotic treatment in rhesus macaques as mentioned in the NIAID webpage below:

National Institute of Allergy and Infectious Diseases "Chronic Lyme Disease"

#### https://www.niaid.nih.gov/diseases-conditions/chronic-lyme-disease

Other researchers are finding the same results in humans; current antimicrobials are not working as described in the letter below addressed to Dr. Mark Klempner.

700 articles LYME Evidence of Persistence (Personal Dropbox storage area)

 $\frac{\text{https://www.dropbox.com/s/n09sk90eo6xz7ua/700\%20articles\%20LYME\%20EvidenceofPersistence-}{\text{V2.pdf?dl=0}}$ 

This pathogen requires an all-out Manhattan project to find a cure! Time to step up the efforts here and STOP propagating the false Lyme disease narrative through omission of the truth, facts, and opposing scientific references. End the debate and find a cure for persistent/chronic infection.

Letter to Dr. Mark Klempner: (For the record, there was no response)

It should be noted that Klempner is presently looking for his piece of the Lyme vaccine pie.

https://theconversation.com/a-lyme-disease-vaccine-doesnt-exist-can-a-seasonal-shot-help-slow-the-epidemic-spread-by-ticks-138230

----- Original Message -----

From: Carl Tuttle

To: mark.klempner@umassmed.edu

Cc: michael.collins@umassmed.edu, ddutko@hanszenlaporte.com, ryan.kantor@usdoj.gov, michelle.seltzer@usdoj.gov, william.rinner@usdoj.gov, makan.delrahim@usdoj.gov, tickbornedisease@hhs.gov, "Elias, John", officeofthechancellor@umassmed.edu

Date: 04/27/2018 7:53 AM

Subject: Persistent Borrelia Infection in Patients with Ongoing Symptoms of Lyme Disease

April 27, 2018

University of Massachusetts Medical School

55 Lake Avenue North

Worcester, Massachusetts 01655

Attn: Mark S. Klempner, MD, Executive Vice Chancellor, MassBiologics

Dr. Klempner,

I would like to call attention to the attached study recently identifying chronic Lyme disease in twelve patients from Canada.

Persistent Borrelia Infection in Patients with Ongoing Symptoms of Lyme Disease

http://www.mdpi.com/2227-9032/6/2/33

All of these patients were culture positive for infection (genital secretions, skin "Morgellons" and blood) even after multiple years on antibiotics so there was no relief from current antimicrobials. Some of these patients had taken as many as eleven different types of antibiotics.

In contrast, your 2001 antibiotic treatment study found; "no evidence of B. burgdorferi in a total of more than 700 different blood and cerebrospinal fluid samples from the 129 patients in these studies."

Two Controlled Trials of Antibiotic Treatment in Patients with Persistent Symptoms and a History of Lyme Disease

http://www.nejm.org/doi/full/10.1056/NEJM200107123450202#article\_references#t=references

Not a single positive Dr. Klempner? Doesn't this statistically prove that your methodology was fatally flawed?

Did you culture skin and genital secretions as the Middelveen paper reports? It would appear that you conveniently stopped looking after your results supported the existing thirty year dogma; chronic Lyme does not exist.

Persistent Lyme disease is not new and has been intentionally/deceitfully suppressed for decades as described in the Vicki Logan case identified in the following letter to past CDC Director Barbara Fitzgerald:

https://www.dropbox.com/s/xaul84dqmqgbre0/Brenda%20Fitzgerald%20MD%20Director%20CDC.docx ?dl=0

In 1991 B. burgdorferi had been isolated in culture from Vicki Logan's CSF (CDC's laboratory in Fort Collins CO.) despite prior treatment with 21 days of IV cefotaxime and 4 months of oral minocycline.

The dishonest science here in the U.S. has denied chronic Lyme which stifled research to find a curative approach. Now the rest of the world is suffering.

We have lost nearly four decades to this 21st century plague due to the racketeering scheme identified in the RICO lawsuit filed by SHRADER & ASSOCIATES, LLP against the Infectious Disease Society of America, seven IDSA Panelists and eight insurance companies. The U.S. Centers for Disease Control has aligned itself with the seven IDSA Panelists identified in this lawsuit.

#### Court Document:

# https://www.courthousenews.com/wp-content/uploads/2017/11/LymeDisease.pdf

Lyme is an incurable disease when not treated immediately which is spreading across North America and deceitfully misclassified as a low-risk and non-urgent health issue. Patient experience is describing a disease that is destroying lives, ending careers, causing death and disability while leaving victims in financial ruin. Current antimicrobials are ineffective for eradicating all forms of the Borrelia spirochete.

Public outcry has been ignored for decades while the Centers for Disease Control sat on evidence that this infection was not easily treated with a one size fits all treatment approach as dictated by the Infectious Diseases Society of America.

Once again your studies were fatally flawed while supporting the controlling dogma leaving hundreds of thousands if not millions worldwide with a persistent infection and absolutely no relief. We have anOther AIDS on our hands.

Carl Tuttle

Independent Researcher

Lyme Endemic Hudson, NH

Cc: -Michael F. Collins, Chancellor

- -The Tick Borne Disease Working Group
- -US Department of Justice
- -Daniel R. Dutko, HANSZEN LAPORTE
- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

Email: runagain@comcast.net

I am responding to this RFI: On behalf of an organization

Name: Greg Tananbaum

Name of Organization: Open Research Funders Group

Type of Organization: Other

Type of Organization-Other: Philanthropic Network

Role: Institutional official

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

The proposed NIH guidance promotes compliance via the archiving of articles in agency-designated repositories (PubMed Central, in the case of NIH). This guidance wisely balances the broad freedom that funded researchers enjoy in deciding where to publish their results with the taxpayers' interest in ensuring federal funds don't inadvertently exacerbate research ecosystem inequities. Paywalls limit access to knowledge, limit replication and reproducibility, and stifle civic engagement in science. Replacing paywalls with exorbitant open access article processing charges (APCs) would potentially trade one set of inequities for anOther, creating a two-tiered system in which authors outside of wellfunded R1 institutions lack the financial wherewithal to publish in some prestigious, brand-name journals. A repository-mediated ("green") route to federal policy compliance, as NIH allows/supports through manuscript deposit in PubMed Central, is an effective way to reduce the impact on younger researchers, women, scholars at minority-serving institutions, and Others who are more likely to be disadvantaged by an APC-dominant publishing system (see, for example, the AAAS survey "Exploring the Hidden Impacts of Open Access Financing Mechanisms"). Note that this input is also intended to address the "Monitoring Evolving Costs & Impacts" request for information proffered by NIH. We also encourage the NIH to explore strategies to support preprints as a mechanism for ensuring equitable, low-cost, and timely access to federally funded research.

Additionally, the NIH should consider providing funded researchers clear guidance on rights retention, building on guidance developed by Other funder groups (e.g., cOAlition S) and the larger academic community. Expecting scientists to be experts not only in biomedicine, but also in the labyrinthine world of copyright law, presents an undue burden. The NIH should make it as easy as possible for grantees to retain sufficient rights to make copies of their papers available and reusable in PubMed Central. We appreciate NIH's inclusion of rights retention considerations in this RFI as a signal of this issue's centrality to a comprehensive public access strategy.

# 2. Steps for improving equity in access and accessibility of publications.

One area of potential improvement for the NIH's draft plan is with respect to reuse rights for shared research, which the OSTP guidance includes as an important consideration. While the draft plan does say, "NIH will continue to promote the broadest possible reuse of its supported articles", it does not include an open licensing requirement that would codify and maximize reuse rights. This lack of specificity means researchers could potentially deposit both articles (and data) under a variety of licenses or conditions that could significantly restrict how these materials can be built upon by

researchers and the broader community. A CC BY license or functional equivalent is the best way to enable text and data mining computational uses, and educational reuse. Importantly, from an inclusivity standpoint, this form of licensing is the best way to ensure content accessibility via assistive devices. The ORFG also appreciates the NIH's expansive definition of "accessibility" to emphasize that a range of individuals and communities - including those needing assistive devices and community members not well-versed in scientific jargon - are not presently able to fully engage with federally funded research. We would be pleased to engage with the NIH to identify practical solutions to these limitations.

# 3. Methods for monitoring evolving costs and impacts on affected communities.

Please see "How to best ensure equity in publication opportunities for NIH-supported investigators" response.

#### 4. Early input on considerations to increase findability and transparency of research.

The NIH should include specific, actionable guidance on persistent identifiers (PIDs) and metadata to its funded researchers. The ORFG encourages the NIH and Other federal agencies to embrace de facto community standards where they exist. These include digital object identifiers (DOIs) for articles, datasets and data management plans, ORCIDs for authors, and RORs for institutions. In the interest of making policy compliance as easy as possible for individual researchers, the NIH should coordinate with Other agencies and the National Science and Technology Council's (NSTC) Subcommittee on Open Science, to align on PID and metadata best practices. The ORFG would welcome the engagement of the NIH and Other federal agencies in the community we have nurtured since fall 2022 to improve research output tracking. This group is uniquely positioned - with its cross-sector expertise drawing from funders, higher education, technology providers, publishers, standards bodies, and international organizations - to provide such guidance on best practices.

# **Uploaded File:**

ORFG-NIH-2023-Public-Access-RFI-response.pdf

**Description:** Full response to the "NIH Plan to Enhance Public Access to the Results of NIH-Supported Research" request for public input submitted on behalf of the Open Research Funders Group

Email: greg@orfg.org

This response to the "NIH Plan to Enhance Public Access to the Results of NIH-Supported Research" request for public input is submitted on behalf of the Open Research Funders Group. The Open Research Funders Group (ORFG) is a partnership of 25 philanthropic organizations committed to the open sharing of research outputs. We believe openness is better for philanthropy, better for research, and better for society. Open research accelerates the pace of discovery, reduces information–sharing gaps, encourages innovation, and promotes reproducibility. Collectively, the ORFG members hold assets in excess of \$250 billion, with total annual giving in the \$12 billion range. Members' interests range the entirety of the disciplinary spectrum, including life sciences, physical sciences, social sciences, and the humanities. This response has been prepared by Greg Tananbaum and Dr. Erin McKiernan, Director and Community Manager (respectively) of the ORFG, in conjunction with representatives of the ORFG membership.

The Open Research Funders Group applauds both the substance of the NIH's draft plan and the added step of making it available for public comment. From a process perspective, the NIH's approach reinforces the federal government's stated desire to co-develop practical public access strategies in a transparent and inclusive manner. The plan itself identifies practical mechanisms for the timely sharing of scholarly publications and research data. The draft plan wisely builds upon the lessons learned by NIH through both their long-term stewardship of PubMed Central and their recent rollout of the 2023 Data Management and Sharing Policy. In this regard, the plan articulates clear, easy-to-follow guidance for grantees.

The NIH has requested feedback on four specific areas, which the ORFG provides below. Our perspective is that this guidance should be considered by all federal agencies and departments as they draft plans to address the OSTP's "Ensuring Free, Immediate, and Equitable Access to Federally Funded Research" memorandum at scale. Consistency across federal funding bodies with respect to best practices and standards will make it easier for (a) adjacent sectors (including private philanthropies and higher education institutions) to align their incentive structures to reinforce the key principles of the OSTP memo; and (b) funded researchers to understand and adhere to emerging research sharing norms and good practices.

• **Equity in Publication Opportunities.** The proposed NIH guidance promotes compliance via the archiving of articles in agency-designated repositories (PubMed Central, in the case of NIH). This guidance wisely balances the broad freedom that

funded researchers enjoy in deciding where to publish their results with the taxpayers' interest in ensuring federal funds don't inadvertently exacerbate research ecosystem inequities. Paywalls limit access to knowledge, limit replication and reproducibility, and stifle civic engagement in science. Replacing paywalls with exorbitant open access article processing charges (APCs) would potentially trade one set of inequities for another, creating a two-tiered system in which authors outside of well-funded R1 institutions lack the financial wherewithal to publish in some prestigious, brand-name journals. A repository-mediated ("green") route to federal policy compliance, as NIH allows/supports through manuscript deposit in PubMed Central, is an effective way to reduce the impact on younger researchers, women, scholars at minority-serving institutions, and others who are more likely to be disadvantaged by an APC-dominant publishing system (see, for example, the AAAS survey "Exploring the Hidden Impacts of Open Access Financing Mechanisms"). Note that this input is also intended to address the "Monitoring Evolving Costs & Impacts" request for information proffered by NIH. We also encourage the NIH to explore strategies to support preprints as a mechanism for ensuring equitable, low-cost, and timely access to federally funded research.

Additionally, the NIH should consider providing funded researchers clear guidance on rights retention, building on guidance developed by other funder groups (e.g., coalitions) and the larger academic community. Expecting scientists to be experts not only in biomedicine, but also in the labyrinthine world of copyright law, presents an undue burden. The NIH should make it as easy as possible for grantees to retain sufficient rights to make copies of their papers available and reusable in PubMed Central. We appreciate NIH's inclusion of rights retention considerations in this RFI as a signal of this issue's centrality to a comprehensive public access strategy.

• Equity in Access and Accessibility of Publications. One area of potential improvement for the NIH's draft plan is with respect to reuse rights for shared research, which the OSTP guidance includes as an important consideration. While the draft plan does say, "NIH will continue to promote the broadest possible reuse of its supported articles", it does not include an open licensing requirement that would codify and maximize reuse rights. This lack of specificity means researchers could potentially deposit both articles (and data) under a variety of licenses or conditions that could significantly restrict how these materials can be built upon by researchers and the broader community. A <u>CC BY license</u> or functional equivalent is the best way

to enable text and data mining computational uses, and educational reuse. Importantly, from an inclusivity standpoint, this form of licensing is the best way to ensure content accessibility via assistive devices. The ORFG also appreciates the NIH's expansive definition of "accessibility" to emphasize that a range of individuals and communities – including those needing assistive devices and community members not well-versed in scientific jargon – are not presently able to fully engage with federally funded research. We would be pleased to engage with the NIH to identify practical solutions to these limitations.

• Increasing Findability and Transparency of Research. The NIH should include specific, actionable guidance on persistent identifiers (PIDs) and metadata to its funded researchers. The ORFG encourages the NIH and other federal agencies to embrace de facto community standards where they exist. These include digital object identifiers (DOIs) for articles, datasets and data management plans, ORCIDs for authors, and RORs for institutions. In the interest of making policy compliance as easy as possible for individual researchers, the NIH should coordinate with other agencies and the National Science and Technology Council's (NSTC) Subcommittee on Open Science, to align on PID and metadata best practices. The ORFG would welcome the engagement of the NIH and other federal agencies in the community we have nurtured since fall 2022 to improve research output tracking. This group is uniquely positioned – with its cross-sector expertise drawing from funders, higher education, technology providers, publishers, standards bodies, and international organizations – to provide such guidance on best practices.

The Open Research Funders Group wishes to again express our gratitude and support for the work of the NIH, the OSTP, and other federal agencies to advance a more open, equitable, and inclusive research ecosystem. We appreciate the opportunity to comment on this draft plan, and we are eager to assist in its eventual rollout.

I am responding to this RFI: On behalf of an organization

Name: John Willinsky

Name of Organization: Public Knowledge Project

Type of Organization: University

Role: Scientific researcher

# 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

I have attached a letter, signed by a publisher, legal scholars, librarian, program manager, and two researchers, outlining why I believe that the NIH can better serve its mission and the progress of science, capitalize on its public access experience, and build on its leadership in this area by moving beyond an enhancement of its current policies to engage in discussions aimed at sustainable, universal public access on a global scale.

2. Steps for improving equity in access and accessibility of publications.

The letter addresses equity in access.

3. Methods for monitoring evolving costs and impacts on affected communities.

The letter addresses cost management.

4. Early input on considerations to increase findability and transparency of research.

The letter addresses transparency of research.

**Uploaded File:** 

Willinsky-Response.pdf

**Description:** An Open Letter on NIH's Request for Information on Public Access

Email: willinsk@stanford.edu

# AN OPEN LETTER

Re: Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research, <u>Notice Number: NOT-OD-23-091</u>

Few organizations have done as much to increase public access to research as the National Institutes of Health. Recognizing the digital-era potential for much wider access to biomedical research – as well as the scientific and public benefits of such access – the NIH has set the standard globally for research funding agencies. The agency has strengthened its policies to overcome the inertia of researchers on matters of public access. It has stood up to large corporate publishers that have actively lobbied against its public access measures. The NIH has led the way in achieving what is now a consensus among scholarly communication stakeholders on the value of public access for research and its benefits to humankind.

That the NIH is now reaching out for public input on "the NIH plan to enhance public access to the results of NIH-supported research" is another admirable demonstration, at least in principle, of its commitment to promoting the progress of science through greater access. For it may, in fact, be time to consider whether *enhancing* the NIH's pioneering methods of the last two decades is the best possible path forward for this Year of Open Science, as federal agencies have designated it.

One indication of the changes afoot has recently been made clear by Dr. Alondra Nelson in her role as Director of the White House Office of Science and Technology Policy (OSTP) and Assistant to the President for Science and Technology. In the August 25, 2022 OSTP policy directive, now known as the Nelson Memo, she sets the tone by stating that "when research is widely available to other researchers and the public, it can save lives, provide policymakers with the tools to make critical decisions, and drive more equitable outcomes across every sector of society." There are three ways in which this consequential statement suggests that the NIH should consider a substantial change in direction, one that goes well beyond the policy document's position on introducing zero embargoes for NIH-sponsored research:

- 1. First, Dr. Nelson's statement reminds readers of the public access benefits, rather than stating the government's policy. It reflects what is now a consensus, reinforced by the pandemic, among researchers, societies, librarians, publishers, and funding agencies on the value of public access. To arrive, then at a time "when research is widely available" will require a leveraging of that consensus. Consider, for example, the Nelson Memo's elimination of embargoes on public access to federally financed research. The NIH first introduced an embargo period (before public access is provided to federally funded research) in the 2000s, it seems fair to say, as a concession to the publishers' subscription model in exercising their copyright. To eliminate embargoes may further enhance NIH's public access policy but it places a further check on publishers' intellectual property rights. The consensus alternative is to find a way to align stakeholders' interests with sustainable public access through copyright reform.
- 2. Secondly, Dr. Nelson's statement recognizes the benefits of public access to the whole

of research without qualification. In introducing a public access policy in the early 2000s, the NIH understandably based it on the public's research investment. Yet as Dr. Nelson makes clear, the reason for public access is not that the public paid for it, but that public access promotes the progress of science to the benefit of humankind. That progress is not being well served today nor are policymakers and physicians by the fragmented, partial, and unpredictable nature of public access to research publications. To paraphrase John Donne, no study is an island entire of itself; every study is a piece of the continent, a part of the main body of the literature. Having done so much to establish the benefits within its sponsored research, the NIH needs to now look at supporting the far broader goal implicit in Dr. Nelson's vision.

3. Thirdly, Dr. Nelson brings home the vital urgency of public access. It can help to save lives, make critical policy decisions, and "drive more equitable outcomes across every sector of society." This stands in contrast to public access' current rate of progress. In 2021, 88% of the global scholarly journal revenues were from exclusive subscriptions, according to the market research company Simba. The current scholarly publishing market, despite a great deal of experimentation, is not delivering this commonly agreed upon good of public access in a timely manner or, many would argue, at a fair market price. Rather than reducing embargoes, the NIH needs to join with other stakeholders in considering how copyright, which so aptly facilitates subscription revenues, can provide comparable incentives to speed the move to public access.

Now, some are bound to object that the NIH should stay in its lane. Yet, it can readily be argued that the NIH has made public access its lane over the last two decades, just as the effective pursuit of its mission calls for improving access to the whole of the research literature. Rather than steer clear of copyright, the NIH could be said to have a responsibility to bring its accumulated expertise to bear on a digital-era copyright update for science. This is all well in advance of any subsequent initiatives by the Copyright Office or Congress

Almost every other cultural enterprise, from video games to music streaming, has instigated copyright reforms, since the onset of the internet. The current law served scholarly publishing's Age of Print. It does not, however, offer an equivalent means of recouping publisher investments in public access. Article processing charges have had limited success, while "read and publish" agreements still depend on subscriptions. As the largest biomedical research funder, who better to initiate a national conversation among stakeholders sharing this common goal of promoting the progress of science through public access.

Nor need such deliberations start from scratch, as considerable work on science and copyright has been undertaken, whether on strengthening limitations and exceptions,<sup>1</sup> offering secondary publishing rights to authors,<sup>2</sup> introducing statutory licensing for research publications,<sup>3</sup> or

<sup>&</sup>lt;sup>1</sup> Flynn, S., et al. (2020). <u>Implementing user rights for research in the field of artificial intelligence: A call for international action</u>. *Joint PIJIP/TLS Research Paper Series*, (48).

<sup>&</sup>lt;sup>2</sup> A position statement from Knowledge Rights 21 on secondary publishing rights (2022). Knowledge Rights 21.

<sup>&</sup>lt;sup>3</sup> Willinsky, J. (2023). <u>Copyright's broken promise: How to restore the law's ability to promote the progress of science.</u> MIT Press.

removing research from copyright's domain.<sup>4</sup> There is also precedent for the necessarily international scope of this endeavor, with the WIPO Copyright Treaties and the TRIPPS Agreement, while the Marrakesh Treaty (2013) provides a particularly encouraging example of bringing human rights to bear on access to knowledge on a global scale.

Since the NIH began on this public access path, the alignment around public access makes possible a reform of copyright to facilitate public access. Such a change will not only enable the benefits Dr. Nelson has set out, it could free up the inordinate amount of energy spent on pursuing public access by working around copyright with limited success. Our hope is that the NIH will consider expressing a willingness to join with others to consider how a digital-era copyright law can serve this common goal of an open science.

John Willinsky Khosla Family Professor of Education Emeritus, Stanford University

Curtis T. Bundy Iowa State University Library

Richard Gallagher
President & Editor-in-Chief, Annual Reviews

Peter B. Kaufman Senior Program Officer, MIT Open Learning

Michael J. Madison Professor, University of Pittsburgh School of Law

Jefferson Pooley
Professor, Media & Communication, Muhlenberg College

<sup>&</sup>lt;sup>4</sup> Shavell, S. (2010). Should copyright of academic works be abolished? Journal of Legal Analysis 2, no. 1.

I am responding to this RFI: On behalf of myself

Name: Steven D. Smith

Name of Organization: Frontiers

Type of Organization: Other

Type of Organization-Other: Publisher

Role: Member of the public

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

NIH public access plan is welcome. Universities and institutes can continue to be encouraged to strike institutional agreements with publishers to ensure cost-neutral access, with relevant discounts etc. Right now, there are surprisingly few institutional partnerships, although a few well-publicized so-called transformative agreements have been signed, such as Wiley's with California (CDL?) in recent days. There is a potential inequity here with the so-called 'free rider' effect: that research universities pay for the publication of open access research through APCs (gold OA), but universities and colleges that do not produce research effectively get 'free access' to the research. So costs are spread unevenly.

Publishers should be encouraged to offer discounts and waivers, which are typically country-specific, but could in theory vary based on Other factors.

#### 2. Steps for improving equity in access and accessibility of publications.

It may be publicly available, but needs CC-BY license. Clarification seem, rights are important.

# 3. Methods for monitoring evolving costs and impacts on affected communities.

There should be better clarity around the APCs so that people understand the reason for investment, sustainability and transparency; such as the journal-checker tool / database with Plan S.

#### 4. Early input on considerations to increase findability and transparency of research.

Use of ORCiD should be encouraged.

Grant ID numbers and PIDs for grants.

But much is not currently interoperable or universal.

Making and collecting meta-data and making sure open review is captured.

Data citations, links to resources.

In 2017 Cross-Ref does offer Fund Ref and global PIDs for grants and facilities. Interoperable identifiers are necessary!

Email: steve.smith@frontiersin.org

I am responding to this RFI: On behalf of myself

Name: Mark Peifer

Name of Organization: University of North Carolina

Type of Organization: University

Role: Scientific researcher

# 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

If you simply mandate posting of the author version of the PMC article with NO embargo, that would be awesome. We all already produce author versions for review and you have forced journals to accept the posting of PMC versions--removing the embargo will be a simple solution for researchers and will also undercut some of the outrageous fees some for-profit journals are charging for "Gold Open Access"

# 2. Steps for improving equity in access and accessibility of publications.

I think this is a good investment, but the burden needs to be on NIH, not the researcher.

# 3. Methods for monitoring evolving costs and impacts on affected communities.

I love the fact that you are stepping in here. The outrageous fees charged by some for profit journals (the Nature family has gotten the most attention) are creating inequities. fees at most non-profit society journals are much lower. I would speak with those smaller publishers to get their input.

4. Early input on considerations to increase findability and transparency of research.

This is outside my expertise

I am responding to this RFI: On behalf of an organization

Name: Hilary Davis

Name of Organization: NC State University Libraries

Type of Organization: University

Role: Institutional official

# 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

We applaud the NIH's recent efforts to engage with stakeholders on topics such as current policy on public access to the results of federally funded research, the evolution of scholarly communications, and access to data and code. We thank the NIH for the opportunity to provide feedback and recommendations. These comments are provided on behalf of NC State University.

Many publishers are actively promoting that the primary path to open access (and public access) is via the payment of an Article Processing Charge, or APC. We are concerned that many grantees will assume that the publishers are correct and will not feel confident in choosing Other options available to them, including green open access. We strongly encourage the NIH to be explicit early and often about the nocost options for compliance when working with grantees (at grant submission stage, at grant award stage, during progress reports, etc.).

Feedback from some authors is that the author-initiated process of submitting articles to PMC is confusing and creates additional burden. Ideally, the NIH will establish a role/unit that will streamline publisher-initiated deposits or NIH-mediated deposits of articles into PMC and make researcher/author involvement in the PMC deposit process optional. This would alleviate a burden that authors/grantees currently bear.

Some NIH-funded researchers have and will publish their articles as Open Access (e.g., via payment of an Article Processing Charge, or APC). In these cases, the researchers may not realize that they still must comply with the requirement to submit articles to PubMed Central (PMC).

Therefore, we recommend that the NIH make it clear via the FAQs and on the public access policy website that even though an article may have been published as Open Access (e.g., via payment of an Article Processing Charge, or APC), authors must also submit the article to PubMed Central (PMC).

Many publishers who used to submit articles on behalf of authors to PubMed Central (PMC) do not follow that practice anymore. Some publishers will only submit articles to PMC if an APC is paid by the author, creating further confusion and placing more burden on the researcher/author. We encourage the NIH to provide clear guidance on a situation that we expect to be common: if a publisher refuses to deposit an article into a repository (PMC) without a paid APC, the "final peer-reviewed manuscript" should still qualify as eligible for deposit In PMC, and this action will not be in violation of copyright.

# 2. Steps for improving equity in access and accessibility of publications.

We hope the NIH will encourage and explain to grantees the need to provide alternative text for images, figures, and tables written by subject matter experts rather than editors or publishers. In an effort to

make any visual content in a publication accessible to readers who use assistive technology, descriptive alternative text is key and is best created by the subject matter experts who understand not only the visual content but also its relationship to the surrounding textual content.

We are glad to see the use of the most recent American National Standards Institute (ANSI) NISO Journal Article Tag Suite (JATS) XML format to create accessible documents in PubMed Central (PMC). It would be useful to continue to encourage authors to consider accessibility in manuscript creation by using word processing programs' headings, formatting, and tagging features. Using NIH's influence to encourage authors to make born-accessible manuscripts can improve the accessibility of not only NIH-funded manuscripts but also manuscripts in general, particularly when it comes to preprints and Other manuscripts without formal editing or curation.

Actively encouraging the use of accessible markup languages for formulas, such as MathML, may save time for PMC's JATS markup by ensuring that manuscripts are coming in with accessible formulas.

We hope the NIH will consider extended engagement with or investment in the infrastructure needed to support the PID (or DPI) ecosystem that currently makes research outputs discoverable and accessible (see Section 4 below).

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

We encourage the NIH to think expansively about costs and fees associated with publication, taking into consideration the costs associated with publishing data, including data curation and long-term storage of research data. We acknowledge that these costs can be highly variable, particularly among disciplines, but we hope that the NIH can develop some general guidance for estimating these costs and exhibit some flexibility when awarding funds.

We additionally encourage the NIH to consider further investment in the cyberinfrastructure required to publish and preserve research outputs, including data. Recommending specific disciplinary repositories for researchers seeking to make available NIH-funded research is an excellent step. Acknowledging the cost and ongoing effort required to maintain these systems will help to further the discussion around support.

We are glad to see that the NIH does not propose requiring authors to publish in journals with any specific type of business model, e.g., publishing their articles Open Access in gold or hybrid journals which requires the payment of an Article Processing Charge (APC).

We are concerned that publishers may introduce new fees for publishing NIH-funded research or require NIH-funded researchers to publish their articles as Open Access. We recommend that the NIH keep a close watch on fees or APC charges that are being levied specifically against NIH-funded authors.

AnOther concern is whether publishers will begin flipping hybrid journals to Gold OA in response to the NIH's and Other federal agencies' updated public access policies and/or raise APC costs. To monitor costs and provide transparency, It may be helpful for the NIH to ask publishers who have Participation Agreements with PMC to make available up-to-date pricing models. Alternatively, the NIH or a collaborator organization can track APCs paid out of research funding to see if these costs increase over time.

#### 4. Early input on considerations to increase findability and transparency of research.

We are glad to see the emphasis placed on persistent identifiers and robust metadata, as these play a key role in making research FAIR. Common standards, such as DOIs, ORCiDs, and RORs, have gained traction and are well regarded. We recommend that the NIH endorse the usage of these PIDs to the community, and, to the extent possible, require the use of ORCiDs. We also recommend that the NIH discourage the proliferation of new PIDs, ensure any new systems where necessary are interoperable with existing systems, and consider supporting efforts to sustain existing and well-established PIDs.

We additionally encourage the NIH to continue to expand the use of existing identifiers into new contexts, like machine-actionable DMSPs, to facilitate better metrics and tracking of research outputs.

There is an opportunity to sustain, grow, or improve efforts around Other PIDs (e.g., instrumentation), and we recommend that the NIH remain aware and supportive of these efforts.

Email: hmdavis4@ncsu.edu

I am responding to this RFI: On behalf of myself

Name: john vaughen

Name of Organization: stanford University

Type of Organization: University

Role: Scientific researcher

1. How to best ensure equity in publication opportunities for NIH-supported investigators.

I 100% approve of making NIH-funded work available on PMC immediately without paywall/embargo!! Could we retroactively make currently paywalled articles done w/ research historically funded by NIH available on PMC as well?

2. Steps for improving equity in access and accessibility of publications.

Consult with smaller journals and users not affiliated with large R1 institutes. Is there a mechanism for ensuring authors comply with PMC upload?

- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

**Submit date: 4/20/2023** 

I am responding to this RFI: On behalf of an organization

Name: Geeta Swamy, MD, Associate Vice President for Research, Duke University Office of Research and

Innovation

Name of Organization: Duke University

Type of Organization: University

**Role:** Institutional official

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

As NIH has already indicated, certain publishing models (such as those that charge fees to authors to publish their research) shift barriers to access from readers (and their proxies) to authors (and their proxies), and are likely to create inequity for researchers who lack funds to cover article processing charges (APCs) or Other publishing fees. NIH should endeavor to make very clear to researchers that they are not required to pay APCs to publishers in order to be in compliance with the public access requirement, and should make clear to publishers and organizations that they should not try to convince authors that paying an APC is the only way (or the best way) to comply with the requirements, publisher behavior that we already see happening. It is not in the interest of NIH, taxpayers, researchers, and research organizations for APC-based publishing to become the dominant model, so NIH should ensure that its public access policies do not inadvertently help establish paying APCs as a norm.

To the extent possible, public access deposit and compliance processes should be integrated into existing researcher workflows, so that public access compliance does not become an additional burden that may create further inequity and potential resistance to the policy and its intentions. The complexity of the current process for depositing publications requires significant infrastructure, training, and time that often falls on lower paid administrative staff at major research institutions, especially administrative assistants, grant coordinators, and librarians. Many smaller institutions, including those that serve primarily rural populations and communities of color, may not have support staff available to assist with policy compliance. This is harmful to the research landscape as these constraints make it even harder to perform research that reflects the needs of vulnerable populations. It is in the best interest of the scientific community to limit the complexity of compliance processes that fall to investigators and their support staff, and instead leverage or mandate the resources of publishers. For example, publishers could make final versions of manuscripts available to PubMed Central when sending records to PubMed for indexing. Managing this complexity should be of primary concern when executing Section III.A.3.b.

#### 2. Steps for improving equity in access and accessibility of publications.

NIH can improve equity in access and accessibility of publications by requiring that NIH-funded research be openly licensed for re-use, through a license such as CC-BY (Creative Commons Attribution), which unambiguously enables a variety of re-use possibilities while still allowing authors to retain those associated rights and the rewards. This would concretely clarify concerns raised in section III.C.1. As noted above, NIH should monitor whether publishers are attempting to charge authors for public access or use of open licenses, and push back by asserting a pre-existing open license for NIH-funded research -

in Other terms, a rights-retention policy such as that being used by "Plan S" funding agencies in Europe. Duke University has had an open access policy since 2010 that retains for Duke and Duke Faculty authors a non-exclusive license to make their scholarly articles open access via Duke's repository. This has enabled Duke research to be made open access, while allowing authors to continue to publish in the venues of their choosing - even if publishers pressure them to sign over Other rights in order to be published, a pre-existing non-exclusive license remains in place to enable them to make their work available through open access, at no cost to them.

While senior researchers who are already established in their careers may feel confident about negotiating with publishers to retain their rights, early career researchers and researchers from historically disadvantaged communities may fear a punitive reaction, and as a result may be reluctant to advocate on their own behalf. When funders like NIH and institutions like Duke establish a baseline of rights retention for their researchers, this levels the playing field and provides a more equitable benefit to all researchers, enabling them to retain control over their own research outputs, make them widely available, improve the reach and impact of their research, and support maximum benefit to the public and their own careers.

NIH has already established expectations for machine-readable publications with high quality metadata, and Duke supports these efforts, as they should enable research to be findable and accessible to people using assistive technologies, researchers who wish to do "distant reading" analysis via software, or Other potential uses that may emerge in the future.

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

NIH can monitor which journals grantees are publishing in, whether they charge fees to authors, and what these fees are. Neither authors nor NIH should be expected to pay high fees simply to publish. As Harvard scholar Peter Suber has noted, high publication fees are essentially a "prestige tax" that are set at the level of what researchers might be willing to pay for the benefit of being associated with a particular journal's brand, and have no connection to the actual costs of publication. As noted above, high publication fees are a barrier to entry for any authors, and especially early career researchers or those from smaller institutions or historically disadvantaged communities.

NIH can use its influence as a major funder of research to lead efforts to transition journals away from charging either authors or readers, by partnering with Other research funding agencies in the US and abroad and with research organizations like universities and libraries to expand initiatives like SCOAP3 to journals in more disciplines. Such initiatives allow costs to be controlled and sponsors to have greater influence in scholarly publishing processes and outputs, while removing barriers for researchers and readers.

NIH should invest in open and community owned infrastructure to help develop and sustain research infrastructure that is aligned with the research mission of funders and universities rather than primarily with a profit motive. Organizations such as Invest in Open Infrastructure, Lyrasis, CrossRef, ORCID, and Dryad are non-profit member organizations that develop and manage essential research infrastructure, and are more transparent about their costs and the value they provide to the community, as well as having a more direct connection to the research community and a clearly defined mission to serve its needs rather than primarily to extract value.

#### 4. Early input on considerations to increase findability and transparency of research.

High quality metadata and standard persistent identifiers for both research outputs and researchers and their organizations will significantly assist findability and transparency of research, as well as accurately providing credit to researchers for their contributions and funders for their sponsorship.

Wherever possible, NIH should require use of existing identifiers such as DOI, ORCID, ROR, and existing taxonomies such as CRediT (Contributor Roles Taxonomy), and work in conjunction with NISO and Other standards bodies to ensure integration of NIH processes with identifiers and infrastructure already widely used by the research community.

Additionally, Duke encourages the NIH to pursue a more open and accessible API for the MeSH Database and materials categorized using this metadata to further increase the findability and bibliometric analysis of medical information. More standardized metadata for NIH funding can assist institutions in developing and maintaining compliance reports, enable better discovery of published manuscripts based on funding information, and may help in crosswalking with Other PIDs, such as NCT numbers. Enhanced metadata for data availability will also assist with connecting published manuscripts with data.

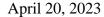
It's important to recognize that full implementation and integration of these standards, identifiers, and functionality involves significant technical challenges, as well as commitment of staff and resources. NIH should aim to support research institutions and smaller organizations in implementing the necessary functionality in their own systems, to avoid potential inequities where systems supported by large publishers and corporations are more easily equipped to develop and support this functionality, leaving less resourced institutions and smaller organizations at a disadvantage.

#### **Uploaded File:**

Duke-University-response-to-NIH-Public-Access-RFI-April-2023.pdf

**Description:** PDF attachment includes an introductory paragraph about why Duke supports this effort, some links in the text body, and information about leaders of multiple Duke University organizations that are signatories in support of these comments.

Email: paolo.mangiafico@duke.edu





## Response from Duke University to the Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research

https://osp.od.nih.gov/nih-plan-to-enhance-public-access-to-the-results-of-nih-supported-research/

Submitted on behalf of:

Geeta Swamy, MD

Associate Vice President for Research, Duke University Office of Research and Innovation

Vice Dean for Scientific Integrity, Duke University Office of Scientific Integrity Haywood Brown, MD Distinguished Professor of Women's Health, Duke University School of Medicine

Joseph Salem, Ph.D.

Rita DiGiallonardo Holloway University Librarian and Vice Provost for Library Affairs, Duke University Libraries

Megan Von Isenburg, MSLS Associate Dean for Library Services & Archives, Duke University Medical Center Library & Archives

#### Introduction

Duke University is a major research university and economic engine for our region, with nearly \$650 million per year from federal government agencies. With ~8000 faculty and staff engaged in research, and an additional ~7500 graduate and professional students engaged in research, Duke produces over 13,000 research publications every year. One of Duke University's key strategic goals is using knowledge in the service of society. We encourage our researchers to make the results of their research (publications, data, and code) as broadly available as possible, and to translate their research into modes that can be effectively consumed by the public and quickly generate economic and social value. As an institution, we have put in place a number of initiatives to support this, including open access policies for faculty and graduate student publications, multiple open repositories with services to integrate them into researcher workflows and staff to provide support in using them, and funding to support open access journals, monographs, and publishers. Duke University leadership is deeply engaged with the HELIOS initiative, and provides support for a variety of programs and training opportunities that encourage our researchers to make open scholarship practices a key part of their research workflows

and culture. We are pleased that the NIH is working to expand public and equitable access and benefit for more federally funded research, and offer these comments in support of that effort.

## Question 1: How can NIH best ensure equity in publication opportunities for its investigators?

As NIH has already indicated, certain publishing models (such as those that charge fees to authors to publish their research) shift barriers to access from readers (and their proxies) to authors (and their proxies), and are likely to create inequity for researchers who lack funds to cover article processing charges (APCs) or other publishing fees. NIH should endeavor to make very clear to researchers that they are not required to pay APCs to publishers in order to be in compliance with the public access requirement, and should make clear to publishers and organizations that they should not try to convince authors that paying an APC is the only way (or the best way) to comply with the requirements, publisher behavior that we already see happening. It is not in the interest of NIH, taxpayers, researchers, and research organizations for APC-based publishing to become the dominant model, so NIH should ensure that its public access policies do not inadvertently help establish paying APCs as a norm.

To the extent possible, public access deposit and compliance processes should be integrated into existing researcher workflows, so that public access compliance does not become an additional burden that may create further inequity and potential resistance to the policy and its intentions. The complexity of the current process for depositing publications requires significant infrastructure, training, and time that often falls on lower paid administrative staff at major research institutions, especially administrative assistants, grant coordinators, and librarians. Many smaller institutions, including those that serve primarily rural populations and communities of color, may not have support staff available to assist with policy compliance. This is harmful to the research landscape as these constraints make it even harder to perform research that reflects the needs of vulnerable populations. It is in the best interest of the scientific community to limit the complexity of compliance processes that fall to investigators and their support staff, and instead leverage or mandate the resources of publishers. For example, publishers could make final versions of manuscripts available to PubMed Central when sending records to PubMed for indexing. Managing this complexity should be of primary concern when executing Section III.A.3.b.

## Question 2: What steps can NIH take to improve equity in access and accessibility of publications?

NIH can improve equity in access and accessibility of publications by requiring that NIH-funded research be openly licensed for re-use, through a license such as CC-BY (Creative Commons Attribution), which unambiguously enables a variety of re-use possibilities while still allowing authors to retain those associated rights and the rewards. This would

concretely clarify concerns raised in section III.C.1. As noted above, NIH should monitor whether publishers are attempting to charge authors for public access or use of open licenses, and push back by asserting a pre-existing open license for NIH-funded research – in other terms, a rights-retention policy such as that being used by "Plan S" funding agencies in Europe. Duke University has had an open access policy since 2010 that retains for Duke and Duke Faculty authors a non-exclusive license to make their scholarly articles open access via Duke's repository. This has enabled Duke research to be made open access, while allowing authors to continue to publish in the venues of their choosing – even if publishers pressure them to sign over other rights in order to be published, a pre-existing non-exclusive license remains in place to enable them to make their work available through open access, at no cost to them.

While senior researchers who are already established in their careers may feel confident about negotiating with publishers to retain their rights, early career researchers and researchers from historically disadvantaged communities may fear a punitive reaction, and as a result may be reluctant to advocate on their own behalf. When funders like NIH and institutions like Duke establish a baseline of rights retention for their researchers, this levels the playing field and provides a more equitable benefit to all researchers, enabling them to retain control over their own research outputs, make them widely available, improve the reach and impact of their research, and support maximum benefit to the public and their own careers.

NIH has already established expectations for machine-readable publications with high quality metadata, and Duke supports these efforts, as they should enable research to be findable and accessible to people using assistive technologies, researchers who wish to do "distant reading" analysis via software, or other potential uses that may emerge in the future.

## Question 3: How can NIH best monitor evolving costs, specifically publication fees, and impacts on affected communities?

NIH can monitor which journals grantees are publishing in, whether they charge fees to authors, and what these fees are. Neither authors nor NIH should be expected to pay high fees simply to publish. As Harvard scholar <u>Peter Suber has noted</u>, high publication fees are essentially a "prestige tax" that are set at the level of what researchers might be willing to pay for the benefit of being associated with a particular journal's brand, and have no connection to the actual costs of publication. As noted above, high publication fees are a barrier to entry for any authors, and especially early career researchers or those from smaller institutions or historically disadvantaged communities.

NIH can use its influence as a major funder of research to lead efforts to transition journals away from charging either authors or readers, by partnering with other research funding agencies in the US and abroad and with research organizations like universities and libraries to expand initiatives like <u>SCOAP3</u> to journals in more disciplines. Such

initiatives allow costs to be controlled and sponsors to have greater influence in scholarly publishing processes and outputs, while removing barriers for researchers and readers.

NIH should invest in open and community owned infrastructure to help develop and sustain research infrastructure that is aligned with the research mission of funders and universities rather than primarily with a profit motive. Organizations such as <a href="Invest in Open Infrastructure">Invest in Open Infrastructure</a>, <a href="Lyrasis">Lyrasis</a>, <a href="CrossRef">CrossRef</a>, <a href="ORCID">ORCID</a>, and <a href="Dryad">Dryad</a> are non-profit member organizations that develop and manage essential research infrastructure</a>, and are more transparent about their costs and the value they provide to the community, as well as having a more direct connection to the research community and a clearly defined mission to serve its needs rather than primarily to extract value.

## Question 4: Early input on considerations to increase findability and transparency of research.

High quality metadata and standard persistent identifiers for both research outputs and researchers and their organizations will significantly assist findability and transparency of research, as well as accurately providing credit to researchers for their contributions and funders for their sponsorship.

Wherever possible, NIH should require use of existing identifiers such as DOI, ORCID, ROR, and existing taxonomies such as CRediT (Contributor Roles Taxonomy), and work in conjunction with NISO and other standards bodies to ensure integration of NIH processes with identifiers and infrastructure already widely used by the research community.

Additionally, Duke encourages the NIH to pursue a more open and accessible API for the MeSH Database and materials categorized using this metadata to further increase the findability and bibliometric analysis of medical information. More standardized metadata for NIH funding can assist institutions in developing and maintaining compliance reports, enable better discovery of published manuscripts based on funding information, and may help in crosswalking with other PIDs, such as NCT numbers. Enhanced metadata for data availability will also assist with connecting published manuscripts with data.

It's important to recognize that full implementation and integration of these standards, identifiers, and functionality involves significant technical challenges, as well as commitment of staff and resources. NIH should aim to support research institutions and smaller organizations in implementing the necessary functionality in their own systems, to avoid potential inequities where systems supported by large publishers and corporations are more easily equipped to develop and support this functionality, leaving less resourced institutions and smaller organizations at a disadvantage.

**Submit date: 4/20/2023** 

I am responding to this RFI: On behalf of an organization

Name: Katherine B. McGuire and Jasper Simons

Name of Organization: American Psychological Association

Type of Organization: Professional org association

- 1. How to best ensure equity in publication opportunities for NIH-supported investigators.
- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

#### **Uploaded File:**

Letter-in-Reponse-to-NIH-RFI-on-Enhancing-Access-April-18-2023.pdf

Email: ahill@apa.org



April 20, 2023

Lawrence A. Tabak, DDS, PhD Acting Director, National Institutes of Health

Submitted electronically at <a href="https://osp.od.nih.gov/nih-plan-to-enhance-public-access-to-the-results-of-nih-supported-research/">https://osp.od.nih.gov/nih-plan-to-enhance-public-access-to-the-results-of-nih-supported-research/</a>

# RE: NIH Plan to Enhance Public Access to the Results of NIH-Supported Research (Notice Number NOT-OD-23-091)

Dear Dr. Tabak:

The American Psychological Association (APA) applauds NIH for its efforts to enhance public access to the results of the research the agency funds, and we appreciate the opportunity to respond to this request for information.

APA is a scientific and professional organization composed of more than 146,000 members, affiliates, and students. APA's mission is to promote the advancement, communication, and application of psychological science and knowledge to benefit society and improve lives. Among the organization's aims are to elevate the public's understanding of, regard for, and use of psychology and to prepare the discipline and profession of psychology for the future.

As a means of achieving these goals, APA has built a reputable publishing program. Through the program's output as a nonprofit society publisher, APA balances the needs of scholars, members, and the organization while seeking to apply psychology broadly in society. APA's publishing program is dedicated to producing high-quality, evidence-based content that informs the discipline of psychology; the program also publishes the journals of many other scholarly societies in psychology and related disciplines. APA creates publishing standards through its collaboration with the community the publishing program serves. APA's publishing program supports psychologist members and funds the work of the organization, which applies the scholarship of psychology to improve everyday life.

Given the APA publishing program's frequent publication of research funded by NIH, we have a number of recommendations regarding NIH's plan to enhance public access to the research NIH funding supports. Those recommendations include the following:

- Continue to protect researchers' freedom to publish where they choose to do so, without payment. As we said in our letter of April 17, 2020, in response to the request for information titled <u>Public Access to Peer-Reviewed Scholarly Publications, Data, and Code Resulting From Federally Funded Research</u>, protecting researchers' freedom to publish without payment is imperative. Federal intervention that privileges a mandatory pay-to-publish model (e.g., gold open access) disadvantages researchers from historically excluded groups and those without access to funding, particularly early-career researchers and those from historically excluded groups.
- Ensure sufficient funding for research and discovery; for sharing both the elements necessary to validate and replicate the results of this research; and for metadata and infrastructure required to label, host, and link these elements.
- Protect against the misuse of research that is harmful to the public and to public trust in science.

Next, we offer comments on and questions regarding each of the topics identified in your request for information.

# 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

We are glad to see that NIH is retaining the policy of allowing authors to choose where to publish. APA supports the academic freedom to publish in researchers' chosen venues without payment, because requiring payment to publish disadvantages underfunded researchers. Given this freedom to publish without payment, we wonder what "reasonable publishing costs" would be considered allowable in research budgets. NIH proposes to require that the peer-reviewed manuscript be made available without embargo, but the costs of managing peer review; providing additional author and reader services, such as the creation and preservation of a permanent record; and providing the infrastructure necessary to provide metadata, machine-readability, and interoperability are not delineated in the <u>Code of Federal Regulations</u> or in the <u>NIH Grants Policy Statement</u>.

Because APA is a member organization, we are concerned about additional burdens for researchers, especially those who now expect to share data, code, and materials. We appreciate that NIH allows research funds to be allocated for data deposit and publication fees. Without a corresponding increase in allocated funds, though, this represents a transfer of funds away from research and discovery and could perpetuate known inequities. It is likely that researchers at well-funded institutions will be the ones paying article processing charges (APCs); moreover, researchers who pay APCs are likely to be White men at advanced stages in their careers (Olejniczak & Wilson, 2020).

Assessing the costs of publication is difficult. NIH's proposed plan seems to imply that the burdens of reporting and tracking these costs will fall on publishers and NIH, respectively.

Evolving infrastructure requirements and changes arising from new policies might encourage more publishers to move toward pay-to-publish or read-and-publish open access models. In addition, pressure to lower publication costs could compromise quality and standards, thereby undermining public trust. One example of a potential threat to trust in science is articles being published after going through fraudulent peer review. This threat is arguably looming: Hindawi <u>recently retracted</u> more than 500 articles because of fraudulent peer review; IOP Publishing, approximately 500 articles; and PLOS, 100 articles (Kincaid, 2022).

### 2. Steps for improving equity in access and accessibility of publications.

We applaud NIH's steps to improve accessibility; we share the agency's commitment to equity in access to publications by diverse communities of users.

Public access policies should protect researchers' freedom to choose the licenses that are appropriate for their specific works.

We are concerned about the protection of research outputs from misuse that could harm the public and damage trust in science. The request for information promises that, for articles with permissive licenses, NIH will "continue to promote the broadest possible reuse of its supported articles, while limiting inappropriate uses, such as redistribution of PMC content for sale" (Section III.C.2). How will NIH limit reuse of material shared under permissive licenses (e.g., CC BY)? Will use limitations extend to artificial intelligence products like ChatGPT or applications that purport to offer medical or mental health advice but inappropriately cull that advice from research articles? We are concerned that clinical data and research conclusions provided out of context in a diagnostic setting by unqualified practitioners could produce negative and harmful results.

How will reuse in service of misinformation and/or misappropriation be prevented?

We also note that the plan specifies that machine-readable text will be provided by publishers, but the costs of providing such text are not accounted for in the proposed plan. As we detail next, APA distributes metadata to facilitate the linking and interoperability of research artifacts, but the existing infrastructure that enables this distribution relies on funding from larger publishers.

### 3. Methods for monitoring evolving costs and impacts on affected communities.

We share and commend NIH's commitment to equity in fees, policies, and publication opportunities. APA is committed to promoting diversity and to advancing equity and inclusion. To eliminate structural barriers and scientific practices that have prevented the full participation of those who are societally marginalized and historically underrepresented in the field of psychology, APA focuses on finding ways to encourage and enable representation, fair treatment, access, opportunity, and advancement.

In reference to the NIH plan, we call attention to the importance of defining how equity will be determined. What inequities will NIH identify and what steps will NIH consider to remedy them? For example, will NIH be assessing outcomes such as who publishes where and who accesses publicly available material? Among the equity concerns we have are the consequences of increasing the number of open access articles being published and attracting both readers and citations. Will publishers be asked to report statistics on these metrics?

Allowing researchers to choose where they publish means that equity will be assessed across multiple publishing models. Will NIH look at models in which publication fees are reapplied to transformative agreements (e.g., read-and-publish agreements wherein journal subscription fees are reallocated for open access publishing)? We note also that society publishers who move to open access models may lose revenues that would have funded other scientific activities.

And, finally, as a member organization, we note that it is important for researchers to have clear expectations as they navigate NIH's revised plan. Will there be limits on publication fees? How many publications may be paid for under one grant? Can the funds to cover publication costs be requested from the agency that awarded the grant after the grant is no longer active? Practices differ among scientific communities, and psychological science is an especially diverse discipline whose practitioners range from researchers doing basic experimental and observational studies to therapists engaging in clinical interventions.

## 4. Early input on considerations to increase findability and transparency of research.

We support this first step in developing an updated plan for persistent identifiers and metadata. APA already recommends linking practices through our publishing program's journals and through <u>APA Style</u> and uses identifiers such as researcher IDs, grant numbers, data set IDs, article DOIs, and institution IDs. In our publishing program, we consider the linking and interoperability of these research artifacts to be important, and we distribute metadata to facilitate these connections. Managing metadata standards and interoperability is difficult, however, and requires resources that are not available to all institutions and publishers. Has NIH analyzed these difficulties alongside potential compliance costs to researchers, publishers, and the agency? Existing infrastructure relies on funding from large publishers, and developing new standards and approaches will be expensive.

Compliance with identification and metadata standards presents additional administrative burdens for authors, and this will likely have the largest effect on researchers who are already underresourced. How will NIH monitor, for example, access via persistent identifiers outside of posting in PubMed or open access status if not deposited?

Finally, APA encourages NIH to continue to allow NIH-supported researchers to choose appropriate repositories for their data, code, and materials provided that agency criteria are met, consistent with 2020's <u>Supplemental Information to the NIH Policy for Data Management and Sharing: Selecting a Repository for Data Resulting From NIH-Supported Research</u>. As we noted in our March 6, 2020, letter, written in response to the draft titled <u>Desirable</u> <u>Characteristics of Repositories for Managing and Sharing Data Resulting From Federally Funded Research</u> (Document Number 2020-00689), psychologists work with a wide range of data from surveys, laboratory experiments, government statistics, administrative records, imaging, genomics, social media, and other sources (Alter & Gonzalez, 2018), and these data are suited to different types of repositories. Guidance on how to pay for these repositories is needed, given the costs of curating protected data sets and storing large data sets.

### **Closing Comments**

APA supports NIH's goal of enhancing public access to the results of and data from federally funded research. The ultimate objective of funders, researchers, and publishers should be advancing the quality and pace of scientific research. APA also shares NIH's aim of encouraging greater scientific integrity and enabling future inquiry, discovery, and translation.

Improving the availability of peer-reviewed articles, data, and code is one of many means to those ends, but care must be taken to avoid potential unintended consequences that could diminish the quality and pace of scientific research. Continuing to protect the freedom to publish without payment is crucial, because inequities will disproportionally affect the researchers who are less likely to have sufficient funding, namely, those who are members of historically excluded groups (e.g., racial or ethnic minorities), early-career researchers, researchers at underfunded universities, and researchers in the Global South with whom American researchers collaborate to advance science.

We agree that transparency increases scientific integrity and bolsters public trust in scientific research. At the same time, we also maintain that funding is needed for all phases of the research life cycle, not just research and discovery. The sharing of data and results entails costs, and metadata and infrastructure necessary to link and make them accessible, enable machine-readability, and ensure interoperability.

In keeping with APA's mission to promote the advancement and communication of psychological science to benefit society and improve lives, we end this letter by highlighting the need to protect against misuse of research that is harmful to the public and to public trust in science.

APA thanks NIH for this opportunity to share comments on the *Plan to Enhance Public Access to the Results of NIH-Supported Research*. If you have any questions or if we can provide any further information, please contact us at kmcguire@apa.org or jsimons@apa.org.

### Sincerely,

Katherine B. McGuire Chief Advocacy Officer

Jasper Simons

**Chief Publishing Officer** 

#### References

Alter, G., & Gonzalez, R. (2018). Responsible practices for data sharing. *American Psychologist*, 73(2), 146–156. https://doi.org/10.1037/amp0000258

Olejniczak, A. J., & Wilson, M. J. (2020). Who's writing open access (OA) articles? Characteristics of OA authors at Ph.D.-granting institutions in the United States. *Quantitative Science Studies*, 1(4), 1429–1450. https://doi.org/10.1162/qss\_a\_00091

Kincaid, E. (2022, September 28). Exclusive: Hindawi and Wiley to retract over 500 papers linked to peer review rings.

\*Retraction Watch.\* https://retractionwatch.com/2022/09/28/exclusive-hindawi-and-wiley-to-retract-over-500-papers-linked-to-peer-review-rings/

**Submit date:** 4/20/2023

I am responding to this RFI: On behalf of an organization

Name: Angela Cochran

**Name of Organization:** American Society of Clinical Oncology, American College of Physicians, NEJM Group, American Heart Association, American Diabetes Association, American Society of Anesthesiologists, American Urological Association, American Thoracic Society, American Gastr

Type of Organization: Professional orgassociation

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

The medical societies represented in this response to the National Institutes of Health (NIH) Request for Information (RFI) on the NIH Plan to Enhance Public Access to the Results of NIH Supported Research thank the NIH for the opportunity to comment on the proposed public access plan.

As the NIH works to incorporate feedback and refine a draft policy, we recognize that the NIH proposed plan has a path for compliance whether a funded author chooses to publish in journals with an open access model, a subscription model, or Other publishing model. It would be extraordinary and detrimental to non-profit organizations for a US agency to develop policies that force one business model over anOther with no consideration for the economic harm and/or impact to societies and science communication overall.

In recognition of our continued support in aiding researcher compliance with NIH requirements and to make peer-reviewed content accessible without an embargo, and we ask that the NIH policy refrain from requiring reuse rights under licenses that restrict our ability to establish copyright and preserve the downstream revenue associated with the final version of record.

Beyond whether a journal is subscription access, open access, or hybrid, there are supplementary revenue streams that society journals use to remain sustainable including licensing, commercial reprints, permissions, and advertising. Broad reuse licenses that do not respect publisher copyright rights jeopardize those revenue streams and the sustainability of society publishers. The value we provide to our research communities is at risk. Under copyright provisions, we guard against misuse of author content by requiring third parties to follow our policies regarding appropriate use of published content.

Maintaining scientific integrity is paramount.

The societies represented in these comments take seriously the scientific integrity of research published in our journals. The reputations of our societies and journals rely on being a provider of trusted content.

Our clinical journals focus on expedient but thorough review and publication of research that affects patient care—not in a matter of years, but sometimes hours. Our societies use our journals to disseminate clinical practice guidelines that impact research practice or clinical decisions, rules of hospitals and clinics, spending by government and insurers, and ultimately public health. The guidelines are developed at great expense and with significant resource burden. Utmost care is taken that they are current on the research, provide appropriate guidance based on proper methods and analysis of evidence, and bar any industry influence.

Maintaining this trusted role in society, at a time when disinformation is rampant, requires a significant investment. Vigilance in publication research integrity and conflict of interest management not only aligns with our missions but, more importantly, gives confidence to clinicians and researchers that information we publish has been verified and is reliable.

Diligent peer review, management and public disclosures of conflicts, and data and figure integrity checks are vital parts of the process. Threats such as plagiarism, "paper mills," and fraudulent data are increasingly present and require steady attention.

These services are critical to production of a final product researchers and clinicians can rely upon as they conduct vital research and deliver evidence-based care—but they also require direct and substantial expense. Significant staff training and resources could be endangered if publishers lose revenue in the form of cancelled subscriptions, insufficient total article processing charge (APC) income, and lost licensing fees for approved reuse of content, among Others. Each publisher will have their own budgetary tipping point when decreased revenues impact our ability to provide services that now protect the integrity of research published in our journals.

1. How to best ensure equity in publication opportunities for NIH-supported investigators.

The subscription model is largely accessible to researchers submitting their work and thus the most financially equitable for authors. Free to read (via gold or green OA) is most equitable to the readers.

The NIH proposed plan to mandate zero-embargo and allow green access appears equitable for both authors and readers. However, that assumption does not consider that many subscription and hybrid journals will have a large quantity of content that they invested in freely accessible. Under this zero-embargo proposal continuing subscription revenue may be implausible for some journals. Libraries have begun and will continue to cancel subscriptions to journals with large amounts of content that are free to access.

In such an environment, journals with high numbers of papers reporting on NIH funded research may need to convert to an author-pays open access (gold OA) model. While the NIH portends that NIH funded authors will have the ability to pay Article Processing Charges (APCs) to these now newly flipped journals, this creates an equity issue for NIH authors who have minimal funding or their funding is expended on necessary research expenses.

This proposed plan will be mandated for all NIH funded authors regardless of how much funding they received or how small a role any given individual plays in a research project or manuscript. The NIH should apply a minimum threshold of funding and/or level of participation by authors and researchers before subjecting the papers to the proposed mandate. Also, minimal contributions to studies (or use of funded shared resources) made by NIH-funded authors should not qualify a paper for the proposed mandate.

It is commonly understood that there is significant overreporting of federal support on submitted manuscripts as a component of research grants. We are aware that grantees, or Others working on their behalf at their institutions, have deposited articles in PubMed Central in error. In light of this—and the impact of proposed changes—we urge NIH to publish clear guidance, both on circumstances that qualify submitted papers to claim NIH funding, and the conditions that invoke a requirement to comply with the

public access mandate. More and better communication to grantees and Other stakeholders regarding the administration of compliance is essential with the planned zero embargo policy.

Regardless of whether NIH funded authors intend to pursue a green OA option and reserve their funds for Other research purposes, a concerning number of scientific journals will be vulnerable to library subscription cancelations given the amount of content that will be accessible without embargo on PubMed Central and Other government repositories. Not all journals will be able to offer a green route. We do not believe OSTP or federal funding agencies fully appreciate the extent to which zero-embargo public access policies will disrupt the entire ecosystem of the research enterprise.

#### 2. Steps for improving equity in access and accessibility of publications.

The work of converting Word files into machine readable, highly tagged extensible markup language (XML) is important, particularly for readers in need of assistive devices. It also aids in search and discovery. One efficiency and savings of taxpayer dollars we can do today is to remove the redundancy of this being done twice—once by the publisher and once by the National Library of Medicine (NLM). This is not a good use of taxpayer money.

The NIH could reduce their expenses in performing duplicate tasks. We call on the NIH to engage publishers in possible private-public partnerships to avoid duplication of work and excess spending.

Our organizations invest in development of journal hosting platforms with capabilities for ensuring that content is tagged and optimized for adaptive devices needed by users with visual and auditory disabilities. We are concerned that by taking users off our platforms to read our content on PubMed Central, the value of this investment will be diminished.

Medical societies routinely produce infographics, visual abstracts, context summaries, plain-language summaries, and patient pages for individuals outside the typical subscriber or society member. Currently the NLM refuses to link references to the publisher site, and users on PMC have little chance to discover this content. A zero-embargo policy is likely to further diminish existing usage.

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

The NIH must engage the researcher community to understand their concerns about public access policy changes. While the NIH asserts authors can use grant money to pay publication fees, our members overwhelmingly tell us that they do not have enough money in their grants to cover publication fees for multiple papers likely to arise from a single grant. Further, researchers tell us their proposals for funding are typically cut in review.

The likelihood of large increases in government funding of agencies is low and researchers are concerned that publication fees will not be adequately covered by their research grants.

#### 4. Early input on considerations to increase findability and transparency of research.

Publishers are very interested in and have been early adopters of persistent identifiers (PIDs) in the scholarly communication life cycle. We encourage the NIH to engage with publishers and the PID community of partners to use or adapt what has already been created. We strongly recommend the NIH both employ digital object identifiers (DOIs) for grants and require them for datasets published. By

adopting persistent identifiers already in use in scholarly publishing, journals can include persistent links to critical pieces of research for users to access.

Lastly, a commitment from the NIH to adopt PIDs already in use should end the current NLM practice of replacing publisher DOIs in the references of papers in PubMed. The NLM does not have permission from publishers or authors to make material changes to the deposited manuscripts. By stripping the DOIs from reference links or choosing to include links to the PMC version instead of the version of record (VOR), the NLM is depriving the user of access to associated editorials, letters to the editor, podcasts, infographics, etc. The NIH has shown strong interest in understanding how journals make content more accessible to non-subscribers and non-specialists; it makes no sense for the NLM to refuse to link to the VOR for the discovery of this content via references.

We urge the NIH, OSTP, and OMB to carefully consider the points raised and we thank you for the opportunity to comment.

#### **Uploaded File:**

Medical-Societies-RFI-Response.pdf

**Description:** Full letter in PDF form included.

Email: angela.cochran@asco.org

The medical societies represented in this response to the National Institutes of Health (NIH) Request for Information (RFI) on the NIH Plan to Enhance Public Access to the Results of NIH Supported Research thank the NIH for the opportunity to comment on the proposed public access plan.

As the NIH works to incorporate feedback and refine a draft policy, we recognize that the NIH proposed plan has a path for compliance whether a funded author chooses to publish in journals with an open access model, a subscription model, or other publishing model. It would be extraordinary and detrimental to non-profit organizations for a US agency to develop policies that force one business model over another with no consideration for the economic harm and/or impact to societies and science communication overall.

In recognition of our continued support in aiding researcher compliance with NIH requirements and to make peer-reviewed content accessible without an embargo, and we ask that the NIH policy **refrain from requiring reuse rights under licenses that restrict our ability to establish copyright** and preserve the downstream revenue associated with the final version of record.

Beyond whether a journal is subscription access, open access, or hybrid, there are supplementary revenue streams that society journals use to remain sustainable including licensing, commercial reprints, permissions, and advertising. Broad reuse licenses that do not respect publisher copyright rights jeopardize those revenue streams and the sustainability of society publishers. **The value we provide to our research communities is at risk.** Under copyright provisions, we guard against misuse of author content by requiring third parties to follow our policies regarding appropriate use of published content.

#### Maintaining scientific integrity is paramount.

The societies represented in these comments take seriously the scientific integrity of research published in our journals. The reputations of our societies and journals rely on being a provider of trusted content.

Our clinical journals focus on expedient but thorough review and publication of research that affects patient care—not in a matter of years, but sometimes hours. Our societies use our journals to disseminate clinical practice guidelines that impact research practice or clinical decisions, rules of hospitals and clinics, spending by government and insurers, and ultimately public health. The guidelines are developed at great expense and with significant resource burden. Utmost care is taken that they are current on the research, provide appropriate guidance based on proper methods and analysis of evidence, and bar any industry influence.

Maintaining this trusted role in society, at a time when disinformation is rampant, requires a significant investment. **Vigilance in publication research integrity and conflict of interest management** not only aligns with our missions but, more importantly, gives confidence to clinicians and researchers that information we publish has been verified and is reliable.

Diligent peer review, management and public disclosures of conflicts, and data and figure integrity checks are vital parts of the process. Threats such as plagiarism, "paper mills," and fraudulent data are increasingly present and require steady attention.

These services are critical to production of a final product researchers and clinicians can rely upon as they conduct vital research and deliver evidence-based care—but they also require direct and substantial expense. Significant staff training and resources could be endangered if publishers lose revenue in the form of cancelled subscriptions, insufficient total article processing charge (APC) income, and lost licensing fees for approved reuse of content, among others. Each publisher will have their own budgetary tipping point when decreased revenues impact our ability to provide services that now protect the integrity of research published in our journals.

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

The subscription model is largely accessible to researchers submitting their work and thus the most financially equitable for authors. Free to read (via gold or green OA) is most equitable to the readers.

The NIH proposed plan to mandate zero-embargo and allow green access *appears* equitable for both authors and readers. However, that assumption does not consider that many subscription and hybrid journals will have a large quantity of content that they invested in freely accessible. Under this zero-embargo proposal continuing subscription revenue may be implausible for some journals. Libraries have begun and will continue to cancel subscriptions to journals with large amounts of content that are free to access.

In such an environment, **journals with high numbers of papers reporting on NIH funded research may need to convert to an author-pays open access** (gold OA) model. While the NIH portends that NIH funded authors will have the ability to pay Article Processing Charges (APCs) to these now newly flipped journals, this creates an equity issue for NIH authors who have minimal funding or their funding is expended on necessary research expenses.

This proposed plan will be mandated for all NIH funded authors regardless of how much funding they received or how small a role any given individual plays in a research project or manuscript. The **NIH should apply a minimum threshold of funding** and/or level of participation by authors and researchers before subjecting the papers to the proposed mandate. Also, minimal contributions to studies (or use of funded shared resources) made by NIH-funded authors should not qualify a paper for the proposed mandate.

It is commonly understood that there is significant overreporting of federal support on submitted manuscripts as a component of research grants. We are aware that grantees, or others working on their behalf at their institutions, have deposited articles in PubMed Central in error. In light of this—and the impact of proposed changes—we urge NIH to publish clear guidance, both on circumstances that qualify submitted papers to claim NIH funding, and the conditions that invoke a requirement to comply with the public access mandate. More and better communication to grantees and other stakeholders regarding the administration of compliance is essential with the planned zero embargo policy.

Regardless of whether NIH funded authors intend to pursue a green OA option and reserve their funds for other research purposes, a concerning number of scientific journals will be vulnerable to library subscription cancelations given the amount of content that will be accessible without embargo on PubMed Central and other government repositories. **Not all journals will be able to offer a green** 

**route.** We do not believe OSTP or federal funding agencies fully appreciate the extent to which zeroembargo public access policies will disrupt the entire ecosystem of the research enterprise.

#### 2. Steps for improving equity in access and accessibility of publications.

The work of converting Word files into machine readable, highly tagged extensible markup language (XML) is important, particularly for readers in need of assistive devices. It also aids in search and discovery. One efficiency and savings of taxpayer dollars we can do today is to remove the redundancy of this being done twice—once by the publisher and once by the National Library of Medicine (NLM). This is not a good use of taxpayer money.

The NIH could reduce their expenses in performing duplicate tasks. We call on the NIH to engage publishers in possible private-public partnerships to avoid duplication of work and excess spending.

Our organizations invest in development of journal hosting platforms with capabilities for **ensuring that content is tagged and optimized for adaptive devices needed by users with visual and auditory disabilities**. We are concerned that by taking users off our platforms to read our content on PubMed Central, the value of this investment will be diminished.

Medical societies routinely produce infographics, visual abstracts, context summaries, plain-language summaries, and patient pages for individuals outside the typical subscriber or society member. Currently the **NLM refuses to link references to the publisher site, and users on PMC have little chance to discover this content**. A zero-embargo policy is likely to further diminish existing usage.

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

The NIH must engage the researcher community to understand their concerns about public access policy changes. While the NIH asserts authors can use grant money to pay publication fees, our members overwhelmingly tell us that they do not have enough money in their grants to cover publication fees for multiple papers likely to arise from a single grant. Further, researchers tell us their proposals for funding are typically cut in review.

The likelihood of large increases in government funding of agencies is low and researchers are concerned that publication fees will not be adequately covered by their research grants.

#### 4. Early recommendations for increasing findability and transparency of research.

Publishers are very interested in and have been early adopters of persistent identifiers (PIDs) in the scholarly communication life cycle. We **encourage the NIH to engage with publishers and the PID community of partners** to use or adapt what has already been created. We strongly recommend the NIH both employ digital object identifiers (DOIs) for grants and require them for datasets published. By adopting persistent identifiers already in use in scholarly publishing, journals can include persistent links to critical pieces of research for users to access.

Lastly, a commitment from the NIH to adopt PIDs already in use should end the current NLM practice of replacing publisher DOIs in the references of papers in PubMed. The NLM does not have permission from publishers or authors to make material changes to the deposited manuscripts. By stripping the DOIs from reference links or choosing to include links to the PMC version instead of the version of record (VOR), the NLM is depriving the user of access to associated editorials, letters to the editor, podcasts, infographics, etc. The NIH has shown strong interest in understanding how journals make content more accessible to non-subscribers and non-specialists; it makes no sense for the NLM to refuse to link to the VOR for the discovery of this content via references.

We urge the NIH, OSTP, and OMB to carefully consider the points raised and we thank you for the opportunity to comment.

American Society of Clinical Oncology
American College of Physicians
NEJM Group
American Heart Association
American Diabetes Association
American Society of Anesthesiologists
American Urological Association
American Thoracic Society
American Gastroenterological Association
Endocrine Society
American Academy of Neurology
American Society of Nephrology
American College of Rheumatology

**Submit date: 4/21/2023** 

I am responding to this RFI: On behalf of an organization

Name: Joanna L. Groden, PhD

Name of Organization: University of Illinois Chicago

Type of Organization: University

Role: Institutional official

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

We appreciate the NIH's leadership in areas of open science and open data sharing and are delighted to see the move towards coordinating with the requirements of the Office of Science Technology Policy memo. We support the continued move towards open research which will benefit researchers, the general public, and communities around the world to improve their health and knowledge. As we actively support sharing all forms of scholarship equitably and responsibly, we are pleased to have the opportunity to provide feedback and raise a few ongoing concerns about sharing both publications and the underlying research data.

In order to ensure equity in publication opportunities for NIH researchers, there are significant challenges related to misunderstandings related to fee expectations and across different publishing models. We encourage the NIH to state explicitly that the researcher needs to pay no charge to comply with NIH's policy. It should be possible for a researcher to deposit the final peer-reviewed manuscript of any work funded by the NIH in PubMedCentral free of any payment to a publisher as a way to ensure that the researcher is in full compliance.

Journal business models requiring authors to pay a fee for journal publication (APC) present significant publication barriers for many researchers. Any open access fees charged by a publisher should only be for the standard APC for publications in verified fully open access journals. No additional fees should be required for compliance with NIH's Public Access Policy, either to make the article open or to submit it to PubMedCentral on the authors' behalf.

In addition to concerns about fees, creating a standard template for copyright agreements would improve equitable change for authors across disciplines. NIH should also offer clear language and processes that investigators can use upon submission to publishers to retain rights to make their peer-reviewed manuscript freely available and fully reusable post-publication in PMC without an embargo period. Specific instructions for doing this effectively would help authors comply with the policies, make federally funded research reusable, and further support NIH's goal to ensure equity in publishing.

Institutional repositories run by libraries and Other research institutions generally do not charge authors to deposit articles or manuscripts, and NIH should work with the U. S. Repository Network to encourage investigators to utilize options that are interoperable and free to use for deposit. For example, it would be great if researchers could deposit in one repository such as PMC through the NIHMS and have a way to also deposit the same material in the repository of their home institution. This allows universities to highlight and provide access to their organization and communities and allows researchers to display

their impact more effectively. Adding this secondary path for discovery also improves opportunities for access both by researchers and members of the public.

Finally, equitable publishing considerations will require that the NIH work with the higher education community to align research assessment and career advancement incentives to support scientific channels that actively promote equity in publication opportunities.

#### 2. Steps for improving equity in access and accessibility of publications.

To fulfill the reusability requirement, all publications resulting from NIH-funded research should carry open licenses, and NIH or authors should explicitly retain the rights needed to authorize those open licenses. As part of this, NIH should offer clear language that investigators can use to specifically retain rights to make their final peer-reviewed manuscript freely available and fully reusable post-publication in PMC without an embargo period. Placing a creative commons attribution-only license or its functional equivalent on a publication is the best way to ensure that publications can be freely accessed and fully reused.

As part of the grant development process, NIH should provide guidance for researchers on budgeting for publication costs, though we recognize that this will be highly variable across disciplines and programs. It is also advised that NIH have a cap on the amount that can be paid towards APC funds. Publishers' APC costs are increasing without added value to the services they provide. They will continue to increase their costs unless a cap is put on how much can be charged per article.

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

Models requiring authors to pay an Article Processing Charge (APC) fee for journal publication present significant publication barriers for most researchers. There is a significant risk that scholarly publishers will attempt to use the embargo removal to attempt to further extract increased funding from researchers seeking to comply. Publishers should be required to be transparent when charging hybrid OA fees that the researcher has the option to pay no fee in order to comply with the NIH Public Access Policy by submitting to PubMedCentral.

Additionally, the rising cost of APCs has already proven prohibitive to individuals and their institutions, resulting in fewer opportunities for publications and increasing disparities. Even this week colleagues such as the entire board of NeuroImage (an Elsevier journal) resigned over the high new publication fee. As research is already funded primarily by federal agencies and the majority of scholarly publishing labor including editing and peer review work is done on an unpaid volunteer basis or with modest stipends, the current practice of article processing fees on top of page fees and Other charges can only be seen as exploitative by the scholarly publishers.

Studies, such as Exploring the Hidden Impacts of Open Access Financing Mechanisms by AAAS, have documented that APC costs disproportionately affect younger researchers, female researchers, and those at less well-funded institutions. APCs also require a diversion of funds away from the research process; investigators often must use money originally intended for materials and equipment, supporting trainees, and professional development opportunities including presenting research results at conferences. This is in addition to the hundreds of millions of dollars annually spent by academic libraries attempting to provide and maintain access to subscription-based journals.

We recommend that the NIH monitor costs associated with APCs to ensure that federal research dollars are being spent as intended on research and that the costs of publishing are not creating arbitrary barriers to entry for researchers, and the ultimate availability of publishing opportunities for researchers at traditionally underrepresented institutions and in less-well-resourced disciplines. The NIH should monitor the cost of APCs levied on its investigators. Data collection on the amount spent to publish NIH-funded research regardless of the source would increase transparency and insight into how these fees affect various communities - including the potential impacts of publishing opportunities - on traditionally underrepresented institutions and in less-well-resourced disciplines.

#### 4. Early input on considerations to increase findability and transparency of research.

The NIH should ensure that the results of NIH-funded research along with metadata containing information about who conducted the research, where it was done, and with what resources. This requires NIH to articulate clear expectations about the use of Persistent Identifiers (PIDs) throughout the research process. Where possible, NIH should explicitly name and require the use of existing external identifiers (DOIs for data sets, DMPs, and publications, ORCIDs for authors, RORs for institutions, etc.) along with continued requirements for internal identifiers (PMCIDs, GeneBank Accession numbers, etc.).

Similar identifiers are required to be used by all federal agencies as a result of the OSTP Memorandum. The NIH should coordinate its efforts with Other participants in interagency working groups, including the National Science and Technology Council's (NSTC) Subcommittee on Open Science, to identify best practices and potential standards and announce these as soon as possible to allow institutions to advise researchers. NIH should also consider collaboration with a standards body, such as the National Information Standards Organization (NISO), to help to begin to develop a set of standards and framework for a national PIDs strategy to facilitate smooth implementation. NIH should consider mechanisms for increasing the findability and transparency of research, including exploring the use of the DOI system to overlay NIH's current unique identifiers for awards, and current best practices for assigning PIDs and collecting metadata for articles, clinical trials, and genetic sequences.

We appreciate that the NIH has recently implemented the new Data Management and Sharing Policy, however, we have remaining concerns about its impact on researchers and whether it will be sufficient guidance for researchers to be able to meet the goals stated in the OSTP memo.

There are several areas where the NIH and the DMSP implementation team have failed to provide requested guidance including providing a recommended minimum duration for data retention, have not yet addressed the challenges of the need for controlled data repositories, and have not yet made a public statement about managing intellectual property rights alongside meeting data sharing goals.

In regards to the duration of data preservation, the policy guides researchers to keep data as long as is appropriate or necessary. However, this does not provide a preservation baseline. As a result, it is difficult for researchers and institutions to appropriately budget for retaining and managing data, which may further compound inequities for smaller grants or institutions that are unable to provide local repositories. We recommend that the NIH establish a standard minimum timeline for preservation and collaborate with professional associations to identify best practices for data retention standards that address the complexity of data captured across the institutes and centers.

The DMSP implementation team has repeatedly touted the availability of NIH repositories and generalist repositories as mechanisms for researchers to use when planning to share their data. While these are a solution for very specific grant proposals, these resources cannot meet the general broad need for data sharing as required by the data policy. This can be seen when reviewing the NIH repositories in that many of them are not currently accepting new data. Further, the reliance on generalist repositories does not address the significant challenge of providing access to sensitive and controlled data. Instead, it has the potential to create a two-tier system for researchers whose institutions do not have a data repository and who must therefore use vendor-controlled generalist repositories and runs the risk of researchers inadvertently exposing sensitive data in order to meet data sharing expectations. We encourage the NIH to invest in and support the development of non-profit databases and repositories that will not only meet the generalist repository initiative but also engage further with the challenges related to controlled data access and preserving the privacy of sensitive data that we get from research participants.

We wish to ensure that data can be shared in a controlled fashion that does not inadvertently create further situations of harm where minority groups' data sets are mined and Otherwise used against their wishes in order to pursue research interests. As a minority-serving institution, we have encountered frequently challenges with interest in extracting data and the value of it from the individuals we serve without mechanisms that allow those communities to actively participate in the work that is being done and without returning specific value to those communities either in the forms of knowledge, education, job force training, or Other active development. This additionally runs the risk of becoming a target for vendors who seek to gatekeep data or charge exorbitant fees to manage access, further exacerbating data-sharing inequities.

The storage and preservation issues in particular must be addressed in order to ensure equitable participation in open sharing opportunities that do not further replicate historical inequities in what data can be retained and shared. This is of critical importance to fund funding for women and minority health programs and we encourage the reevaluation and the reallocation of funding to these programs to ensure access.

Email: pearsong@uic.edu

**Submit date: 4/21/2023** 

I am responding to this RFI: On behalf of an organization

Name: Matthew Thakur

Name of Organization: European Bioinformatics Institute EMBL-EBI

Type of Organization: Nonprofit research organization

Role: Institutional official

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

On Section III.D.1. NIH 'will continue to allow reasonable publication costs for all NIH-supported or authored scholarly publications consistent with current policy and guidance' - this section of the policy should state a more explicit intent to learn from, and in some areas align with, similar open access mandates of funders elsewhere, as in the case of the EU's Plan S.

### 2. Steps for improving equity in access and accessibility of publications.

On Section III.B.1. - this gives NIH's intent to 'continue the current practice of making manuscript files and Other article files submitted with permissive licenses available...'. To maximise the utility of manuscript files, can the use of permissive licenses also be mandated such that all of these files become machine readable? Guidance on licenses that reach the required level of permissiveness should be specified or even mandated (similarly to how explicit guidance is suggested below for preferred repositories)

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

On Section III.D.2. 'monitoring trends in publication fees', as all funders with open access mandates are likely to require or indeed already be carrying out similar monitoring, the policy should state a more explicit intent to coordinate with Other funders eg EU Plan S. Beyond monitoring, early indication of how NIH plans to ensure that publication costs remain reasonable would be helpful.

#### 4. Early input on considerations to increase findability and transparency of research.

The decision to remove embargos is a welcome one, and consistent with broader trends in scientific publishing which have accelerated in recent years eg the exponentially growing popularity of preprints.

On the requirement for DMS Plans - is the aim simply to cue researchers to spend more time planning their data management - or is the intent for DMS plans to become useful digital objects in their own right? Making these openly accessible, machine readable and, if possible, linked to a subsequent review of their implementation and any outputs would provide greater transparency and a means of monitoring how the realisation of data sharing measures up to intent at the award stage.

A big step forward would be the requirement to submit, at project end, together with the financial report, a DMS Report, which provides a point-by-point report on the intentions stated in the DMS Plan.

On 'Maximising Sharing' - is the intent to maximise the likely utility of the shared data to Other researchers, or to maximise its findability - or both? If the intent is that data should be both findable and

of greatest utility, this would place a greater obligation on investigators, but also align more with the desired impact of data sharing.

On Section IIA (Scientific data) the current proposals use 'use of PIDs and metadata' as a catch all for many types of research outputs which funders may need to make more findable. The policy is explicit about which outputs are exempt from the expectation for open sharing. Among these exemptions, laboratory specimens stand out as a data type which does currently have mechanisms in operation for findability/PIDs (eg through the BioSamples database), which suggests that an exemption may be unnecessary. Following the exemptions, similarly explicit expectations should be stated about the outputs which are to be shared - for example, whether making data alone findable/accessible is sufficient, or whether Other outputs such as software and beyond-preliminary analysis should also be made findable. Some of these output types have relatively well developed systems for persistent identifiers eg accessions and DOIs for datasets. Others are further behind eg software. The roadmap and lessons learned from institutions' experiences with making each findbale are likely to be very different.

On Section II.D - guidance of repositories used - the policy should take account of already existing systems for recognising databases of greatest value to the research community - such as the Core Trust Seal, Global Core Biodata Resource and ELIXIR Core Data Resource systems - rather than attempting to invent any new accreditation system.

On Section IV.A - metadata associated with data - in addition to the minimum metadata fields listed, the point on 'referencing digital persistent identifiers' should be developed further to highlight opportunities to build on existing PIDs such as ORCID for researchers, ROR IDs for research organizations, Accessions and Data DOIs for research data.

On Section IV.B. "Instruct federally funded researchers to obtain a digital persistent identifier..." While a requirement for individuals receiving NIH support to have Open Researcher and Contributor Identifiers (ORCID iDs) is laudable, additional policies may be required to ensure these are then used and linked to subsequent outputs, in order for the benefits of the PID to be realised.

The FAQ notes state that preprints are excluded from the public access policy - yet the data preprints refer to is included as per Section II.C. This seems rather inconsistent - does this not imply that the preprint based on the data should also be within scope for public access?

Email: mthakur@ebi.ac.uk

**Submit date: 4/21/2023** 

I am responding to this RFI: On behalf of an organization

Name: Kevin Wilson

Name of Organization: The American Society for Cell Biology

Type of Organization: Professional org association

- 1. How to best ensure equity in publication opportunities for NIH-supported investigators.
- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

#### **Uploaded File:**

ASCB-letter-to-Larry-Tabek-re-Open-Access-Policy.pdf

**Description:** ASCB position on NIH Open Access policy

Email: kwilson@ascb.org



6120 executive boulevard, suite 750 • rockville, maryland 20852, usa tel: 301-347-9300 • fax: 301-347-9310 • email: ascbinfo@ascb.org • website: www.ascb.org

January 30, 2023

Lawrence Tabak, DDS, PhD Acting Director U.S. National Institutes of Health Bethesda, Maryland 20892

Dear Dr. Tabek:

The American Society for Cell Biology (ASCB) is a professional society with more than 7,000 basic biomedical researchers in all 50 states and more than 60 countries around the world. Our membership ranges from first-year graduate students to Nobel laureates and, together, they help form the backbone of basic biomedical research in the United States and around the world. U.S. federal funding supports much of the research carried out by our members, which is communicated to the world via publication in peer-reviewed journals.

As an organization, ASCB fully supports the goal outlined in a recent memorandum from the Office of Science and Technology Policy (OSTP) of making "all federally funded research freely available without delay." Such public access to scientific research is important to ensure: (1) free and equitable distribution of knowledge gained through taxpayer support; (2) economic benefits of continual innovation; and (3) rapid progress of the scientific community, especially when working together to solve pressing problems.

We note, however, that some mechanisms for promoting access are better than others. For example, requiring that all federally funded research be published under a "Gold Open Access" policy (as defined as one where the final peer-reviewed, edited, formatted, and typeset article is made freely and immediately available to everyone with few restrictions on re-use) could substantially increase the amount of money that laboratories would have to pay to disseminate their work. For example, authors are charged more than \$11,000 to publish an article in the high-profile journal Nature under an arrangement similar to the Gold open access model described above.2 This publication cost burden would fall disproportionately on smaller laboratories with less overall funding,

#### OFFICERS

ERIKA HOLZBAUR President

MARTIN CHALFIE
Past President

SANDRA A. MURRAY President-Elect

KERRY BLOOM

A. MALCOLM CAMPBELL Treasurer

REBECCA ALVANIA Chief Executive Officer

#### COUNCIL

CASPER HOOGENRAAD ANDREW G. CAMPBELL AMY S. GLADFELTER ERIN D. GOLEY RONG LI BEATA EDYTA MIERZWA VERONICA A. SEGARRA LESLEY NICOLE WEAVER DEREK A. APPLEWHITE ELIZABETH H. CHEN WILLIAM C. EARNSHAW CHANTELL SKYE EVANS

#### COMMITTEE CHAIRS

SARA WONG
Co-Chair
SCOTT WILKINSON
Co-Chair
Committee for Postdocs & Students

MARINA CROWDER Co-Chair ALISON DELL Co-Chair

A. MALCOLM CAMPBELL Chair

Finance & Audit

Education

JORDAN RAFF Chair International Affairs

KERRY BLOOM Chair Membership

MARIAELENA B. ZAVALA Co-Chair LATANYA HAMMONDS-ODIE Co-Chair

Minorities Affairs

SANDRA A. MURRAY Co-Chair ANDREW G. CAMPBELL Co-Chair Nominating

AMY GLADFELTER Co-Chair SARA WICKSTROM Co-Chair Program

QUYEN AOH Co-Chair RYOMA "PUCK" OHI Co-Chair Public Information

HOLLY GOODSON Public Policy

EMILY MACE
Co-Chair
MARY MUNSON
Co-Chair
Women in Cell Biology

BRUNO DA ROCHA-AZEVED Co-Chair FRED CHANG Co-Chair *LGBTO*+

CBE—LIFE SCIENCES EDUCATION KIMBERLY TANNER & JEFF SCHINSKE Editor-in-Chief

MOLECULAR BIOLOGY OF THE CELL MATTHEW D. WELCH Editor-in-Chief and could dramatically exacerbate existing geographical and demographic disparities within the biomedical research community. Moreover, since the new mandate covers all federally funded research, it is also important to consider effects on other fields: the publication cost burden of Gold Open Access could have a broader negative impact on fields with less funding than the biomedical sciences.

Rather than mandating publication via Gold Open Access we urge the U.S. National Institutes of Health (NIH) to adopt a more effective, and more equitable policy, and allow authors to deposit the accepted version of their manuscript on a publicly accessible site simultaneously with publication of the final (edited, formatted, and typeset) article.

As for where to post the manuscripts, most funding agencies have an already-established repository for collecting the peer-reviewed results of funded work. For example, PubMed Central (PMC) is the repository for biomedical research funded by the NIH, while NSF uses the NSF Public Access Repository (NSF-PAR). NIH-and NSF-funded investigators have been working for many years with a mandate to submit their publications to PMC or NSF-PAR. Thus, the only change required to meet the new public access policy would be to mandate that these manuscripts be released into PMC, NSF-PAR, or other relevant repository immediately upon publication, without the currently allowed embargo/delay (presently up to one year).

We look forward to working with you on this important issue.

Sincerely,

Holly Goodson, PhD

Chair, Public Policy Committee American Society for Cell Biology

- Dr. Alondra Nelson. "MEMORANDUM FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES" Re: Guidance to Make Federally Funded Research Freely Available Without Delay. August 25, 2022. Office of Science and Technology Policy (OSTP). <a href="https://www.whitehouse.gov/ostp/news-updates/2022/08/25/ostp-issues-guidance-to-make-federally-funded-research-freely-available-without-delay/">https://www.whitehouse.gov/ostp/news-updates/2022/08/25/ostp-issues-guidance-to-make-federally-funded-research-freely-available-without-delay/</a>
- 2. <a href="https://www.nature.com/articles/d41586-020-03324-y">https://www.nature.com/articles/d41586-020-03324-y</a>

**Submit date:** 4/21/2023

I am responding to this RFI: On behalf of an organization

Name: Casey Rojas

Name of Organization: Massachusetts Medical Society - New England Journal of Medicine (NEJM)

Type of Organization: Not applicable

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

The New England Journal of Medicine (NEJM) is the most widely read, cited, and influential general medical journal and website in the world and the oldest continuously published medical periodical. Widely recognized as the gold standard for current research and best practices in medicine, NEJM publishes peer-reviewed research and interactive clinical content for physicians, educators, and the global medical community. Our mission is to bring health care professionals the most reliable biomedical research and clinical information to inform their practice and improve outcomes for patients. NEJM is a publication of NEJM Group, a division of the Massachusetts Medical Society, a non-profit corporation.

Thank you for the opportunity to respond.

We are writing to express our concern over the NIH implementation of policies in response to the 2022 Office of Science and Technology Policy (OSTP) Public Access Memorandum. We call on the NIH to remain neutral with respect to publishing business models, honor copyright, and not place additional burdens on researchers and small society publishers by mandating license requirements with overly broad reuse rights.

Patient-care professionals and the patients they serve rely on medical journal content that is vetted by medical experts, peer reviewed, revised, edited, and enhanced through the editorial process to provide them with results that are appropriately measured for making evidence-based clinical decisions. Each year, our editors filter through over 5,000 research manuscripts submitted and select only the best. Our editors are experts in their fields, most of whom are practicing clinicians, who work to ensure that conclusions are not overstated or misleading and that results are put into the proper context for treating patients. We strive to uphold standards around rigor and reproducibility, and we are investing in programs to improve equity not only in research but also in patient care and outcomes. Considering the medical misinformation that has spread over the last several years, the need for top quality and highly credible medical information has become even more apparent.

Each manuscript accepted for publication benefits from hundreds of hours of work by medical editors, statistical experts, manuscript editors, illustrators, proofreaders, and production staff, who work to ensure that every paper meets exacting standards before it becomes a published article. Our reader-pays subscription model allows us to continuously invest in subject-matter experts, statistical reviews, innovations in science communication, professional publishing talent, and editorial and production systems to ensure that NEJM meets the need of physicians and health care professionals for trusted, rigorously peer-reviewed research and review articles.

We fully believe the reader-pays business model is the best approach to serve our readers and their patients and to sustain our publication. The reader pays model maintains editorial independence and

protects against bias. Furthermore, this model also spreads the costs of publishing across many institutions and large number of readers rather than smaller number of authors.

We caution the NIH against requiring a CC-BY license, which by permitting derivative works allows for the misrepresentation and misuse of research results, increasing the risk for patient harm and leading to greater mistrust in science. In addition, forcing a specific CC-BY or similar requirement will severely diminish our ability to recover the substantial investments made in ensuring that NEJM articles meet our exacting quality standards and can be trusted by medical professionals to bring them the most impactful advancements in clinical care. We ask that the NIH policy refrain from requiring one size fits all licenses that permit broad commercial and derivative reuse rights.

Our reader-pays subscription model is the most equitable approach for ensuring that all authors have the opportunity to publish in our pages regardless of their financial means. And we firmly believe that authors should be able to choose where to publish.

We acknowledge that Other business models may work for Other publishers. However, we remain committed to a subscription-based publishing model, as that best fits the standards that we have set for ourselves and that our readers expect of NEJM. Further, for the reasons mentioned above, we believe that mandating a single approach to publishing — particularly one that favors high volume, rapid publication of medical research with less rigorous or no peer-review — will not result in a more equitable publishing ecosystem or better care for patients.

Thank you for this opportunity to provide information relevant to this important issue. NEJM looks forward to staying engaged and stands ready to assist in any way that we may be of assistance. Please feel free to reach out to Casey Rojas, Federal Relations and Health Equity Manager at crojas@mms.org with any questions or to continue this discussion.

- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

#### **Uploaded File:**

NIH-RFI-submission-21Apr2023\_.pdf

Email: crojas@mms.org

#### **NEJM Group NIH RFI submission**

How to best ensure equity in publication opportunities for NIH-supported investigators.

The NIH Public Access Plan aims to maintain the existing broad discretion for researchers and authors to choose how and where to publish their results. Consistent with current practice, the NIH Public Access Plan allows the submission of final published articles to PMC (in cases where a formal agreement is in place) to minimize the compliance burden on NIH-supported researchers and also maintains the flexibility of NIH-supported researchers to submit the final peer-reviewed manuscript. These submission routes are allowed regardless of whether or not the journal uses an open access model, a subscription model of publishing, or other publication model. This flexibility aims to protect against concerns that have been raised about certain publishing models potentially disadvantaging early career researchers and researchers from limited-resourced institutions or under-represented groups. NIH policy already allows supported researchers to charge reasonable publishing costs against their awards. NIH seeks information on additional steps it might consider taking to ensure that proposed changes to implementation of the NIH Public Access Policy do not create new inequities in publishing opportunities or reinforce existing ones.

The New England Journal of Medicine (NEJM) is the most widely read, cited, and influential general medical journal and website in the world and the oldest continuously published medical periodical. Widely recognized as the gold standard for current research and best practices in medicine, NEJM publishes peer-reviewed research and interactive clinical content for physicians, educators, and the global medical community. Our mission is to bring health care professionals the most reliable biomedical research and clinical information to inform their practice and improve outcomes for patients. NEJM is a publication of NEJM Group, a division of the Massachusetts Medical Society, a non-profit corporation.

Thank you for the opportunity to respond.

We are writing to express our concern over the NIH implementation of policies in response to the 2022 Office of Science and Technology Policy (OSTP) Public Access Memorandum. We call on the NIH to remain neutral with respect to publishing business models, honor copyright, and not place additional burdens on researchers and small society publishers by mandating license requirements with overly broad reuse rights.

Patient-care professionals and the patients they serve rely on medical journal content that is vetted by medical experts, peer reviewed, revised, edited, and enhanced through the editorial process to provide them with results that are appropriately measured for making evidence-based clinical decisions. Each year, our editors filter through over 5,000 research manuscripts submitted and select only the best. Our editors are experts in their fields, most of whom are practicing clinicians, who work to ensure that conclusions are not overstated or misleading and that results are put into the proper context for treating patients. We strive to uphold standards around rigor and reproducibility, and we are investing in programs to improve equity not only in research but also in patient care and outcomes. Considering the medical misinformation that has spread over the last several years, the need for top quality and highly credible medical information has become even more apparent.

Each manuscript accepted for publication benefits from hundreds of hours of work by medical editors, statistical experts, manuscript editors, illustrators, proofreaders, and production staff, who work to

ensure that every paper meets exacting standards before it becomes a published article. Our reader-pays subscription model allows us to continuously invest in subject-matter experts, statistical reviews, innovations in science communication, professional publishing talent, and editorial and production systems to ensure that NEJM meets the need of physicians and health care professionals for trusted, rigorously peer-reviewed research and review articles.

We fully believe the reader-pays business model is the best approach to serve our readers and their patients and to sustain our publication. The reader pays model maintains editorial independence and protects against bias. Furthermore, this model also spreads the costs of publishing across many institutions and large number of readers rather than smaller number of authors.

We caution the NIH against requiring a CC-BY license, which by permitting derivative works allows for the misrepresentation and misuse of research results, increasing the risk for patient harm and leading to greater mistrust in science. In addition, forcing a specific CC-BY or similar requirement will severely diminish our ability to recover the substantial investments made in ensuring that NEJM articles meet our exacting quality standards and can be trusted by medical professionals to bring them the most impactful advancements in clinical care. We ask that the NIH policy refrain from requiring one size fits all licenses that permit broad commercial and derivative reuse rights.

Our reader-pays subscription model is the most equitable approach for ensuring that all authors have the opportunity to publish in our pages regardless of their financial means. And we firmly believe that authors should be able to choose where to publish.

We acknowledge that other business models may work for other publishers. However, we remain committed to a subscription-based publishing model, as that best fits the standards that we have set for ourselves and that our readers expect of NEJM. Further, for the reasons mentioned above, we believe that mandating a single approach to publishing — particularly one that favors high volume, rapid publication of medical research with less rigorous or no peer-review — will not result in a more equitable publishing ecosystem or better care for patients.

Thank you for this opportunity to provide information relevant to this important issue. NEJM looks forward to staying engaged and stands ready to assist in any way that we may be of assistance. Please feel free to reach out to Casey Rojas, Federal Relations and Health Equity Manager at <a href="mailto:crojas@mms.org">crojas@mms.org</a> with any questions or to continue this discussion.

**Submit date:** 4/21/2023

I am responding to this RFI: On behalf of an organization

Name: Shawna Sadler

Name of Organization: ORCID

Type of Organization: Nonprofit research organization

Role: Institutional official

- 1. How to best ensure equity in publication opportunities for NIH-supported investigators.
- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

### **Uploaded File:**

ORCID\_Response\_to\_NIH-Signed.pdf

Email: <a href="mailto:s.sadler@orcid.org">s.sadler@orcid.org</a>



21 April 2023

National Institutes of Health 9000 Rockville Pike, Bethesda, Maryland 20892, United States

**RE: Request for Information;** NIH Plan to Enhance Public Access to the Results of NIH-Supported Research

To the National Institutes of Health,

Thank you for the opportunity to provide feedback to the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research, ORCID would like to respond to Questions 1 and 4.

ORCID appreciates NIH's various efforts to adopt persistent identifiers, specifically the ORCID policy for individuals supported by research training, fellowship, research education, and career development awards, and the work to include ORCID in the SciENcv system, populating data into the various forms in an effort to reduce administrative burden on researchers. We believe increasing the interoperability between NIH systems and ORCID's repository will improve the quality and timeliness of data about researchers, which in turn improves research integrity.

#### Question 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

The NIH Public Access Plan aims to maintain the existing broad discretion for researchers and authors to choose how and where to publish their results. Consistent with current practice, the NIH Public Access Plan allows the submission of final published articles to PubMed Central (PMC) (in cases where a formal agreement is in place) to minimize the compliance burden on NIH-supported researchers and also maintains the flexibility of NIH-supported researchers to submit the final peer-reviewed manuscript. NIH seeks information on additional steps it might consider taking to ensure that proposed changes to implementation of the NIH Public Access Policy do not create new inequities in publishing opportunities or reinforce existing ones.





#### **ORCID Response**

ORCID would like to encourage NIH to optimize and accelerate the adoption of persistent identifiers for organizations, people, funding awards and research outputs across internal systems and funded organizations and individuals. This will create the technical infrastructure required to support exchanging data across universities, funders and publishers in order to establish the necessary evidence base to monitor any unintended effects of the policy change.

#### Question 4. Early input on considerations to increase findability and transparency of research.

Section IV of the NIH Public Access Plan is a first step in developing the NIH's updated plan for persistent identifiers (PIDs) and metadata, which will be submitted to OSTP by December 31, 2024. NIH seeks suggestions on any specific issues that should be considered in efforts to improve use of PIDs and metadata, including information about experiences institutions and researchers have had with adoption of different identifiers.

#### **ORCID Response**

The specific issues that ORCID would like to propose for the NIH to consider when updating its plan for persistent identifiers (PIDs) and metadata relate to its level of ORCID adoption, as well as the adoption plans for other stakeholders in the research ecosystem.

A fully PID-optimized research ecosystem will deliver more accurate, and more timely data to stakeholders in the research community, realizing the vision for research with integrity.

Many of the recommended points below were developed by the Funders Interest Group that ORCID hosts, which includes representatives from 37 organizations from around the world including NIH, NSF and DOE.

NIH should optimize its adoption of ORCID by:

- Developing policy encouraging all applicants to have an ORCID iD
- Collecting authenticated ORCID iDs from all applicants during the grant application process
- Collecting authenticated ORCID iDs from all grant reviewers
- Assigning Grant DOIs to awarded grants and deposit all associated metadata, including the awardees' ORCID iDs, with the relevant PID service provider
- Writing the funding award metadata to researchers' ORCID records, which will help researchers
  meet the new funding disclosure requirements





- Publicly acknowledging the work of your peer reviewers by writing the peer review activity to their ORCID records
- Improving the ability to measure the impact of your research funds by utilizing persistent identifiers and their associated metadata.
- Encouraging the <u>Generalist Repository Ecosystem Initiative</u> (GREI) to include full ORCID integrations into its best practices

NIH should encourage **Research Institutions** like universities that receive NIH funding to:

- Ensure that funded researchers have an ORCID iD
- Collect authenticated ORCID iDs from researchers in their internal administrative and compliance tools
- Write relevant employment data to affiliated researchers' ORCID records so they can meet the new affiliation disclosure requirements
- Write relevant education data to affiliated students' ORCID records
- Write relevant data to visiting researchers' ORCID records for all "Invited Positions"

NIH should encourage Publishers of NIH-funded research outputs to:,

- Collect authenticated ORCID iDs for all authors and co-authors
- Collect authenticated ORCID iDs from all peer reviewers
- Assign DOIs to all publications
- Include ORCID iDs in the published and publicly available versions of the work
- Write the publication's metadata to the ORCID records of all authors and co-authors
- Publicly acknowledge the work of their peer reviewers by writing the peer review activity to their peer-reviewer's ORCID records
- Deposit all relevant persistent identifiers and metadata related to the publication, including its DOI, all ORCID iDs, Grant DOIs, ROR IDs, data DOIs, and research instrument identifiers with the relevant PID service providers

NIH should encourage **Developers and Managers of Repositories** that host NIH-funded research outputs to:,

- Collect authenticated ORCID iDs from researchers
- Integrate their repositories with the ORCID Registry
- Write the metadata describing the deposited scholarly publications, data sets, and other research outputs to researchers' ORCID records to help compliance with the new disclosure requirements.





We have elaborated on the above recommendations in our recent blog post, "ORCID Poised to Support Research Institutions in New Era of Public Access and Research Security" <a href="https://info.orcid.org/orcid-poised-to-support-research-institutions-in-new-era-of-public-access-and-research-security/">https://info.orcid.org/orcid-poised-to-support-research-institutions-in-new-era-of-public-access-and-research-security/</a>

I stand ready to speak with you further about our recommendation should this be useful,

Thank you,

DocuSigned by:

Chris Shillum

Executive Director, ORCID

## **DocuSign**

#### **Certificate Of Completion**

Envelope Id: 0500F040E78243B79AF0873E41BF568B

Subject: Complete with DocuSign: ORCID Response to NIH.pdf

Source Envelope:

Document Pages: 4 Signatures: 1
Certificate Pages: 1 Initials: 0

AutoNav: Enabled

Envelopeld Stamping: Enabled

Time Zone: (UTC-08:00) Pacific Time (US & Canada)

Status: Completed

Envelope Originator: Shawna Sadler

s.sadler@orcid.org IP Address: 174.0.33.147

#### **Record Tracking**

Status: Original

4/21/2023 9:48:43 AM

Holder: Shawna Sadler s.sadler@orcid.org

Location: DocuSign

### **Signer Events**

Chris Shillum c.shillum@orcid.org Executive Director

ORCID

Security Level: Email, Account Authentication

(None)

**Signature** 

DocuSigned by:

7AF6654653DF4F0...

Signature Adoption: Uploaded Signature Image

Using IP Address: 74.65.196.183

Signed using mobile

#### **Timestamp**

Sent: 4/21/2023 9:50:34 AM Viewed: 4/21/2023 11:52:23 AM Signed: 4/21/2023 11:52:42 AM

#### **Electronic Record and Signature Disclosure:**

Not Offered via DocuSign

In Danson Cinner France

Payment Events	Status	Timestamps
Completed	Security Checked	4/21/2023 11:52:42 AM
Signing Complete	Security Checked	4/21/2023 11:52:42 AM
Certified Delivered	Security Checked	4/21/2023 11:52:23 AM
Envelope Sent	Hashed/Encrypted	4/21/2023 9:50:34 AM
Envelope Summary Events	Status	Timestamps
Notary Events	Signature	Timestamp
Witness Events	Signature	Timestamp
Carbon Copy Events	Status	Timestamp
Certified Delivery Events	Status	Timestamp
Intermediary Delivery Events	Status	Timestamp
Agent Delivery Events	Status	Timestamp
Editor Delivery Events	Status	Timestamp
In Person Signer Events	Signature	Timestamp

**Submit date:** 4/21/2023

I am responding to this RFI: On behalf of an organization

Name of Organization: Association of American Medical Colleges

Type of Organization: Professional org association

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

As NIH is well aware, the existence of multiple publishing models as well as varied journal policies create inherent difficulties for researchers as they seek publication opportunities, navigate the processes for making articles publicly available, and access scholarly publications. The current system of publication and its increasing use of article processing charges can disparately impact early-career scientists and researchers in lower-resourced institutions or underfunded disciplines.

We appreciate the intent of NIH to implement an approach to public access which "maintains the flexibility of NIH-supported researchers to publish in the journal of their choice and submit the peerreviewed manuscript, regardless of whether or not the journal uses an open access model, a subscription model of publishing, or Other publication model." In order to achieve this goal, we request that NIH state clearly in the public access plan that researchers will be in full compliance with the requirement to make publications freely available and publicly accessible by depositing the peerreviewed manuscript into PubMed Central (PMC) and emphasize that this is an option which is available to researchers at no charge. Communicating this detail is an essential element so that researchers understand that NIH is not requiring that grantees publish in a journal that requires authors to pay a fee to enable access to their work, which may exacerbate disparities in publication opportunities. This point is particularly important given the diversity of language and statements used in publisher policies for open access and public access. While NIH does not set publisher policies, we believe there is value to the agency identifying and publicly noting those publishers and journals, such as JAMA and Science, with clear policies that support the NIH public access plan by allowing immediate deposition of the authoraccepted manuscript into a public repository. Finally, we also request that information on PMC submission methods, as well as the Public Access Compliance Monitor, be clearly linked in the plan to assist institutions and researchers with this requirement.

AAMC appreciates the clear assertion that "NIH reinforces that NIH-supported authors should retain rights to the final peer-reviewed manuscript, regardless of the pathway to publication." We ask that proposed language for rights retention be included in the draft plan and released for public comment. We also refer NIH to the language developed by many funders within cOAlition S for researchers to submit to publishers along with their manuscript. The suggested language from NIH will not only be critical for researchers to be able to submit their manuscript to PMC, but also for use and re-use of information contained in and across publications, an essential component to maximize the benefit of the growing number of publications available on PMC.

#### 2. Steps for improving equity in access and accessibility of publications.

AAMC believes that access to publications by diverse communities of users, including researchers, clinicians and public health officials, students and educators, and patients and Other members of the public, should be the driving goal of the public access plan, and considered in any decisions the NIH

makes. We recognize the historical inequity in access to publications, especially for individuals not associated with a well-resourced institution.

We appreciate the current practice of making scholarly publications available in accessible and machine-readable formats through PMC. We encourage NIH to continue to work with the broader community on improving article accessibility as well as the PMC interface, particularly to ensure that standards adapt to the latest technology, and also that the agency consider the many factors and broad definition of disability which may impact accessibility, to include physical, sensory, learning, psychological, and chronic health conditions.

AAMC notes the NIH assertion it will "provide additional educational materials and resources to assist the investigator community in improving the accessibility of articles." We request that any resources and educational materials regarding accessibility be directly linked in the final policy and easily findable by NIH grantees.

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

We believe that NIH is uniquely positioned to understand the nature and amounts of publications costs for NIH-funded researchers. We encourage NIH to develop a systematic effort to collect this information and to understand how these costs impact grant budgets and may differentially affect under resourced investigators and institutions. Given the different mechanisms for funding publication costs (grant-based, departmental, library funds, etc.), we suggest that NIH look beyond the grant budget line item for publications to capture publication expenses more fully. Potential methods for capturing this information include surveying researchers at closeout for additional information on publication costs or through a commissioned study. We also ask that NIH commit to sharing the findings of this research back out to the research community.

As stated in the plan, NIH "proposes to continue to monitor trends in publication fees and policies to ensure that they remain reasonable and do not serve as an impediment to publishing by researchers from limited-resourced institutions or under-represented groups." While AAMC supports the efforts to understand publication costs, this statement does not adequately assure the research community that the NIH will be in a position to address the fees and policies that may prevent some researchers from publishing in certain journals. There is a substantial gap between monitoring costs and ensuring that they remain reasonable. This cannot be accomplished without collaborating and reaching consensus across a wide range of publishers, an undertaking which has proved challenging. We urge NIH to provide additional information regarding the actions that NIH is able to take and would pursue in the case that publication costs are found to serve as an impediment to publishing.

Although NIH has made efforts to uncouple compliance with the public access plan from any particular publication model, we note that the plan, along with many similar changes and requirements from Other funders, will have an upstream impact on journals, whether owned by major publishers or small societies. Changes to how articles are accessed will feed into an ongoing and important conversation about the sustainability of current models of publication and how journals are funded, that will have broader consequences than what is discussed in this RFI. Academic researchers are impacted by the publishing process at multiple steps, not only by their ability to submit to certain journals and access articles, but also the entrenched role that publications in high-impact journals, long held as the gold standard in quality, have in determining tenure and promotion.

Finally, we appreciate NIH's intent to develop supplemental information that elaborates on and clarifies allowable costs for publication and believe this would be most useful for the grantee community if developed and released along with the draft plan to allow time for feedback. We also note the longstanding issue that current publication timelines often do not fit within the closeout period for an NIH grant and urge the agency to take this into consideration.

#### 4. Early input on considerations to increase findability and transparency of research.

The AAMC strongly supports the use of persistent identifiers (PIDs) and metadata, not only to increase the findability of research, but also to link researchers to their research outputs, whether this be publications, data, code, or any Other products. AAMC supports a requirement for NIH grantees to have an ORCID ID, as well as DOIs for publications and data resulting from NIH-funded research. As the agency develops these policies, we refer NIH to the considerations for PID adoption from our fellow higher education organizations. Additionally, as AAMC has long been invested in tracking trainee career outcomes, we support the requirement for individuals receiving research training, fellowship, research education, and career development awards to also have an ORCID ID.

As NIH notes, PIDs are most useful when they can be linked in standardized ways, and we encourage NIH to look not only to Other federal agencies, but also to community organizations, institutions, and societies. Cross-stakeholder groups such as the Research Data Alliance and FORCE11 have spent years developing suggested protocols and standards for both PIDs and metadata that align with the FAIR and TRUST principles. We also emphasize that being able to find and use the shared data resulting from the NIH Data Management and Sharing Policy will require significant investment in infrastructure and agency guidance on metadata standards. AAMC recommends that PIDs for research outputs can be easily linked and found when searching grants on NIH RePORTER.

#### **Uploaded File:**

AAMC-Comments-on-NIH-NOT-OD-23-091.pdf

Email: adev@aamc.org



**Association of** 

American Medical Colleges 655 K Street, N.W., Suite 100, Washington, D.C. 20001-2399 T 202 828 0400 F 202 828 1125

April 21, 2023

NIH Office of Science Policy 6705 Rockledge Drive, Suite 630, Bethesda, MD 20892

Re: Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-**Supported Research (NOT-OD-23-091)** 

Submitted electronically at https://osp.od.nih.gov/nih-plan-to-enhance-public-access-to-the-resultsof-nih-supported-research/.

The Association of American Medical Colleges (AAMC) appreciates the opportunity to provide feedback to the National Institutes of Health (NIH) on the NIH plan to enhance public access to the results of NIH-supported research.

The AAMC is a nonprofit association dedicated to improving the health of people everywhere through medical education, health care, medical research, and community collaborations. Its members are all 157 U.S. medical schools accredited by the Liaison Committee on Medical Education; 13 accredited Canadian medical schools; approximately 400 teaching hospitals and health systems, including Department of Veterans Affairs medical centers; and more than 70 academic societies. Through these institutions and organizations, the AAMC leads and serves America's medical schools and teaching hospitals and the millions of individuals across academic medicine, including more than 193,000 full-time faculty members, 96,000 medical students, 153,000 resident physicians, and 60,000 graduate students and postdoctoral researchers in the biomedical sciences. Following a 2022 merger, the Alliance of Academic Health Centers and the Alliance of Academic Health Centers International broadened the AAMC's U.S. membership and expanded its reach to international academic health centers.

The AAMC continues to support federal efforts to increase access to publications and research data resulting from federally funded research. As previously noted in comments to the White House Office of Science and Technology Policy (OSTP)<sup>1</sup> and NIH<sup>2</sup>, "Making these outputs more readily

<sup>&</sup>lt;sup>1</sup>AAMC Comments to OSTP. Request for Information: Public Access to Peer-Reviewed Scholarly Publications, Data and Code Resulting from Federally Funded Research (85 FR 9488). May 6, 2020. https://www.aamc.org/media/44641/download?attachment.

<sup>&</sup>lt;sup>2</sup> AAMC Comments to NIH. Re: NOT-OD-20-013: Request for Public Comments on a DRAFT NIH Policy for Data

available advances science by enabling further validation of experimental results, facilitating reuse of hard to-generate data, catalyzing new research and scientific collaboration, and generally promoting more responsible stewardship of federal resources." We also understand that these efforts are complex and resource-intensive and in order to be effective and equitable, must engage the whole of the research enterprise. This includes federal agencies, academic institutions, and the publishing community, as well as community organizations which have been instrumental in creating standards and practices for effective research dissemination. We provide feedback below on the specific topics requested by NIH.

#### How to best ensure equity in publication opportunities for NIH-supported investigators.

As NIH is well aware, the existence of multiple publishing models as well as varied journal policies create inherent difficulties for researchers as they seek publication opportunities, navigate the processes for making articles publicly available, and access scholarly publications. The current system of publication and its increasing use of article processing charges can disparately impact early-career scientists and researchers in lower-resourced institutions or underfunded disciplines.

We appreciate the intent of NIH to implement an approach to public access which "maintains the flexibility of NIH-supported researchers to publish in the journal of their choice and submit the peerreviewed manuscript, regardless of whether or not the journal uses an open access model, a subscription model of publishing, or other publication model." In order to achieve this goal, we request that NIH state clearly in the public access plan that researchers will be in full compliance with the requirement to make publications freely available and publicly accessible by depositing the peer-reviewed manuscript into PubMed Central (PMC) and emphasize that this is an option which is available to researchers at no charge. Communicating this detail is an essential element so that researchers understand that NIH is not requiring that grantees publish in a journal that requires authors to pay a fee to enable access to their work, which may exacerbate disparities in publication opportunities. This point is particularly important given the diversity of language and statements used in publisher policies for open access and public access. While NIH does not set publisher policies, we believe there is value to the agency identifying and publicly noting those publishers and journals, such as JAMA<sup>3</sup> and Science<sup>4</sup>, with clear policies that support the NIH public access plan by allowing immediate deposition of the author-accepted manuscript into a public repository. Finally, we also request that information on PMC submission methods, as well as the Public Access Compliance Monitor, be clearly linked in the plan to assist institutions and researchers with this requirement.

Management and Sharing and Supplemental Draft Guidance. Jan. 10, 2020. https://www.aamc.org/media/40536/download?attachment.

<sup>&</sup>lt;sup>3</sup> Bibbins-Domingo K, Shields B, Ayanian JZ, et al. Public Access to Scientific Research Findings and Principles of Biomedical Research—A New Policy for the JAMA Network. *JAMA*. 2023;329(1):23–24. doi:10.1001/jama.2022.23451.

<sup>&</sup>lt;sup>4</sup> Parikh, S. Malcolm SM, and B Moran. Public access is not equal access. *Science* 377, 1361-1361(2022). doi:10.1126/science.ade8028

AAMC appreciates the clear assertion that "NIH reinforces that NIH-supported authors should retain rights to the final peer-reviewed manuscript, regardless of the pathway to publication." We ask that proposed language for rights retention be included in the draft plan and released for public comment. We also refer NIH to the language developed<sup>5</sup> by many funders within cOAlition S for researchers to submit to publishers along with their manuscript. The suggested language from NIH will not only be critical for researchers to be able to submit their manuscript to PMC, but also for use and re-use of information contained in and across publications, an essential component to maximize the benefit of the growing number of publications available on PMC.

#### Methods for monitoring evolving costs and impacts on affected communities.

We believe that NIH is uniquely positioned to understand the nature and amounts of publications costs for NIH-funded researchers. We encourage NIH to develop a systematic effort to collect this information and to understand how these costs impact grant budgets and may differentially affect under resourced investigators and institutions. Given the different mechanisms for funding publication costs (grant-based, departmental, library funds, etc.), we suggest that NIH look beyond the grant budget line item for publications to capture publication expenses more fully. Potential methods for capturing this information include surveying researchers at closeout for additional information on publication costs or through a commissioned study. We also ask that NIH commit to sharing the findings of this research back out to the research community.

As stated in the plan, NIH "proposes to continue to monitor trends in publication fees and policies to ensure that they remain reasonable and do not serve as an impediment to publishing by researchers from limited-resourced institutions or under-represented groups." While AAMC supports the efforts to understand publication costs, this statement does not adequately assure the research community that the NIH will be in a position to address the fees and policies that may prevent some researchers from publishing in certain journals. There is a substantial gap between monitoring costs and ensuring that they remain reasonable. This cannot be accomplished without collaborating and reaching consensus across a wide range of publishers, an undertaking which has proved challenging. We urge NIH to provide additional information regarding the actions that NIH is able to take and would pursue in the case that publication costs are found to serve as an impediment to publishing.

Although NIH has made efforts to uncouple compliance with the public access plan from any particular publication model, we note that the plan, along with many similar changes and requirements from other funders, will have an upstream impact on journals, whether owned by major publishers or small societies. Changes to how articles are accessed will feed into an ongoing and important conversation about the sustainability of current models of publication and how journals are funded, that will have broader consequences than what is discussed in this RFI. Academic researchers are impacted by the publishing process at multiple steps, not only by their ability to

<sup>&</sup>lt;sup>5</sup> https://www.coalition-s.org/wp-content/uploads/2023/04/cOAlitionSresponseForNIH.pdf

submit to certain journals and access articles, but also the entrenched role that publications in high-impact journals, long held as the gold standard in quality, have in determining tenure and promotion.

Finally, we appreciate NIH's intent to develop supplemental information that elaborates on and clarifies allowable costs for publication and believe this would be most useful for the grantee community if developed and released along with the draft plan to allow time for feedback. We also note the longstanding issue that current publication timelines often do not fit within the closeout period for an NIH grant and urge the agency to take this into consideration.

#### Steps for improving equity in access and accessibility of publications.

AAMC believes that access to publications by diverse communities of users, including researchers, clinicians and public health officials, students and educators, and patients and other members of the public, should be the driving goal of the public access plan, and considered in any decisions the NIH makes. We recognize the historical inequity in access to publications, especially for individuals not associated with a well-resourced institution.

We appreciate the current practice of making scholarly publications available in accessible and machine-readable formats through PMC. We encourage NIH to continue to work with the broader community on improving article accessibility as well as the PMC interface, particularly to ensure that standards adapt to the latest technology, and also that the agency consider the many factors and broad definition of disability which may impact accessibility, to include physical, sensory, learning, psychological, and chronic health conditions.

AAMC notes the NIH assertion it will "provide additional educational materials and resources to assist the investigator community in improving the accessibility of articles." We request that any resources and educational materials regarding accessibility be directly linked in the final policy and easily findable by NIH grantees.

#### Early input on considerations to increase findability and transparency of research.

The AAMC strongly supports the use of persistent identifiers (PIDs) and metadata, not only to increase the findability of research, but also to link researchers to their research outputs<sup>6</sup>, whether this be publications, data, code, or any other products. AAMC supports a requirement for NIH grantees to have an ORCID ID, as well as DOIs for publications and data resulting from NIH-funded research. As the agency develops these policies, we refer NIH to the considerations for PID adoption from our fellow higher education organizations<sup>7</sup>. Additionally, as AAMC has long been invested in

<sup>&</sup>lt;sup>6</sup> Pierce, H.H., Dev, A., Statham, E., Bierer, B.E. Credit Data Generators for Data Reuse. *Nature*. 2019 June; 570 (7759): 30-32. doi: <a href="https://doi.org/10.1038/d41586-019-01715-4">https://doi.org/10.1038/d41586-019-01715-4</a>

<sup>&</sup>lt;sup>7</sup> Implementing Effective Data Practices: Stakeholder Recommendations for Collaborative Research Support. September 23, 2020. <a href="https://www.arl.org/wp-content/uploads/2020/09/2020.09.25-implementing-effective-data-practices.pdf">https://www.arl.org/wp-content/uploads/2020/09/2020.09.25-implementing-effective-data-practices.pdf</a>

tracking trainee career outcomes, we support the requirement for individuals receiving research training, fellowship, research education, and career development awards to also have an ORCID ID.

As NIH notes, PIDs are most useful when they can be linked in standardized ways, and we encourage NIH to look not only to other federal agencies, but also to community organizations, institutions, and societies. Cross-stakeholder groups such as the Research Data Alliance and FORCE11 have spent years developing suggested protocols and standards for both PIDs and metadata that align with the FAIR and TRUST principles. We also emphasize that being able to find and use the shared data resulting from the NIH Data Management and Sharing Policy will require significant investment in infrastructure and agency guidance on metadata standards. AAMC recommends that PIDs for research outputs can be easily linked and found when searching grants on NIH RePORTER.

The AAMC looks forward to continued engagement with the NIH during the development of the agency's public access plan. We are happy to work with the NIH to identify AAMC member institutions or societies to participate in conversations regarding any of these specific topics. Please feel free to contact me or my colleague Anurupa Dev, PhD, Director of Science Policy and Strategy (adev@aamc.org) with any questions about these comments.

Sincerely,

Ross McKinney, Jr., MD Chief Scientific Officer

Ross M. Limy (. WW)

cc: David J. Skorton, MD, AAMC President and Chief Executive Officer

**Submit date:** 4/21/2023

I am responding to this RFI: On behalf of an organization

Name: Caroline Trupp Gil

Name of Organization: American Chemical Society

Type of Organization: Professional org association

Role: Institutional official

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

The American Chemical Society (ACS) is a congressionally chartered non-for-profit organization and the world's largest scientific society with more than 173,000 individuals in our global membership community across 140 countries. ACS advances knowledge and research through scholarly publishing, scientific conferences, information resources for education and business, and professional development efforts.

As a socially responsible organization deeply rooted in the scholarly community, we share NIH's goal to ensure equity in publication opportunities. The best way to achieve this goal is to ensure that all stakeholders in the process of transitioning to immediate open access, e.g., researchers, funders, and institutions, understand that every method of open access publication has a cost that must be funded and budgeted - and that competition and diversity in publication outlets is the best way to maximize the efficiency, and therefore the cost, of those outlets.

Researchers need specific guidance on planning for and budgeting any new requirements: including budgeting during the grant application process to account for anticipated publications costs. We suggest that NIH work with organizations like ACS to help develop budgeting guidance. Encouragement and education should be provided at the start of the grant process to make sure that appropriate planning takes place. For recent open science examples, the NIH Data Guidance (which explicitly directs grantees to budget for data sharing and curation costs) and NASA SPD-41a (which encourages researchers to fund open access articles) could be used as models.

Of the different methods that can ensure equity at scale in publication opportunities, direct funder support for publishing, i.e., Gold Open Access (Gold OA), is the most financially sustainable. This is because researchers can be secure in the knowledge that they have the funds needed to support publication in the outlet of their choice and the outlets themselves have a reliable source of funding with which to continue their operations and ensure the integrity of the content published. Gold OA at the ACS, as with many Other society publishers, is a dynamic and customizable option for researchers to enable immediate OA. We have a robust waiver and discount program that helps researchers from lowand middle-income countries to publish at highly discounted rates; currently a minimum of 75% discount, rising to a complete waiver for low-income countries.

Immediate access to an accepted manuscript version of an article, i.e., immediate Green OA, has not proven to work at scale, even if it may work for a very small number of publishers or disciplines. It often appears cost free to researchers and Others, but in fact it is reliant on subscriptions to cover the cost of peer review and publication. A widespread use of this method, in conjunction with tools such as

Unsub.org that explicitly encourages institutions to cancel subscriptions where alternative free versions of articles are accessible, threatens the viability of the subscription funding on which Green OA methods of providing public access rely. The loss of subscription funding in this context, means depleted resources available for publications to ensure the quality and integrity of the scientific record. This will directly result in erosion of public trust in science and a dampening effect on innovation, job growth, and scientific progress. It will also increase the likelihood that important publication outlets will cease operations due to lack of funds, creating new barriers to access and equity in publication opportunities. Smaller and not-for-profit publishers, including those associated with learned societies, are most likely to be at risk from this practice that could easily result in increased market consolidation. This, in turn, is likely to reduce author choice and market competition, stifling innovation and undermining equity in publication opportunities.

We recommend that NIH avoid creating these barriers, especially for scientists from traditionally marginalized communities, as well as early career researchers, by ensuring that all its grantees have the funding support necessary to enable their research and choose the publishing option that best suits their needs.

We encourage NIH to read and reference the position statements (https://www.stm-assoc.org/stm-oa-position/) by STM on this subject, representing much of the publishing industry.

#### 2. Steps for improving equity in access and accessibility of publications.

NIH can improve equity in access and accessibility of publications by helping to educate researchers that the publication cost of immediate open access is as much a part of the dissemination of research reports as attendance at scientific conferences and gatherings. They can achieve this by ensuring that adequate funds are available to researchers to enable them to support immediate open access and by advocating for the long-term funding support from Congress needed to enable equity in access and accessibility. NIH is also encouraged to initiate public-private partnerships with organizations like ACS that provide discovery tools widely used by scientists globally to seamlessly identify research reports, data, and analyses that fuel innovation, economic prosperity, and scientific progress.

Of the different methods designed to achieve equity at scale in access, Gold OA has the greatest chance of success and NIH initiatives such as the Cancer Moonshot and Helping to End Addiction Long-term (HEAL) Initiative are both examples of programs that provide financial support to achieve their OA goals. Gold Open Access at the ACS, as with many Other society publishers, is a dynamic and customizable option for researchers to enable immediate OA. We have a robust waiver and discount program that helps researchers from low- and middle-income countries to publish at highly discounted rates; currently a minimum of 75% discount, rising to a complete waiver for low-income countries. Gold OA is a powerful model for enabling universal access to the most authoritative publications reporting on the results of scientific research, the Version of Record (VoR). The VoR is the authoritative version for researchers and the public, and is more cited and used, and garners more attention and trust than Other versions. It can link bi-directionally to research objects like data and code, is continually updated, and is hosted on the publisher's platform where it can be integrated with Other relevant content and analytical tools.

We are aware of NIH's desire to be business model agnostic and therefore caution against the promotion of immediate access to accepted manuscript versions of an article, i.e., immediate Green OA,

especially through the so-called "rights retention strategy" (RRS) that some have observed in the NIH plan, e.g., at section III.C.1. Immediate Green OA has not proven to work at scale, even if it may work for a very small number of publishers or disciplines. It often appears cost free to researchers and Others, but in fact it is reliant on subscriptions to cover the cost of peer review and publication. A widespread use of this method, in conjunction with tools such as Unsub.org that explicitly encourages institutions to cancel subscriptions where alternative free versions of articles are accessible, threatens the viability of the subscription funding on which Green OA methods of providing public access rely. The loss of subscription funding in this context, means depleted resources available for publications to ensure the quality and integrity of the scientific record. This will directly result in erosion of public trust in science and a dampening effect on innovation, job growth, and scientific progress. It will also increase the likelihood that important publication outlets will cease operations due to lack of funds, creating new barriers to access and equity in publication opportunities. Smaller and not-for-profit publishers, including those associated with learned societies, are most likely to be at risk from this practice that could easily result in increased market consolidation. This, in turn, is likely to reduce author choice and market competition, stifling innovation and undermining equity in publication opportunities.

We recommend that NIH avoid creating these barriers, especially for scientists from traditionally marginalized communities, as well as early career researchers, by ensuring that all its grantees have the funding support necessary to enable their research and choose the publishing option that best suits their needs. We encourage NIH to read and reference the position statements (<a href="https://www.stm-assoc.org/stm-oa-position/">https://www.stm-assoc.org/stm-oa-position/</a>) by STM on this subject, representing much of the publishing industry.

Immediate Green OA also contributes to version-control issues and potential confusion because, although there can be important and even critical differences in the text, an accepted manuscript and a VoR can look the same in their raw versions - implying trust when this could be misplaced. It risks slowing the move towards full open access because it is not a publishing model in itself but is primarily supported via subscriptions which leave the most valuable version of an article, the VoR, subject to access controls. Finally, the "rights retention strategy" approach to immediate Green OA restricts rather than expands a scientist's ability to choose how best to maximize the benefits of their work. For these reasons, immediate Green OA cannot deliver on the promise of an easily accessible, navigable, and interconnected Open Research ecosystem.

ACS instead recommends that researchers be allowed to publish under rights consistent with their vision and needs, including non-commercial, non-derivative licenses. We support access methods that are most consistent with academic freedom of expression globally based upon the responsible exercise of independent editorial control.

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

In our answer to question one, we addressed how NIH can best ensure equity in publication opportunities. Here we will respond to the question of monitoring publication fees. The simplest and most effective way for NIH to keep itself informed about publication fees is to partner with publishers and organizations like ACS whose fees are publicly posted on their websites. This practice would not only ensure transparency around costs, but also enable NIH to confirm that grantees are paying a fair market price for the services and value provided. We note that cost structures are very different for different organizations - medicine, physical sciences, social sciences, and humanities - and for different types of journals based on selectivity, services, technology, and Other features. A diverse, financially

sustainable, and robust publishing system which provides authors with broad choice is the most effective way to control cost. We caution against inflexible cost caps which will likely drive existing industry trends toward publisher consolidation and volume-based models which compromise integrity, quality, and author choice.

One constant, regardless of the field of research endeavor, is that rigorous publications are essential to support scientific communication and public trust in science. Researchers and policy makers must be able to rely on the integrity of the scientific publications that inform their decisions. The public, in turn, must be able to feel confident that practitioners' and policymakers' scientific and technical decisions are grounded in accurate information. Organizations like ACS are deeply committed to supporting this integrity and trust in science by building and maintaining infrastructure that enables the widespread production and communication of validated and reliable reports on scientific research. Among Other things, this involves creating scientific journals and staffing their editorial boards with experts that read and evaluate thousands of submitted manuscripts for quality and relevance. ACS also spends significant resources to ensure the integrity of journal articles by verifying author and content integrity, assessing articles for ethical considerations, managing and underscoring authors' potential conflicts of interest, and conducting plagiarism, ghost and gift authorship checks to combat paper mills, image manipulation, and the use of artificial intelligence tools like ChatGPT in inappropriate ways.

Our investments in support of scientific communication do not end when a peer reviewed article is published. We update articles for correction and addenda, update links, and conduct ongoing plagiarism and copyright protection to safeguard the integrity of the work and ensure articles are not modified or pirated in misleading and harmful ways. Upholding the version of record and providing the clarity necessary to easily distinguish between the version of record and earlier, less reliable versions of an article, is a key principle of scientific integrity. In order to build trust in science, readers must be able to easily identify and discover trusted peer reviewed content. To facilitate this process, we assign digital identifiers, provide metadata, conduct search engine optimization, track citations and Other important metrics, and submit articles to abstracting, indexing, and discovery services. These valuable services support scientific integrity by pointing readers to the highest quality scientific publications and data.

At a time when concerns around misinformation — including on critical issues of science and medicine — have become a national priority, there is an urgent need for stakeholders that support scientific integrity to work together and uphold the role of objective, trusted information in a democratic society. Therefore, it is essential that federal policies related to publications ensure that scientists and publishers can continue producing and disseminating the trusted, peer reviewed, VoR of scientific articles by providing sufficient funding for researchers who choose to publish OA to support investments in publishing their works in high-quality journals that uphold scientific integrity.

#### 4. Early input on considerations to increase findability and transparency of research.

ACS is a participant in the Open Research Funders Group's persistent identifiers (PIDs) discussions. We regularly engage with developments around new PIDs and support best practice in ensuring the accurate and enduring tracking of all relevant aspects of the research cycle. It may be helpful to NIH to know how we are taking steps to increase the findability and transparency of research data, perhaps the most challenging object of PID activity. We have data policies and guidelines, consistent with principles of open science, to ensure results reported in ACS journals are verifiable, reproducible, and easily accessible to researchers. The ACS Research Data Policy (https://publish.acs.org/publish/data\_policy)

provides best practice recommendations for data citation, data availability statements, and the use of appropriate data repositories. An evolving set of Data Guidelines by sub-field and data-type provides authors with specific instructions on how to make data available and comply with discipline-specific standards.

We are members of the Research Data Alliance (<a href="https://rd-alliance.org/">https://rd-alliance.org/</a>), a community-driven initiative by the European Commission, the National Science Foundation and National Institute of Standards and Technology, and Australia's Department of Innovation to build the social and technical infrastructure to enable open sharing and re-use of data. We have endorsed the Joint Declaration of the Data Citation Principles (<a href="https://force11.org/info/joint-declaration-of-data-citation-principles-final/">https://force11.org/info/joint-declaration-of-data-citation-principles-final/</a>) that provides a set of guiding principles for data within scholarly literature, anOther dataset, or any Other research object. ACS Publications has also signed the Declaration on Research Assessment (DORA - <a href="https://sfdora.org/">https://sfdora.org/</a>) and made citation data for all ACS journals openly available.

Finally, we have created the ACS Research Data Center (<a href="https://acsopenscience.org/open-science/acs-research-data-center/">https://acsopenscience.org/open-science/acs-research-data-center/</a>) as part of ACS Publications evolution, experimentation, and innovation with new models of OA. It is designed to help researchers forge new partnerships, improve the visibility of their research findings, and facilitate the means by which they can disseminate their work to a wider audience.

#### **Uploaded File:**

ACS-Supplemental-comments-2023-04-20-NIH.pdf

**Description:** Additional comments not included in the answers to the four questions posed in the RFI.

Email: c\_truppgil@acs.org



# Office of Secretary and General Counsel Anthony Pitagno

Senior Director, Government Affairs Outreach & Alliances

April 21, 2023

RE: Additional ACS Comments on the Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research

The American Chemical Society (ACS) is a congressionally chartered non-for-profit organization and the world's largest scientific society with more than 173,000 individuals in our global membership community across140 countries. ACS advances knowledge and research through scholarly publishing, scientific conferences, information resources for education and business, and professional development efforts. ACS' Publications Division exists to provide ACS members and the worldwide scientific community with a comprehensive collection, in any medium, of high-quality information products and services that advance the practice of chemistry and related sciences. In addition to our Chemical & Engineering News magazine and Symposia Series and ACS in Focus e-books program, we publish over 60,000 research papers annually in 80 journals and provide searchable access to over 1.3 million original chemistry articles dating back to 1879.

We welcome the opportunity to respond to the Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research and would like to share the following additional comments not included in the answers to the four questions posed in the consultation.

#### Research Data Sharing (Section II)

We support the goal of maximizing reuse and re-analysis of research data. As noted in our response to question 4 we have standardized our journal data sharing policies in line with <u>current standards</u> <u>from the Research Data Alliance publisher working group</u>. In the absence of a widely supported institutional infrastructure for data sharing, it seems likely that at some point there will be a need for funding support in this area. This can already be seen with commercial services like figshare where datasets over a certain size and supported a fee. These fees may scale in relation to the duration of sharing which should be considered in section II.C. We note that <u>Supplemental Information to the NIH Policy for Data Management and Sharing: Allowable Costs for Data Management and Sharing makes it clear that such costs are considered allowable under grant conditions. However, to avoid replicating some of the issues from the early days of OA publications, we recommend creating a mechanism for reimbursement after the grant has ended.</u>

Thank you for consideration of these points. Further inquiries regarding this comment may be directed to Anthony Pitagno, Senior Director, Government Affairs Global Outreach and Alliances, American Chemical Society, 1155 Sixteenth Street, NW, Washington DC 20036; 202.872.4394; a pitagno@acs.org.

Sincerely,

Anthony Pitagno Senior Director

**Submit date: 4/21/2023** 

I am responding to this RFI: On behalf of an organization

Name: Josh Caplan

Name of Organization: AcademyHealth

Type of Organization: Nonprofit research organization

- 1. How to best ensure equity in publication opportunities for NIH-supported investigators.
- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

#### **Uploaded File:**

A cade my Health-response-to-NIH-on-Public-Access-Plan.pdf

Email: josh.caplan@academyhealth.org



April 24, 2023

Lawrence A. Tabak, D.D.S, Ph.D. National Institutes of Health 9000 Rockville Pike Bethesda, Maryland 20892

Re: Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research Notice Number: NOT-OD-23-091

Dear Dr. Tabak:

We are grateful for this opportunity to respond to the National Institutes of Health (NIH) Request for Information regarding the NIH Public Access Plan. AcademyHealth is the professional home of over 4,000 researchers, policy experts, and practitioners and we strongly support research transparency, evidence-based policy, and equity throughout all phases of research. Our membership is highly engaged in NIH and its work, with many having served on study sections and as grantees themselves. We believe in the critical importance of evidence-based healthcare and policymaking, but that only works if evidence is accessible to clinicians and policymakers. If we are not proactively supporting access and dissemination of NIH funded research at every level to policymakers, practitioners, researchers, clinicians, and patients, then "evidence-based healthcare and policymaking" is little more than a slogan.

In these comments, we provide three recommendations for NIH to support equitable access to publicly supported research and data:

- Focus on enhancing the diversity of researchers producing the data and research findings as well as diversity of the gatekeepers of publication opportunities to enhance the depth, breadth, and relevance of NIH-supported research;
- Expand access to publications for organizations and individuals historically excluded from access, including community-based organizations, safety net providers, patients and policymakers at every level; and
- 3. Actively monitor open access policies to ensure access is realized for those historically excluded.

Public Access to NIH-supported research results impacts our work frequently as well as the work of our colleagues. AcademyHealth staff will often synthesize research for policy audiences, such as this issue brief on <a href="Understanding Pre-Hospital Diagnostic Delays">Understanding Pre-Hospital Diagnostic Delays</a>, and we have found that we can spend significant amounts of time searching for and attempting to access scientific journal articles that are reporting findings from publicly-funded research. Unless we contact the article's author directly, we routinely face delays in obtaining the needed information and have gaps in our final product. As we work to translate the knowledge that is built through research that NIH has supported, many of our colleagues and collaborators seek to utilize current research for their own organizational goals but are unable to access it. This is also true for individual patients seeking to gain understanding of, and agency over, their care needs and are met with journal paywalls. The lack of easy and ready access to research data and study findings impacts the research, practitioner, and patient communities alike and disproportionately affects lower resourced settings such as safety net hospitals and community health centers.

The Office of Science and Technology Policy (OSTP) recently alerted all federal agencies that agencies will be required by the end of 2025 to update their public access policies to include more forms of research and data. These changes update the guidance to include making research findings publicly



available sooner and including other publication types than journal articles as applicable to these new policies. This welcome development could remove barriers to accessing evidence for those most looking to use it if implemented with that goal in mind.

AcademyHealth believes the NIH Public Access Plan is an important step forward in ensuring all those interested and willing to learn of the research being funded by the NIH can obtain access to these results, but more can be done. For the most successful implementation of this plan, we believe meaningful diversity, funding, and support is needed. The NIH can play a significantly more active role in safeguarding equitable practices for their current and future grantees. Below, we will outline our comments addressing the issues identified in the Request for Information.

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Ensuring equity in publication begins with supporting researchers from underrepresented groups and meaningfully engaging people with diverse backgrounds and perspectives in all aspects of the publication process. AcademyHealth is driven by its efforts towards diversity, equity and inclusion in health services research (HSR). We strive to increase workforce diversity through mentoring, training, and publicly committing to diversity within our own organization. These efforts to create a more inclusive space for racially and ethnically diverse researchers are still underway, but we have learned a great deal that is applicable to publication opportunities for NIH-supported investigators.

To allow for equitable publication opportunities for a diverse pool of researchers, the NIH must play a more active role in ensuring there is diversity in those behind the publication as well. This is relevant to those serving as gatekeepers to evidence, including manuscript reviewers as well as those editing the journal as a whole. Meaningful diversity in this space not only applies to race, ethnicity, geographic location, and disability status, but also but also to methodologies. Health equity researchers and those focused on social determinants of health (SDOH) often use qualitative and sociological methods that clinical and biomedical researchers may lack familiarity. Matching these manuscripts to reviewers with the appropriate knowledge base to provide meaningful feedback is integral to ensuring equity in this space. Editors from a variety of backgrounds also allow for a greater spectrum of research to be published in journals that would otherwise turn away a manuscript due to its unfamiliar methodology.

AcademyHealth conducts workshops and trainings on eliminating bias in health services research methods as well as measuring and reporting health disparities, but ultimately, journals have a responsibility to do more to aid underrepresented researchers. Some journals are making strides towards this such as Health Affairs publishing a Racism and Health issue with nearly three-quarters of the articles in this issue featuring a first author who had not previously published in their journal. Other journal editors have recognized the gap in what research is publicly available through open access but acknowledge that some journals are restricted by their minimal profit margins. Yet, we believe more consistent and persistent strides can and should be made by journals to engage diverse readers, researchers, reviewers, and editors.

The NIH could impose more rigorous equity standards on the publishing industry. Below are examples of how we believe the NIH can positively affect the research publication process and industry:

- Safeguarding diversity: In order for federally funded research to be permissibly published in a scientific journal, the NIH could require a threshold of diversity on both editorial boards and reviewer pools.
- Transparent internal practices: In order to model best practices that journals may emulate, the NIH could pledge meaningful diversity in the makeup of their study sections and be transparent when doing so.

 Training and development: In order to develop skills among under-represented minority researchers, the NIH could offer more training programs in peer-review and study section participation.

#### 2. Steps for improving equity in access and accessibility of publications.

Expedient access to research findings is an issue we face constantly at AcademyHealth as we seek evidence-based solutions to challenges in health services and health policy. These barriers are also faced by others within and adjacent to the research field including patients, community-based organizations, and those in policy-making positions. While those in academic settings typically have no issue accessing scientific publications, they are only a small proportion of the HSR community and the broader community of potential end-users. The issue of access is of frequent concern to both our members and collaborators.

As part of AcademyHealth's <u>Paradigm Project</u>, we used human-centered design to better understand and find solutions to aid those unable to access research findings. Through these efforts, federal and congressional agency staff informed us of their inability to consistently access the most current research to inform the creation of policies designed to address healthcare costs, coverage, and experiences.

AcademyHealth held an Open Science Expert Meeting in May of 2021 where participants indicated that free and equitable access to scientific knowledge and quicker access to research findings were the primary concerns that make open science principles important for HSR (over concerns about selective nonreporting and replicability of findings, which may be more relevant in other, more hypothesis-driven fields). Participants also noted that the preprint process or publicly posting research findings could likely be a good avenue for better dissemination of the research, but also engaging more diverse potential reviewers.

This NIH Public Access Plan will benefit many including:

- Community-based organizations and nonprofits: Community-based organizations and community leaders are increasingly involved in, co-leading, or principally leading research in their communities and may not be affiliated with academic institutions that could provide access to research findings. Patients and communities may not always come into their research roles with technical training, but could more easily develop the expertise with access to timely evidence. OSTP's push for immediate access to research data will allow these groups to apply research findings more readily to the communities that have been historically excluded and marginalized and are often the most underserved. Additionally, increasing the accessibility of research results with machine readability and assistive devices allows for an equitable and patient-centered approach to sharing these findings.
- Local and state-level policymakers: This is especially true for those located in smaller towns, rural areas, and who otherwise lack partnerships with academic centers who can share/extend their access, as well as those with fewer or no in-house research resources and who fully rely on others' analyses to inform policy decisions. We saw this to be especially true during the COVID-19 pandemic when there was a need for immediate access to research data, but no avenue to efficiently get it into the hands of policymakers.
- Early-career researchers: These are individuals who are forming their careers, potentially transitioning frequently between roles, especially those on "non-traditional" pathways into and within the field. These are people who want to build careers in community-based research, knowledge translation, embedded/health system research, or contract research. This group

includes many who are less connected to academic institutions and may be well positioned to bring new and creative ideas to the field—but are working from a 12-month-old literature base.

To improve equity in the publication space, we urge the NIH to emulate global examples of fortified open access policies. Given the potential implications for equity in the Public Access Plan's implementation, we must take into consideration the ease with which a layperson will be able to find and read publications in agency-designated repositories.

Fortunately, many other countries have already made strides in operationalizing these ideals. The Scholarly Publishing and Academic Resources Coalition (SPARC) <u>observes</u> that the guidance brings the U.S. policy closer to that of the European Union with strong open access policies to facilitate innovation and aligns with recent <u>open science guidance</u> from UNESCO. The European Union created <u>an official portal</u> for European data through which any of the 27 countries within the EU can "maximize access" to research findings. This repository promotes greater parity in data and research access, propelling science progress globally.

Per the 2003 <u>Berlin Declaration</u>, the international consortium cOAlition S was created to advocate for publication open access and push for <u>"fair and equitable publication fees."</u> cOAlition S launched Plan S, which is a set of guidelines on what an open access research world may look like. Currently, with the support of the European Commission, cOAlition S is actively pushing for journals to limit open access fees and is advocating for researchers to publicly release their peer-reviewed article as their intellectual property. AcademyHealth believes we can adapt the knowledge and experiences of our European counterparts to our American research enterprise.

#### 3. Methods for monitoring evolving costs of publication and impacts on affected communities.

Open access fees are high, and junior researchers, those from smaller institutions, minority serving institutions, and Historically Black Colleges and Universities may be less likely to have the resources to support open access publication.

The current OSTP policy guidance includes "reasonable publication costs and costs associated with submission, curation, management of data, and special handling instruction," but we urge for more support to be considered for open access publication fees. Without additional support, those seeking NIH funding are forced to dedicate thousands of dollars within their grant budgets to open access fees instead of their actual research needs. This could come in the form of administrative supplements available for current NIH grantees or specific grants for this purpose for those with grants that are no longer active. We encourage the NIH to take an active role in assisting researchers with affording the costs of article processing fees to guarantee the importance of open access is understood by new and veteran researchers alike.

Though there are academic institutions actively prohibiting their researchers from including funds for open access publication fees in their extramural grants, there is also the slowly growing existence of Campus Open Access Funds to assist with paying the cost of publication fees. SPARC has monitored the creation of <a href="Campus Open Access Funds">Campus Open Access Funds</a> across the U.S. finding there are currently only 51 universities with these mechanisms. The NIH should encourage institutions to allow for these funds or support other avenues for researchers to fund public access to their research.

Alternatively, the NIH could require all federally funded research to be open-access without restrictions. There is potential for publication fees to be included in grant funding, but for governmental entities to negotiate reasonable rates at what is necessary and fair for publishers. Though negotiating these prices does not fall under the NIH's purview, there is precedent to suggest the feasibility of this process. The <a href="Waxman-Hatch Amendments">Waxman-Hatch Amendments</a> in 1984 brought about fair pricing of drugs developed using taxpayer

dollars. The Office of Management and Budget may be best equipped for these negotiations, but the result will greatly affect the accessibility of research results.

Proactively supporting more equitable practices in publication opportunities, accessibility, and access could greatly impact every field of research. NIH-funded health services research is driven by taxpayer dollars and should therefore be made available to the public. In parallel, data resulting from publicly funded research also should be publicly available once properly protected. The issues of inaccessible publicly funded research and prohibitive publication costs can be addressed with more comprehensive open access policies and additional funder support for publication efforts. The NIH Public Access Plan has the potential to advance equity in research and knowledge accessibility while changing the current pay-to-play structure. While these recommendations require a great deal of further planning and determination, NIH's Public Access Plan is one step towards a more equitable future.

Thank you for the opportunity to provide guidance and assistance on the Public Access Plan. For further comment, clarification, or inquiry, please email Josh Caplan at <a href="mailto:Josh.Caplan@AcademyHealth.org">Josh.Caplan@AcademyHealth.org</a>.

**Submit date: 4/21/2023** 

I am responding to this RFI: On behalf of an organization

Name: Michael Boock

Name of Organization: Oregon State University Libraries and Press

**Type of Organization:** University

Role: Institutional official

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Oregon State University Libraries and Press (OSU) is writing to provide our input on the Plan to Enhance Public Access to the Results of NIH-Supported Research. OSU had more than \$471 million in competitive research grants and contracts in 2022 and has an R1 Carnegie Classification. It is one of only two universities in the U.S. with Land-, Sea-, Space- and Sun-Grant designations and is one of a select group of 28 universities in the United States and its territories to earn the community engagement classification in 2020 and also hold a "very high research activity" classification from the Carnegie Foundation for the Advancement of Teaching. As such, the institution and its libraries place a high value on the importance of open access to research: through a faculty-led University Open Access Policy passed in 2013, providing open access to the University's corpus of extension and experiment station publications, and open access to every thesis and dissertation ever produced by OSU students. We commend the NIH for taking important steps toward ensuring equitable access to the research it funds in the form of research articles and datasets as well as equity in publication opportunities.

The OSTP memo is deliberately "neutral" on the topic of business models for scholarly publishing. But, of course, we know that neutrality favors those who benefit from existing systems. We believe that federal agencies need to make a clear and unambiguous statement in their implementation plans that there is a pathway for researchers to comply with these mandates without paying Article Processing Charges (APCs). It is important that the NIH does not inadvertently entrench the APC system by remaining neutral on it.

We strongly support the proven, most-equitable, manuscript deposit method of policy conformance that permits authors to publish articles in whatever journals they choose so long as they deposit author-accepted manuscript versions to PubMed Central (PMC) or a trusted institutional repository that is able to share requisite metadata with PMC. Our library agrees wholeheartedly with the IVY Plus Libraries letter sent to the White House Office of Science & Technology Policy (<a href="https://libraries.mit.edu/news/libraries-support-3/34036/">https://libraries.mit.edu/news/libraries-support-3/34036/</a>) in their rejection of an article processing charge model that requires direct payment of APCs from authors, libraries, and universities to ensure open access. This model goes against the OSTP goal of providing an equitable system of publication by disadvantaging those who are not fortunate enough to be associated with institutions of higher education that can afford to pay such fees. The model also disadvantages those who do not receive adequate funding to pay such fees.

If the NIH does endorse or choose a gold OA model, we encourage the NIH, or the NIH in collaboration with Other federal agencies, to conduct an independent analysis to determine what a transparent price, based on actual value delivered, would be for an APC. This cannot be left to the publishers to determine.

Their behavior to date shows that they will prioritize neither transparency nor equity in their price-setting. Is a reasonable price \$995 (PeerJ), \$2100-6500 (PLOS), or \$11,000 (Springer Nature)? What does the research tell us? Why does it vary by discipline/publisher/journal? If an APC-based model is endorsed, it must be made clear that it is not an end goal, but a transitional step towards a more sustainable journal publishing system.

#### 2. Steps for improving equity in access and accessibility of publications.

It appears that the current NIH Public Access Plan (III.C.2) argues for the equivalent of CC-BY-NC rather than the equivalent of CC-BY. We do not understand the restriction on commercial reuse. Such a restriction may have the effect of restricting access to publications (and research data) by not allowing commercial and Other interests to create and make available value-added discovery and access tools that include ads.

We also encourage the NIH to require authors to provide a structured abstract for all research articles that can be understood by citizens with less education in the field of study. At a minimum, this might consist of context/background, objectives, design, setting, participants, interventions, main outcome measures, results, and conclusions.

- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

Consider the adoption of the Library of Congress Classification (LCC) system as a means of collocating funded research by subject. The LCC system is a widely used classification scheme in academic and research libraries in the United States and many Other countries that assigns a unique call number to an item based on the subject matter of the item. The assignment of a LCC to research outputs helps users easily locate items on particular topics and browse related materials on the topic. Classification systems can be an important component of artificial intelligence and machine learning algorithms that categorize or group data into specific classes or categories based on their attributes, characteristics, or features.

Email: michael.boock@oregonstate.edu

**Submit date: 4/21/2023** 

I am responding to this RFI: On behalf of an organization

Name: Meagan Phelan

Name of Organization: American Association for the Advancement of Science

Type of Organization: Nonprofit research organization

Role: Institutional official

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

AAAS applauds the NIH for the leadership it has assumed in emphasizing equity as a key consideration in public access policy development. Through experimentation and analysis, AAAS has found that vast differences exist in how different open access models impact the ability and equity of opportunity for scientists aiming to publish their work for wide dissemination. Some models of open access lock in place and exacerbate existing inequities in the scientific enterprise. Finding the right balance between enabling access to published work and publishing opportunity will be crucial as NIH and Other federal agencies move forward with revision of their public access policies. AAAS further wishes to express support for NIH's plan to allow for submission of "the final peer-reviewed manuscript to the NIHMS System at the time of acceptance for publication in a journal" as a means of complying with the Public Access Policy. Allowing for submission of this version of the manuscript is critical to mitigate issues associated with author- and institution-borne costs for publishing open access, including article processing charges (APCs).

To strengthen its policy as relates to equity, AAAS recommends that NIH explicitly define and recognize the "author accepted manuscript" as the version that should be submitted to the NIHMS system, to create public access. This would directly address growing challenges that alternative public access models - for which authors pay to make their work open - create for early-career scientists, scientists at smaller schools, and scientists in underfunded disciplines, among Others. It would help to ensure a diverse universe of scientists can publish important work, regardless of their economic circumstances.

This step would also require the agency to more directly consider the role of business models - some of which do not foster inclusivity - in its efforts around public access. While NIH's Plan for Scholarly Publications does not address business models Other than to state that "NIH does not propose requiring authors to publish in any particular type of journal or journal with any specific type of business model (e.g., subscription model, open access model)," it is essential to recognize that if journal policies do not allow for deposition of the author-accepted manuscript in the NIHMS system at the time of publication, this policy will limit authors' publishing options - driving scientists to publish in open access journals to which they must pay an APC (fees for which only stand to increase as the publishing market consolidates) or in journals with which their institution has a transformative agreement. This may temporarily work well for senior scientists who are (routinely) well-funded, tenured, and overwhelmingly male and white, but it will freeze in place and exacerbate inequities for many Others, including a new generation of scientists. By channeling researchers to a limited number of commercial publishers, it will also drive further consolidation in a market that is already heavily concentrated, and where APC fees will only increase with time. The resultant heavy cost burden will be borne not only by

researchers and their institutions, including at a time when institutional research budgets are increasingly challenged, but by funders of research (including taxpayers). We urge the agency to proactively communicate with publishers about their policies to ensure they allow authors to deposit the AAM in the NIHMS system. This is essential to ensuring that, regardless of a scientist's geographic location, institutional affiliation, academic rank, or identity, they can publish world-changing science.

Finally, as addressed in the response to the third question within this RFI, AAAS also believes that monitoring implementation of changes to the public access policy and publishing costs paid by researchers and institutions will be critical to ensuring that these changes do not create new inequities or reinforce existing ones. It may be valuable for NIH to conduct a survey, as AAAS did on a smaller scale in 2022 (<a href="https://www.aaas.org/news/aaas-survey-many-researchers-face-difficulties-paying-open-access-fees">https://www.aaas.org/news/aaas-survey-many-researchers-face-difficulties-paying-open-access-fees</a>), and/or develop a public reporting scheme about scientist-borne publishing-associated costs and related tradeoffs.

#### 2. Steps for improving equity in access and accessibility of publications.

AAAS supports open-research initiatives, including text and data mining, that use technology to meet the needs of researchers. However, appropriate limitations are important to ensure such offerings remain sustainable; we have seen some initiatives lead to unintended consequences when the necessary rights have not been secured to enable their sustainability. Given the fast pace of artificial intelligence development, it is critically important to monitor the creation and adoption of guidelines for tools that can be trained on full text journal articles, including for the purposes of replicating scholarly journal content, to ensure a focus on responsible and ethical development.

Science journal articles, and specifically the author accepted manuscript (AAM) versions of such articles, may be used for text and data mining by individuals and by nonprofit, noncommercial subscribing institutions. Sustainably increasing accessibility to publications via this route requires that publisher reuse policies are followed by federally funded researchers. AAAS encourages NIH to consider how adherence to related policies will be monitored and what administrative burdens this might create for researchers, institutions, and the agency. NIH should also endeavor to monitor how changes resulting from the open access policy, including a breadth of open license types, might facilitate and incentivize reuse that adversely impacts the integrity and accuracy of the downstream communication of research published by federally funded researchers.

Regarding Other avenues by which to improve accessibility to publications, including for people with disabilities, NIH may wish to consider implementing guidelines around adherence to the Web Content Accessibility Guidelines, with a concerted focus on making text and data available.

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

Careful and continued study of publication fees and policies will be essential for understanding the nearand long-term effects of changes in public access policies. Study of costs effects at the researcher, institution, and enterprise levels is needed. Adaptation of federal grant agreements to require reporting on the payment of publication fees and reliance on transformative agreements (in instances where authors avoid payment of a fee because their institution has a transformative agreement with their journal of choice) represents one logical approach to monitoring fees. AAAS also encourages NIH to consider a study or studies that engage institutional leadership to estimate and report on publishing costs across institutions.

In addition to developing methods for monitoring costs, AAAS encourages NIH to develop and adopt a public reporting scheme to ensure visibility and transparency into publishing costs borne by scientists, their institutions, and ultimately the NIH. This will allow for future course correction.

All analysis of and reporting on publication costs should examine potential variability in costs across disciplines, career stages, and institution type, as well as based on researchers' backgrounds and characteristics. Analysis and reporting should assess if and how changes in the Public Access Policy may affect the volume of research publications authored by scientists who are early career or are from smaller, lesser-funded, and historically underrepresented institutions, including Historically Black Colleges and Universities; Hispanic-Serving Institutions, EPSCoR, and Other Minority-Serving institutions; where researchers choose to publish; and potential variability in effects across different research disciplines including, but not limited to, the life sciences, physical sciences, social sciences, humanities, mathematics, and engineering.

#### 4. Early input on considerations to increase findability and transparency of research.

Access and transparency are foremost considerations at AAAS, where our mission includes communicating science accurately, broadly, and in such a way to ensure the scientific community can reanalyze and reproduce new works. In recognition, AAAS supports the final peer-reviewed authoraccepted version of a paper being broadly and immediately shared and the flexibility afforded by NIH's intention to accept the final peer-reviewed version of the article as a means of complying with its updated public access policy. At AAAS, however, we believe that publisher oversight of a final version (the version of record, or VOR) is essential not only to maintaining the quality and accuracy of scientific research but also to advancing the subsequent work from which new research stems. Only the final version of a manuscript overseen by a publisher committed to maintaining the accuracy of the scientific record can be counted on to be corrected, retracted or Otherwise updated with clear notation for the global scientific research community. Ensuring that publication repositories clearly distinguish between multiple versions of articles (i.e., ensuring that singular publication records point to the VoR, where the AAM is deposited first) will be critical, as NIH moves forward. The NIH may wish to implement guidelines requiring that authors depositing their AAMs provide a DOI (digital object identifier) pointing to the VOR. Indeed, at AAAS, our instructions for authors depositing AAMs require them to include a link to the VOR.

With respect to metadata, linkages between publishers and organizations such as the Research Organization Registry (ROR), Open Researcher and Contributor ID (ORCID), Crossref, and data repositories are aimed at increasing robustness of metadata by providing persistent identifiers and connecting them to research outputs. As a publisher, AAAS monitors and implements best practices for both metadata collection (e.g., on institutions and funders) and metadata propagation in the VOR and associated research objects.

All Science journal papers include details about funding, author contributions, competing interests, data and materials availability, and license information. The publisher oversees accuracy of important associated metadata after publication, including in cases where authors request to change their names in previously published papers, as one example. As a criterion to publish, AAAS requires authors to make

their data publicly accessible. AAAS has also piloted a partnership with Dryad, an international openaccess data repository; we encourage such partnerships because they help ensure that publishers and repositories share the same metadata, thus providing better linkage between the data and the research paper. NIH may wish to consider implementing guidelines for data availability in publications. These guidelines could include a clear set of criteria for data deposition and ease of linking to that data, which publishers could help enforce. As a best practice, NIH could also encourage connections between publishers and data repositories of various kinds (general or field-specific, or both).

#### **Uploaded File:**

AAAS-NIH-RFI-Response-\_-April-2023.pdf

**Description:** Attached, please find the full AAAS response to the NIH RFI on OA. This includes information outside of responses to the four questions listed above.

Email: mphelan@aaas.org

## AAAS Response to RFI on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research

The American Association for the Advancement of Science (AAAS) welcomes the NIH's efforts to enhance public access, in line with the recent OSTP guidance aimed at making federally funded research publications and supporting data publicly available. Open and accessible data are essential to scientific integrity and reproducibility.

AAAS, a multi-disciplinary non-profit association of over 100,000 scientists at all levels of the scientific enterprise, publishes the *Science* family of journals. Our mission is to advance science and innovation throughout the world for the benefit of all.

The Science family of journals is open to the public without embargo using green open access models for five of our journals and a gold open access model for one.

Our journals require published authors to make their data immediately accessible in approved repositories and authors may share their author accepted manuscripts immediately upon publication.

AAAS applauds the NIH for emphasizing equity in its approach to public access policy development and for considering how to balance access to published work with the ability to publish, as well as the unintended consequences that focusing primarily on accelerating access for readers may have. AAAS is committed to collaborating with NIH, other federal research agencies, and OSTP to develop public access policies and supportive publishing models that achieve this balance and is pleased to offer its response to the NIH's RFI in this document. Responses to NIH's four key questions follow, with additional comments on data sharing considerations.

# Question 1: How to best ensure equity in publication opportunities for NIH-supported investigators.

**Prompt:** The NIH Public Access Plan aims to maintain the existing broad discretion for researchers and authors to choose how and where to publish their results. Consistent with current practice, the NIH Public Access Plan allows the submission of final published articles to PubMed Central (PMC) (in cases where

a formal agreement is in place) to minimize the compliance burden on NIH-supported researchers and also maintains the flexibility of NIH-supported researchers to submit the final peer-reviewed manuscript. NIH seeks information on additional steps it might consider taking to ensure that proposed changes to implementation of the NIH Public Access Policy do not create new inequities in publishing opportunities or reinforce existing ones.

### **AAAS Response:**

NIH has AAAS applauds the for the leadership it assumed in emphasizing equity as a key consideration in public access policy development. Through experimentation and analysis, AAAS has found that vast differences exist in how different open access models impact the ability and equity of opportunity for scientists aiming to publish their work for wide dissemination. Some models of open access lock in place and exacerbate existing inequities in the scientific enterprise. Finding the right balance between enabling access to published work and publishing opportunity will be crucial as NIH and other federal agencies move forward with revision of their public access policies. AAAS further wishes to express support for NIH's plan to allow for submission of "the final peer-reviewed manuscript to the NIHMS System at the time of acceptance for publication in a journal" as a means of complying with the Public Access Policy. Allowing for submission of this version of the manuscript is critical to mitigate issues associated with author- and institution-borne costs for publishing open access, including article processing charges (APCs).

To strengthen its policy as relates to equity, AAAS recommends that NIH explicitly define and recognize the "author accepted manuscript" as the version that should be submitted to the NIHMS system, to create public access. This would directly address growing challenges that alternative public access models – for which authors pay to make their work open – create for early-career scientists, scientists at smaller schools, and scientists in underfunded disciplines, among others. It would help to ensure a diverse universe of scientists can publish important work, regardless of their economic circumstances.

This step would also require the agency to more directly consider the role of business models – some of which do not foster inclusivity – in its efforts around public access. While NIH's Plan for Scholarly Publications does not address business models other than to state that "NIH does not propose requiring authors to publish in any particular type of journal or journal with any specific type of business model (e.g., subscription model, open access model)," it is essential to recognize that if journal policies do not allow for deposition of the authoraccepted manuscript in the NIHMS system at the time of publication, this policy

will limit authors' publishing options – driving scientists to publish in open access journals to which they must pay an APC (fees for which only stand to increase as the publishing market consolidates) or in journals with which their institution has a transformative agreement. This may temporarily work well for senior scientists who are (routinely) well-funded, tenured, and overwhelmingly male and white, but it will freeze in place and exacerbate inequities for many others, including a new generation of scientists. By channeling researchers to a limited number of commercial publishers, it will also drive further consolidation in a market that is already heavily concentrated, and where APC fees will only increase with time. The resultant heavy cost burden will be borne not only by researchers and their institutions, including at a time when institutional research budgets are increasingly challenged, but by funders of research (including taxpayers). We urge the agency to proactively communicate with publishers about their policies to ensure they allow authors to deposit the AAM in the NIHMS system. This is essential to ensuring that, regardless of a scientist's geographic location, institutional affiliation, academic rank, or identity, they can publish worldchanging science.

Finally, as addressed in the response to the third question within this RFI, AAAS also believes that monitoring implementation of changes to the public access policy and publishing costs paid by researchers and institutions will be critical to ensuring that these changes do not create new inequities or reinforce existing ones. It may be valuable for NIH to conduct a survey, as AAAS did on a smaller scale in 2022 (<a href="https://www.aaas.org/news/aaas-survey-many-researchers-face-difficulties-paying-open-access-fees">https://www.aaas.org/news/aaas-survey-many-researchers-face-difficulties-paying-open-access-fees</a>), and/or develop a public reporting scheme about scientist-borne publishing-associated costs and related tradeoffs.

# Question 2: Steps for improving equity in access and accessibility of publications.

**Prompt:** Removal of the currently allowable 12-month embargo period for NIH-supported publications will improve access to these research products for all. As noted in the NIH Public Access Plan, NIH also plans to continue making articles available in human and machine-readable forms to support automated text processing. NIH will also seek ways to improve the accessibility of publications via assistive devices. NIH welcomes input on other steps that could be taken to improve equity in access to publications by diverse communities of users, including researchers, clinicians and public health officials, students and educators, and other members of the public.

### **AAAS** Response:

AAAS supports open-research initiatives, including text and data mining, that use technology to meet the needs of researchers. However, appropriate limitations are important to ensure such offerings remain sustainable; we have seen some initiatives lead to unintended consequences when the necessary rights have not been secured to enable their sustainability. Given the fast pace of artificial intelligence development, it is critically important to monitor the creation and adoption of guidelines for tools that can be trained on full text journal articles, including for the purposes of replicating scholarly journal content, to ensure a focus on responsible and ethical development.

Science journal articles, and specifically the author accepted manuscript (AAM) versions of such articles, may be used for text and data mining by individuals and by nonprofit, noncommercial subscribing institutions. Sustainably increasing accessibility to publications via this route requires that publisher reuse policies are followed by federally funded researchers. AAAS encourages NIH to consider how adherence to related policies will be monitored and what administrative burdens this might create for researchers, institutions, and the agency. NIH should also endeavor to monitor how changes resulting from the open access policy, including a breadth of open license types, might facilitate and incentivize reuse that adversely impacts the integrity and accuracy of the downstream communication of research published by federally funded researchers.

Regarding other avenues by which to improve accessibility to publications, including for people with disabilities, NIH may wish to consider implementing guidelines around adherence to the Web Content Accessibility Guidelines, with a concerted focus on making text and data available.

# Question 3: Methods for monitoring evolving costs and impacts on affected communities.

**Prompt:** NIH proposes to actively monitor trends in publication fees and policies to ensure that they remain reasonable and equitable. NIH seeks information on effective approaches for monitoring trends in publication fees and equity in publication opportunities.

### **AAAS Response:**

Careful and continued study of publication fees and policies will be essential for understanding the near- and long-term effects of changes in public access policies. Study of costs effects at the researcher, institution, and enterprise levels is needed. Adaptation of federal grant agreements to require reporting on the

payment of publication fees and reliance on transformative agreements (in instances where authors avoid payment of a fee because their institution has a transformative agreement with their journal of choice) represents one logical approach to monitoring fees. AAAS also encourages NIH to consider a study or studies that engage institutional leadership to estimate and report on publishing costs across institutions.

In addition to developing methods for monitoring costs, AAAS encourages NIH to develop and adopt a public reporting scheme to ensure visibility and transparency into publishing costs borne by scientists, their institutions, and ultimately the NIH. This will allow for future course correction.

All analysis of and reporting on publication costs should examine potential variability in costs across disciplines, career stages, and institution type, as well as based on researchers' backgrounds and characteristics. Analysis and reporting should assess if and how changes in the Public Access Policy may affect the volume of research publications authored by scientists who are early career or are from smaller, lesser-funded, and historically underrepresented institutions, including Historically Black Colleges and Universities; Hispanic-Serving Institutions, EPSCoR, and other Minority-Serving institutions; where researchers choose to publish; and potential variability in effects across different research disciplines including, but not limited to, the life sciences, physical sciences, social sciences, humanities, mathematics, and engineering.

# Question 4: Early input on considerations to increase findability and transparency of research.

**Prompt:** Section IV of the NIH Public Access Plan is a first step in developing the NIH's updated plan for persistent identifiers (PIDs) and metadata, which will be submitted to OSTP by December 31, 2024. NIH seeks suggestions on any specific issues that should be considered in efforts to improve use of PIDs and metadata, including information about experiences institutions and researchers have had with adoption of different identifiers.

### **AAAS** Response:

Access and transparency are foremost considerations at AAAS, where our mission includes communicating science accurately, broadly, and in such a way to ensure the scientific community can reanalyze and reproduce new works. In recognition, AAAS supports the final peer-reviewed author-accepted version of a paper being broadly and immediately shared and the flexibility afforded by NIH's intention to accept the final peer-reviewed version of the article as a means of complying with its updated public access policy. At AAAS, however, we believe

that publisher oversight of a final version (the version of record, or VOR) is essential not only to maintaining the quality and accuracy of scientific research but also to advancing the subsequent work from which new research stems. Only the final version of a manuscript overseen by a publisher committed to maintaining the accuracy of the scientific record can be counted on to be corrected, retracted or otherwise updated with clear notation for the global scientific research community. Ensuring that publication repositories clearly distinguish between multiple versions of articles (i.e., ensuring that singular publication records point to the VoR, where the AAM is deposited first) will be critical, as NIH moves forward. The NIH may wish to implement guidelines requiring that authors depositing their AAMs provide a DOI (digital object identifier) pointing to the VOR. Indeed, at AAAS, our instructions for authors depositing AAMs require them to include a link to the VOR.

With respect to metadata, linkages between publishers and organizations such as the Research Organization Registry (ROR), Open Researcher and Contributor ID (ORCID), Crossref, and data repositories are aimed at increasing robustness of metadata by providing persistent identifiers and connecting them to research outputs. As a publisher, AAAS monitors and implements best practices for both metadata collection (e.g., on institutions and funders) and metadata propagation in the VOR and associated research objects.

All *Science* journal papers include details about funding, author contributions, competing interests, data and materials availability, and license information. The publisher oversees accuracy of important associated metadata after publication, including in cases where authors request to change their names in previously published papers, as one example. As a criterion to publish, AAAS requires authors to make their data publicly accessible. AAAS has also piloted a partnership with Dryad, an international open-access data repository; we encourage such partnerships because they help ensure that publishers and repositories share the same metadata, thus providing better linkage between the data and the research paper. NIH may wish to consider implementing guidelines for data availability in publications. These guidelines could include a clear set of criteria for data deposition and ease of linking to that data, which publishers could help enforce. As a best practice, NIH could also encourage connections between publishers and data repositories of various kinds (general or field-specific, or both).

# **Additional Comments Not Addressed in Responses to Question Above**

With the ongoing shifts in public access policy, researchers and institutions will face new challenges for data management. These require fundamental research

infrastructure that is lacking. Database infrastructure oversight is resourceintensive, and in a way that is only increasing with time. The ability to protect data – even anonymized data – is also increasingly challenging in the age of sophisticated artificial intelligence tools. As more databases come online to meet the needs of various parts of the scientific community, it is crucially important that these databases follow best practices designed to ensure the data are maintained and protected to highest current standards.

AAAS is among publishers requiring that data supporting research papers be available at the time of publication, either in the manuscript or supplementary material or through a public repository. This may include nontrivial costs, as some repositories require payment for data archiving. These costs will only increase as the capacity for high-throughput research increases. At the *Science* journals, we found that questions around complying with our data policies arose mainly for datasets that lack a field-specific repository. To help support equity among authors, AAAS is piloting a partnership with Dryad, an international open-access general data repository, and covering costs for data publication for *Science* family journal authors.

However, covering data archiving costs is not viable on a larger scale. At AAAS, we can foresee the challenges in this space, even as efforts to ensure all data underlying new publications are available via repositories are, in many ways, at their earliest stages in the scientific publishing ecosystem. How to manage the cost for publishing in data repositories and for maintenance of data infrastructure - be it through grants or other means - is an important question. How this can be achieved while also ensuring the protection of data, especially sensitive equally pressing and resourcedata. is an intensive consideration. As NIH updates its public access policy, therefore, AAAS encourages its leadership to consider these issues and how NIH will monitor and manage inequities in data deposition and sharing. Without a strategy at the federal level, data curation and access could become a focus of the private sector. A sophisticated data access strategy that best serves the research community may also need to consider tradeoffs in data deposition that ensure only the data most important to analyzing and replicating research is deposited, helping to reasonably manage data input streams.

**Submit date:** 4/21/2023

I am responding to this RFI: On behalf of an organization

Name: Lindsay Morton, Senior Manager, Open Science Community Engagement

Name of Organization: Public Library of Science (PLOS)

Type of Organization: Other

Type of Organization-Other: Scholarly Publisher

Role: Member of the public

### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

As a nonprofit Open Science publisher, PLOS aims to ensure that Open Science is practicable for the entire scholarly community, and that the reward system of science appropriately acknowledges and honors Open Science practices as contributing to a common good. We wish to express our enthusiastic support for the vision articulated in the OSTP memo of August 25 2022 and our appreciation of the NIH's thoughtful steps toward realizing that vision. We appreciate the opportunity to share our perspective on the NIH's plan to enhance access to the results of research, and improve equity in scholarly publishing for authors and readers alike.

The rapid dissemination and widespread availability of research and underlying data through Open Science is key to meeting major challenges—from the Sars-CoV-2 pandemic to the climate crisis—with effective, evidence-based solutions grounded in rigorous reproducible science. We see Openness as more than the ability to read research articles. Openness includes unrestricted access to the tools and information necessary to understand research results in context, to verify and reproduce results, and to reuse data and methods. True Openness also means equitable opportunities for publication and participation in the peer review process.

How to best ensure equity in publication opportunities for NIH-supported investigators.

We believe that equitable access to Open publishing opportunities requires a shift away from a volumetric 'pay per publication' model reliant on Article Processing Charges (APCs). APCs have demonstrated that Open Access is viable—but they are exclusionary and they create incentives for publishers to increase volume or price. Waivers, while a useful stop-gap, are not a sustainable solution.

Instead, we should work together to evolve new models based on partnership, collaboration, and community. Already, publishers, including PLOS, are experimenting with new ways to finance Open Access, including Community Action Publishing, Subscribe to Open programs, and more.

In the short term, and in parallel with developing and honing new solutions, we can implement simple changes to better meet author and stakeholder needs.

- 1. Establish funding mechanisms specifically for research dissemination. Researchers shouldn't have to choose between using their grant to pay a publication fee, or to conduct more experiments.
- 2. Aggregate funding for publication services fees through a University library or similar body, rather than allocating small amounts through individual research grants. Centralizing administrative functions

increases efficiency, reduces the administrative burden on individual researchers and the administrative costs to publishers, and makes it possible to more fairly distribute the cost of publication, putting Open Access within reach for more of the research community.

In the US, libraries and consortia have shown that they are open to testing new methods, and that these types of partnership can be effective both in increasing transparency, and addressing cost inflation.

# 2. Steps for improving equity in access and accessibility of publications.

PLOS supports the NIH's efforts to increase and accelerate access to publicly funded research. Eliminating the embargo will meaningfully benefit researchers, practitioners, and patients alike, and is reasonable and feasible for publishers as well. PLOS has always deposited research with indexing and archiving services as soon as possible following publications.

We also appreciate the emphasis placed on machine readability, which is essential to discoverability, reuse, and reanalysis. However, because the NIH policy provides for access alone, without the legal right to reuse that true Open Access licensing provides, its utility is limited—especially in this era of big data and rich text data mining. Reuse and redistribution are key to maximizing the reach and impact of research.

Equally vital to reproducibility is ensuring access to research outputs Other than articles, such as data and methods documentation, including study designs, code, and protocols. The NIH can help to drive change in this area by encouraging, reinforcing, and rewarding the sharing of a broader range of research outputs in line with best practices for reproducibility, transparency, and inclusivity, in the grant application process.

### 3. Methods for monitoring evolving costs and impacts on affected communities.

The expectation of transparency in pricing policy will encourage continued experimentation with more equitable scholarly communications business models, helping to drive positive change.

In the short term, we recommend that the NIH take advantage of the considerable public information on pricing already available, by aligning with established systems (like those of Coalition S). Gathering similar information independently in a new system will create additional administrative tasks and unnecessary expense.

In developing any new monitoring or measurement frameworks, it's crucial to recognize that individual article fees are not an essential part of an Open system. Future monitoring efforts must be structured in a way that allows for the evolution of business models, which is key to increasing equity in publication opportunities. In order to be successfully adopted, any new monitoring framework must also be broadly applicable beyond the NIH, or the US context alone.

### 4. Early input on considerations to increase findability and transparency of research.

The development and consistent application of shared metadata standards is key to discoverability and credibility. We encourage the NIH to invest and participate in community-based metadata initiatives in order to build systems that work for all, and to prioritize systems with the broadest potential impact, focusing on utility and reuse with the aim of increasing system-wide efficiency and accelerating scientific advancement.

To be effective, metadata and persistent identifiers (PIDs) must be interoperable and follow some level of standardization. Therefore the NIH should recognize the benefits of making specific recommendations in this area, to accelerate harmonization around emerging standards adopted by the scientific community. In developing guidelines for grantees and publishers, the NIH should:

- Specify clear and detailed metadata standards and provide recommendations about which PIDs to use to describe diverse research artifacts and the links between them, both in a machine-readable way at scale, and as human readers accessing individual research elements.
- Set expectations for PIDs and metadata to understand the individual contributions of authors, editors, and peer reviewers, and provide a digital infrastructure to support credit for all contributions.

#### Conclusion

Although Open Access has made great strides over the past two decades, the majority of research outputs are still not accessible, either because they are behind a paywall (according to a recent analysis of Web of Science and Dimensions data, 53-56% of published research remains closed), or because they have not been shared at all (e.g. datasets, protocols, negative and null results).

In order to actively move away from paywalled research, we need to change the reward system of science, ensuring that researchers receive meaningful credit and recognition for all kinds of contributions. This includes both acknowledgement for a wider variety of research creation and assessment roles, from protocol development through peer review, and a more representative range of research outputs, including information that contextualizes research articles and enables reproducibility.

**Uploaded File:** PLOS-response-to-the-NIH-Public-Access-RIF.pdf

**Description:** Please find the above attached in PDF format, with hyperlinks

Email: Imorton@plos.org

# PLOS response to the NIH Public Access RIF

As a nonprofit Open Science publisher, PLOS aims to ensure that Open Science is practicable for the entire scholarly community, and that the reward system of science appropriately acknowledges and honors Open Science practices as contributing to a common good. We wish to express our enthusiastic support for the vision articulated in the OSTP memo of August 25 2022 and our appreciation of the NIH's thoughtful steps toward realizing that vision. We appreciate the opportunity to share our perspective on the NIH's plan to enhance access to the results of research, and improve equity in scholarly publishing for authors and readers alike.

The rapid dissemination and widespread availability of research and underlying data through Open Science is key to meeting major challenges—from the Sars-CoV-2 pandemic to the climate crisis—with effective, evidence-based solutions grounded in rigorous reproducible science. We see Openness as more than the ability to read research articles. Openness includes unrestricted access to the tools and information necessary to understand research results in context, to verify and reproduce results, and to reuse data and methods. True Openness also means equitable opportunities for publication and participation in the peer review process.

1. How to best ensure equity in publication opportunities for NIH-supported investigators.

We believe that equitable access to Open publishing opportunities requires a shift away from a volumetric 'pay per publication' model reliant on Article Processing Charges (APCs). APCs have demonstrated that Open Access is viable—but they are exclusionary and they create incentives for publishers to increase volume or price. Waivers, while a useful stop-gap, are not a sustainable solution.

Instead, we should work together to evolve new models based on partnership, collaboration, and community. Already, publishers, including PLOS, are experimenting with new ways to finance Open Access, including <a href="Community Action Publishing">Community Action Publishing</a>, Subscribe to Open programs, and more.

In the short term, and in parallel with developing and honing new solutions, we can implement simple changes to better meet author and stakeholder needs.

- Establish funding mechanisms specifically for research dissemination. Researchers shouldn't have to choose between using their grant to pay a publication fee, or to conduct more experiments.
- Aggregate funding for publication services fees through a university library or similar body, rather than allocating small amounts through individual research grants.
   Centralizing administrative functions increases efficiency, reduces the administrative burden on individual researchers and the administrative costs to publishers, and makes it possible to more fairly distribute the cost of publication, putting Open Access within reach for more of the research community.

In the US, libraries and consortia have shown that they are open to testing new methods, and that these types of partnership can be effective both in increasing transparency, and addressing cost inflation.

### 2. Steps for improving equity in access and accessibility of publications.

PLOS supports the NIH's efforts to increase and accelerate access to publicly funded research. Eliminating the embargo will meaningfully benefit researchers, practitioners, and patients alike, and is reasonable and feasible for publishers as well. PLOS has always deposited research with indexing and archiving services as soon as possible following publications.

We also appreciate the emphasis placed on machine readability, which is essential to discoverability, reuse, and reanalysis. However, because the NIH policy provides for access alone, without the legal right to reuse that true <a href="Open Access licensing">Open Access licensing</a> provides, its utility is limited—especially in this era of big data and rich text data mining. Reuse and redistribution are key to maximizing the reach and impact of research.

Equally vital to reproducibility is ensuring access to research outputs other than articles, such as data and methods documentation, including study designs, code, and protocols. The NIH can help to drive change in this area by encouraging, reinforcing, and rewarding the sharing of a broader range of research outputs in line with best practices for reproducibility, transparency, and inclusivity, in the grant application process.

# 3. Methods for monitoring evolving costs and impacts on affected communities.

The expectation of transparency in pricing policy will encourage continued experimentation with more equitable scholarly communications business models, helping to drive positive change.

In the short term, we recommend that the NIH take advantage of the considerable public information on pricing already available, by aligning with established systems (like those of Coalition S). Gathering similar information independently in a new system will create additional administrative tasks and unnecessary expense.

In developing any new monitoring or measurement frameworks, it's crucial to recognize that individual article fees are <u>not</u> an essential part of an Open system. Future monitoring efforts must be structured in a way that allows for the evolution of business models, which is key to increasing equity in publication opportunities. In order to be successfully adopted, any new monitoring framework must also be broadly applicable beyond the NIH, or the US context alone.

### 4. Early input on considerations to increase findability and transparency of research.

The development and consistent application of shared metadata standards is key to discoverability and credibility. We encourage the NIH to invest and participate in

community-based metadata initiatives in order to build systems that work for all, and to prioritize systems with the broadest potential impact, focusing on utility and reuse with the aim of increasing system-wide efficiency and accelerating scientific advancement.

To be effective, metadata and persistent identifiers (PIDs) must be interoperable and follow some level of standardization. Therefore the NIH should recognize the benefits of making specific recommendations in this area, to accelerate harmonization around emerging standards adopted by the scientific community. In developing guidelines for grantees and publishers, the NIH should:

- Specify clear and detailed metadata standards and provide recommendations about which PIDs to use to describe diverse research artifacts and the links between them, both in a machine-readable way at scale, and as human readers accessing individual research elements.
- Set expectations for PIDs and metadata to understand the individual contributions of authors, editors, and peer reviewers, and provide a digital infrastructure to support credit for all contributions.

### Conclusion

Although Open Access has made great strides over the past two decades, the majority of research outputs are still not accessible, either because they are behind a paywall (according to a <u>recent analysis</u> of Web of Science and Dimensions data, 53-56% of published research remains closed), or because they have not been shared at all (e.g. datasets, protocols, negative and null results).

In order to actively move away from paywalled research, we need to change the reward system of science, ensuring that researchers receive meaningful credit and recognition for all kinds of contributions. This includes both acknowledgement for a wider variety of research creation and assessment roles, from protocol development through peer review, and a more representative range of research outputs, including information that contextualizes research articles and enables reproducibility.

**Submit date:** 4/21/2023

I am responding to this RFI: On behalf of an organization

Name: Tom Ciavarella

Name of Organization: Frontiers Media Inc

Type of Organization: Other

Type of Organization-Other: Scientific and academic publishing company

Role: Institutional official

### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

On public repositories, we believe the NIH Public Access Plan rightly encourages and prioritizes the widest possible choices for researchers as they relate to publishing venue, as well as the principles of academic freedom. We think the Plan strikes the right balance by making PubMed Central (PMC) a convenient and compliant repository for research without privileging or mandating it.

On the fairness of the article processing charge (APC), we believe it is both fair and effective as it is a fee for a service. But although it is the most efficient and transparent method, it is worth noting this charge is not the only way to finance Gold Open Access (Gold OA) publishing. Indeed, we recognize that in some cases, it is not the preferred or most sustainable price structure for researchers, funders, libraries, and research institutions. And while we, like Others in the publishing industry, think the APC model is a good one, we are continually in touch with institutional partners to find solutions that meet their needs. And we are seeking new models to help authors cover the fair and actual cost of publishing, to make scientific knowledge accessible to the widest possible audience.

Within an APC framework, we have expanded our portfolio of institutional models to meet the tailored needs of our customers (recognizing, for example, the distinct needs of research-intensive "publish" organizations as well as high consumption "read" institutions and societies). Our success indicates a range of pricing options can meet the needs of a range of customers and institutions.

On the additional steps the NIH might take to ensure new inequities are not created, or existing ones reinforced, we believe the NIH should encourage researchers to publish in the Gold OA model - on the basis that the public funding of public access is efficient, scalable, and delivers value for money.

In our view, Gold OA publishing is the most effective way of securing that outcome. It offers a simple, transparent, and competitive way to unlock the benefits of fully accessible science; and it enables researchers, agencies, universities, libraries, and repositories to fulfil both the NIH Public Access Policy and the OSTP guidance.

Publishing in a Gold OA journal immediately facilitates the transfer of articles to a repository, with metadata in machine-readable formats. In this model, there are no embargoes and no superfluous or costly bundled services that are common in "hybrid" or "transformative" subscription options offered by legacy commercial publishers.

On public value for money, new federal guidelines seek public access but do not specify delivery models. We agree that openly accessible science can - and should - be delivered by more than one publishing

model. We welcome competition if it spurs innovation and the amount of rigorous science accessible to all.

But in judging those delivery models, federal agencies must make a robust and transparent assessment and a comparison for efficiency, scalability, and public value for money - guided by the objective of discoverability that underpins public access.

For example, public access known as "Green Open Access (Green OA)" clearly removes some barriers and does not create or perpetuate inequity. But the mechanisms for finding, reading, and sharing Green OA files vary widely. Substantial new funding will be required just to bring that variance down and lift standards for discoverability, with new investment in infrastructure for metadata enrichment. Those institutions unable to fund that investment are likely to face the continued cost pressure of paywall subscriptions that might only minimally ease search and discovery.

So, it is vital that the funding of public access is as efficient, scalable, and as good a value for money as possible, and in our view, Gold OA publishing is the most effective way of securing that outcome. It offers a simple, transparent, and competitive way to unlock the benefits of fully accessible science.

### 2. Steps for improving equity in access and accessibility of publications.

On the 12-month embargo, we strongly welcome the NIH's decision to end it on publications. We believe that so-called Transformative Agreements (TAs) were worthwhile in their conception as a means of smoothing the transition to fully open access science, but in their execution have not effectively led to transformation and have instead become a blunt instrument.

TAs lack transparency, have complex bundles of often unnecessary services making it all but impossible to judge value for money, and come with no contractual commitment to a move to full open access (Green, Gold, or Otherwise) within a binding deadline.

Most of these TAs are large scale "read and publish" or hybrid deals. Publishers will often allow authors' work to appear in hybrid journals without being charged (if their institutions pay), while at the same time maintaining the amount of science they publish behind paywalls.

We believe TAs help subsidize the market dominance of legacy publishers by controlling the pace of transition to fully open access science.

The worldwide scientific publishing oligopoly is a market estimated to be around US \$27 billion by revenues in 2021, as per Outsell Inc., Segment View: Scientific, Technical and Medical, 2021, cited in STM Global Brief 2021 - Economics & Market Size: <a href="https://www.stm-assoc.org/wp-content/uploads/2021">https://www.stm-assoc.org/wp-content/uploads/2021</a> 10 19 STM Global Brief 2021 Economics and Market Size-1.pdf

Furthermore, the five largest paywall publishing houses (Elsevier, Wiley, Springer Nature, Taylor & Francis, and SAGE) have captured more than half of that market, as per the Livres Hebdo/Publishers Weekly 2021 ranking of top global publishers: <a href="https://www.publishersweekly.com/pw/by-topic/international/Frankfurt-Book-Fair/article/87466-frankfurt-book-fair-preview-2021-relx-rules.html">https://www.publishersweekly.com/pw/by-topic/international/Frankfurt-Book-Fair/article/87466-frankfurt-book-fair-preview-2021-relx-rules.html</a>

On the basis the NIH seeks equity in access as well as transparency in costs, backed by financial sustainability, we believe Gold OA publishers can deliver.

On automated text processing, assistive devices, and Other inclusionary measures, we fully support the NIH's position. We consistently invest in measures that improve the accessibility of our publications. Many such requirements were mandated by the Coalition S initiative, which Frontiers fully supported, and which saw wide-ranging and progressive open access policies adopted in the United Kingdom and across Europe.

We firmly back public policies that promote equity of opportunity, the ability both to read and to publish research, and the scientific rigor, academic freedom, institutional values, and personal and professional recognition that underpin success.

We are committed to increasing research access, knowledge resources, and educational opportunities for all, especially for those groups, nations, and individuals who are historically marginalized, underrepresented, or disadvantaged.

On institutional success, we work to build communities and tackle the inadequacies and inequities often characterizing research dissemination. The shift toward open access represents an opportunity to expand access to knowledge in a significant way across academic institutions of all stripes, as well as to small businesses and the public.

We urge the NIH to draw on its influence to see that library, research, and educational institutions commit to investing in open access so that all parties can source sufficient funding for publishing. Several equitable open publishing models are readily available. It cannot be right if colleges and universities are encouraged to maintain robust publications budgets for subscriptions and then asked to make cuts to open access.

We believe there is enough funding in the system to make the transition to open access complete. But that funding can only be unlocked with public sector, policymaker, and buyer leadership, on the basis we look beyond legacy publishing models that have been responsible for a decades-long cost explosion in scholarly publishing.

See for example the University of Missouri analysis (<a href="https://library.missouri.edu/news/lottes-health-sciences-library/scholarly-publishing-and-the-health-sciences-library">https://scholarly-publishing-and-the-health-sciences-library</a>); the University of California San Francisco analysis (<a href="https://www.library.ucsf.edu/about/subscriptions/journals-costs/">https://www.library.ucsf.edu/about/subscriptions/journals-costs/</a>); and the Guardian analysis (<a href="https://www.theguardian.com/science/2017/jun/27/profitable-business-scientific-publishing-bad-for-science">https://www.theguardian.com/science/2017/jun/27/profitable-business-scientific-publishing-bad-for-science</a>).

With the right policies and incentives, agencies can help drive the value of taxpayer-funded investment and spur innovation.

### 3. Methods for monitoring evolving costs and impacts on affected communities.

On financial costs, we welcome the NIH's interest in the commercial drivers of scholarly publishing, particularly in matters of access or equity.

Since our inception as a born-digital publisher, we have positioned ourselves as a researcher-centric organization focused on quality, speed, collaboration, and innovation. The governing principle of all scholarly publishing should be that the researchers have the most freedom possible to focus on their research. And so, all publishers compete to lower administrative and process-based burdens.

While the dissemination of research requires a complex ecosystem, we believe a wide-scale shift to open access would allow libraries and research institutions to free substantial resources now tied up in (paywalled) subscriptions, and to apply those resources to researchers' publishing costs.

A strong signal or directive from the NIH that research institutions should commit these freed-up funds - as well as grant money ringfenced for publication - to the widespread and immediate sharing of research would have a profound and positive impact on the drive to fully open access science.

On the perceived relative fairness of pricing regimes, and as we say in response to Question 1, it is worth noting the article processing charge (APC) is not the only way to finance Gold Open Access (Gold OA) publishing. Indeed, we recognize that in some cases, it is not the preferred or most sustainable price structure for researchers, funders, libraries, and research institutions. And while we, like Others in the publishing industry, think the APC model is a good one, we are continually in touch with institutional partners to find solutions that meet their needs. And we are seeking new models to help authors cover the fair and actual cost of publishing, to make scientific knowledge accessible to the widest possible audience.

Within an APC framework, we have expanded our portfolio of institutional models to meet the tailored needs of our customers (recognizing, for example, the distinct needs of research-intensive "publish" organizations as well as high consumption "read" institutions and societies). Our success indicates a range of pricing options can meet the needs of a range of customers and institutions.

The publishing industry at large is experimenting with pricing models and introducing new ones in its drive to innovate. Though the nomenclature varies - advance annual payment, fixed fee, flat fee, multipayer, Subscribe 2 Open, waivers - all of these seek to offer more cost-efficient and sustainable alternatives to libraries' subscription expenditure.

### 4. Early input on considerations to increase findability and transparency of research.

On data sharing, we fully back the NIH's effort through its Public Access Plan to spur a better and more consistent use of persistent identifiers (PIDs) and metadata. In driving this effort, the NIH is providing critical leadership in the scholarly publishing ecosystem.

Moreover, we welcome the NIH's focus on the findability and transparency of research. Open data drives scientific and technological innovation and spurs collaboration; is critical to driving efficiency and scaling innovation; and in uniform standards can be verified, reproduced, and built upon.

If data is transparent and open to scrutiny and evaluation, it follows that trust and confidence in science are more likely to be sustainable. The infrastructure for open data is readily available and an increasingly frequent resource; what's more, many large-scale repositories already exist to make data open.

Examples include Figshare (<a href="https://figshare.com/">https://figshare.com/</a>), a commercial, field-agnostic repository; field-specific, non-profit databases like the society-supported FlowRepository for cytometry data (<a href="https://sharing.nih.gov/occessing-data/accessing-genomic-data/accessing-genomic-data-from-nih-repositories">https://sharing.nih.gov/accessing-data/accessing-genomic-data/accessing-genomic-data-from-nih-repositories</a>).

On data repositories, substantial funding will be required for operation and upgrades. And in the absence of funding committed to scaling up PMC, Frontiers would back a federated approach that focuses on shared standards and access across multiple repositories. By way of illustration, we deposit the full text or metadata of our 230-plus journals in more than 20 repositories when we publish articles.

As a Gold OA publisher, we have made thousands of peer-reviewed articles available online immediately, without embargo. Our starting point - and end point - is ease of discovery.

In simple terms, an article that cannot be found, cannot be shared, and cannot be cited also cannot spur vital collaboration and breakthrough. Publishing in a Gold OA journal unlocks discoverability. The articles and underlying data are transferred to a repository such as PubMed Central (https://www.ncbi.nlm.nih.gov/pmc/) or stored in commercial or Other non-profit databases.

Moreover, the metadata from Gold OA journals come in XML files and Other machine-readable formats to meet FAIR data standards of findability, accessibility, interoperability, and reuse.

(<a href="https://sharing.nih.gov/data-management-and-sharing-policy/data-management#:~:text=NIH%20encourages%20data%20management%20and,repurposing%20datasets%2">https://sharing.nih.gov/data-management-and-sharing-policy/data-management#:~:text=NIH%20encourages%20data%20management%20and,repurposing%20datasets%2</a>
Ofor%20secondary%20research.)

The metadata includes PIDs such as that of ORCID for author identification (https://info.orcid.org/what-is-orcid/), a Digital Object Identifier (DOI) for the article itself, and tags to the relevant grant funding or research institution. And compliance with JATS DTD for XML and Other PMC-recommended tagging enables an even more efficient search and discovery experience.

The new federal guidelines seek public access without specifying delivery models, and we agree that openly accessible science can - and should - be delivered by more than one publishing model. We welcome competition if it spurs innovation and the amount of rigorous science accessible to all.

But in judging delivery models, we believe federal agencies must make a robust and transparent assessment and comparison across efficiency, scalability, and public value for money - guided by the objective of discoverability that underpins public access.

For example, public access known as "Green Open Access (Green OA)" clearly removes some barriers and does not create or perpetuate inequity. But the mechanisms for finding, reading, and sharing Green OA files vary widely. Substantial new funding will be required just to bring that variance down and lift standards for discoverability, with new investment in infrastructure for metadata enrichment. Those institutions unable to fund that investment are likely to face the continued cost pressure of paywall subscriptions that might only minimally ease search and discovery.

So, it is vital that the funding of public access is as efficient, scalable, and as good a value for money as possible, and in our view, Gold OA publishing is the most effective way of securing that outcome. It offers a simple, transparent, and competitive way to unlock the benefits of fully accessible science.

We think it is possible to achieve the fullest possible access to our collective knowledge - for fairer outcomes in all parts of society - in a business model that is cost-effective, commercially sustainable, and underpinned by private sector innovation.

# **Uploaded File:**

Frontiers\_response\_NIH\_RFI\_2023-04-24.pdf



# Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research

April 24, 2023

# Summary

We welcome the chance to respond to this important <u>request for information</u> from the National Institutes of Health (NIH). Frontiers is a leading research publisher and open science platform. It is the third most-cited and sixth largest in the world. The science we publish is peer-reviewed, globally shared, and free to read.

Our mission is to make all science open – so that we can collaborate better and innovate faster, for fairer and more equitable outcomes in all parts of society. That is our social purpose as a business.

So, we fully support the August 2022 OSTP (Office of Science and Technology Policy) guidelines. We strongly <u>welcomed them</u> at the time. And we think the NIH has posed critical questions in this request for information, not least about the findability and transparency of research.

As a Gold Open Access (OA) publisher, we have made thousands of peerreviewed articles available online immediately, without embargo. Our starting point – and end point – is ease of discovery.

We face global, existential threats. From health emergencies to climate change, we see and feel them now. We can manage and reverse these threats, to live healthy lives on a healthy planet. But that will require political will, global collaboration, and scientific breakthrough at a scale not yet seen.

On all those counts, success will depend on the widespread sharing of the latest scientific knowledge. All of it. We think scale matters. Tackling these threats will require more than incremental change. Good research published at scale and shared globally, with machine readability across large volumes of information, will accelerate scientific discovery and grow our chances of success.

In simple terms, an article that cannot be found, cannot be shared, and cannot be cited also cannot spur collaboration and breakthrough. Publishing in a Gold OA journal unlocks discoverability. The articles and underlying data are transferred to a repository such as <a href="PubMed Central">PubMed Central</a> or stored in commercial or other non-profit databases. The metadata come in XML files and other machine-readable formats to meet <a href="FAIR data standards">FAIR data standards</a> of findability, accessibility, interoperability, and reuse. And that data includes persistent identifiers (PIDs) such as that of <a href="ORCID">ORCID</a> for author identification, a Digital Object Identifier (DOI) for the article itself, and tags to the relevant grant funding or research institution.

The new federal guidelines seek public access but do not specify delivery models. We agree that openly accessible science can – and should – be delivered by more than one publishing model. We welcome competition if it spurs innovation and the amount of rigorous science accessible to all.



But in judging those delivery models, federal agencies must make a robust and transparent assessment to compare them for efficiency, scalability, and public value for money – guided by the objective of discoverability that underpins public access.

For example, public access known as "Green Open Access (Green OA)" clearly removes some barriers and does not create or perpetuate inequity. But the mechanisms for finding, reading, and sharing Green OA files vary widely. Substantial new funding will be required just to bring that variance down and lift standards for discoverability, with new investment in infrastructure for metadata enrichment. Those institutions unable to fund that investment are likely to face the continued cost pressure of paywall subscriptions that might only minimally ease search and discovery.

So, it is vital that the funding of public access is as efficient, scalable, and as good a value for money as possible, and in our view, Gold OA publishing is the most effective way of securing that outcome. It offers a simple, transparent, and competitive way to unlock the benefits of fully accessible science and does so more effectively than the Green OA option.

As such we believe that the NIH, if it chooses to allow for compliance through either a Green OA or Gold OA model, should express a preference for compliance through Gold OA.

We think it is possible to achieve the fullest possible access to our collective knowledge – for fairer outcomes in all parts of society – in a business model that is cost-effective, commercially sustainable, and underpinned by private sector innovation. That is possible only in a Gold OA model.

We stand ready to support the NIH and its partners in the federal government. It is vital we back this effort for open science and meet the public appetite for accountability, transparency, and trust.

### Full response

Our detailed responses to the NIH's framing (in italics) are set out here.

# 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

On public repositories, we believe the NIH Public Access Plan rightly encourages and prioritizes the widest possible choices for researchers as they relate to publishing venue, as well as the principles of academic freedom. We think the Plan strikes the right balance by making PubMed Central (PMC) a convenient and compliant repository for research without privileging or mandating it.

On the fairness of the article processing charge (APC), we believe it is both fair and effective as it is a fee for a service. But although it is the most efficient and transparent method, it is worth noting this charge is not the only way to finance



Gold Open Access (Gold OA) publishing. Indeed, we recognize that in some cases, it is not the preferred or most sustainable price structure for researchers, funders, libraries, and research institutions. And while we, like others in the publishing industry, think the APC model is a good one, we are continually in touch with institutional partners to find solutions that meet their needs. And we are seeking new models to help authors cover the fair and actual cost of publishing, to make scientific knowledge accessible to the widest possible audience.

Within an APC framework, we have expanded our portfolio of institutional models to meet the tailored needs of our customers (recognizing, for example, the distinct needs of research-intensive "publish" organizations as well as high consumption "read" institutions and societies). Our success indicates a range of pricing options can meet the needs of a range of customers and institutions.

On the additional steps the NIH might take to ensure new inequities are not created, or existing ones reinforced, we believe the NIH should encourage researchers to publish in the Gold OA model – on the basis that the public funding of public access is efficient, scalable, and delivers value for money.

In our view, Gold OA publishing is the most effective way of securing that outcome. It offers a simple, transparent, and competitive way to unlock the benefits of fully accessible science; and it enables researchers, agencies, universities, libraries, and repositories to fulfil both the NIH Public Access Policy and the OSTP guidance.

Publishing in a Gold OA journal immediately facilitates the transfer of articles to a repository, with metadata in machine-readable formats. In this model, there are no embargoes and no superfluous or costly bundled services that are common in "hybrid" or "transformative" subscription options offered by legacy commercial publishers.

On public value for money, new federal guidelines seek public access but do not specify delivery models. We agree that openly accessible science can – and should – be delivered by more than one publishing model. We welcome competition if it spurs innovation and the amount of rigorous science accessible to all.

But in judging those delivery models, federal agencies must make a robust and transparent assessment and a comparison for efficiency, scalability, and public value for money – guided by the objective of discoverability that underpins public access.

For example, public access known as "Green Open Access (Green OA)" clearly removes some barriers and does not create or perpetuate inequity. But the mechanisms for finding, reading, and sharing Green OA files vary widely. Substantial new funding will be required just to bring that variance down and lift standards for discoverability, with new investment in infrastructure for metadata enrichment. Those institutions unable to fund that investment are likely to face



the continued cost pressure of paywall subscriptions that might only minimally ease search and discovery.

So, it is vital that the funding of public access is as efficient, scalable, and as good a value for money as possible, and in our view, Gold OA publishing is the most effective way of securing that outcome. It offers a simple, transparent, and competitive way to unlock the benefits of fully accessible science.

# 2. Steps for improving equity in access and accessibility of publications.

On the 12-month embargo, we strongly welcome the NIH's decision to end it on publications. We believe that so-called Transformative Agreements (TAs) were worthwhile in their conception as a means of smoothing the transition to fully open access science, but in their execution have not effectively led to transformation and have instead become a blunt instrument.

TAs lack transparency, have complex bundles of often unnecessary services making it all but impossible to judge value for money, and come with no contractual commitment to a move to full open access (Green, Gold, or otherwise) within a binding deadline.

Most of these TAs are large scale "read and publish" or hybrid deals. Publishers will often allow authors' work to appear in hybrid journals without being charged (if their institutions pay), while at the same time maintaining the amount of science they publish behind paywalls.

We believe TAs help subsidize the market dominance of legacy publishers by controlling the pace of transition to fully open access science. The worldwide scientific publishing oligopoly is a market estimated to be around US \$27 billion.1 The five largest paywall publishing houses<sup>2</sup> have captured more than half of it.<sup>3</sup>

On the basis the NIH seeks equity in access as well as transparency in costs, backed by financial sustainability, we believe Gold OA publishers can deliver.

On automated text processing, assistive devices, and other inclusionary measures, we fully support the NIH's position. We consistently invest in measures that improve the accessibility of our publications. Many such requirements were mandated by the Coalition S initiative, which Frontiers fully supported, and which saw wide-ranging and progressive open access policies adopted in the United Kingdom and across Europe.

We firmly back public policies that promote equity of opportunity, the ability both to read and to publish research, and the scientific rigor, academic freedom, institutional values, and personal and professional recognition that underpin success.

<sup>&</sup>lt;sup>1</sup> By revenues. In 2021. Outsell Inc., Segment View: Scientific, Technical and Medical, 2021 (cited in STM Global Brief 2021 - Economics & Market Size).

Elsevier, Wiley, Springer Nature, Taylor & Francis, and SAGE.
 Livres Hebdo/Publishers Weekly 2021 ranking of top global publishers.



We are committed to increasing research access, knowledge resources, and educational opportunities for all, especially for those groups, nations, and individuals who are historically marginalized, underrepresented, or disadvantaged.

On institutional success, we work to build communities and tackle the inadequacies and inequities often characterizing research dissemination. The shift toward open access represents an opportunity to expand access to knowledge in a significant way across academic institutions of all stripes, as well as to small businesses and the public.

We urge the NIH to draw on its influence to see that library, research, and educational institutions commit to investing in open access so that all parties can source sufficient funding for publishing. Several equitable open publishing models are readily available. It cannot be right if colleges and universities are encouraged to maintain robust publications budgets for subscriptions and then asked to make cuts to open access.

We believe there is enough funding in the system to make the transition to open access complete. But that funding can only be unlocked with public sector, policymaker, and buyer leadership, on the basis we look beyond legacy publishing models that have been responsible for a decades-long cost explosion in scholarly publishing.<sup>4</sup> With the right policies and incentives, agencies can help drive the value of taxpayer-funded investment and spur innovation.

# 3. Methods for monitoring evolving costs and impacts on affected communities.

On financial costs, we welcome the NIH's interest in the commercial drivers of scholarly publishing, particularly in matters of access or equity.

Since our inception as a born-digital publisher, we have positioned ourselves as a researcher-centric organization focused on quality, speed, collaboration, and innovation. The governing principle of all scholarly publishing should be that the researchers have the most freedom possible to focus on their research. And so, all publishers compete to lower administrative and process-based burdens.

While the dissemination of research requires a complex ecosystem, we believe a wide-scale shift to open access would allow libraries and research institutions to free substantial resources now tied up in (paywalled) subscriptions, and to apply those resources to researchers' publishing costs.

A strong signal or directive from the NIH that research institutions should commit these freed-up funds – as well as grant money ringfenced for publication – to the widespread and immediate sharing of research would have a profound and positive impact on the drive to fully open access science.

<sup>&</sup>lt;sup>4</sup> See for example: <u>University of Missouri analysis</u>; <u>University of California San Francisco analysis</u>; <u>Guardian analysis</u>.



On the perceived relative fairness of pricing regimes, and as we say in response to Question 1, it is worth noting the article processing charge (APC) is not the only way to finance Gold Open Access (Gold OA) publishing. Indeed, we recognize that in some cases, it is not the preferred or most sustainable price structure for researchers, funders, libraries, and research institutions. And while we, like others in the publishing industry, think the APC model is a good one, we are continually in touch with institutional partners to find solutions that meet their needs. And we are seeking new models to help authors cover the fair and actual cost of publishing, to make scientific knowledge accessible to the widest possible audience.

Within an APC framework, we have expanded our portfolio of institutional models to meet the tailored needs of our customers (recognizing, for example, the distinct needs of research-intensive "publish" organizations as well as high consumption "read" institutions and societies). Our success indicates a range of pricing options can meet the needs of a range of customers and institutions.

The publishing industry at large is experimenting with pricing models and introducing new ones in its drive to innovate. Though the nomenclature varies – advance annual payment, fixed fee, flat fee, multi-payer, Subscribe 2 Open, waivers – all of these seek to offer more cost-efficient and sustainable alternatives to libraries' subscription expenditure.

# 4. Early input on considerations to increase findability and transparency of research.

On data sharing, we fully back the NIH's effort through its Public Access Plan to spur a better and more consistent use of persistent identifiers (PIDs) and metadata. In driving this effort, the NIH is providing critical leadership in the scholarly publishing ecosystem.

Moreover, we welcome the NIH's focus on the findability and transparency of research. Open data drives scientific and technological innovation and spurs collaboration; is critical to driving efficiency and scaling innovation; and in uniform standards can be verified, reproduced, and built upon.

If data is transparent and open to scrutiny and evaluation, it follows that trust and confidence in science are more likely to be sustainable. The infrastructure for open data is readily available and an increasingly frequent resource; what's more, many large-scale repositories already exist to make data open. Examples include <a href="Figshare">Figshare</a>, a commercial, field-agnostic repository; field-specific, non-profit databases like the society-supported <a href="FlowRepository">FlowRepository</a> for cytometry data and the commercial <a href="Protein Data Bank">Protein Data Bank</a>; and federally backed databases like NIH's <a href="databases">databases</a> like NIH's <a href="

On data repositories, substantial funding will be required for operation and upgrades. And in the absence of funding committed to scaling up PMC, Frontiers would back a federated approach that focuses on shared standards and access across multiple repositories. By way of illustration, we deposit the full text or



metadata of our 230-plus journals in more than 20 repositories when we publish articles.

As a Gold OA publisher, we have made thousands of peer-reviewed articles available online immediately, without embargo. Our starting point – and end point – is ease of discovery.

In simple terms, an article that cannot be found, cannot be shared, and cannot be cited also cannot spur vital collaboration and breakthrough. Publishing in a Gold OA journal unlocks discoverability. The articles and underlying data are transferred to a repository such as <a href="PubMed Central">PubMed Central</a> or stored in commercial or other non-profit databases.

Moreover, the metadata from Gold OA journals come in XML files and other machine-readable formats to meet <u>FAIR data standards</u> of findability, accessibility, interoperability, and reuse. The metadata includes PIDs such as that of <u>ORCID</u> for author identification, a Digital Object Identifier (DOI) for the article itself, and tags to the relevant grant funding or research institution. And compliance with JATS DTD for XML and other PMC-recommended tagging enables an even more efficient search and discovery experience.

The new federal guidelines seek public access without specifying delivery models, and we agree that openly accessible science can – and should – be delivered by more than one publishing model. We welcome competition if it spurs innovation and the amount of rigorous science accessible to all.

But in judging delivery models, we believe federal agencies must make a robust and transparent assessment and comparison across efficiency, scalability, and public value for money – guided by the objective of discoverability that underpins public access.

For example, public access known as "Green Open Access (Green OA)" clearly removes some barriers and does not create or perpetuate inequity. But the mechanisms for finding, reading, and sharing Green OA files vary widely. Substantial new funding will be required just to bring that variance down and lift standards for discoverability, with new investment in infrastructure for metadata enrichment. Those institutions unable to fund that investment are likely to face the continued cost pressure of paywall subscriptions that might only minimally ease search and discovery.

So, it is vital that the funding of public access is as efficient, scalable, and as good a value for money as possible, and in our view, Gold OA publishing is the most effective way of securing that outcome. It offers a simple, transparent, and competitive way to unlock the benefits of fully accessible science.

We think it is possible to achieve the fullest possible access to our collective knowledge – for fairer outcomes in all parts of society – in a business model that is cost-effective, commercially sustainable, and underpinned by private sector innovation.

 $\textbf{Description:} \ \textbf{A PDF version of the comments submitted via this webform, with a summary of the} \\$ 

Frontiers position at the top of the document

Email: tom.ciavarella@frontiersin.org

**Submit date:** 4/21/2023

I am responding to this RFI: On behalf of an organization

Name: Lauren Gross, J.D.

Name of Organization: The American Association of Immunologists (AAI)

Type of Organization: Professional org association

# 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

The American Association of Immunologists (AAI) appreciates this opportunity to submit comments in response to NOT-OD-23-091: "Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research." AAI is the nation's largest association of professionally trained scientists dedicated to advancing the knowledge of immunology and its related disciplines, fostering the interchange of ideas and information among investigators, and addressing the potential integration of immunologic principles into clinical practice. Founded in 1913, AAI serves its members and the global immunology community by providing a center for the dissemination of information relevant to the field and its practices, organizing and sponsoring educational and professional opportunities, planning and hosting scientific meetings, addressing members' issues and opinions, and advocating for funding and policy priorities that strengthen the biomedical research enterprise, particularly for immunologists. Central to AAI's mission is its role as a scientific publisher: AAI owns and publishes The Journal of Immunology (The JI), the most highly cited journal in the field, as well as ImmunoHorizons (IH), a fully open-access, peer-reviewed journal dedicated to the science of immunology. As a not-for-profit scholarly scientific society, AAI invests the net revenue it receives from publications in programs and activities that advance immunology/related fields or that support AAI members' research and work lives.

AAI understands and appreciates the Administration's goal of increasing public access to the results of taxpayer-funded research. It is imperative, however, that NIH, as the nation's premier and largest funder of biomedical research, adopt and implement a plan that will foster access to accurate, peer-reviewed, reliable scientific information, while also helping to limit the potential for unintended proliferation of poor-quality or unreliable scientific content. Public access for its own sake, without the safeguards provided by professional scientific societies like AAI (as described herein) and Other responsible publishers, could increase public distrust of science, delay scientific advancement, damage public health, and/or undermine the competitive edge the U.S. has long had in scientific research and development. In addition to the comments below, AAI calls to NIH's attention important comments submitted by the Federation of American Societies for Experimental Biology (FASEB).

AAI supports the intent of the NIH Public Access Plan to maintain the existing broad discretion that allows authors to choose how and where to publish their research. Until recently, this was in fact the author's choice: authors could use their grant (or Other) funds to publish in the journal best suited to their needs and their research findings. However, the White House Office of Science and Technology Policy (OSTP) memorandum on "Ensuring Free, Immediate, and Equitable Access to Federally Funded Research" ("Nelson memo"), published on August 25, 2022, has accelerated a trend set by European funders and a small group of U.S.-based funders: requiring authors to publish only in journals with specific open access models. As a result, fewer authors are submitting to hybrid or subscription-only journals, many of which are owned and/or published by not-for-profit professional scientific societies,

and some of which could be in financial jeopardy as a result of this impending policy. This "thumb on the scale" by the federal government has left researchers and authors in a bind: they may no longer be able to choose the journal that might best support them as authors or showcase their work as broadly and responsibly as they would like. Instead, they must find a publisher that satisfies their funder requirements, based on the model of the journal or a contractual agreement and not necessarily on its quality, mentoring, publication record, or any Other feature.

AAI does not believe that authors should be required to publish in journals with specific business models. As a not-for-profit professional society, AAI's scholarly journals offer two different models (The JI is hybrid, IH is open access) and a shared commitment to peer review and mentoring. In keeping with AAI's educational mission and in order to maintain the integrity of AAI journals' scientific content, all AAI reviewers are Ph.D.-level scientists conducting active research in their fields. AAI staff scientists use a database of thousands of potential reviewers to find subject matter experts to serve as reviewers for each manuscript submission. This database, developed and maintained at AAI expense, includes but is not limited to members' self-identified areas of expertise and information about the perceived usefulness and timeliness of past reviews. In the past five years, AAI secured more than 12,000 reviewers who were qualified and available to undertake a review. AAI also invests in preventing both real and apparent conflicts of interest (COI) with respect to research activities and collaborative or personal interactions. The careful solicitation of reviewers, managing the peer-review process, ensuring research integrity, and avoiding COI are essential steps toward ensuring that reviews are scientifically sound, impartial, professional, and equitable to all submitting authors. These activities require extensive time commitments from AAI staff as well as access to expensive software and tools.

Unlike most publishers, AAI peer reviews 94% of submissions; only articles considered out-of-scope are rejected before peer review. Offering this peer review is part of the AAI educational mission and ensures that high-quality peer review is available to virtually all immunological researchers regardless of laboratory, University/institution, or country of origin. It may be particularly helpful to early career scientists, some of whom may have little or no relevant mentoring at their institutions, who learn how to prepare a scientific paper for publication and are able to publish in a respected scholarly journal, which is necessary for career advancement.

Beyond funder restrictions is the matter of publication costs. AAI urges NIH to develop clear guidance on all ways in which investigators may charge these costs. In addition to allowing authors to charge reasonable publication costs to the direct portion of their grants, NIH should develop novel ways and funding mechanisms, and work with academia and institutions to consider alternatives, including the use of indirect funds, that do not require researchers to utilize grant funds intended for research. NIH should also acknowledge, and consider solutions for, the fact that using direct grant funding for publishing costs reduces the available funding for necessary research costs (including support for personnel, equipment and supplies, funding for experiments, etc.), which may decelerate scientific discovery and will almost certainly place an additional burden on less well-funded investigators and/or institutions. The NIH Public Access Plan's removal of the 12-month embargo period, resulting in a fully open-access model, will likely cause publication fees to increase, perhaps dramatically, disproportionately and negatively impacting under-resourced investigators and institutions, especially those that do not have libraries with the means to enter into transformative agreements or Other arrangements that would not require authors to pay publication costs from the direct portions of their grants. NIH should monitor, and provide guidance on addressing, disparities in publishing opportunities.

NIH should consider ways to alleviate the potential increase in administrative burden that investigators will face if they become responsible for ensuring their publications are publicly and freely available (e.g., deposition of manuscripts to PubMed). Currently, this service is often provided by the publisher. With regard to The JI, AAI has deposited manuscripts on the author's behalf since 2011, a service that may have to be discontinued without the support of revenue currently received from subscriptions. Similarly, NIH should acknowledge and address the fact that not-for-profit scientific societies that publish scholarly journals, which provide tremendous value to the biomedical ecosystem, do not have the same resources as large publishers; NIH should assist these societies during and after the transition to ensure their continued ability to serve their authors, the federal government, and taxpayers, including reviewing and validating the accuracy and rigor of federally funded scientific research.

Finally, although not addressed in this RFI, AAI strongly supports the ability of authors to choose the copyright license that best suits the needs of their funders and themselves. A copyright license that restricts the reuse of derivatives maintains the scientific integrity of a researcher's work that could be misconstrued or misunderstood if presented in partial form. In addition, a copyright license that restricts the reuse for commercial purposes ensures that the work is not misappropriated.

# 2. Steps for improving equity in access and accessibility of publications.

All scientists and physicians conducting (or training to conduct) research in immunology or related fields are welcome to apply for membership in AAI. Members receive immediate access to The JI (and like the public, can access IH at no cost online). Nonmember scientists, physicians, and public health officials who wish to view The JI content before the 12-month embargo period ends often access it at their institution's (or government) library. As the most highly cited journal in the field, The JI is widely available, and as a publication of a not-for-profit professional society, it is reasonably priced and affordable to smaller institutions.

AAI is acutely aware of the importance of sharing scientific and medical information with the general public but believes that immediate and free dissemination of full-length scholarly journal articles is not the most effective or efficient way to accomplish this goal. (Scientific journal articles are tailored to experts in a specific field and are sometimes not well understood even by experts in a different subdiscipline of the same field, much less by a lay audience.) AAI has a long track record of programs intended to accomplish the goal of making scientific content accessible to the public. AAI develops educational materials for the public and for Congress and offers immediate and free online access to abstracts of all scholarly articles published in AAI's journals. Furthermore, AAI has invested in developing accessibility tools like "Key Points" (three-sentence lay summaries) and visual abstracts (lay-friendly graphic representations of the main points of articles) for published articles, free of charge and publicly available on The JI website (<a href="https://journals.aai.org/jimmunol/issue">https://journals.aai.org/jimmunol/issue</a>).

Additionally, AAI was a responsible contributor to the sharing of critically important scientific information throughout the COVID-19 pandemic and intends to respond with equal commitment to any future public health emergency. Most recently, the association launched a new initiative through which AAI members have explained the importance of vaccination and how it works to protect from illness, among Other immunology topics, on television, radio, social media, and in print media.

AAI has been able to provide these important services to the public only because of the revenue generated by the AAI journal subscription model. Should AAI lose revenue as a result of the new public

access policy, the association may not be able to continue to provide programs and services that expand access and information to both scientists and the public.

# 3. Methods for monitoring evolving costs and impacts on affected communities.

AAI supports NIH's effort to monitor and share information regarding trends in publication fees. However, AAI recommends against any action that may inadvertently lead to inequities in publishing opportunity, favor high-volume rather than high-quality publishing, and/or negatively affect the quality of publications. Peer review, in which AAI heavily invests and which is essential to upholding scientific integrity, cannot be undertaken or accomplished at no cost, and any model that does not sufficiently compensate for providing peer review and ensuring Other critical aspects of scientific integrity (e.g., ethics, rigor, reproducibility, etc.) will inevitably lead to a reduction in publication quality and will ultimately slow, or could even reverse, the very scientific progress that NIH wishes to speed.

# 4. Early input on considerations to increase findability and transparency of research.

AAI encourages all authors to provide an ORCID ID, a unique, persistent identifier that can be obtained free of charge by researchers, with their article submission. In an effort to successfully capture AAI authors' funding information, AAI further customized - at additional expense - the AAI manuscript submission system to include funder(s), grant reference numbers, and investigators' name.

At considerable expense, AAI also added digital object identifiers (DOIs) to nearly 100,000 articles from its journal archive, dating back to 1916, and continues to utilize them for all publications. A DOI is a unique and never-changing alphanumeric string assigned to online journal articles, which makes it easier to search for and retrieve published works, and makes content more accessible to researchers, clinicians and public health officials, students and educators, and Other members of the public. AAI supports the adoption of DOIs for NIH grants; this would allow for efficient and consistent tracking of investigators' grants, publications, and research data.

AAI appreciates that NIH is asking about, and urges NIH to allow continued use of, persistent identifiers (PIDs) and metadata that have been commonly used by scholarly scientific societies. This is important to avoid unnecessary disruption, confusion, and cost.

AAI appreciates NIH's willingness to hear the concerns of scholarly scientific societies that wish to continue publishing high-quality, peer-reviewed scientific articles designed for experts in their discipline, and to engage in an iterative process to achieve a policy with broad consensus. AAI believes that there is a way forward to address the widespread desire for more public access to needed scientific information that can still preserve the unique and essential role of scholarly scientific society publishers to conduct the necessary review, editing, dissemination, monitoring (including corrections and retractions), and archiving of the manuscripts/articles that AAI publishes. AAI looks forward to continuing to work with NIH to ensure that that the association can continue to advance the field of immunology through publication and Other educational activities in the years to come.

#### **Uploaded File:**

NIH-Public-Access-Plan-RFI.Final-AAI-Comments.4.21.23.pdf

Email: <a href="mailto:lgross@aai.org">lgross@aai.org</a>



Submission by The American Association of Immunologists to the National Institutes of Health (NIH) Request for Information (RFI) on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research

April 21, 2023

#### Introduction

The American Association of Immunologists (AAI) appreciates this opportunity to submit comments in response to NOT-OD-23-091: "Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research." AAI is the nation's largest association of professionally trained scientists dedicated to advancing the knowledge of immunology and its related disciplines, fostering the interchange of ideas and information among investigators, and addressing the potential integration of immunologic principles into clinical practice. Founded in 1913, AAI serves its members and the global immunology community by providing a center for the dissemination of information relevant to the field and its practices, organizing and sponsoring educational and professional opportunities, planning and hosting scientific meetings, addressing members' issues and opinions, and advocating for funding and policy priorities that strengthen the biomedical research enterprise, particularly for immunologists.

Central to AAI's mission is its role as a scientific publisher: AAI owns and publishes *The Journal of Immunology (The JI)*, the most highly cited journal in the field, as well as *ImmunoHorizons (IH)*, a fully open-access, peer-reviewed journal dedicated to the science of immunology. As a not-for-profit scholarly scientific society, AAI invests the net revenue it receives from publications in programs and activities that advance immunology/related fields or that support AAI members' research and work lives.

AAI understands and appreciates the Administration's goal of increasing public access to the results of taxpayer-funded research. It is imperative, however, that NIH, as the nation's premier and largest funder of biomedical research, adopt and implement a plan that will foster access to accurate, peer-reviewed, reliable scientific information, while also helping to limit the potential for unintended proliferation of poor-quality or unreliable scientific content. Public access for its own sake, without the safeguards provided by professional scientific societies like AAI (as described herein) and other responsible publishers, could increase public distrust of science, delay scientific advancement, damage public health, and/or undermine the competitive edge the U.S. has long had in scientific research and development. In addition to the comments below, AAI calls to NIH's attention important comments submitted by the Federation of American Societies for Experimental Biology (FASEB).

1. How to best ensure equity in publication opportunities for NIH-supported investigators.

The NIH Public Access Plan aims to maintain the existing broad discretion for researchers and authors to choose how and where to publish their results. Consistent with current practice, the NIH Public Access

Plan allows the submission of final published articles to PubMed Central (PMC) (in cases where a formal agreement is in place) to minimize the compliance burden on NIH-supported researchers and also maintains the flexibility of NIH-supported researchers to submit the final peer-reviewed manuscript. NIH seeks information on additional steps it might consider taking to ensure that proposed changes to implementation of the NIH Public Access Policy do not create new inequities in publishing opportunities or reinforce existing ones.

AAI supports the intent of the NIH Public Access Plan to maintain the existing broad discretion that allows authors to choose how and where to publish their research. Until recently, this was in fact the author's choice: authors could use their grant (or other) funds to publish in the journal best suited to their needs and their research findings. However, the White House Office of Science and Technology Policy (OSTP) memorandum on "Ensuring Free, Immediate, and Equitable Access to Federally Funded Research" ("Nelson memo"), published on August 25, 2022, has accelerated a trend set by European funders and a small group of U.S.-based funders: requiring authors to publish only in journals with specific open access models. As a result, fewer authors are submitting to hybrid or subscription-only journals, many of which are owned and/or published by not-for-profit professional scientific societies, and some of which could be in financial jeopardy as a result of this impending policy. This "thumb on the scale" by the federal government has left researchers and authors in a bind: they may no longer be able to choose the journal that might best support them as authors or showcase their work as broadly and responsibly as they would like. Instead, they must find a publisher that satisfies their funder requirements, based on the model of the journal or a contractual agreement and not necessarily on its quality, mentoring, publication record, or any other feature.

AAI does not believe that authors should be required to publish in journals with specific business models. As a not-for-profit professional society, AAI's scholarly journals offer two different models (*The JI* is hybrid, *IH* is open access) and a shared commitment to peer review and mentoring. In keeping with AAI's educational mission and in order to maintain the integrity of AAI journals' scientific content, all AAI reviewers are Ph.D.-level scientists conducting active research in their fields. AAI staff scientists use a database of thousands of potential reviewers to find subject matter experts to serve as reviewers for each manuscript submission. This database, developed and maintained at AAI expense, includes but is not limited to members' self-identified areas of expertise and information about the perceived usefulness and timeliness of past reviews. In the past five years, AAI secured more than 12,000 reviewers who were qualified and available to undertake a review. AAI also invests in preventing both real and apparent conflicts of interest (COI) with respect to research activities and collaborative or personal interactions. The careful solicitation of reviewers, managing the peer-review process, ensuring research integrity, and avoiding COI are essential steps toward ensuring that reviews are scientifically sound, impartial, professional, and equitable to all submitting authors. These activities require extensive time commitments from AAI staff as well as access to expensive software and tools.

Unlike most publishers, AAI peer reviews 94% of submissions; only articles considered out-of-scope are rejected before peer review. Offering this peer review is part of the AAI educational mission and ensures that high-quality peer review is available to virtually all immunological researchers regardless of laboratory, university/institution, or country of origin. It may be particularly helpful to early career scientists, some of whom may have little or no relevant mentoring at their institutions, who learn how to prepare a scientific paper for publication and are able to publish in a respected scholarly journal, which is necessary for career advancement.

Beyond funder restrictions is the matter of publication costs. AAI urges NIH to develop clear guidance on all ways in which investigators may charge these costs. In addition to allowing authors to charge

reasonable publication costs to the direct portion of their grants, NIH should develop novel ways and funding mechanisms, and work with academia and institutions to consider alternatives, including the use of indirect funds, that do not require researchers to utilize grant funds intended for research. NIH should also acknowledge, and consider solutions for, the fact that using direct grant funding for publishing costs reduces the available funding for necessary research costs (including support for personnel, equipment and supplies, funding for experiments, etc.), which may decelerate scientific discovery and will almost certainly place an additional burden on less well-funded investigators and/or institutions. The NIH Public Access Plan's removal of the 12-month embargo period, resulting in a fully open-access model, will likely cause publication fees to increase, perhaps dramatically, disproportionately and negatively impacting under-resourced investigators and institutions, especially those that do not have libraries with the means to enter into transformative agreements or other arrangements that would not require authors to pay publication costs from the direct portions of their grants. NIH should monitor, and provide guidance on addressing, disparities in publishing opportunities.

NIH should consider ways to alleviate the potential increase in administrative burden that investigators will face if they become responsible for ensuring their publications are publicly and freely available (e.g., deposition of manuscripts to PubMed). Currently, this service is often provided by the publisher. With regard to *The JI*, AAI has deposited manuscripts on the author's behalf since 2011, a service that may have to be discontinued without the support of revenue currently received from subscriptions. Similarly, NIH should acknowledge and address the fact that not-for-profit scientific societies that publish scholarly journals, which provide tremendous value to the biomedical ecosystem, do not have the same resources as large publishers; NIH should assist these societies during and after the transition to ensure their continued ability to serve their authors, the federal government, and taxpayers, including reviewing and validating the accuracy and rigor of federally funded scientific research.

Finally, although not addressed in this RFI, AAI strongly supports the ability of authors to choose the copyright license that best suits the needs of their funders and themselves. A copyright license that restricts the reuse of derivatives maintains the scientific integrity of a researcher's work that could be misconstrued or misunderstood if presented in partial form. In addition, a copyright license that restricts the reuse for commercial purposes ensures that the work is not misappropriated.

# 2. Steps for improving equity in access and accessibility of publications.

Removal of the currently allowable 12-month embargo period for NIH-supported publications will improve access to these research products for all. As noted in the NIH Public Access Plan, NIH also plans to continue making articles available in human and machine-readable forms to support automated text processing. NIH will also seek ways to improve the accessibility of publications via assistive devices. NIH welcomes input on other steps that could be taken to improve equity in access to publications by diverse communities of users, including researchers, clinicians and public health officials, students and educators, and other members of the public.

All scientists and physicians conducting (or training to conduct) research in immunology or related fields are welcome to apply for membership in AAI. Members receive immediate access to *The JI* (and like the public, can access *IH* at no cost online). Nonmember scientists, physicians, and public health officials who wish to view *The JI* content before the 12-month embargo period ends often access it at their institution's (or government) library. As the most highly cited journal in the field, *The JI* is widely available, and as a publication of a not-for-profit professional society, it is reasonably priced and affordable to smaller institutions.

AAI is acutely aware of the importance of sharing scientific and medical information with the general public but believes that immediate and free dissemination of full-length scholarly journal articles is not the most effective or efficient way to accomplish this goal. (Scientific journal articles are tailored to experts in a specific field and are sometimes not well understood even by experts in a different subdiscipline of the same field, much less by a lay audience.) AAI has a long track record of programs intended to accomplish the goal of making scientific content accessible to the public. AAI develops educational materials for the public and for Congress and offers immediate and free online access to abstracts of all scholarly articles published in AAI's journals. Furthermore, AAI has invested in developing accessibility tools like "Key Points" (three-sentence lay summaries) and visual abstracts (lay-friendly graphic representations of the main points of articles) for published articles, free of charge and publicly available on The JI website (https://journals.aai.org/jimmunol/issue).

Additionally, AAI was a responsible contributor to the sharing of critically important scientific information throughout the COVID-19 pandemic and intends to respond with equal commitment to any future public health emergency. Most recently, the association launched a new initiative through which AAI members have explained the importance of vaccination and how it works to protect from illness, among other immunology topics, on television, radio, social media, and in print media.

AAI has been able to provide these important services to the public only because of the revenue generated by the AAI journal subscription model. Should AAI lose revenue as a result of the new public access policy, the association may not be able to continue to provide programs and services that expand access and information to both scientists and the public.

# 3. Methods for monitoring evolving costs and impacts on affected communities.

NIH proposes to actively monitor trends in publication fees and policies to ensure that they remain reasonable and equitable. NIH seeks information on effective approaches for monitoring trends in publication fees and equity in publication opportunities.

AAI supports NIH's effort to monitor and share information regarding trends in publication fees. However, AAI recommends against any action that may inadvertently lead to inequities in publishing opportunity, favor high-volume rather than high-quality publishing, and/or negatively affect the quality of publications. Peer review, in which AAI heavily invests and which is essential to upholding scientific integrity, cannot be undertaken or accomplished at no cost, and any model that does not sufficiently compensate for providing peer review and ensuring other critical aspects of scientific integrity (e.g., ethics, rigor, reproducibility, etc.) will inevitably lead to a reduction in publication quality and will ultimately slow, or could even reverse, the very scientific progress that NIH wishes to speed.

### 4. Early input on considerations to increase findability and transparency of research.

Section IV of the NIH Public Access Plan is a first step in developing the NIH's updated plan for persistent identifiers (PIDs) and metadata, which will be submitted to OSTP by December 31, 2024. NIH seeks suggestions on any specific issues that should be considered in efforts to improve use of PIDs and metadata, including information about experiences institutions and researchers have had with adoption of different identifiers.

AAI encourages all authors to provide an ORCID ID, a unique, persistent identifier that can be obtained free of charge by researchers, with their article submission. In an effort to successfully capture AAI authors' funding information, AAI further customized – at additional expense – the AAI manuscript submission system to include funder(s), grant reference numbers, and investigators' name.

At considerable expense, AAI also added digital object identifiers (DOIs) to nearly 100,000 articles from its journal archive, dating back to 1916, and continues to utilize them for all publications. A DOI is a unique and never-changing alphanumeric string assigned to online journal articles, which makes it easier to search for and retrieve published works, and makes content more accessible to researchers, clinicians and public health officials, students and educators, and other members of the public. AAI supports the adoption of DOIs for NIH grants; this would allow for efficient and consistent tracking of investigators' grants, publications, and research data.

AAI appreciates that NIH is asking about, and urges NIH to allow continued use of, persistent identifiers (PIDs) and metadata that have been commonly used by scholarly scientific societies. This is important to avoid unnecessary disruption, confusion, and cost.

#### Conclusion

AAI appreciates NIH's willingness to hear the concerns of scholarly scientific societies that wish to continue publishing high-quality, peer-reviewed scientific articles designed for experts in their discipline, and to engage in an iterative process to achieve a policy with broad consensus. AAI believes that there is a way forward to address the widespread desire for more public access to needed scientific information that can still preserve the unique and essential role of scholarly scientific society publishers to conduct the necessary review, editing, dissemination, monitoring (including corrections and retractions), and archiving of the manuscripts/articles that AAI publishes. AAI looks forward to continuing to work with NIH to ensure that that the association can continue to advance the field of immunology through publication and other educational activities in the years to come.

**Submit date:** 4/21/2023

I am responding to this RFI: On behalf of an organization

Name: Douglas Kondziolka

Name of Organization: Congress of Neurological Surgeons

Type of Organization: Professional org association

**Role:** Scientific researcher

1. How to best ensure equity in publication opportunities for NIH-supported investigators.

see attached letter.

2. Steps for improving equity in access and accessibility of publications.

see attached letter.

3. Methods for monitoring evolving costs and impacts on affected communities.

see attached letter.

4. Early input on considerations to increase findability and transparency of research.

see attached letter.

**Uploaded File:** 

CNS-Response-to-NIH-RFI.docx\_FINAL\_4.17.23.pdf

**Description:** CNS response to NIH RFI

Email: Douglas.Kondziolka@nyulangone.org



10 N. Martingale Road, Suite 190 Schaumburg, Illinois 60173 Phone: 847.240.2500

Email: neurosurgerypubs@cns.org

April 17, 2023

Response to Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research

**Dear NIH Office of Science Policy,** 

I write to you today as Editor-in-Chief of **Neurosurgery Publications**, which is the **Congress of Neurological Surgeons** (CNS) suite of journals published by Wolters Kluwer (WK). As a physician scientist I am supportive of the goals of NIH in funding research and development in order to drive scientific discovery and appreciate the opportunity to respond to this Request for Information.

Founded in 1951, CNS is one of the world's largest scientific and educational associations of neurological surgeons. Neurological surgery is the medical specialty concerned with the prevention, diagnosis, treatment and rehabilitation of disorders that affect the nervous system, including the spinal column, spinal cord, brain and peripheral nerves. CNS' mission is to "enhance health and improve lives through innovative neurosurgical education, advancement of clinical practice & scientific exchange." Neurosurgery Publications is integral to the CNS' broader mission, the continued advancements in neurosurgical research, and improving patient outcomes.

Neurosurgery Publications includes *Neurosurgery, Operative Neurosurgery*, and *Neurosurgery Practice*. Neurosurgery Publications fully supports the open science initiative demonstrated through the launch of *Neurosurgery Practice* (*née Neurosurgery Open*) in late 2019. *Neurosurgery Practice* provides a pathway to publication of neurosurgical research supported by funders who mandate publication in a fully, open access journal. *Neurosurgery* and *Operative Neurosurgery*, which publish under a more traditional business model, offer a hybrid open access option and permit embargoed, green open access, in compliance with the 2013 memorandum from the White House Office of Science and Technology Policy (OSTP).

We understand the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research outlines the organization's response to the 2022 OSTP Memorandum "Ensuring Free, Immediate, and Equitable Access to Federally Funded Research". We also understand the merits of the 2022 OSTP Memorandum are not the topic of this RFI but would like to state that we had previously expressed our concerns around the elimination of the 12month post-publication embargo in response to the 2020 OSTP RFI "Public Access to Peer-Reviewed Scholarly Publications, Data and Code Resulting from Federally Funded Research". In that response we cited concerns that such a policy change would materially impact the CNS' ability to continue advancing neurosurgical research, as these efforts are supported in large part by the Neurosurgery Publications portfolio. These same concerns are shared with the NIH Plan, which calls for the elimination of the 12-month embargo. We believe this change has the potential to disrupt the scholarly publishing industry broadly by compromising key revenue streams and could in turn have a detrimental impact on Neurosurgery Publications, specifically subscription and advertising revenues. The revenue received from these journals are critical to the continued editorial operations, included but not limited to the Editorial Office's financial support, operation of peer review, solicitation of manuscripts, and the costly investments into video production and infographic support-elements which enhance the scientific and clinical nature of our publications and CNS' ability to fulfill its mission broadly. Given the potential for disruption were the NIH's plan to proceed as outlined we would suggest NIH consider an approach that meets its objectives but also supports publication models that allow publishers and their society partners to sustainably support their operations including the wide dissemination of scientific content.

As a stakeholder in this space Neurosurgery Publications has a vested interest in how the NIH's plan is developed and implemented and we are very interested in being a part of the policy development process. We thank you for the opportunity to submit this response and look forward to continued discussions between the NIH, societies, and publishers toward achieving the goals of the NIH Plan that allow for the accelerated access to scholarly work.

Signed,

Douglas Kondziolka, MD, MSc, FRCSC, FACS

Editor-in-Chief, Neurosurgery Publications

Dougla Kongrier

**Submit date: 4/22/2023** 

I am responding to this RFI: On behalf of an organization

Name: Gerald C. Blazey

Name of Organization: Northern Illinois University

Type of Organization: University

Role: Institutional official

# 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

NIH policy already allows supported researchers to charge reasonable publishing costs against their awards. However, because grants are often subject to an informal funding cap, less-resourced institutions, such as ERIs, will still be disadvantaged because the research authors are presented with the choice of diverting resources from research. NIH should consider mechanisms to avoid this, for instance, by allocating funding for a minimum number of publications from a special funding source AFTER the award is made. That is, ensure that the APC funds are truly and rigorously added "on top" of the research funds.

- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.

Authors at ERIs and early career authors are most likely to have limited resources for APCs. NIH should consider an extension or longitudinal study of the AAAS study Exploring the Hidden Impact of Open Access Financing Mechanisms. Informative extensions of the study would be to discriminate between large and emerging research institutions and low and high diversity institutions. Increased statistics for gender and race would also be helpful. NIH should also consider direct institutional inquiries to compare institutional APC contributions at large and emerging research institutions and low and high diversity institutions. Funding models should be adjusted to mitigate any equities observed.

4. Early input on considerations to increase findability and transparency of research.

### **Uploaded File:**

NIU-Position-Paper-NSF-Reauthorization.pdf

**Description:** Northern Illinois University Response to RFI

Email: gblazey@niu.edu

# Increasing diversity through broadened distribution of research funding Northern Illinois University July 22, 2020

### **Summary:**

Historically, the majority of federal research funding has been distributed to a fraction of our Nation's research universities. Students at the remaining universities, including one half of the Nation's underrepresented minorities (URM) students, have limited or no opportunity to engage in research. Participation in research is considered a high impact practice for student retention and graduation and workforce diversification. Promoting partnerships between the two sets of universities will broaden opportunity and diversity while serving the Nation's research needs. This proposal addresses a concern expressed across the spectrum of research universities.

# **Legislative Proposal:**

Insert at the end of the National Science Foundation reauthorization bill text:

"All funding announcements with anticipated grant awards of \$1 million or more made under authorizations in this bill require the recipient to direct at least 10% of the grant funding to building meaningful partnerships with emerging research institutions. Emerging research institutions are defined as institutions of higher education that have less than \$30 million in annual federal science and engineering research and development expenditures as reported by the National Center for Science and Engineering Statistics."

# **Background:**

Taken together, data from two separate sources show that Federal research support is concentrated at a fraction of the Nation's research universities. The concentration presents a structural impediment to diversification. The Higher Education Research and Development (HERD)¹ survey collected and maintained by the National Center for Science and Engineering Statistics (NCSES) at the National Science Foundation provides comprehensive information on national and institution investments in science and engineering. The Carnegie Classification of Institutions of Higher Education² recognizes *very high research doctoral universities* and *high research doctoral universities*. The classification utilizes data from NCSES and the Integrated Postsecondary Education Data System (IPEDS)³ from the Department of Education. Traditionally, and commonly, the two doctoral university classifications are referred to as R1 and R2 universities; all other universities with research programs are traditionally designated R3 universities.

<sup>&</sup>lt;sup>1</sup> https://ncsesdata.nsf.gov/herd/2018/

<sup>&</sup>lt;sup>2</sup> https://carnegieclassifications.iu.edu/index.php

<sup>3</sup> https://nces.ed.gov/ipeds/

According to the 2018 HERD survey data (the most recent year reported) the 131 R1 institutions received 91.7% (\$32.4 billion) of federal science and engineering research and development dollars. However, IPEDS data shows that those same institutions serve 46.6% of the nation's URM college students, and 52.8% of all college students. Said another way, **over half of our nation's students of color see about 8% (\$2.9 billion) of federal research dollars on their campus**. This leaves the other nearly 300 predominately R2 and R3 institutions listed in the HERD survey with limited research funding and opportunities for over one-half of our URM students. Anecdotally, there is also uneven geographic distribution, with a majority of the R1s located in urban areas and R2s and R3s in peri-urban and rural areas. As a result, students in more rural settings also see less opportunity. These structural characteristics have been evident for decades.

One way to quantify the concentration of federal research resources is to consider the number of awards totaling \$1 million or more made by the National Science Foundation (NSF). These awards often support the creation of new centers, nodes, hubs, or other large-scale research operations. Only 7% of the 33,509 active standard grants and cooperative agreements at the NSF exceed \$1 million. The prevalence of awards of \$1 million or more has increased over time. Nearly 35% of all standard grants and 20% of all cooperative agreements totaling \$1 million or more ever issued by the NSF are currently active, meaning they were likely awarded in the last five years. Looking back over the past two decades, the NSF budget has doubled since fiscal year 2000 while the number of awards over \$1 million has tripled during the same time period, so the concentration of resources in large grants is increasing faster than the NSF budget. This tracks with the trends seen in federal policymaking, where there is increasing interest in deploying federal research resources to create large new hubs, nodes, or centers to address important research topics.

Participation in research is extremely effective for the retention of students and the diversification of STEM fields. Recommendations from The National Academies of Sciences, Engineering and Medicine (NASEM,2016<sup>4</sup>; 2018<sup>5</sup>; 2019<sup>6</sup>;2019<sup>7</sup>), National Academy of Engineering (2018<sup>8</sup>) and National Survey of Student Engagement (2016<sup>9</sup>) support URM student research to increase student engagement and foster a sense of belonging and self-efficacy, which, in turn, leads to higher student interest and graduation rates (NASEM, 2017<sup>10</sup>,2018<sup>11</sup>).

<sup>&</sup>lt;sup>4</sup> National Academies of Sciences, Engineering, and Medicine (2016). Quality in the Undergraduate Experience: What Is It? How Is It Measured? Who Decides? Summary of a Workshop. Washington, DC: The National Academies Press. https://doi.org/10.17226/23514.

<sup>&</sup>lt;sup>5</sup> National Academies of Sciences, Engineering, and Medicine (2018). How People Learn II: Learners, Contexts, and Cultures. Washington, DC: The National Academies Press. https://doi.org/10.17226/24783

<sup>&</sup>lt;sup>6</sup> National Academies of Sciences, Engineering, and Medicine, (2019). The Science of Effective Mentorship in STEM. Washington, DC: The National Academies Press. https://doi.org/10.17226/25568

<sup>&</sup>lt;sup>7</sup> National Academies of Science, Engineering, and Medicine (2019). Minority Serving Institutions: America's Underutilized Resource for Strengthening the STEM Workforce. Washington, DC: The National Academies Press. Doi: https://doi.org.org/10.17226/25257

<sup>&</sup>lt;sup>8</sup> National Academy of Engineering, (2018). Understanding the Educational and Career Pathways of Engineers. Washington, DC: The National Academies Press. doi:https://doi.org/10.17226/25284 https://www.nap.edu/catalog/25284/understanding-the-educational-and-career-pathways-of-engineers

<sup>9</sup> National Survey of Student Engagement. (2016). Retrieved from https://nsse.indiana.edu/html/engagement indicators.cfm

<sup>&</sup>lt;sup>10</sup> National Academies of Sciences, Engineering, and Medicine (2017). Undergraduate Research Experiences for STEM Students: Successes, Challenges, and Opportunities. Washington, DC: The National Academies Press. <a href="https://doi.org/10.17226/24622">https://doi.org/10.17226/24622</a>.

<sup>&</sup>lt;sup>11</sup> National Academies of Sciences, Engineering, and Medicine (2018). Indicators for Monitoring Undergraduate STEM Education. Washington, DC: The National Academies Press. https://doi.org/10.17226/24943.

Participation in research prepares students to think critically, communicate their ideas, and apply their knowledge in the field (NASEM, 2016<sup>4</sup>) and is identified as a high-impact practice by the Association of American Colleges and Universities (Kuh, 2008)<sup>12</sup>.

The concentration of federal research dollars at R1s may provide the expertise and resources for addressing complex problems, however it creates a structural barrier for URM STEM students' participation in high-impact practices. Any proposal must maintain the excellence of the R1s while leveraging the strength of the R1s to broaden opportunity. Both can be achieved by requiring R1 institutions that are hosting new initiatives, research centers, and other large grants to partner with non-R1 institutions.

To fully benefit the nation and broaden participation, these partnerships must ensure that expertise is shared and sustained at the non-R1 institutions. As an example, a quantum information sciences or artificial intelligence center established at a large R1 university could provide fellowships for faculty from their non-R1 partners. Upon return to their home universities these fellows could continue their research and engage students with the support of the R1 center.

Absent legislation requiring structural change, non-R1 institutions with over half our students of color, and nearly half of all U.S. college students, and a strong presence in peri-urban and rural areas will be unable to fully participate in the new research directions under consideration by Congress. Their limited involvement will hamper efforts to diversify opportunity and broaden science literacy. Ensuring partnerships between institutions offers a path forward.

-

<sup>12</sup> https://secure.aacu.org/imis/ItemDetail?iProductCode=E-HIGHIMP&Category=

**Submit date: 4/22/2023** 

I am responding to this RFI: On behalf of an organization

Name: Jessica Polka

Name of Organization: ASAPbio

Type of Organization: Other

Type of Organization-Other: Advocacy organization

Role: Member of the public

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

ASAPbio is a 501(c)(3) organization working to promote innovation and transparency in life sciences communication.

We are fully supportive of the 2022 OSTP directive to make all federally-funded research immediately accessible upon publication. Based on the public access plan the NIH has proposed in response to this memo, we appreciate the NIH's desire to ensure equitable access to research for diverse stakeholders, and to ensure that this is provided at reasonable costs that do not exacerbate existing disparities. Furthermore, we support the need to ensure that research outputs are findable and discoverable through robust infrastructure and standards.

Many of these goals can be supported by moving toward a model where preprints are the primary form of sharing; this would also provide a strong foundation for aligning researchers' incentives with the goals set out in the RFI. Many researchers now experience a disconnect between wanting to share work with the community and existing incentives for keeping data private. In a preprint-centric model, researchers would be recognized for sharing their work early and completely, which would also accelerate scientific discovery. Preprints also support rigor, reproducibility, and integrity by allowing broad engagement in public commenting and peer review. Given these benefits, we offer the following suggestions for using preprints to promote equitable, cost-effective, and discoverable publishing.

We appreciate the prioritization of equitable publication opportunities for researchers as well as access to research articles. Preprints provide a mechanism to meet both goals. Unlike many journal publishing models, preprints are free to post and free to access. Given racial disparities in federal funding, preprints create equity by including those who do not have access to funds for journal publication costs. We call on NIH to recognize preprints that are identical in substance to the latest article version as an option for compliance with its open access policy.

#### 2. Steps for improving equity in access and accessibility of publications.

Preprints need to be open access, meaning licensed for reuse. NIH has already taken a positive step by recommending the CC BY license for preprints (<a href="https://grants.nih.gov/grants/guide/notice-files/NOT-OD-17-050.html">https://grants.nih.gov/grants/guide/notice-files/NOT-OD-17-050.html</a>). However, many preprints on popular servers are still not being published under these licenses, risking the creation of walled gardens. To remedy this, we urge NIH to require that supported investigators publish their preprints and Other publications under a CC BY or less restrictive license.

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

Operating costs for preprint servers are much lower than the average ~\$1,600 article-processing charge at journals that require publication fees (Morrison, Heather et al., 2021, "2011 - 2021 OA APCs", <a href="https://doi.org/10.5683/SP2/84PNSG">https://doi.org/10.5683/SP2/84PNSG</a>, Scholars Portal Dataverse, V1). However, the sustainability of preprint servers is a critical question. They are currently supported by private funders, publishers, institutions and library consortia without long-term commitments. A publicly funded preprint infrastructure offers a sustainable way to achieve equitable access to publishing. We suggest that NIH directly fund the community-owned preprint servers that support the communication needs of its researchers.

#### 4. Early input on considerations to increase findability and transparency of research.

The NIH could make preprints more discoverable by extending the NLM preprint pilot to all preprints, not just those that are NIH-funded. Furthermore, an increasing number of preprints now are being reviewed outside of journals (see groups listed at sciety.org). These reviews should be indexed and connected to preprints on NLM's databases, and they should be visible on the SciENcv profiles of the reviewers who authored them. In addition, metadata for preprints and preprint reviews should be made freely available through appropriate infrastructures, such as the Crossref infrastructure.

Finally, we urge the NIH to move forward with an international focus. Scientific progress is a global endeavor, and implementation needs to be in line with broader frameworks rather than reinventing existing infrastructure. There is support for broad and equitable access to research works via government and funder initiatives in Latin America (e.g. SciELO and AmeliCA) and Europe (e.g. Open Research Europe (ORE)), and also mature infrastructure to enable the use of persistent identifiers (DOI, ORCID, ROR) and appropriate metadata. The NIH should ensure any new infrastructures make use of these common standards and are interoperable with these existing projects. Now is the time for global collaboration to make rapid progress on improving scientific communications infrastructure.

Email: jessica.polka@asapbio.org

**Submit date:** 4/23/2023

I am responding to this RFI: On behalf of an organization

Name: Diane Gern

Name of Organization: American Thoracic Sociey

Type of Organization: Professional org association

Role: Institutional official

- 1. How to best ensure equity in publication opportunities for NIH-supported investigators.
- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

#### **Uploaded File:**

ATS-Letter-4.23.pdf

**Description:** Official letter from the American Thoracic Society (ATS)

Email: dgern@thoracic.org



25 Broadway, 4th Floor, New York, NY 10004 Phone: 212.315.8600 Fax: 212.315.6498 thoracic.org

**Gregory P. Downey, MD, ATSF**President

M. Patricia Rivera, MD, ATSF President-Elect

Lynn M. Schnapp, MD, ATSF Immediate Past President

Irina Petrache, MD, ATSF Secretary

Jesse Roman, MD Treasurer

Karen J. Collishaw, MPP, CAE Chief Executive Officer

American Journal of Respiratory and Critical Care Medicine

American Journal of Respiratory Cell and Molecular Biology

Annals of the American Thoracic Society

ATS Scholar

ATS 2023 International Conference April 23, 2023

We are writing to address the Request for Information on the NIH Plan to Enhance Public Access to NIH-Supported Research (NOT-OD-23-091).

Established in 1905, the American Thoracic Society (ATS), is a nonprofit medical organization focused on improving care for pulmonary diseases, critical illnesses, and sleep-related breathing disorders. The society publishes 4 journals: the *American Journal of Respiratory and Critical Care Medicine*, (AJRCCM), the *American Journal of Respiratory Cell and Molecular Biology* (AJRCMB), *AnnalsATS*, and *ATS Scholar*. All journal content is free to our 15,000 society members, and free to all readers after one year. For the past 3 years, all COVID content, as well as any content that changes practice, is open access (OA) and immediately accessible by all readers. In addition, all authors are offered an OA choice in how their articles are published.

## Our major concerns with the OSTP policy, are addressed below:

#### Apply the OSTP Policy to Original Research Articles Only

There is significant confusion within the academic publishing community about how broadly the OSTP memo will be applied. Clearly, original research articles and their accompanying data are the primary focus of the OSTP policy. However, some have suggested that the immediate public access requirement may apply to other types of academic publications including workshop reports, case reports, review articles, commentaries, and editorials. We strongly urge that the immediate public access policy apply only to federally funded original manuscripts. Applying the immediate public access mandate to other types of academic manuscripts is beyond the reasonable application of the OSTP memo.

#### **Ensure Publication Fees are Covered Grant Expenses**

As many academic scientific journals have noted, implementation of an immediate public access mandate will require significant changes in our business models and will likely result in significantly higher publication fees. We are pleased that federal research programs have recognized this reality and several federal research granting entities have explicitly stated that publication fees currently are and will continue to be permissible grant expenses. We strongly encourage all federal research granting agencies to explicitly state that *full* publication fees are permissible and expected grant expenses.

#### A Flexible Business Model

As you are aware, the push to increase public access to federally funded research results has given rise to a wide variety of publication models (Gold OA, Green OA, etc). Eliminating the embargo shifts the cost of maintaining peer review and publishing efforts to authors, researchers, and institutions. Researchers need to know that their grant awards will

be increased to cover these costs. As NIH considers how to implement the OSTP immediate public access policy, we urge you to craft policy that is "publication model neutral." This will continue to allow the academic publication community the freedom to develop and try a wide variety of publication models that are both responsive to the needs of the immediate public access model and other needs of the scientific community.

#### **Original Manuscript**

There continues to be much discussion about public posting of what version of federally funded research manuscripts will satisfy the public access mandate. We urge you to use public posting (by the authors) of the non-peer-reviewed version of the manuscript as meeting the public access mandate. We further recommend that NIH not require posting in any specific public repository, but rather describe the requirements that an eligible public repository must meet and allow journals to decide which repository best meets their academic community needs, while also meeting the public posting mandate.

#### **Copyright Protection**

Given that the protection of intellectual property is an author's first line of defense, we ask that you consider upholding copyright protection. A Creative Commons license (CC BY) that allows for the free use of the material without any parameters may misrepresent the science and lead to erroneous conclusions. A Creative Commons license that does not allow commercial or derivate use (CC BY-NC-ND) would allow the science to be publicly available while protecting the author's intellectual property against commercial exploitation.

#### **The Solution**

In lieu of requiring deposit of the original, non-peer-reviewed manuscript in PubMed Central, the NIH should consider, instead, mandating the deposit of the author's original manuscript in a *pre-print repository* and maintaining the current, one-year embargo for the peer-reviewed, edited, typeset version. This would provide taxpayers access to the original version funded by the NIH grant while making clear that the article has not been peer reviewed. Consequently, scientific discoveries will be available in real time, while ensuring that all *published* content is peer-reviewed, ensuring that patients will not be potentially injured by non-peer-reviewed, published content.

Thank you for the opportunity to respond to the NIH. We would be glad to join efforts to come to a reasonable plan for government-funded content.

Sincerely,

Karen Collishaw,

CEO, American Thoracic Society

Diane gun

Kan f. Collect

Diane Gern

Chief, Journals, American Thoracic Society

**Submit date:** 4/23/2023

I am responding to this RFI: On behalf of an organization

Name: Stefano Bertuzzi

Name of Organization: American Society for Microbiology

Type of Organization: Professional org association

Role: Institutional official

- 1. How to best ensure equity in publication opportunities for NIH-supported investigators.
- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

#### **Uploaded File:**

ASM-Response-to-NIH-RFI-on-Public-Access-to-Publications\_April-2023.pdf

**Description:** The American Society for Microbiology (ASM) appreciates the opportunity to respond to the National Institutes of Health's (NIH) request for information on the agency's plan to enhance public access to results of NIH-funded research. As one of the oldest a

Email: mwatts@asmusa.org



April 24, 2023

Dr. Larry Tabak National Institutes of Health 9000 Rockville Pike Bethesda, MD 20892

# Response to NOT-OD-23-091: Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-funded Research

Dear Dr. Tabak:

The American Society for Microbiology (ASM) appreciates the opportunity to respond to the National Institutes of Health's (NIH) request for information on the agency's plan to enhance public access to results of NIH-funded research. As one of the oldest and largest life science societies with more than 30,000 members in the United States and around the globe, our mission is to promote and advance the microbial sciences. ASM has a long-standing commitment to equity in science and recognizes that making research more widely accessible is a step in that direction. ASM has been an open access (OA) leader and advocate for many years and supports the fundamental principles of open science.

ASM's fifteen peer-reviewed journals, six of which are fully Gold open access and all of which publish open access content, are fundamental to ASM's mission and provide a critical service not only to our members, but also to the advancement of the microbial sciences globally. As you consider input on this public access plan, we stand ready to work with you to ensure a thoughtful, balanced approach. In the spirit of open science and open access, we have embarked on a journey to transform our publication business model to allow this important transition to happen.

#### **General Comments**

Given the scope, size, unique and indispensable function of the nonprofit, scientific society ecosystem in the United States, policy changes need to be made in a transparent, flexible and stepwise fashion. This is important to avoid unintended consequences that could result in reduced access to quality published scientific research provided by non-profit scientific societies in the United States. A key strength of our community is the diversity of its publishing operations. Each organization must be afforded the opportunity to find its own path forward and have the flexibility to adjust its business model to accommodate OA accordingly. We appreciate that your plan states your commitment to business model flexibility, and we expand more on this and other concepts below.

In addition to embracing the concept of OA, ASM believes that data availability and data sharing are critical to our mission to advance the microbial sciences. In October 2019, we expanded our data policy to be more comprehensive and to apply across all of our journals, not just those that are open access. For over three years now, authors have been required to make their data publicly available (except in rare circumstances) in order to publish in any ASM journal, preferably by depositing it in publicly accessible, curated and sustainable data repositories. While our policy has not been implemented without challenges, we believe the open data policy benefits both authors and readers in the long run.

Below are more specific answers to the topics outlined in the RFI.

#### How to best ensure equity in publication opportunities for NIH-supported investigators

ASM commends NIH for its commitment to equity and its work to achieve equitable access to publishing and to research in setting forth this plan. ASM continues to be concerned about the unintended consequences of enacting federal policies that might shift costs to researchers, or otherwise result in significant additional costs related to publication, repositories, data management, and staffing. There remains a cost to publishing good science. Peer reviewers are not paid, and maintaining the peer review system, which is integral to upholding scientific integrity and rigor, demands human time and adoption of innovative technologies. These standards and the integrity and rigor they uphold in science should not be compromised in the pursuit of public access.

As the costs of publishing increase, the costs are increasingly falling on individual researchers and institutions. We believe that if researchers are forced to make up for lost subscription revenue, a new kind of inequity will result. For example, author processing charges (APCs), which have evolved as an open access alternative revenue model to page charges and library subscriptions, have become increasingly expensive and created financial barriers for researchers from underserved populations, including early career researchers, those from historically excluded backgrounds, those at less research-intensive institutions and with limited resources and those living in the global south. We know that the scientific community does not want to disadvantage our colleagues in these institutions and countries by this approach.

It is critical that NIH policies support alternatives means for funding public access. We urge you to work with Congress and the research community to identify appropriate financial support to address these unequal additional burdens in future spending bills and through other strategies. Investing in infrastructure and services that are directly aligned with the research mission will be critical to laying the foundation for a more open and equitable system. We are pleased to see NIH's commitment to convening the community throughout the process to work through

<sup>&</sup>lt;sup>1</sup> https://journals.asm.org/content/open-data-policy? ga=2.31103164.223548841.1577910900-1577609744.1550589292

these challenges and share ideas, and we look forward to participating in these discussions. For example, to address equity concerns ASM is exploring a novel publication business model called "Subscribe to Open." We envision this model will include incentives for our customers to invest in ASM as a publisher and a partner in curating and disseminating science in accordance with open access policies.

#### Methods for monitoring evolving costs and impacts on affected communities

We thank NIH for taking the approach of establishing a public access policy and not a publishing policy. When it comes to scientific publishing, one size does not fit all, and it is imperative that federal policies respect the business of publishing and allow for each organization to take an approach that works for its members and customers. We appreciate the need to monitor the costs as publishing entities move to public access models given it is a new policy; however, we urge you to proceed with caution when evaluating "reasonable cost." For example, monitoring publication fees could result in a system that favors quantity over quality. There should remain bright lines between public access policies and publisher business models, and transparency in pricing should not be confused with transparency in business operations.

#### Conclusion

The American Society for Microbiology thanks the NIH for the opportunity to provide input on this important issue, and we encourage you to continue to engage stakeholders as you move from the plan to policy over the next two years. We appreciate your willingness to consider additional convenings with the community to work through these complex issues. Through open communication and by working together, ASM is confident that we can move boldly toward a world of open science, while also preserving the critical organizational infrastructure, including the financial health of nonprofit scientific societies, that has been foundational to publishing research in this country.

ASM and its members look forward to continuing to work with you as NIH implements this new public in this endeavor. For more information, please contact Allen Segal, ASM Director of Public Policy and Advocacy, at <a href="mailto:asegal@asmusa.org">asegal@asmusa.org</a> or 202-942-9294.

Sincerely,

Stefano Bertuzzi, PhD

SHP Pt.

**ASM Chief Executive Officer** 

**Submit date: 4/24/2023** 

I am responding to this RFI: On behalf of an organization

Name: Katherine Eve

Name of Organization: Elsevier

Type of Organization: Other

Type of Organization-Other: Publisher

Role: Member of the public

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Elsevier shares the White House Office of Science and Technology Policy's (OSTP's) and NIH's goals of ensuring the wide availability of trustworthy and impactful research findings, as well as equity in publication opportunities for NIH-supported investigators. At Elsevier, we look forward to working collaboratively with NIH and Other key stakeholders to achieve these goals principally via the gold open access model. With support from NIH, we believe this will best ensure equity in publication opportunities for all.

We recognize that there is currently no 'one-size fits all' publication model that meets all the diverse needs, preferences and circumstances of authors, institutions, funders in the US or indeed globally. This is why we have long offered both the gold open access, or pay-to-publish, model as well as the subscription, or pay-to-read, model, so that institutions and authors can choose the right route for them depending on their funding environment, discipline, and research goals. We therefore respect - and generally reflect - NIH's agnostic stance in its draft policy as to publication model, we understand the need for choice, and we support free market dynamics to sustainably achieve shared objectives on public access.

Consistent with the above principles, we agree that publicly funded research outputs should be publicly accessible. We fully support and enable researchers to freely and immediately share research outputs that have not benefitted from publishers' investments - for example, datasets and preprints. Where, under the terms of NIH's draft public access policy, researchers will be required by NIH to make peer-reviewed article versions immediately available, and asked to retain copyright, we will enable this through the gold open access (pay-to-publish) model.

Gold open access is a well-established and sustainable mechanism that ensures publishers are recompensed for the substantial value-added investments they make in these versions. These cover services that we and Other publishers provide, which include ensuring the quality, discoverability, and accessibility of research in perpetuity, safeguarding the integrity of published research by effectively managing editorial and peer review processes, and applying innovative technology towards continually expanding and enhancing all these services. Additionally, Elsevier is increasingly playing a critical role in tackling misinformation and fraud of unprecedented scale in science, as we validate the rigor of the research we publish in our journals. Sustainable funding models are vital if publishers are to continue providing these services to safeguard trust in science into the future, and for us to reinvest and innovate in a range of areas - including the examples related to equity outlined under question 2 - to advance

knowledge for society in the long term. Commensurately, we are committed to providing researchers with value for money in relation to our services, and to pricing fairly and transparently - themes we explore further under question 3.

We are supportive of choice and flexibility. Different publishers will provide different choices, services and business models. We will be unable to support publication models which rely on subscription-funded content being made freely and immediately accessible, and which also include requirements for authors to retain copyright via 'rights-retention'-like strategies, as we believe these models will prove unsustainable in the long-term. These measures do not provide a mechanism to recover our investments that enable us to continue innovating and ultimately providing value for NIH and the public. This position is shared by the vast majority of journals and publishers [Ref: <a href="https://www.stm-assoc.org/stm-oa-position/">https://www.stm-assoc.org/stm-oa-position/</a>].

We therefore welcome that NIH's draft policy enables researchers to charge reasonable costs for publishing gold open access against their awards. Availability and take up of this funding will be critical for grantees to be able to comply with NIH's immediacy policy across the full spectrum of available journals, so they are supported to publish in the journal that will provide the best visibility for their research. This will meet NIH's goal for equitable publication opportunities: without funding, grantees seeking to comply with NIH's policy would only be able to publish in journals that allow researchers to immediately share research they publish under the subscription model (just 4% of Health Science journals according to recent research published by JISC [Ref:

https://research.jiscinvolve.org/wp/2022/12/14/subject-analysis-of-routes-to-compliance-for-ukri-funded-authors/]), or that offer free open access publishing (8% of total articles across all subject areas were published in diamond journals based on 2021 Scopus data), which may be lower quality journals and regionally or institutionally focused titles.

Furthermore, to ensure equity in publication opportunities for all NIH-supported investigators, we suggest that all grantees should be provided with clear and consistent guidance on budgeting for the full cost of disseminating their research, and funds for publication should remain available after the end of the grant period. In so doing, all grantees will be afforded the same benefits of gold open access, including increased readership to maximize the reach of their work, a policy goal shared by NIH and OSTP.

There is much we can learn from Other markets as we work together to achieve the goals of OSTP's memo for immediate open access. The gold open access model is already widely adopted by the research community and successfully implemented across various countries [Ref: <a href="https://www.stm-assoc.org/oa-dashboard/open-access-uptake-for-the-top-30-article-producing-countries-and-Other-geographical-groupings/">https://www.stm-assoc.org/oa-dashboard/open-access-uptake-for-the-top-30-article-producing-countries-and-Other-geographical-groupings/</a>]. These include research-intensive countries such as the UK, Germany, Italy, Spain, Poland, and the Netherlands, where so-called 'combined' or 'read and publish' agreements with publishers have contributed to achieving immediate access to research through gold open access. All stakeholders have a role to play in developing solutions to enable gold open access in practice. At Elsevier, we draw on our experiences of co-creating agreements that already enable gold open access publishing across more than 2,100 institutions globally.

Finally, consistent with our commitment to evolving publishing practices, we welcome innovation in the marketplace. As publishers, we will continue to analyze, monitor and experiment with different publication models to ensure that we are serving our diverse communities as effectively as possible.

Notwithstanding, we have a responsibility to science and society to ensure that any approaches we endorse safeguard integrity, quality, discoverability, and accessibility of research in perpetuity. Thinking pragmatically about already proven mechanisms to support the OSTP's and NIH's policy goals for immediate public access, and acknowledging the limited time available to develop scalable solutions, at Elsevier we will therefore support the gold open access, pay-to-publish, model. This does not preclude continued experimentation to understand different publication models, or flexibility to test new models over time, in line with our long-standing tradition of working creatively with and for the scientific community to advance scientific knowledge sharing for the benefit of science and society.

Finding a solution that meets all OSTP's policy objectives, including equity, requires a collaborative and cooperative approach. We are committed to working with the research community, including NIH, towards finding workable solutions that will achieve these objectives for all.

#### 2. Steps for improving equity in access and accessibility of publications.

We share OSTP's Equity and Excellence Vision and would welcome opportunities to collaborate with NIH and Others in the research community to leverage our equity work and to exchange insights and experience, towards our shared ambitions for helping both individuals and science to achieve their full potential.

As outlined in our response to point 1, at Elsevier, we will enable NIH's grantees to meet its policy goals, and fully support equity in access, by offering the gold open access model, which is a well-established mechanism to achieve access, integrity, and quality at scale. With measures in place to ensure consistent guidance for all grantees on budgeting for the full cost of gold open access publication, equity can be safeguarded.

As discussed above, there is currently no one-size-fits all model that will best resolve all issues in relation to equity. Ultimately, there are trade-offs to consider between equity in access and equity in the ability to publish. We've done much to address inequities in the pay-to-read or subscription model. We have a range of initiatives in place to provide access to subscription content, which are made possible in part through the revenue generated by our sustainable publishing models. These include: our participation in Research4Life through which we provide free or discounted reading and publishing to researchers in over 120 low- and middle-income countries [Ref: https://www.elsevier.com/about/corporateresponsibility/research4life]; providing free access to health-related articles for patients and caregivers and establishing dedicated emergency resource and information centers, most recently for the novel coronavirus (COVID-19) [Ref: https://www.elsevier.com/connect/coronavirus-information-center] and Mpox [Ref: https://www.elsevier.com/connect/monkeypox-information-center]; supporting authors to share their publications peer-to-peer [Ref: https://www.elsevier.com/authors/submit-yourpaper/sharing-and-promoting-your-article]; and supporting interlibrary loans. Unfortunately, an open access world presents new kinds of inequities, and we are now working to develop solutions to mitigate these. These include: our vast programs of waivers/discounts on publishing, where appropriate; our work with institutions to fairly and equitably transition costs for reading to publishing as part of commercial agreements so institutions can fund publishing; and our piloting of new commercial models to address issues of equity head on. By way of an example, our pilot with California Digital Library works to meet gaps in funding for publication fees in an equitable manner [Ref: https://www.elsevier.com/about/press-releases/corporate/University-of-california-and-elsevier-signground-breaking-transformative-agreement].

We understand that mitigating inequities in the global research community requires that we look beyond publication models, and we have therefore undertaken a range of actions to identify issues and develop solutions towards equity in research. We have done this both as an individual publisher [Ref: <a href="https://www.elsevier.com/about/inclusion-and-diversity">https://www.elsevier.com/about/inclusion-and-diversity</a>], supported by our I&D Advisory Board [Ref: <a href="https://www.elsevier.com/about/inclusion-diversity-board">https://www.elsevier.com/about/inclusion-diversity-board</a>], and as a sector via the Joint Commitment for Action on Inclusion and Diversity in Publishing [Ref:

https://scholarlykitchen.sspnet.org/2021/04/21/joint-commitment-for-action-on-inclusion-and-diversity-in-publishing-an-interview-with-laura-norton-and-nicola-nugent-of-the-rsc/]. We work with our editors and reviewers, and the broader publishing community, to nurture inclusion and diversity, to widen participation in journals at all levels, and to ensure that researchers' work is assessed fairly on its scientific merits. We also employ innovative approaches, such as Registered Reports and Results Masked Review, to ensure research is judged on the merits of the research question and methodology. This aims to minimize the risk of publishing bias and supports accessibility to all federally funded research output, not only that which delivers a positive result.

With regards to accessibility, our accessibility policy ensures that we consistently and proactively endeavor to make our products fully accessible to all users, regardless of physical abilities [Ref: <a href="https://www.elsevier.com/about/accessibility">https://www.elsevier.com/about/accessibility</a>]. We are thrilled that the 2023 WebAIM million report ranks ScienceDirect as the #1 most accessible home page on the internet, ensuring an optimized experience for individuals with disabilities and impairments [Ref:

https://webaim.org/projects/million/lookup?domain=sciencedirect.com]. This incredible achievement is supported in part through the insights from a collaborative working group we have convened since 2011 comprising University leaders in assistive technology and web accessibility from six US institutes [Ref: https://www.elsevier.com/solutions/sciencedirect/librarian-resource-center/web-accessibility].

We are also proud to support health equity, and as part of recent additions to our 3D platform, Complete Anatomy [Ref: <a href="https://www.elsevier.com/solutions/complete-anatomy">https://www.elsevier.com/solutions/complete-anatomy</a>], we have introduced a full female model, and a range of skin tones and facial feature options. These enhancements allow educators to visualize, edit and teach anatomy from diverse perspectives.

These activities all require substantial investments. At Elsevier, we will continue to make a wide range of research outputs more accessible to a greater group of potential readers, to help researchers' work achieve the greatest impact, and to help advance research progress and efficiency so that funders such as NIH can maximize the value of their investment in research. We would welcome discussing these ideas and collaborating on further initiatives with NIH regarding both accessibility and initiatives or models for equitable access to content and publishing.

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

Feedback from researchers demonstrates that they value the publishing process and feel that the work we do has a material impact [Ref: <a href="https://www.elsevier.com/connect/how-scientific-publishing-supports-research-what-authors-are-telling-us">https://www.elsevier.com/connect/how-scientific-publishing-supports-research-what-authors-are-telling-us</a>]. We are heartened that 90% of researchers tell us the changes made by our journals' teams to their articles improved the clarity of their research. We want to continue to serve the research community by maintaining and building on this work, which is why we

will continue to seek researchers' input on how we can improve our services and their experiences with us.

We strive to offer researchers real value, and we are continuing our commitment to pricing our journals competitively with an underlying principle of pricing lower than the market for like-for-like quality.

Moreover, we follow this pricing principle even though our commitment to quality means we must invest resources to assess many more articles than we eventually publish. Elsevier journal articles account for around 18% of global research output and 28% of citations, further demonstrating our commitment to quality, significantly ahead of the industry average. We further recognize the importance of providing the research community with transparent and straightforward information about our journals and pricing on our public-facing pages, to help them make data-led decisions. As a responsible business we take care to ensure we work within the parameters permitted by law, and to a degree that avoids market alignment, that would Otherwise risk disadvantaging customers.

Key demonstrations of this commitment include:

- Our pricing policy page, covering the components that factor into our pricing, details of our strict no double dipping policy, and links to our subscription and APC list prices [Ref: <a href="https://www.elsevier.com/about/policies/pricing">https://www.elsevier.com/about/policies/pricing</a>].
- Sharing journal-level metrics for many of our journals, including acceptance rates, and average review and publication times, via Journal Insights pages (e.g., <a href="https://journalinsights.elsevier.com/journals/1072-7515">https://journalinsights.elsevier.com/journals/1072-7515</a>) and our Journal Finder tool [Ref: <a href="https://journalfinder.elsevier.com">https://journalfinder.elsevier.com</a>].
- Analysis of our publishing volumes under subscription and open access business models for individual journals (e.g., <a href="https://journalinsights.elsevier.com/journals/0021-9991">https://journalinsights.elsevier.com/journals/0021-9991</a>) and the whole of Elsevier [Ref: <a href="https://www.elsevier.com/data/assets/pdf">https://www.elsevier.com/data/assets/pdf</a> file/0010/616474/elsevier-journal-and-article-ecosystem-2021-summary.pdf].

We hold ourselves accountable for continuing to build on this transparency across the more than 2,800 journals we publish. We welcome views and will continue to ask for feedback from the research community, including partners such as NIH, as we enhance this offering, to provide helpful and meaningful insights to the communities that we serve.

#### 4. Early input on considerations to increase findability and transparency of research.

We support NIH's goals to increase the discoverability and transparency of research. Below are examples of platforms and initiatives that we provide to enable these. We welcome further dialogue and collaboration with partners in the research community, including NIH, to continue to build on this work.

Example 1: Improving research discovery via our ScienceDirect platform

All the content Elsevier publishes, including both journals and books content, is hosted on the ScienceDirect platform [Ref: <a href="https://www.sciencedirect.com/">https://www.sciencedirect.com/</a>]. ScienceDirect is completely free to search and browse in a number of ways; it serves around 50 million unique monthly users of which over 60% are not institutional customers, demonstrating that its use extends far beyond subscribers. Key

elements of an article published under the pay-to-read model are available to all readers irrespective of their access status e.g., the abstract and reference list. The introduction and 'section snippets' are in the process of being rolled out across all articles. All readers are further signposted to related relevant articles to help them continue their search and deepen their understanding of a particular topic. Furthermore, our dedicated Topic pages support researchers with gaining easily digestible introductions to new subjects, drawing from subject matter expert insights, and content highlights from our foundational resources [Ref: https://www.sciencedirect.com/topics].

#### Example 2: Enabling and encouraging transparent research data sharing

Transparent sharing of the data underlying research output enables research to be validated, supporting the quality and integrity of research. Data sharing also promotes greater reuse of research outputs, supporting research efficiency, reproducibility and maximizing the value of funders' investments by avoiding duplication of efforts and engendering new discoveries and research developments beyond the scope of the original study. This ultimately brings benefits for wider society and helps build trust in science.

We are committed to collaborating with stakeholders from across the research community, and to playing our role in enhancing data sharing practices to support and enable researchers and institutions to store, share, discover and effectively (re-)use data. At Elsevier we provide infrastructure and workflows in support of this: our research data management solutions support the end-to-end research data management workflow [Ref: <a href="http://www.elsevier.com/rdm">http://www.elsevier.com/rdm</a>], from providing Mendeley Data, an NIH Generalist Repository Ecosystem Initiative (GREI) supported open and free generalist repository [Ref: <a href="https://datascience.nih.gov/news/nih-office-of-data-science-strategy-announces-new-initiative-to-improve-data-access">https://datascience.nih.gov/news/nih-office-of-data-science-strategy-announces-new-initiative-to-improve-data-access</a>], to Data Monitor, which enables institutions, and ultimately funders, to track and monitor compliance with data sharing mandates. During our submission process we prompt and enable authors to share links to their datasets, made available in a repository of their choice, and to provide data availability statements in their publication.

#### Example 3: Surfacing metadata fields and persistent identifiers

Elsevier surfaces metadata fields and persistent identifiers (PIDs) to support discoverability, access, and compliance monitoring by research institutes and funders. We are actively participating in community discussions and initiatives on these topics, such as those led by the Open Research Funders Group. We would welcome further discussion with NIH and Other stakeholders on ways to improve on discoverability and transparency of research.

We already open a number of metadata fields for articles and their references within Crossref. In terms of identifiers, we use industry standards, such as article DOI and Fundref, and where there are a range of identifiers in use across the industry, we enable interoperability, for example, users can import their Scopus profiles into ORCID or link ORCID identifiers to Scopus profiles.

#### Example 4: Nurturing research integrity

The OSTP memo pointed to the role that metadata and PIDs can play in nurturing research integrity. We thus wish to highlight the broader role that publishers, including Elsevier, and learned societies play to ensure research integrity throughout all stages of submission and publication so that researchers and readers are assured of the quality and trustworthiness of research outputs. We do this by: screening

submissions for integrity issues; carefully managing the editorial and peer review process; supporting authors to develop and share transparency statements which are published alongside the published manuscript; and maintaining the integrity of the scholarly record through post-publication updates. We develop screening tools ourselves, as well as contribute to industry-wide approaches to nurture research integrity, for example via the STM Association's Integrity Hub [Ref: <a href="https://www.stm-assoc.org/stm-integrity-hub/">https://www.stm-assoc.org/stm-integrity-hub/</a>].

As you would expect, in all these aspects we seek to maintain the highest industry standards and best practice, as developed and maintained by the Committee on Publication Ethics (COPE), International Committee of Medical Journal Editors (ICMJE) and the like. We are keen to share our learnings and would welcome further dialogue with NIH and stakeholders regarding transparency and integrity of research.

#### **Uploaded File:**

Elseviers-Response\_Request-for-Information-on-the-NIH-Plan-to-Enhance-Public-Access-to-the-Results-of-NIH-Supported-Research.pdf

**Description:** Elsevier's Response: Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research

Email: k.eve@elsevier.com

### Elsevier's Response: Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research

#### 24 April 2023

Elsevier helps researchers and healthcare professionals to advance science and improve health outcomes by combining quality information and data sets with analytical tools to facilitate insights and critical decision-making. We enable researchers to disseminate their scientific findings through our more than 2,800 journals, applying tools and services, as well as coordinating editorial and peer review assessment in collaboration with 32,000 editors and almost 1.4 million reviewers each year, to ensure the integrity and quality of the research we publish.

Elsevier welcomes the opportunity to work alongside the White House Office of Science & Technology (OSTP), NIH and other federal agencies, and the research community, to advance open science, including through open access publishing. Elsevier seeks to advance the recommendations of the Public Access Memo issued by OSTP via mechanisms that are durable and sustainable for the entire research community. We endorse approaches that realize the clear benefits of widening public access while avoiding unintended consequences. Specifically, we applaud NIH's focus on equity and reiterate our willingness to share our experience and insights to support NIH with achieving equitable publishing opportunity and impact for all its grantees. Critically, we must work together to pursue models that safeguard the impact, quality, discoverability, and accessibility of research.

We appreciate your consideration of our comments at this critical juncture as NIH seeks to expand its public access policy, centered on our shared commitment to quality, trust and equity. We look forward to collaborating with NIH and other stakeholders as we lay the foundations for how to best address the emerging global societal challenges of our times.

# 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Elsevier shares the White House Office of Science and Technology Policy's (OSTP's) and NIH's goals of ensuring the wide availability of trustworthy and impactful research findings, as well as equity in publication opportunities for NIH-supported investigators. At Elsevier, we look forward to working collaboratively with NIH and other key stakeholders to achieve these goals principally via the gold open access model. With support from NIH, we believe this will best ensure equity in publication opportunities for all.

We recognize that there is currently no 'one-size fits all' publication model that meets all the diverse needs, preferences and circumstances of authors, institutions, funders in the US or indeed globally. This is why we have long offered both the gold open access, or pay-to-publish, model as well as the subscription, or pay-to-read, model, so that institutions and authors can choose the right route for them depending on their funding environment, discipline, and research goals. We therefore respect – and generally reflect – NIH's agnostic stance in its draft policy as to publication model, we understand the need for choice, and we support free market dynamics to sustainably achieve shared objectives on public access.

Consistent with the above principles, we agree that publicly funded research outputs should be publicly accessible. We fully support and enable researchers to freely and immediately share research outputs that have not benefitted from publishers' investments – for example, datasets and

preprints. Where, under the terms of NIH's draft public access policy, researchers will be required by NIH to make peer-reviewed article versions immediately available, and asked to retain copyright, we will enable this through the gold open access (pay-to-publish) model.

Gold open access is a well-established and sustainable mechanism that ensures publishers are recompensed for the substantial value-added investments they make in these versions. These cover services that we and other publishers provide, which include ensuring the quality, discoverability, and accessibility of research in perpetuity, safeguarding the integrity of published research by effectively managing editorial and peer review processes, and applying innovative technology towards continually expanding and enhancing all these services. Additionally, Elsevier is increasingly playing a critical role in tackling misinformation and fraud of unprecedented scale in science, as we validate the rigor of the research we publish in our journals. Sustainable funding models are vital if publishers are to continue providing these services to safeguard trust in science into the future, and for us to reinvest and innovate in a range of areas – including the examples related to equity outlined under question 2 – to advance knowledge for society in the long term. Commensurately, we are committed to providing researchers with value for money in relation to our services, and to pricing fairly and transparently – themes we explore further under question 3.

We are supportive of choice and flexibility. Different publishers will provide different choices, services and business models. We will be unable to support publication models which rely on subscription-funded content being made freely and immediately accessible, and which also include requirements for authors to retain copyright via 'rights-retention'-like strategies, as we believe these models will prove unsustainable in the long-term. These measures do not provide a mechanism to recover our investments that enable us to continue innovating and ultimately providing value for NIH and the public. This position is shared by the vast majority of journals and publishers.

We therefore welcome that NIH's draft policy enables researchers to charge reasonable costs for publishing gold open access against their awards. Availability and take up of this funding will be critical for grantees to be able to comply with NIH's immediacy policy across the full spectrum of available journals, so they are supported to publish in the journal that will provide the best visibility for their research. This will meet NIH's goal for equitable publication opportunities: without funding, grantees seeking to comply with NIH's policy would only be able to publish in journals that allow researchers to immediately share research they publish under the subscription model (just 4% of Health Science journals according to recent research published by JISC), or that offer free open access publishing (8% of total articles across all subject areas were published in diamond journals based on 2021 Scopus data), which may be lower quality journals and regionally or institutionally focused titles.

Furthermore, to ensure equity in publication opportunities for all NIH-supported investigators, we suggest that all grantees should be provided with clear and consistent guidance on budgeting for the full cost of disseminating their research, and funds for publication should remain available after the end of the grant period. In so doing, all grantees will be afforded the same benefits of gold open access, including increased readership to maximize the reach of their work, a policy goal shared by NIH and OSTP.

There is much we can learn from other markets as we work together to achieve the goals of OSTP's memo for immediate open access. The gold open access model is already widely adopted by the research community and <u>successfully implemented across various countries</u>. These include research-intensive countries such as the UK, Germany, Italy, Spain, Poland, and the Netherlands, where so-called 'combined' or 'read and publish' agreements with publishers have contributed to achieving

immediate access to research through gold open access. All stakeholders have a role to play in developing solutions to enable gold open access in practice. At Elsevier, we draw on our experiences of co-creating agreements that already enable gold open access publishing across more than 2,100 institutions globally.

Finally, consistent with our commitment to evolving publishing practices, we welcome innovation in the marketplace. As publishers, we will continue to analyze, monitor and experiment with different publication models to ensure that we are serving our diverse communities as effectively as possible. Notwithstanding, we have a responsibility to science and society to ensure that any approaches we endorse safeguard integrity, quality, discoverability, and accessibility of research in perpetuity. Thinking pragmatically about already proven mechanisms to support the OSTP's and NIH's policy goals for immediate public access, and acknowledging the limited time available to develop scalable solutions, at Elsevier we will therefore support the gold open access, pay-to-publish, model. This does not preclude continued experimentation to understand different publication models, or flexibility to test new models over time, in line with our long-standing tradition of working creatively with and for the scientific community to advance scientific knowledge sharing for the benefit of science and society.

Finding a solution that meets all OSTP's policy objectives, including equity, requires a collaborative and cooperative approach. We are committed to working with the research community, including NIH, towards finding workable solutions that will achieve these objectives for all.

#### 2. Steps for improving equity in access and accessibility of publications.

We share OSTP's Equity and Excellence Vision and would welcome opportunities to collaborate with NIH and others in the research community to leverage our equity work and to exchange insights and experience, towards our shared ambitions for helping both individuals and science to achieve their full potential.

As outlined in our response to point 1, at Elsevier, we will enable NIH's grantees to meet its policy goals, and fully support equity in access, by offering the gold open access model, which is a well-established mechanism to achieve access, integrity, and quality at scale. With measures in place to ensure consistent guidance for all grantees on budgeting for the full cost of gold open access publication, equity can be safeguarded.

As discussed above, there is currently no one-size-fits all model that will best resolve all issues in relation to equity. Ultimately, there are trade-offs to consider between equity in access and equity in the ability to publish. We've done much to address inequities in the pay-to-read or subscription model. We have a range of initiatives in place to provide access to subscription content, which are made possible in part through the revenue generated by our sustainable publishing models. These include: our participation in Research4Life through which we provide free or discounted reading and publishing to researchers in over 120 low- and middle-income countries; providing free access to health-related articles for patients and caregivers and establishing dedicated emergency resource and information centers, most recently for the <u>novel coronavirus</u> (COVID-19) and <u>Mpox</u>; supporting authors to share their publications peer-to-peer; and supporting interlibrary loans. Unfortunately, an open access world presents new kinds of inequities, and we are now working to develop solutions to mitigate these. These include: our vast programs of waivers/discounts on publishing, where appropriate; our work with institutions to fairly and equitably transition costs for reading to publishing as part of commercial agreements so institutions can fund publishing; and our piloting of new commercial models to address issues of equity head on. By way of an example, our pilot with California Digital Library works to meet gaps in funding for publication fees in an equitable manner.

We understand that mitigating inequities in the global research community requires that we look beyond publication models, and we have therefore undertaken a range of actions to identify issues and develop solutions towards equity in research. We have done this both as an <u>individual publisher</u>, supported by our <u>I&D Advisory Board</u>, and as a sector via the <u>Joint Commitment for Action on Inclusion and Diversity in Publishing</u>. We work with our editors and reviewers, and the broader publishing community, to nurture inclusion and diversity, to widen participation in journals at all levels, and to ensure that researchers' work is assessed fairly on its scientific merits. We also employ innovative approaches, such as Registered Reports and Results Masked Review, to ensure research is judged on the merits of the research question and methodology. This aims to minimize the risk of publishing bias and supports accessibility to all federally funded research output, not only that which delivers a positive result.

With regards to accessibility, our <u>accessibility policy</u> ensures that we consistently and proactively endeavor to make our products fully accessible to all users, regardless of physical abilities. We are thrilled that the <u>2023 WebAIM million report</u> ranks ScienceDirect as the #1 most accessible home page on the internet, ensuring an optimized experience for individuals with disabilities and impairments. This incredible achievement is supported in part through the insights from a <u>collaborative working group</u> we have convened since 2011 comprising university leaders in assistive technology and web accessibility from six US institutes.

We are also proud to support health equity, and as part of recent additions to our 3D platform, <a href="Complete Anatomy">Complete Anatomy</a>, we have introduced a full female model, and a range of skin tones and facial feature options. These enhancements allow educators to visualize, edit and teach anatomy from diverse perspectives.

These activities all require substantial investments. At Elsevier, we will continue to make a wide range of research outputs more accessible to a greater group of potential readers, to help researchers' work achieve the greatest impact, and to help advance research progress and efficiency so that funders such as NIH can maximize the value of their investment in research. We would welcome discussing these ideas and collaborating on further initiatives with NIH regarding both accessibility and initiatives or models for equitable access to content and publishing.

3. Methods for monitoring evolving costs and impacts on affected communities. Feedback from researchers demonstrates that they value the publishing process and feel that the work we do has a material impact. We are heartened that 90% of researchers tell us the changes made by our journals' teams to their articles improved the clarity of their research. We want to continue to serve the research community by maintaining and building on this work, which is why we will continue to seek researchers' input on how we can improve our services and their experiences with us.

We strive to offer researchers real value, and we are continuing our commitment to pricing our journals competitively with an underlying principle of pricing lower than the market for like-for-like quality.

Moreover, we follow this pricing principle even though our commitment to quality means we must invest resources to assess many more articles than we eventually publish. Elsevier journal articles account for around 18% of global research output and 28% of citations, further demonstrating our commitment to quality, significantly ahead of the industry average. We further recognize the importance of providing the research community with transparent and straightforward information about our journals and pricing on our public-facing pages, to help them make data-led decisions. As

a responsible business we take care to ensure we work within the parameters permitted by law, and to a degree that avoids market alignment, that would otherwise risk disadvantaging customers.

Key demonstrations of this commitment include:

- Our <u>pricing policy page</u>, covering the components that factor into our pricing, details of our strict no double dipping policy, and links to our subscription and APC list prices.
- Sharing journal-level metrics for many of our journals, including acceptance rates, and average review and publication times, via Journal Insights pages (<u>example</u>) and our <u>Journal Finder tool</u>.
- Analysis of our publishing volumes under subscription and open access business models for individual journals (<u>example</u>) and <u>the whole of Elsevier</u>.

We hold ourselves accountable for continuing to build on this transparency across the more than 2,800 journals we publish. We welcome views and will continue to ask for feedback from the research community, including partners such as NIH, as we enhance this offering, to provide helpful and meaningful insights to the communities that we serve.

#### 4. Early input on considerations to increase findability and transparency of research.

We support NIH's goals to increase the discoverability and transparency of research. Below are examples of platforms and initiatives that we provide to enable these. We welcome further dialogue and collaboration with partners in the research community, including NIH, to continue to build on this work.

#### Improving research discovery via our ScienceDirect platform

All the content Elsevier publishes, including both journals and books content, is hosted on the ScienceDirect platform. ScienceDirect is completely free to search and browse in a number of ways; it serves around 50 million unique monthly users of which over 60% are not institutional customers, demonstrating that its use extends far beyond subscribers. Key elements of an article published under the pay-to-read model are available to all readers irrespective of their access status e.g., the abstract and reference list. The introduction and 'section snippets' are in the process of being rolled out across all articles. All readers are further signposted to related relevant articles to help them continue their search and deepen their understanding of a particular topic. Furthermore, our dedicated Topic pages support researchers with gaining easily digestible introductions to new subjects, drawing from subject matter expert insights, and content highlights from our foundational resources.

#### Enabling and encouraging transparent research data sharing

Transparent sharing of the data underlying research output enables research to be validated, supporting the quality and integrity of research. Data sharing also promotes greater reuse of research outputs, supporting research efficiency, reproducibility and maximizing the value of funders' investments by avoiding duplication of efforts and engendering new discoveries and research developments beyond the scope of the original study. This ultimately brings benefits for wider society and helps build trust in science.

We are committed to collaborating with stakeholders from across the research community, and to playing our role in enhancing data sharing practices to support and enable researchers and institutions to store, share, discover and effectively (re-)use data. At Elsevier we provide infrastructure and workflows in support of this: our <u>research data management solutions</u> support the end-to-end research data management workflow, from providing Mendeley Data, an NIH

<u>Generalist Repository Ecosystem Initiative (GREI)</u> supported open and free generalist repository, to Data Monitor, which enables institutions, and ultimately funders, to track and monitor compliance with data sharing mandates. During our submission process we prompt and enable authors to share links to their datasets, made available in a repository of their choice, and to provide data availability statements in their publication.

#### Surfacing metadata fields and persistent identifiers

Elsevier surfaces metadata fields and persistent identifiers (PIDs) to support discoverability, access, and compliance monitoring by research institutes and funders. We are actively participating in community discussions and initiatives on these topics, such as those led by the Open Research Funders Group. We would welcome further discussion with NIH and other stakeholders on ways to improve on discoverability and transparency of research.

We already open a number of metadata fields for articles and their references within Crossref. In terms of identifiers, we use industry standards, such as article DOI and Fundref, and where there are a range of identifiers in use across the industry, we enable interoperability, for example, users can import their Scopus profiles into ORCID or link ORCID identifiers to Scopus profiles.

#### *Nurturing research integrity*

The OSTP memo pointed to the role that metadata and PIDs can play in nurturing research integrity. We thus wish to highlight the broader role that publishers, including Elsevier, and learned societies play to ensure research integrity throughout all stages of submission and publication so that researchers and readers are assured of the quality and trustworthiness of research outputs. We do this by: screening submissions for integrity issues; carefully managing the editorial and peer review process; supporting authors to develop and share transparency statements which are published alongside the published manuscript; and maintaining the integrity of the scholarly record through post-publication updates. We develop screening tools ourselves, as well as contribute to industry-wide approaches to nurture research integrity, for example via the STM Association's Integrity Hub.

As you would expect, in all these aspects we seek to maintain the highest industry standards and best practice, as developed and maintained by the Committee on Publication Ethics (COPE), International Committee of Medical Journal Editors (ICMJE) and the like. We are keen to share our learnings and would welcome further dialogue with NIH and stakeholders regarding transparency and integrity of research.

Victoria Eva SVP Global Policy and Industry Relations Elsevier Limited v.eva@elsevier.com **Submit date: 4/24/2023** 

I am responding to this RFI: On behalf of an organization

Name: William B. Coleman, PhD

Name of Organization: American Society for Investigative Pathology

Type of Organization: Professional orgassociation

Role: Institutional official

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

It is generally regarded that the NIH's public access policies represent an unfunded mandate that may impose financial hardship on many funded researchers. While NIH policy may allow "...reasonable publishing costs..." it should be recognized that researchers would rather preserve their funds for costs directly related to research projects than to use their funds for open access fees, which are more expensive than typical publication costs. Productive laboratories might publish numerous manuscripts in the course of a year and if these publications carry open access fees, could significantly impact modular budgets. Most researchers would advocate for the NIH to provide for publication costs in a manner that does not impact the budget that directly supports the funded research project (outside the modular budget). Further, the NIH should recognize that requirement to publish open access may force some researchers to publish less due to financial constraints. This represents an unintended consequence of the public access policy. When researchers are forced to choose what they publish (because publication of all research results would be cost prohibitive), their measures of productivity and their impact on the field decreases. This potential consequence of the public access policy would disproportionately affect young investigators who have less research funding and need to prioritize research productivity that reflects generation of results and publication of those results to build a successful research program.

Submission of published papers or final manuscripts to PMC is often accomplished by commercial publishers on behalf of the authors (which is the case for the ASIP journals), reducing the burden to the individual researcher. However, when this service is not provided by the publisher, the requirement does present a burden to the researcher. This burden could be diminished by allowing authors to deposit the pdf version of their final published paper rather than requiring upload to PMC of the manuscript's deconstructed component parts (necessary for PMC format).

The NIH should also be sensitive to the need of journals (and their publishers) to receive data on the numbers of times that their manuscripts are accessed and downloaded from PMC. These metrics are particularly important to non-profit societies that publish journals as these data contribute to the overall measures of their journal's value in the current era where impact factor (based upon citations) is only one dimension of the overall contribution the journal makes to the advancement of science.

#### 2. Steps for improving equity in access and accessibility of publications.

The 12-month embargo is absolutely necessary for the survival of journals that operate on a subscription or hybrid business model. At such time when research becomes immediately available upon publication, the need for institutions and individuals to subscribe to journals will disappear. This would likely force many quality journals that are not among the elite few out of business. That would result in fewer

respected journals in which to publish and a broadening of the impact gap between those who publish in Nature, Science, Cell, etc. and those who have to publish in journals with questionable review practices. The unintended consequence would be a dilution of good science and a widening of the gap between the "haves" and the "have-nots." Again, this would disproportionally affect junior investigators.

Elimination of the embargo period will force journals that primarily publish results from NIH-supported researchers to convert to 100% open access or would force authors to choose the open access option. While both of these scenarios are plausible, both would result in increased costs for publication that would be passed on to the researcher (discussed above). When NIH-funded researchers are required to publish all their work in open access journals or utilizing open access options (to comply with the elimination of the embargo period), the financial burden to individual investigators will increase, which would disproportionately affect young investigators and may negatively impact on the amount of research results that are published.

The NIH should also consider the current practice by many researchers of utilizing preprint servers and how this impacts public access to NIH-funded research. Servers such as arXiv, bioRxiv, and medRxiv, host preprints that reflect a large volume of biomedical research. Preprints are attractive to researchers because it allows them to establish a time-stamp on their work while the results are submitted for peer-reviewed publication. Preprints are accessible at no cost and may provide a more appropriate resource for the public, particularly non-scientists. While most journals do not allow citation of preprints, authors are allowed to submit the work contained in the preprint for publication since preprints by definition are not yet published. We note that the National Library of Medicine is running a pilot study to make preprints resulting from NIH-funded research available on PubMed Central (which includes all of the preprint servers listed above (https://www.ncbi.nlm.nih.gov/pmc/about/nihpreprints/)).

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

Publication costs vary considerably across journals and publishers. This is true for subscription-based journals, hybrid journals and open access journals. With many journals now being online-only (no print), there are cost savings, particularly in the area of color figures and page charges. However, the production of a journal issue (whether print or online) continues to require technology, expertise, and personnel, all of which contribute to the publications current cost structure for any given journal. Commercial publishers have the need to make a profit from their publications in order to continue to provide journals to disseminate research results. If/when journals become less profitable we may see some journals disappear - reducing the number of outlets for publication of research results. Hence, reasonable publication costs must be viewed from the perspective of the publisher, as well as from the perspective of the author and their funders.

We commend the NIH for committing to monitor the costs of publication and how this affects the laboratory finances of their funded researchers. We would encourage the NIH to make monitoring of publication costs a required reporting element of NIH progress reports. This would allow the NIH to assess the total cost of publication of NIH-funded research (collectively and on the basis of the individual researcher) and generate a database of publication costs by publisher, journal, journal-type, and nature of the published work. This reporting requirement would not significantly increase the burden on the investigator beyond what currently exists. NIH grantees routinely report their publications as evidence

of productivity, and the new reporting requirement would simply ask for investigators to disclose the cost of each publication.

#### 4. Early input on considerations to increase findability and transparency of research.

As researchers and publishers, we agree that the use of PIDs (or Other sorts of DOIs) is important and should be implemented to enhance transparency and discoverability of published research. We commend efforts on the part of the NIH to provide appropriate linkages between published research results, investigator/authors (utilizing ORCID IDs), and sources of research funding. In this manner, assigning a PID/DOI to funded research grants would benefit the connectivity and traceability of these elements of research (people-funding-results).

#### **Uploaded File:**

ASIP-RFI-NIH-Plan-to-Enhance-Public-Access-to-NIH-Supported-Research-4-24-2023-FINAL.pdf

**Description:** Letter containing response to RFI on public access.

Email: wbcoleman@asip.org



#### WILLIAM B. COLEMAN, PHD EXECUTIVE OFFICER

1801 Rockville Pike, Suite 350, Rockville, MD 20852 (USA)
Telephone: 240-283-9700 • Fax: 301-984-4047 • Email: <a href="wbcoleman@asip.org">wbcoleman@asip.org</a> • <a href="www.asip.org">www.asip.org</a>

April 20, 2023

Lyric Jorgenson, PhD Acting Director, Office of Science Policy Acting NIH Associate Director for Science Policy The National Institutes of Health 6705 Rockledge Drive, Suite 630 Bethesda, MD 20892

RE: Comments in Response to NOT-OD-23-091, Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research

Dear Dr. Jorgenson,

The American Society for Investigative Pathology (ASIP) appreciates the opportunity to provide comments in response to NOT-OD-23-091 – Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research. The ASIP is a non-profit society of biomedical scientists who investigate mechanisms of disease, many of whom direct academic research laboratories and are supported by NIH funding. The mission of the ASIP is to promote the discovery, advancement, and dissemination of basic and translational knowledge in experimental/investigative pathology and related disciplines. Experimental pathology is an integrative discipline that links the presentation of disease in the whole organism to its fundamental cellular and molecular mechanisms. It uses a variety of structural, functional, and genetic techniques and ultimately applies research findings to the diagnosis and treatment of diseases.

Dissemination of basic and translational knowledge in experimental/investigative pathology and related disciplines is critical to the disease-focused scientific endeavor. Discoveries that are not available to the scientific community do not contribute to the advancement of the field. Hence, dissemination and access to research results are essential, and can be accomplished in several different ways, including through publication and presentation at scientific meetings. Publication in a peer-reviewed journal represents the gold-standard for the scientific community. The ASIP publishes two scientific journals – *The American Journal of Pathology* and *The Journal of Molecular Diagnostics*. Both of these journals have options for open access publishing.

While the ASIP recognizes the value of increased access (or open access) to the results of NIH-sponsored research to the scientific community, we also understand that this is a complex issue with numerous stakeholders and that the ramifications of enhanced public access requirements may create problems that are antithetical to the intended policy. Hence, the NIH needs to carefully consider if furtherance of public access policies will have unintended negative consequences for researchers, journals, and publishers (including non-profit Societies that operate journals).

1. How to best ensure equity in publication opportunities for NIH-supported investigators.

The NIH Public Access Plan aims to maintain the existing broad discretion for researchers and authors to choose how and where to publish their results. Consistent with current practice, the NIH Public Access Plan allows the submission of final published articles to PubMed Central (PMC) (in cases where a formal agreement is in place) to minimize the compliance burden on NIH-supported researchers and also

maintains the flexibility of NIH-supported researchers to submit the final peer-reviewed manuscript. These submission routes are allowed regardless of whether or not the journal uses an open access model, a subscription model of publishing, or other publication model. This flexibility aims to protect against concerns that have been raised about certain publishing models potentially disadvantaging early career researchers and researchers from limited-resourced institutions or under-represented groups. NIH policy already allows supported researchers to charge reasonable publishing costs against their research grants.

NIH seeks information on additional steps it might consider taking to ensure that proposed changes to implementation of the NIH Public Access Policy do not create new inequities in publishing opportunities or reinforce existing ones.

It is generally regarded that the NIH's public access policies represent an unfunded mandate that may impose financial hardship on many funded researchers. While NIH policy may allow "...reasonable publishing costs..." it should be recognized that researchers would rather preserve their funds for costs directly related to research projects than to use their funds for open access fees, which are more expensive than typical publication costs. Productive laboratories might publish numerous manuscripts in the course of a year and if these publications carry open access fees, could significantly impact modular budgets. Most researchers would advocate for the NIH to provide for publication costs in a manner that does not impact the budget that directly supports the funded research project (outside the modular budget). Further, the NIH should recognize that requirement to publish open access may force some researchers to publish less due to financial constraints. This represents an unintended consequence of the public access policy. When researchers are forced to choose what they publish (because publication of all research results would be cost prohibitive), their measures of productivity and their impact on the field decreases. This potential consequence of the public access policy would disproportionately affect young investigators who have less research funding and need to prioritize research productivity that reflects generation of results and publication of those results to build a successful research program.

Submission of published papers or final manuscripts to PMC is often accomplished by commercial publishers on behalf of the authors (which is the case for the ASIP journals), reducing the burden to the individual researcher. However, when this service is not provided by the publisher, the requirement does present a burden to the researcher. This burden could be diminished by allowing authors to deposit the pdf version of their final published paper rather than requiring upload to PMC of the manuscript's deconstructed component parts (necessary for PMC format).

The NIH should also be sensitive to the need of journals (and their publishers) to receive data on the numbers of times that their manuscripts are accessed and downloaded from PMC. These metrics are particularly important to non-profit societies that publish journals as these data contribute to the overall measures of their journal's value in the current era where impact factor (based upon citations) is only one dimension of the overall contribution the journal makes to the advancement of science.

#### 2. Steps for improving equity in access and accessibility of publications.

Removal of the currently allowable 12-month embargo period for NIH-supported publications will improve access to these research products for all. As noted in the NIH Public Access Plan, NIH also plans to continue making articles available in human and machine-readable forms to support automated text processing. NIH will also seek ways to improve the accessibility of publications via assistive devices.

NIH welcomes input on other steps that could be taken to improve equity in access to publications by diverse communities of users, including researchers, clinicians and public health officials, students and educators, and other members of the public.

The 12-month embargo is absolutely necessary for the survival of journals that operate on a subscription or hybrid business model. At such time when research becomes immediately available upon publication, the need for institutions and individuals to subscribe to journals will disappear. This would likely force many quality journals that are not among the elite few out of business. That would result in fewer respected journals in which to publish and a broadening of the impact gap between those who publish in *Nature, Science, Cell,* etc. and those who have to publish in journals with questionable review practices. The unintended consequence would be a dilution of good science and a widening of the gap between the "haves" and the "have-nots." Again, this would disproportionally affect junior investigators.

Elimination of the embargo period will force journals that primarily publish results from NIH-supported researchers to convert to 100% open access or would force authors to choose the open access option. While both of these scenarios are plausible, both would result in increased costs for publication that would be passed on to the researcher (discussed above). When NIH-funded researchers are required to publish all their work in open access journals or utilizing open access options (to comply with the elimination of the embargo period), the financial burden to individual investigators will increase, which would disproportionately affect young investigators and may negatively impact on the amount of research results that are published.

The NIH should also consider the current practice by many researchers of utilizing preprint servers and how this impacts public access to NIH-funded research. Servers such as arXiv, bioRxiv, and medRxiv, host preprints that reflect a large volume of biomedical research. Preprints are attractive to researchers because it allows them to establish a time-stamp on their work while the results are submitted for peer-reviewed publication. Preprints are accessible at no cost and may provide a more appropriate resource for the public, particularly non-scientists. While most journals do not allow citation of preprints, authors are allowed to submit the work contained in the preprint for publication since preprints by definition are not yet published. We note that the National Library of Medicine is running a pilot study to make preprints resulting from NIH-funded research available on PubMed Central (which includes all of the preprint servers listed above (https://www.ncbi.nlm.nih.gov/pmc/about/nihpreprints/)).

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

NIH proposes to actively monitor trends in publication fees and policies to ensure that they remain reasonable and equitable.

# NIH seeks information on effective approaches for monitoring trends in publication fees and equity in publication opportunities.

Publication costs vary considerably across journals and publishers. This is true for subscription-based journals, hybrid journals and open access journals. With many journals now being online-only (no print), there are cost savings, particularly in the area of color figures and page charges. However, the production of a journal issue (whether print or online) continues to require technology, expertise, and personnel, all of which contribute to the publications current cost structure for any given journal. Commercial publishers have the need to make a profit from their publications in order to continue to provide journals to disseminate research results. If/when journals become less profitable we may see some journals disappear — reducing the number of outlets for publication of research results. Hence, reasonable publication costs must be viewed from the perspective of the publisher, as well as from the perspective of the author and their funders.

We commend the NIH for committing to monitor the costs of publication and how this affects the laboratory finances of their funded researchers. We would encourage the NIH to make monitoring of publication costs a required reporting element of NIH progress reports. This would allow the NIH to assess the total cost of publication of NIH-funded research (collectively and on the basis of the individual researcher) and generate a database of publication costs by publisher, journal, journal-type, and nature

of the published work. This reporting requirement would not significantly increase the burden on the investigator beyond what currently exists. NIH grantees routinely report their publications as evidence of productivity, and the new reporting requirement would simply ask for investigators to disclose the cost of each publication.

4. Early input on considerations to increase findability and transparency of research.

Section IV of the NIH Public Access Plan is a first step in developing the NIH's updated plan for persistent identifiers (PIDs) and metadata, which will be submitted to OSTP by December 31, 2024.

NIH seeks suggestions on any specific issues that should be considered in efforts to improve use of PIDs and metadata, including information about experiences institutions and researchers have had with adoption of different identifiers.

As researchers and publishers, we agree that the use of PIDs (or other sorts of DOIs) is important and should be implemented to enhance transparency and discoverability of published research. We commend efforts on the part of the NIH to provide appropriate linkages between published research results, investigator/authors (utilizing ORCID IDs), and sources of research funding. In this manner, assigning a PID/DOI to funded research grants would benefit the connectivity and traceability of these elements of research (people-funding-results).

The American Society of Investigative Pathology is grateful for the opportunity to provide feedback in response to this request for information. The input/feedback assembled here resulted from extensive discussions among the members of our Publications Committee and our Research and Science Policy Committee – many of whom are now NIH-funded researchers or have been NIH-funded researchers in the past. Our members approached these discussions from the perspective of the various stakeholders to provide a well-rounded response that might be of value to the NIH.

Sincerely,

William B. Coleman, PhD ASIP Executive Officer

Allhis. Ih

**Submit date: 4/24/2023** 

I am responding to this RFI: On behalf of an organization

Name: Caroline Sutton

Name of Organization: STM (The International Association of Scientific, Technical and Medical

Publishers)

Type of Organization: Professional org association

Role: Institutional official

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Changing access requirements within the scientific ecosystem are likely to solve inequities from a reader aspect, but concerted and collaborative action will be necessary to ensure sustainability and equity across the ecosystem. Agencies can minimize the risk of creating new inequities, especially for scientists from traditionally marginalized communities, as well as early career researchers, by ensuring that these researchers and institutions have the funding support necessary for their research to flourish and choose the publishing option that best suits their needs. Publishers are doing their part by supporting new approaches, including Read and Publish Agreements, that provide opportunities for all to participate and access scholarly communication. Ultimately, a financially sustainable scientific publishing system is critical to advance trusted and impactful science, and attention to these issues can ensure that this is achieved.

To promote publishing equity, NIH needs to make appropriate and enduring funding available to the researcher and their research institution, together with appropriate and enduring support and guidance on the use of funds and the options for providing access. In order to ensure equity for all researchers, such funding and guidance needs to be provided alongside Other guidance for researchers, and in a manner that ensures author choice for whatever journals they choose to advance their research and impact. This funding also needs to be provided on an equal basis so that researchers who choose to publish in journals that are supported by APCs are not disadvantaged in the resources available for their research, student support, and Other critical needs. Finally, NIH should provide clear and prominent guidance on planning and budgeting and the explicit acknowledgement throughout the guidance that publication has real costs that need to be addressed in the proposal, as it has with the NIH Data Sharing and Management requirements.

Agreements with institutions or funders like Read and Publish Agreements or Other pooled payment agreements have the potential to reduce inequality by making OA publishing available to all researchers. Publishers are actively working to develop and promote these models, which can reduce inequity for researchers at participating institutions and also can help increase compliance with policy and reduce administrative burdens. We have received reports of the success of such efforts, thanks to the real-world experiment of growth of transformative agreements around the world.

AnOther aspect of equity in publishing opportunities relates to the promotion of equity and diversity in the research enterprise. Support for diverse publishing outlets is critical to such efforts, although to proactively drive further change requires input from stakeholders across the research ecosystem. One way in which publishers encourage equity and diversity in the research enterprise is by providing an

objective space in which work can be assessed by peers (though our impartial oversight of an independent peer review process). More specifically, in recent years publishers have established industry-wide initiatives such as the Joint Commitment on Diversity and Inclusion and C4DISC which are developing consensus-based standards and best practice (e.g., developing guidelines around the peer review of articles and data; creating policies to support authors with deadnames; etc.).

Finally, publishers support and invest in various initiatives to enable researchers to participate in the scholarly dialogue. This includes support for educational efforts and funding programs that expand participation to underrepresented groups and ensure quality and integrity. For example, Research4Life, a UN-publisher partnership, supports researcher skill development, provides Research Lifecycle Training Webinars, and enhances the ability of LMIC researchers to publish with participating publishers. Many publishers support and partner with AuthorAID, a global network that provides free resources and training, including in article writing, for researchers in low- and middle-income countries. Publishers offer various funding programs to support the participation of less-well-resourced researchers, including discounts and waivers, both individually and through collective approaches like Research4Life. Publishers also work with Other stakeholders to provide resources to identify trusted outlets to present their work (e.g., Think. Check. Submit. (thinkchecksubmit.org) a cross-industry initiative) and promote integrity in scholarly research and its publication through the Committee on Publication Ethics (COPE, www.publicationethics.org) and Other efforts.

#### 2. Steps for improving equity in access and accessibility of publications.

Publishers invest significantly in efforts to provide access, accessible formats, and accessible modes of dissemination for publications. It is important to note that for access and accessibility to be provided, first the publications and infrastructures must be created and disseminated. Therefore, it is a necessary precondition to improve equity in access and accessibility of publication that NIH work to ensure the viability of a robust ecosystem of scholarly communications that drives innovation, supports quality and integrity, and ensures appropriate infrastructure to enable accessibility to diverse users.

As alluded to in the introduction to this response, steps to improve access and accessibility could be broken down into three requirements: 1) sufficient, enduring, and appropriate funding, 2) encouragement and education of researchers to budget for and choose open science, and 3) flexibility for researchers and organizations to enable diverse modes of communication.

Appropriate and enduring funding is fundamental to achieve the open science goals outlined in the draft NIH plan and in the August OSTP memo and make sure that NIH's revised policy can promote equity in access. This is because the sustainability of publishing is a precondition to the availability, utility, and accessibility functions of scholarly communications.

Encouragement and education of researchers is also key, as they will ultimately be responsible for ensuring that the articles that they write are available to the public. Experience with funder requirements and compliance around the world indicates that researchers are often confused about grant requirements, including on how and when to provide access to publications, and a significant percentage of researchers erroneously believe that it is an inappropriate use of grant funds to pay for publication. STM's members' experience with guidance and education indicates that such efforts can make a big difference in researchers willingness to choose open access and compliance with funder and Other requirements.

Flexibility is needed to promote diversity in publication, ensure author choice, and support access to publishing in ways that work for researchers. As noted earlier, different publishers may offer distinct approaches to provide access, each of which may be appropriate to the communities they serve, and each of which should be allowed as a method for researchers to ensure access to any article they author that reports on NIH-funded research. A diversity of publication outlets, enabled by flexible approaches to implementation of the NIH policy, supports diversity in research.

Publishers invest significantly to ensure that articles are accessible in various human and machine-readable formats and are available to those with diverse needs. Many publishers have invested in technology and infrastructure to build towards, meet, or exceed Section 508 accessibility and have created a diverse ecosystem of accessible resources available to diverse audiences with or without assistive technologies. Some of our members were leaders in developing braille resources in multiple languages, screen reading technology implementation, and Other innovations. These additional infrastructure and formatting investments are enabled by sustainable business models.

STM also notes various initiatives that we or our members have promoted to ensure access and accessibility for diverse audiences. These include Research4Life which provides access to researchers in Low- and Middle- Income countries; efforts to share plain language summaries to broaden the accessibility of cutting-edge research to non-experts; and investments in the promotion of articles to the media and through social media channels.

Finally, STM notes that equity in access requires that publications that are made available are accurate and trustworthy. STM and its members invest significantly in ensuring research integrity and the quality and reliability of the scholarly record. For example, STM Solutions recently launched the Research Integrity Hub (<a href="https://www.stm-assoc.org/stm-integrity-hub/">https://www.stm-assoc.org/stm-integrity-hub/</a>), a robust and holistic set of tools to safeguard the integrity of science through a combination of shared data and experiences and by harnessing technological innovation. Individual publishers are working individually and in partnership with Other organizations to prevent misconduct and ensure the integrity of the system. Safeguarding research integrity can only be done through collaboration with all stakeholders in the scholarly ecosystem, and in an environment where continued investments can be made.

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

STM's members compete in a dynamic environment that drives them to provide the widest possible access to the articles that they publish at the lowest possible cost to the research and user communities. Costs and revenue streams can vary significantly from one publisher to anOther, and even from one journal to anOther, depending on many factors such as audience, circulation/reach, ranking, number of articles published, field/specialty, and distribution method. These differences need to be considered when evaluating the market dynamics and taking a broad average of dissimilar journals is not recommended.

More broadly, it is important to consider the changing dynamics of how scholarly publication is supported when attempting to monitor trends. Historically, publishers' costs have been spread across those that consume the research (readers / subscribers) of which there are many. The NIH plan may move associated costs to Other payers, of which there are fewer. The cost burden will therefore increase for some (e.g., research-intensive universities) while many Others will no longer contribute to

the costs (e.g., commercial industries, which traditionally subscribe to journals without publishing extensively in them).

When considering the budget for supporting public access to high-quality, peer-reviewed articles reporting on NIH-funded research, it is important to look beyond a single aspect of pricing (i.e., APCs) and consider the total investment in scholarly communications, which includes subscriptions, APCs, transformative agreements, and Other inputs. The cost and pricing structures are very different for different disciplines - medicine, physical sciences, social sciences, and humanities - and for different types of journals based on selectivity, services, technology, and Other features.

That said, APC prices are virtually always transparent. Our members are committed to the maximum possible transparency around pricing, in accordance with regulation and antitrust concerns, and note that APCs may vary across journal titles based on a variety of factors. Our members are also committed to ensuring that every researcher - regardless of geographic location, discipline or personal circumstance has relevant and realistic options available to them to publish their work, so that no researcher is left without a voice, regardless of funding source. Consistent with this commitment publishers have developed Read and Publish Agreements with institutions and maintain active waiver and discount programs to serve researchers.

STM is not aware of any Other NIH efforts to monitor expenses for specific research services or outputs and cautions that any efforts to look at trends in publishing must be carefully interpreted in the context of an evolving and dynamic ecosystem. Those who monitor APC prices and perform market analysis are aware that any trends in this data always need to be contextualized with respect to Other trends in publishing (e.g., the growth in the sharing of research outputs) and revenue (e.g., subscription rates and transformative agreements) and with respect with efforts to ensure equity in publication opportunities (e.g., provision of waivers and discounts).

A diverse, financially sustainable, and robust publishing system which provides authors with broad choice is the most effective way to ensure fair and competitive pricing and address any cost concerns. Hard price caps will likely drive existing industry trends toward publisher consolidation and volume-based models which could compromise integrity, quality, and author choice. The research enterprise, and the impact of NIH-funded research on innovation and public health, is best served by diversity that is enabled by flexibility and full support for open access publishing options.

In addition, care must be taken with respect to interventions that seek to ensure fees and policies remain reasonable and equitable, as they may lead to unintended consequences or constitute anti-competitive market interference under antitrust laws. As STM and Others have recommended in Other contexts, NIH should seek legal advice regarding competition law and any undue influence on industry market pricing. Finally, we underline that the goals of the NIH policy are best achieved though NIH efforts to ensure that researchers are budgeting appropriately for publications.

#### 4. Early input on considerations to increase findability and transparency of research.

We divide our response into two sections, as the concepts and needs of findability and transparency, while interrelated, are also quite distinct.

a. Findability (including persistent identifiers (PIDs), metadata, and Other infrastructure).

STM and its member publishers would welcome collaboration with NIH to support approaches to findability that leverage and build on existing standards, technologies, infrastructure, and protocols. Publishers have committed to and invested significantly in ensuring the findability of articles and research data. Our experience suggests that additional efforts to support the use and development of persistent identifiers throughout the research ecosystem would bear additional fruit, including identifiers for articles and research data as well for funding agencies, grant awards, facilities, and the like.

Where possible, NIH should leverage existing standards and systems, as supported by publishers, institutions, and Other stakeholders. The primary existing PID and metadata structure, enabled through organizations including CrossRef and DataCite, should be adopted and adapted as necessary to minimize disruption, promote compliance, and prevent unnecessary duplication of effort and investment in the scholarly communications system.

Publishers already invest heavily in creating persistent identifiers and machine-readable metadata that promote greater visibility of research findings and data, and these help to promote trust, reliability, and transparency for the scientific system. Cross publisher and industry initiatives around PIDs include researcher (ORCID), institutional (Ringgold), and funder (Open Registry of Funders) PIDS embedded in our content workflows as standard across the majority of the scholarly communication ecosystem. Embedding standards supports our infrastructure development to build better links between interrelated research outputs and improve visibility from funding through to publication. In general, PIDs used or recommended by NIH should be those used by the community, as those can be validated and maintained. Where NIH needs additional or bespoke PIDs, efforts need to be made to ensure they map well to Other PIDs that are already well embedded in the ecosystem.

Specifically, STM recommends that NIH support the use of community-adopted PIDs through the grant application process (e.g., ORCIDs for researchers, organization IDs for the institutions(s) affiliated with each researcher, and Funder IDs for the distinct funders of the grant). While organization IDs are not as well-established or robust as researcher IDs (with ORCID), there are several emerging options for organizations, and NIH should consider recommending one of the following PIDs to ensure harmonization and avoid unnecessary duplication in the scholarly record: Ringgold (a global organization identifier system); ISNI (ISO standard name identifier system); ROR (the Research Organization Registry); and Crossref's Funder Registry; along with ORCID. NIH should also ensure there are metadata fields for all of these.

In addition, publishers have invested significantly in discoverability, search engine optimization, and Other efforts to make sure that published articles can be found and used to advance scientific research. To support the findability of both articles and research data, NIH should also engage with and implement community-based standards and infrastructure initiatives that link and promote access to the best available versions of articles and research data. These include open protocols like Scholix, a multi-stakeholder initiative to link scholarly literature and research data, and services like CHORUS, that helps the public find and access articles reporting on federally-funded research. Initiatives such as seamlessaccess.org, a service designed to help foster a more streamlined online access experience by leveraging an existing single-sign-on infrastructure, and GetFTR, a tool that streamlines access to journal articles on discovery tools and collaboration networks, are also available to enable and accelerate access. STM would welcome additional dialogue to discover which existing initiatives could best be

utilized to support findability and access to articles and research data related to NIH-funded research, and to collaboratively develop solutions where services or infrastructures do not already exist.

#### b. Transparency (including reproducibility and trust in science)

Findability is necessary to promote transparency, but it is not sufficient to enable it. Transparency needs to be fostered through education and the research culture and enabled by infrastructure. Publishers continually invest in such systems and infrastructure and promulgate policies that encourage open sharing to promote trust. This includes efforts to promote trust and transparency through the sharing of research data (e.g, STM's Research Data initiative ) and especially the use of FAIR (Findable, Accessible, Interoperable, and Reproducible) principles in sharing research data. Innovations in open peer review, the broadening of publishable articles to include negative results, the introduction of registered reports, and Other efforts to make publication and the publication process more transparent have the potential to improve public trust in science and the utility of research. Many of our members have signed on to Transparency and Openness Promotion (TOP) Guidelines and engaged with Other initiatives to drive transparency.

STM recommends that NIH leverage existing resources to promote transparency and avoid creating duplicative resources. For example, NIH can point to existing resources to support researchers in making their research outputs more transparent. Some potential examples include a manifesto for reproducible science designed to optimize key elements of the scientific process and "STAR Methods: Structured, Transparent, Accessible Reporting," designed to provide a structure for experimental methods that increases reproducibility. Existing, robust infrastructure should be considered before recommending or developing new systems.

We note that new modes of scientific inquiry are providing opportunities to improve scholarly practices, including with respect to transparency and integrity, but these may also carry risks that are not fully understood at this time. NIH's policies must be flexible enough to address any issues that might arise in these new modes of scholarship, as well as provide support for new and existing infrastructure and services that can help provide the review and analysis needed to ensure quality and integrity of both new and existing systems.

Finally, we note that the most important action that NIH can take to ensure transparency, quality and integrity in scholarly communication is to support and encourage the systems and services that currently provide these benefits for the research enterprise. These include, but are not limited to, market incentives that encourage the development of high-quality publication outlets for scholarly communication such as those produced by STM's members.

#### **Uploaded File:**

STM-submission-to-NIH-public-access-RFI-2023.pdf

**Description:** STM full response to RFI

Email: <a href="mailto:caroline@stm-assoc.org">caroline@stm-assoc.org</a>



April 24, 2023

# STM response to Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research (NOT-OD-23-091)

Thank you for the opportunity to comment on the "NIH Plan to Enhance Public Access to the Results of NIH-Supported Research" (NIH Public Access Plan), as issued in the Request for Information NOT-OD-23-091). STM is pleased that NIH is pursuing a robust stakeholder engagement process, as it did for the development of the NIH Policy on Data Management and Sharing (to which STM submitted significant comments). STM hopes that the comments made by stakeholders in the current process will be fully considered in the development of the final policy and its implementation. STM further hopes that we, and our members, will continue to be consulted on the various ways that NIH policy may impact scholarly communications.

STM stands for advancing open and trusted research, where researchers and the rest of society can rely on information that is credible, accessible, linked, and searchable in perpetuity. We therefore share with NIH the goal of increasing access to publications and data, not just for federally funded research, but for all research. More broadly, STM and our members are supportive of the goals of NIH in funding research and development. We therefore hope that STM and its members will have the opportunity to work with NIH to support researchers to advance biomedical research and public health, as well as promote quality, trust in science, equity, and the sustainability of the scholarly communications ecosystem.

Publishers have led and responded to the interest in open science by investing heavily in open science over the last 25 years, broadening and expanding the public's ability to understand and access the work of scientists and scholars. Many of the products necessary for open science were created and maintained by publishers, including online infrastructure, as well as preprinting, archiving, linking, and data management, and we continue to support and grow those efforts today. Our members have also invested in new models and approaches to providing access, including experimentation with a variety of business models to support quality, sustainability, and equity.

These experiences have demonstrated that there is not one best route to providing access. A mixed ecosystem is likely to persist for some time, even as publishers, institutions, and funders move to support open science. That said, STM believes that knowledge-creation, discovery, and sharing is best enabled when the final articles resulting from all stages of the peer-review and publication process are immediately openly available to all. The Version of Record (VoR) is the most thoroughly vetted version of the research publication, having been through all stages of the peer-review and publication process. The VoR is the authoritative version for researchers and the public, and it is more cited, used, and garners more attention than other versions of an

article.¹ For example, the VoR can link bi-directionally to research objects like data and code, has the latest updates on corrections, and sits on the publisher's platform where it can be integrated with other relevant content, allowing the public to better put this information into context. For these reasons, we urge NIH to ensure researchers have the option to make the VoR Open Access upon publication through a fully-funded Gold Open Access route. Our members would be happy to work together with you to provide the guidance and funding necessary for researchers to make this choice.

Regardless of the route to publication and public access, reliable funding needs to be made available to the researcher and their research institution, together with appropriate and enduring support and guidance on the use of funds and the options for providing access. In order to ensure equity for all researchers, such funding and guidance needs to be provided alongside other guidance for researchers, and in a manner that ensures author choice for whatever journals they choose to advance their research and impact. This funding also needs to be provided on an equal basis so that researchers who choose to publish in journals that are supported by APCs are not disadvantaged in the resources available for their research, student support, and other critical needs. All researchers must have options to meet their funder obligations, regardless of the journal they choose or the agreements their institution has with individual journals. Publishers have a wealth of experience in supporting policymakers and researchers with practical aspects of policy implementation and could work with the NIH to cocreate relevant guidance.

Current global efforts to expand open access indicate that direct support for publishing (which includes APC-supported Open Access, Read and Publish Agreements, and other evolving models) provides the most sustainable path to open access. Immediate access to a version of the article funded under subscription models has not proven to work at scale, even if it may temporarily work for some publishers or disciplines, or as a transitional model. While efforts to provide immediate access to articles funded by subscription journal publishers appear cost free to the researcher and funder, they are reliant on subscriptions to support the significant investments publishers make that ensure the quality, discoverability and accessibility of research in perpetuity. Subscription-supported investments include effectively managing the editorial and peer review processes and applying innovative technology to validate the rigor of the research we publish. Subscriptions are put at risk by the immediate availability of a large body of free accepted manuscripts, as demonstrated by widely used resources, such as Unsub.org, that encourage institutions to cancel subscriptions for materials that can be freely accessed. Nor is immediate access to articles funded by subscription journal publishers costfree for funders and institutions, as it causes additional, and duplicative, costs for the dissemination and long-term curation of research outcomes. Without sustainable funding – for a diversity of models for access -- fewer resources are available to ensure the quality and integrity of the scientific record, undermining the ability of scholarly communication to support

<sup>&</sup>lt;sup>1</sup> Researchers prefer the Version of Record, as outlined in a survey undertaken by Springer Nature (<a href="https://www.springernature.com/gp/open-research/version-of-record">https://www.springernature.com/gp/open-research/version-of-record</a>).

public trust in science and a dampening effect on innovation, job growth, and scientific progress. New barriers to access could also be created if important journals that serve critical research communities cease publication.

Flexibility is key to ensuring equity, academic freedom, and ensuring that researchers have the opportunity to best advance their discoveries to support innovation and public health. NIH should continue to allow the accepted manuscript to be shared sustainably, while also encouraging and enabling researchers to choose the VoR where appropriate. Critically, there should also be flexibility in licensing, allowing authors to provide articles under licenses and through agreements that best enable them to publish articles that best serve their research and impact. The draft plan indicates that NIH will provide guidance on how to "retain sufficient rights" to comply with the NIH public access policy, and we urge NIH to focus steps to ensure that researchers can supply a copy of any paper reporting on NIH-funded research to NIH for public availability. Requiring that researchers obtain additional rights risks creating inequities in publication opportunities for NIH-supported investigators, particularly in conjunction with an immediate access requirement. This is because some journals will need exclusive rights to support sustainable business models and continue investments needed for quality, preservation, discoverability, innovation, and impact. These risks can be mitigated by ensuring there is sufficient and enduring funding for Gold Open Access, which also can support the ability of researchers to share articles with the licensing option of their choice. STM therefore recommends that NIH retain the current policy of recommending that researchers ensure their publishing agreements include the right to provide a copy of the final peer-reviewed manuscript to the NIH upon acceptance for Journal publication, for public archiving in PubMed Central, which has served the public and NLM well.

Providing flexibility needs to go hand-in-hand with providing support for compliance. In order to minimize researcher burden, promote equity, and ultimately ensure the success of the NIH policy, this support should not just be financial, but should also include guidance for researchers and institutions as well as collaboration with publishers and research offices. The new NIH policy has the potential to significantly increase the amount of time and effort spent by researchers and institutions on implementation, and researchers will be looking to both publishers and their compliance offices to take on some of the responsibility. Therefore, collaboration and dialogue is key. There is likely to be a diversity of approaches and a mixed ecosystem that develops, and STM recommends that NIH provide flexibility and guidance that allows for diverse approaches to succeed.

This current response is focused on the publication side of the new policy, as STM submitted responses to each of the RFI opportunities for the NIH Data Management and Sharing Policy, as well as on the NLM Strategic Plan, to which we refer you for more details on our thoughts regarding data sharing and open science more broadly. The deliberative process and education of the research community provided for in the implementation of NIH's Data Management and Sharing Policy could be a valuable model for implementation of the new publication policy. In particular, the research community has been well served by the prominent guidance on planning and budgeting and the explicit acknowledgement throughout that data sharing has

real costs that need to be addressed in the proposal. Similar planning and budgeting will be needed for publications as well.

STM and our member publishers have invested significantly in a system of scholarly communication that enables the sharing of the latest discoveries and innovations, supports public trust in science and public health, enables interoperability through standards and infrastructure (metadata, persistent identifiers, etc), and ensures articles and data related to research are findable, accessible, and reusable. Publishers continue to invest and innovate to meet the changing needs of the communities that they serve, and to take advantage of the latest technologies to help research outcomes reach audiences as effectively as possible. STM supports an environment where publishers, in collaboration with NIH and the broad stakeholder communities funded and engaged in research related to NIH-funded projects, can continue to drive quality, integrity, and innovation in scholarly communication. In response to the prompts provided in the RFI, below we expand on some of the ideas mentioned above. It is our hope that this response will lead to further dialogue and engagement.

# 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Changing access requirements within the scientific ecosystem are likely to solve inequities from a reader aspect, but concerted and collaborative action will be necessary to ensure sustainability and equity across the ecosystem. Agencies can minimize the risk of creating new inequities, especially for scientists from traditionally marginalized communities, as well as early career researchers, by ensuring that these researchers and institutions have the funding support necessary for their research to flourish and choose the publishing option that best suits their needs. Publishers are doing their part by supporting new approaches, including Read and Publish Agreements, that provide opportunities for all to participate and access scholarly communication. Ultimately, a financially sustainable scientific publishing system is critical to advance trusted and impactful science, and attention to these issues can ensure that this is achieved.

To promote publishing equity, NIH needs to make appropriate and enduring funding available to the researcher and their research institution, together with appropriate and enduring support and guidance on the use of funds and the options for providing access. In order to ensure equity for all researchers, such funding and guidance needs to be provided alongside other guidance for researchers, and in a manner that ensures author choice for whatever journals they choose to advance their research and impact. This funding also needs to be provided on an equal basis so that researchers who choose to publish in journals that are supported by APCs are not disadvantaged in the resources available for their research, student support, and other critical needs. Finally, NIH should provide clear and prominent guidance on planning and budgeting and the explicit acknowledgement throughout the guidance that publication has real costs that need to be addressed in the proposal, as it has with the NIH Data Sharing and Management requirements.

Agreements with institutions or funders like Read and Publish Agreements or other pooled payment agreements have the potential to reduce inequality by making OA publishing available to all researchers. Publishers are actively working to develop and promote these models, which can reduce inequity for researchers at participating institutions and also can help increase compliance with policy and reduce administrative burdens. We have received reports of the success of such efforts, thanks to the real-world experiment of growth of transformative agreements around the world.<sup>2</sup>

Another aspect of equity in publishing opportunities relates to the promotion of equity and diversity in the research enterprise. Support for diverse publishing outlets is critical to such efforts, although to proactively drive further change requires input from stakeholders across the research ecosystem. One way in which publishers encourage equity and diversity in the research enterprise is by providing an objective space in which work can be assessed by peers (though our impartial oversight of an independent peer review process). More specifically, in recent years publishers have established industry-wide initiatives such as the Joint Commitment on Diversity and Inclusion<sup>3</sup> and C4DISC<sup>4</sup> which are developing consensus-based standards and best practice (e.g., developing guidelines around the peer review of articles and data; creating policies to support authors with deadnames; etc.).

Finally, publishers support and invest in various initiatives to enable researchers to participate in the scholarly dialogue. This includes support for educational efforts and funding programs that expand participation to underrepresented groups and ensure quality and integrity. For example, Research4Life, a UN-publisher partnership, supports researcher skill development, provides Research Lifecycle Training Webinars, and enhances the ability of LMIC researchers to publish with participating publishers. Many publishers support and partner with AuthorAID, a global network that provides free resources and training, including in article writing, for researchers in low- and middle-income countries. Publishers offer various funding programs to support the participation of less-well-resourced researchers, including discounts and waivers, both individually and through collective approaches like Research4Life. Publishers also work with other stakeholders to provide resources to identify trusted outlets to present their work (e.g., Think. Check. Submit. (thinkchecksubmit.org) a cross-industry initiative) and promote integrity in scholarly research and its publication through the Committee on Publication Ethics (COPE, www.publicationethics.org) and other efforts.

<sup>&</sup>lt;sup>2</sup> For example, our member Taylor & Francis notes that the top 10 most published subject areas under their transformative agreements in the past two years have been in humanities, arts, and social sciences, which have traditionally been less likely than those in the physical and biomedical sciences to choose OA. For additional data, see the STM Open Access Dashboard <a href="https://www.stm-assoc.org/oa-dashboard/">www.stm-assoc.org/oa-dashboard/</a>.

<sup>&</sup>lt;sup>3</sup> https://www.rsc.org/new-perspectives/talent/joint-commitment-for-action-inclusion-and-diversity-in-publishing/

<sup>4</sup> https://c4disc.org/

# 2. Steps for improving equity in access and accessibility of publications.

Publishers invest significantly in efforts to provide access, accessible formats, and accessible modes of dissemination for publications. It is important to note that for access and accessibility to be provided, first the publications and infrastructures must be created and disseminated. Therefore, it is a necessary precondition to improve equity in access and accessibility of publication that NIH work to ensure the viability of a robust ecosystem of scholarly communications that drives innovation, supports quality and integrity, and ensures appropriate infrastructure to enable accessibility to diverse users.

As alluded to in the introduction to this response, steps to improve access and accessibility could be broken down into three requirements: 1) sufficient, enduring, and appropriate funding, 2) encouragement and education of researchers to budget for and choose open science, and 3) flexibility for researchers and organizations to enable diverse modes of communication.

Appropriate and enduring funding is fundamental to achieve the open science goals outlined in the draft NIH plan and in the August OSTP memo and make sure that NIH's revised policy can promote equity in access. This is because the sustainability of publishing is a precondition to the availability, utility, and accessibility functions of scholarly communications.

Encouragement and education of researchers is also key, as they will ultimately be responsible for ensuring that the articles that they write are available to the public. Experience with funder requirements and compliance around the world indicates that researchers are often confused about grant requirements, including on how and when to provide access to publications, and a significant percentage of researchers erroneously believe that it is an inappropriate use of grant funds to pay for publication. STM's members' experience with guidance and education indicates that such efforts can make a big difference in researchers willingness to choose open access and compliance with funder and other requirements.

Flexibility is needed to promote diversity in publication, ensure author choice, and support access to publishing in ways that work for researchers. As noted earlier, different publishers may offer distinct approaches to provide access, each of which may be appropriate to the communities they serve, and each of which should be allowed as a method for researchers to ensure access to any article they author that reports on NIH-funded research. A diversity of publication outlets, enabled by flexible approaches to implementation of the NIH policy, supports diversity in research.

Publishers invest significantly to ensure that articles are accessible in various human and machine-readable formats and are available to those with diverse needs. Many publishers have invested in technology and infrastructure to build towards, meet, or exceed Section 508 accessibility and have created a diverse ecosystem of accessible resources available to diverse

STM (The International Association of Scientific, Technical and Medical Publishers) www.stm-assoc.org

<sup>&</sup>lt;sup>5</sup> E.g., nearly 1 in 6 in the 2016 <u>Pay It Forward Report</u> and 1 in 5 in the 2019 <u>Taylor & Francis Researcher Survey</u>

audiences with or without assistive technologies. Some of our members were leaders in developing braille resources in multiple languages, screen reading technology implementation, and other innovations. These additional infrastructure and formatting investments are enabled by sustainable business models.

STM also notes various initiatives that we or our members have promoted to ensure access and accessibility for diverse audiences. These include Research4Life which provides access to researchers in Low- and Middle- Income countries; efforts to share plain language summaries to broaden the accessibility of cutting-edge research to non-experts; and investments in the promotion of articles to the media and through social media channels.

Finally, STM notes that equity in access requires that publications that are made available are accurate and trustworthy. STM and its members invest significantly in ensuring research integrity and the quality and reliability of the scholarly record. For example, STM Solutions recently launched the Research Integrity Hub (<a href="https://www.stm-assoc.org/stm-integrity-hub/">https://www.stm-assoc.org/stm-integrity-hub/</a>), a robust and holistic set of tools to safeguard the integrity of science through a combination of shared data and experiences and by harnessing technological innovation. Individual publishers are working individually and in partnership with other organizations to prevent misconduct and ensure the integrity of the system. Safeguarding research integrity can only be done through collaboration with all stakeholders in the scholarly ecosystem, and in an environment where continued investments can be made.

# 3. Methods for monitoring evolving costs and impacts on affected communities.

STM's members compete in a dynamic environment that drives them to provide the widest possible access to the articles that they publish at the lowest possible cost to the research and user communities. Costs and revenue streams can vary significantly from one publisher to another, and even from one journal to another, depending on many factors such as audience, circulation/reach, ranking, number of articles published, field/specialty, and distribution method. These differences need to be considered when evaluating the market dynamics and taking a broad average of dissimilar journals is not recommended.

More broadly, it is important to consider the changing dynamics of how scholarly publication is supported when attempting to monitor trends. Historically, publishers' costs have been spread across those that consume the research (readers / subscribers) of which there are many. The NIH plan may move associated costs to other payers, of which there are fewer. The cost burden will therefore increase for some (e.g., research-intensive universities) while many others will no

<sup>&</sup>lt;sup>6</sup> E.g., Elsevier (<u>https://www.elsevier.com/about/accessibility</u>) and Taylor and Francis (<u>https://taylorandfrancis.com/about/corporate-responsibility/accessibility-at-taylor-francis/</u>).

<sup>&</sup>lt;sup>7</sup> E.g., Optica's Spotlight on Optics (<a href="https://opg.optica.org/spotlight/about.cfm">https://opg.optica.org/spotlight/about.cfm</a>) and Taylor and Francis Plain Language Summaries (<a href="https://authorservices.taylorandfrancis.com/publishing-your-research/writing-your-paper/how-to-write-a-plain-language-summary/">https://authorservices.taylorandfrancis.com/publishing-your-research/writing-your-paper/how-to-write-a-plain-language-summary/</a>)

longer contribute to the costs (e.g., commercial industries, which traditionally subscribe to journals without publishing extensively in them).

When considering the budget for supporting public access to high-quality, peer-reviewed articles reporting on NIH-funded research, it is important to look beyond a single aspect of pricing (i.e., APCs) and consider the total investment in scholarly communications, which includes subscriptions, APCs, transformative agreements, and other inputs. The cost and pricing structures are very different for different disciplines – medicine, physical sciences, social sciences, and humanities – and for different types of journals based on selectivity, services, technology, and other features.

That said, APC prices are virtually always transparent.<sup>8</sup> Our members are committed to the maximum possible transparency around pricing, in accordance with regulation and antitrust concerns, and note that APCs may vary across journal titles based on a variety of factors. Our members are also committed to ensuring that every researcher – regardless of geographic location, discipline or personal circumstance has relevant and realistic options available to them to publish their work, so that no researcher is left without a voice, regardless of funding source. Consistent with this commitment publishers have developed Read and Publish Agreements with institutions and maintain active waiver and discount programs to serve researchers.

STM is not aware of any other NIH efforts to monitor expenses for specific research services or outputs and cautions that any efforts to look at trends in publishing must be carefully interpreted in the context of an evolving and dynamic ecosystem. Those who monitor APC prices and perform market analysis are aware that any trends in this data always need to be contextualized with respect to other trends in publishing (e.g., the growth in the sharing of research outputs) and revenue (e.g., subscription rates and transformative agreements) and with respect with efforts to ensure equity in publication opportunities (e.g., provision of waivers and discounts).

A diverse, financially sustainable, and robust publishing system which provides authors with broad choice is the most effective way to ensure fair and competitive pricing and address any cost concerns. Hard price caps will likely drive existing industry trends toward publisher consolidation and volume-based models which could compromise integrity, quality, and author choice. The research enterprise, and the impact of NIH-funded research on innovation and public health, is best served by diversity that is enabled by flexibility and full support for open access publishing options.

<sup>&</sup>lt;sup>8</sup> APC price lists are generally public, and transparently shared. Some examples include American Chemical Society: (<a href="https://acsopenscience.org/researchers/oa-pricing/">https://acsopenscience.org/researchers/oa-pricing/</a>), American Physical Society (<a href="https://journals.aps.org/authors/apcs">https://journals.aps.org/authors/apcs</a>), Elsevier (<a href="https://www.elsevier.com/about/policies/pricing">https://journals.aps.org/authors/apcs</a>), Elsevier (<a href="https://www.elsevier.com/about/policies/pricing">https://www.elsevier.com/about/policies/pricing</a>), Springer Nature (<a href="https://www.springernature.com/gp/open-research/journals-books/journals">https://www.springernature.com/gp/open-research/journals-books/journals</a>), Wiley (<a href="https://authorservices.wiley.com/author-resources/Journal-Authors/open-access/article-publication-charges.html">https://authorservices.wiley.com/author-resources/Journal-Authors/open-access/article-publication-charges.html</a>), The Public Library of Science (PLoS) (<a href="https://plos.org/publish/fees/">https://plos.org/publish/fees/</a>).

In addition, care must be taken with respect to interventions that seek to ensure fees and policies remain reasonable and equitable, as they may lead to unintended consequences or constitute anti-competitive market interference under antitrust laws. As STM and others have recommended in other contexts, NIH should seek legal advice regarding competition law and any undue influence on industry market pricing. Finally, we underline that the goals of the NIH policy are best achieved though NIH efforts to ensure that researchers are budgeting appropriately for publications.

4. Early input on considerations to increase findability and transparency of research.

We divide our response into two sections, as the concepts and needs of findability and transparency, while interrelated, are also quite distinct.

a. Findability (including persistent identifiers (PIDs), metadata, and other infrastructure).

STM and its member publishers would welcome collaboration with NIH to support approaches to findability that leverage and build on existing standards, technologies, infrastructure, and protocols. Publishers have committed to and invested significantly in ensuring the findability of articles and research data. Our experience suggests that additional efforts to support the use and development of persistent identifiers throughout the research ecosystem would bear additional fruit, including identifiers for articles and research data as well for funding agencies, grant awards, facilities, and the like.

Where possible, NIH should leverage existing standards and systems, as supported by publishers, institutions, and other stakeholders. The primary existing PID and metadata structure, enabled through organizations including CrossRef and DataCite, should be adopted and adapted as necessary to minimize disruption, promote compliance, and prevent unnecessary duplication of effort and investment in the scholarly communications system.

Publishers already invest heavily in creating persistent identifiers and machine-readable metadata that promote greater visibility of research findings and data, and these help to promote trust, reliability, and transparency for the scientific system. Cross publisher and industry initiatives around PIDs include researcher (ORCID), institutional (Ringgold), and funder (Open Registry of Funders) PIDS embedded in our content workflows as standard across the majority of the scholarly communication ecosystem. Embedding standards supports our infrastructure development to build better links between interrelated research outputs and improve visibility from funding through to publication. In general, PIDs used or recommended by NIH should be those used by the community, as those can be validated and maintained. Where NIH needs additional or bespoke PIDs, efforts need to be made to ensure they map well to other PIDs that are already well embedded in the ecosystem.

Specifically, STM recommends that NIH support the use of community-adopted PIDs through the grant application process (e.g., ORCIDs for researchers, organization IDs for the

institutions(s) affiliated with each researcher, and Funder IDs for the distinct funders of the grant). While organization IDs are not as well-established or robust as researcher IDs (with ORCID), there are several emerging options for organizations, and NIH should consider recommending one of the following PIDs to ensure harmonization and avoid unnecessary duplication in the scholarly record: Ringgold (a global organization identifier system); ISNI (ISO standard name identifier system); ROR (the Research Organization Registry); and Crossref's Funder Registry; along with ORCID. NIH should also ensure there are metadata fields for all of these.

In addition, publishers have invested significantly in discoverability, search engine optimization, and other efforts to make sure that published articles can be found and used to advance scientific research. To support the findability of both articles and research data, NIH should also engage with and implement community-based standards and infrastructure initiatives that link and promote access to the best available versions of articles and research data. These include open protocols like Scholix, a multi-stakeholder initiative to link scholarly literature and research data, and services like CHORUS, that helps the public find and access articles reporting on federally-funded research. Initiatives such as seamlessaccess.org, a service designed to help foster a more streamlined online access experience by leveraging an existing single-sign-on infrastructure, and GetFTR, a tool that streamlines access to journal articles on discovery tools and collaboration networks, are also available to enable and accelerate access. STM would welcome additional dialogue to discover which existing initiatives could best be utilized to support findability and access to articles and research data related to NIH-funded research, and to collaboratively develop solutions where services or infrastructures do not already exist.

## b. Transparency (including reproducibility and trust in science)

Findability is necessary to promote transparency, but it is not sufficient to enable it. Transparency needs to be fostered through education and the research culture and enabled by infrastructure. Publishers continually invest in such systems and infrastructure and promulgate policies that encourage open sharing to promote trust. This includes efforts to promote trust and transparency through the sharing of research data (e.g., STM's Research Data initiative<sup>9</sup>) and especially the use of FAIR (Findable, Accessible, Interoperable, and Reproducible) principles in sharing research data. Innovations in open peer review, the broadening of publishable articles to include negative results, the introduction of registered reports, and other efforts to make publication and the publication process more transparent have the potential to improve public trust in science and the utility of research. Many of our members have signed on to Transparency and Openness Promotion (TOP) Guidelines and engaged with other initiatives to drive transparency.

STM recommends that NIH leverage existing resources to promote transparency and avoid creating duplicative resources. For example, NIH can point to existing resources to support

<sup>9</sup> www.stm-researchdata.org

researchers in making their research outputs more transparent. Some potential examples include <u>a manifesto for reproducible science</u> designed to optimize key elements of the scientific process and "<u>STAR Methods: Structured, Transparent, Accessible Reporting</u>," designed to provide a structure for experimental methods that increases reproducibility. Existing, robust infrastructure should be considered before recommending or developing new systems.

We note that new modes of scientific inquiry are providing opportunities to improve scholarly practices, including with respect to transparency and integrity, but these may also carry risks that are not fully understood at this time. NIH's policies must be flexible enough to address any issues that might arise in these new modes of scholarship, as well as provide support for new and existing infrastructure and services that can help provide the review and analysis needed to ensure quality and integrity of both new and existing systems.

Finally, we note that the most important action that NIH can take to ensure transparency, quality and integrity in scholarly communication is to support and encourage the systems and services that currently provide these benefits for the research enterprise. These include, but are not limited to, market incentives that encourage the development of high-quality publication outlets for scholarly communication such as those produced by STM's members.

## **About STM**

At STM we support our members in their mission to advance trusted research worldwide. Our more than 140 members collectively publish 66% of all journal articles and tens of thousands of monographs and reference works. As academic and professional publishers, learned societies, university presses, start-ups and established players, we work together to serve society by developing standards and technology to ensure research is of high quality, trustworthy and easy to access. We promote the contribution that publishers make to innovation, openness and the sharing of knowledge and embrace change to support the growth and sustainability of the research ecosystem. As a common good, we provide data and analysis for all involved in the global activity of research.

The majority of our members are small businesses and not-for-profit organizations, who represent tens of thousands of publishing employees, editors, reviewers, researchers, authors, readers, and other professionals across the United States and world who regularly contribute to the advancement of science, learning, culture and innovation throughout the nation. They comprise the bulk of a \$25 billion publishing industry that contributes significantly to the U.S. economy and enhances the U.S. balance of trade.

**Submit date: 4/24/2023** 

I am responding to this RFI: On behalf of an organization

Name: Helen Burstin, MD, MPH

Name of Organization: Council of Medical Specialty Societies

Type of Organization: Professional org association

**Role:** Member of the public

# 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

The Council of Medical Specialty Societies (CMSS) is a coalition of 50 specialty societies representing more than 800,000 physicians across the house of medicine. CMSS works to catalyze improvement across specialties through convening, collaborating, and collective action. We are pleased to provide input on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research (NIH Public Access Plan) and the 2022 White House Office of Science and Technology Policy (OSTP) memo on Ensuring Free, Immediate, and Equitable Access to Federally Funded Research.

As non-profit society publishers, we bring our best practices to the peer review of the articles and to wide dissemination of this content in support of the scholarly communication enterprise. Our long history of working together with our research communities has resulted in publication of some of the most impactful and practice changing content. The integrity of peer review is vital to sharing research findings in a way that assures accuracy, integrity, and the transmission of science that promotes new evidence vital to patient care; our comments raise questions that are important to carefully assess in order to preserve that US research enterprise as a source of high-quality scientific information.

# **Shifting of Revenue Streams**

While the proposed policy allows publication in journals with varied publishing models, it does not address the impact that NIH Public Access Plan will have on publishing fees. Opening papers prior to the current 12-month embargo will result in the loss of subscription revenue from institutions and individuals and, for many publishers, a corresponding decrease in advertising revenue. In order for publishers to provide the scientific community with the support it has become accustomed to, including, but not limited to, maintaining the integrity of the science, robust peer review, support for discoverability, reproducibility and dissemination of the science, the financial burden will shift to the authors. Diligent peer review, management and public disclosures of conflicts, and data and figure integrity checks are vital parts of a responsible publication process. Threats to the integrity of the content, such as plagiarism, paper mills, inappropriate Al generated content, and fraudulent data, are always present and require steady attention. While no system is perfect, peer-review increases the opportunity to mitigate these risks and protect the public from ensuing harm.

Publishers also provide additional benefits to their communities by providing educational material, alternative metrics and enhanced metadata that may also suffer due to diminishing revenue. All of this requires resources that are likely to be endangered if publishers lose the revenue that currently sustains this work. Such losses could occur in the form of cancelled subscriptions, insufficient total article processing charge (APC) income, and lost licensing fees for approved reuse of content, among Others.

Each publisher will have their own budgetary tipping point when decreased revenues force a decision to discontinue vital services now protecting the integrity of research published in our journals, but all will face this challenge, and all will have to make cost-saving changes to maintain a viable publishing enterprise.

Policies that restrict publishers' abilities to collaborate with authors to realize their protection of rights under United States copyright law would further limit revenue streams on which we depend, including royalties, licensing, reprints, and advertising. We urge the NIH not to include rights retention language or license requirements in the final policy Other than the grantee's right to deposit the manuscript. Preserving a Green OA route presents a sustainable business model that should be embraced. Expanding rights retention policies beyond the deposition of the manuscript would also erode the publisher's ability to monitor usage of the content in support of the author's intellectual property.

Our specialty societies strongly recommend a two-year delay to adhere to the mandate. This time would allow us to work with you to develop policies that sustain reliable, equitable, high quality scientific content.

## Access to funding

OSTP and NIH state that grants can be used to cover publication costs, which is a positive step; however, it is important that NIH increase the total amount of grant funding per award so that the additional Article Processing Charges, including potential fees to deposit papers into PubMed Central for example, will not reduce the funds available for research.

There are Other concerns to consider. Certain grants do not permit use of funds for publication fees. As such, CMSS recommends that NIH exempt certain types of infrastructure-related grants (e.g., cancer center support grants, CTSAs, NCORPs) and teaching grants (K awards, T awards) from reporting funding to journals and thus requiring deposit.

The broad reach and impact of this proposed plan will be a challenge to implement and enforce if compliance is mandated for all NIH funded authors regardless of how much funding they received or how small a role any given individual plays in a research project or manuscript. The NIH should instead apply a minimum threshold of funding and/or level of participation by authors and researchers before subjecting the papers to the proposed mandate.

#### Copyright protection

Copyright protection is the first line of defense for any author against the misuse of their research, and publishers stand ready to defend investigators' intellectual property. Journals customarily allow authors to post their paper on their institutions' site, make use of their work at conferences, but this policy needs to clearly state that making the content freely accessible does not give anyone the right to create derivative products without permission. Clarification that the rights remain with the copyright holder needs to be articulated. The final guidance should also clarify that authors are obligated to follow the NIH Guidelines only for the papers they author as a result of NIH funding.

#### **Definition of First Publication**

There is confusion in the community concerning the definition of First Publication. We are interpreting NIH's draft language regarding first publication to mean that the manuscript uploaded to PubMed

Central in compliance with this policy will be embargoed until the first appearance of the final typeset article. Are we also correct in understanding that the Pub Med Central first publication will include a link to the publisher's site? Clarification of this matter in the final policy is strongly recommended to avoid confusion in the community.

### 2. Steps for improving equity in access and accessibility of publications.

Removal of the currently allowable 12-month embargo period for NIH-supported publications will improve access to these research products for all. As noted in the NIH Public Access Plan, NIH also plans to continue making articles available in human and machine-readable forms to support automated text processing. NIH will also seek ways to improve the accessibility of publications via assistive devices. NIH welcomes input on Other steps that could be taken to improve equity in access to publications by diverse communities of users, including researchers, clinicians and public health officials, students and educators, and Other members of the public.

## Access and accessibility of publications

Journal publishers have long been collaborating with various stakeholders to develop and implement collaborative projects that enhance the public access, utility, preservation, and discoverability of materials that report on and analyze and interpret results of federally funded research. Publishers participate in a multitude of services that enhance discoverability, including ORCID, Crossref, the Committee on Publication Ethics, and provide guidelines that are not influenced by pharmaceutical companies as well as making sure conflicts of interest are accurately noted. Federal agencies should collaborate with publishers and Other stakeholders to ensure minimum standards, share best practices, and minimize duplication of work.

Providing immediate access to all scientific research comes with significant issues and significant financial/labor costs of compliance. We want to make sure that authors' intellectual property remains accurately presented on the worldwide stage; we are concerned that the research could be pirated by outside bodies that may misinterpret the results to suit their needs. While publisher's efforts to support free, immediate access to COVID-19 research were a boon to scientists, we also saw a rise of misuse and misunderstanding of research among the public. As the medical and research community collectively works to increase the public's trust in health and science, these proposed changes could unintentionally foster misinformation. Strong intellectual property protections are a necessary safeguard against the acceleration of this trend.

We recommend that NIH support publisher's ability to enforce copyright protection by maintaining publishers' rights in and to the content published.

# 3. Methods for monitoring evolving costs and impacts on affected communities.

NIH proposes to actively monitor trends in publication fees and policies to ensure that they remain reasonable and equitable. NIH seeks information on effective approaches for monitoring trends in publication fees and equity in publication opportunities.

# Diverse Publishing Landscape

Publishers continually develop enhancements to the peer review and publishing processes, and this requires constant investment that would be slowed or stopped by a lack of funds. Many publishers

currently provide checks against plagiarism and graphic manipulation which ensure the veracity of the new literature and protect previously published works. Publishers work tirelessly to ensure the reproducibility of science which in turn protects patients. It is also worth noting that requiring all publishers to supply financial information in pursuit of fixed pricing conflicts with fair trade.

# 4. Early input on considerations to increase findability and transparency of research.

Section IV of the NIH Public Access Plan is a first step in developing the NIH's updated plan for persistent identifiers (PIDs) and metadata, which will be submitted to OSTP by December 31, 2024. NIH seeks suggestions on any specific issues that should be considered in efforts to improve use of PIDs and metadata, including information about experiences institutions and researchers have had with adoption of different identifiers.

## Consistent Guidance

There are many examples of advancements already accepted by the industry such as DOIs, ORCID, funder registries, discovery tools for content mining, and use of JATS for structured metadata. If NIH wants to aggregate these data, they should collaborate with various stakeholders to create and engage in guidance for authors and publishers regarding standards to ensure best practices and minimize duplication of work.

Email: hburstin@cmss.org

**Submit date:** 4/24/2023

I am responding to this RFI: On behalf of an organization

Name of Organization: Alliance for Nursing Informatics

Type of Organization: Professional org association

- 1. How to best ensure equity in publication opportunities for NIH-supported investigators.
- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

# **Uploaded File:**

NIH\_Public-Access-Plan\_ANI-Comments\_2023.04.24.pdf

**Description:** Alliance for Nursing Informatics Comment Letter

Email: sharon.giarrizzo-wilson@cuanschutz.edu



April 24, 2023

NIH Office of Science Policy National Institutes of Health 6705 Rockledge Dr #750 Bethesda, MD 20817

Re: NOT-OD-23-091, Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research

Submitted electronically at: <a href="https://osp.od.nih.gov/nih-plan-to-enhance-public-access-to-the-results-of-nih-supported-research">https://osp.od.nih.gov/nih-plan-to-enhance-public-access-to-the-results-of-nih-supported-research</a>

Dear NIH Office of Science Policy,

The Alliance for Nursing Informatics (ANI) appreciates the opportunity to comment as nursing stakeholders on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research.

The Alliance for Nursing Informatics (ANI), co-sponsored by AMIA and HIMSS, advances nursing informatics leadership, practice, education, policy, and research through a unified voice of nursing informatics organizations. We transform health and healthcare through nursing informatics and innovation. ANI is a collaboration of organizations representing more than 25,000 nurse informaticists and bringing together 29 distinct nursing informatics groups globally. ANI crosses academia, practice, industry, and nursing specialty boundaries and collaborates with the more than 4 million nurses in practice today.

We fully support the goals of the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research, aligning with the U.S. Government's directive for "Ensuring Free, Immediate, and Equitable Access to Federally Funded Research" and acknowledge the central role patients and the public play in health care and health research. Recognizing the importance of patient engagement in these initiatives to building the infrastructure for research dissemination and improving care delivery, it is equally important that the plan does not supersede patient privacy and autonomy. We provide the following recommendations for your consideration:

1. Steps for improving equity in access and accessibility of publications.

We applaud the NIH's aims to improve equity in access to publications by diverse communities of users. Our research shows that returning study findings to patients increases trust in the scientific process, especially for underrepresented groups. Therefore, we propose that NIH

<sup>&</sup>lt;sup>1</sup> Mangal S, Niño de Rivera S, Choi J, et al. Returning study results to research participants: Data access, format, and sharing preferences. *Int J Med Inform*. 2023;170:104955.

consider two important features of access and accessibility, namely: (1) findability and (2) comprehensibility.

Many members of the American public may be unfamiliar with scholarly resources and instead rely on mainstream media for their scientific news, which may contribute to misinformation and misinterpretation of findings.<sup>2</sup> To improve findability, NIH should consider partnering with mainstream platforms to drive readers to the source of information on the NIH's website (e.g., PubMed).

Disparities in consumer health literacy, literacy, and numeracy skills decrease comprehension of scientific abstracts.<sup>3</sup> Lay interpretations of abstracts with published manuscripts on publicly accessible platforms (e.g., PubMed) are needed. For example, generative artificial intelligence and machine learning platforms could assist scientists in producing lay abstracts. Providing easy-to-read and interpretable abstracts will promote broader comprehension by the lay public and help reduce misinterpretations associated with scholarly publications.

# 2. Early input on considerations to increase findability and transparency of research.

ANI appreciates the NIH's <u>2023 Data Management & Sharing Policy</u> with the inclusion of justifiable reasons for limiting data sharing. However, in practice, our members note that grant management teams remain unclear on the precise interpretation of these guidelines and what constitutes a strong rationale for limiting data sharing.

Our concern focuses on the number of specific types of health data for which persistent identifiers (PIDs) and metadata would significantly compromise patient confidentiality. First, electronic health records (EHRs) used in research are easily re-identifiable, and the risk of re-identification is higher when external metadata are available. Attempts to fully de-identify EHR datasets (e.g., date removal) render the data meaningless for research purposes. Second, data from wearable devices, such as smartphones and smartwatches, are growing in use for health research to collect sensitive data, such as reproductive health, and granular, continuous data about individuals' locations and behaviors. The exposure of these data through public repositories represents risks for those participating in research for stigmatized conditions (e.g., mental health, sexually transmitted infections) and those that are subject to changing laws, for which the patient and/or their healthcare professionals may be criminally liable (e.g., limited reproductive rights). Third, transcripts from qualitative interviews can never be truly anonymized; even with the omission of names and other overtly identifying information,

<sup>&</sup>lt;sup>2</sup> Funk C, Gottfried J, Mitchell A. Science news and information today. Pew Research Center's Journalism Project. Published September 20, 2017. Accessed April 16, 2023. <a href="https://www.pewresearch.org/journalism/2017/09/20/science-news-and-information-today/">https://www.pewresearch.org/journalism/2017/09/20/science-news-and-information-today/</a>

<sup>&</sup>lt;sup>3</sup> Prince LY, Schmidtke C, Beck JK, Hadden KB. An Assessment of Organizational Health Literacy Practices at an Academic Health Center. *Qual Manag Health Care*. 2018;27(2):93-97. doi:10.1371/journal.pone.0028071

<sup>&</sup>lt;sup>4</sup> El Emam K, Jonker E, Arbuckle L, Malin B. A systematic review of re-identification attacks on health data. *PLoS One*. 2011;6(12):e28071. https://doi.org/10.1371/journal.pone.0028071

<sup>5 ·</sup> Xia W, Liu Y, Wan Z, et al. Enabling realistic health data re-identification risk assessment through adversarial modeling. *J Am Med Inform Assoc.* 2021;28(4):744-752. doi:10.1093/jamia/ocaa327

<sup>6</sup> Dong Z, Wang L, Xie H, Xu G, Wang H. Privacy Analysis of Period Tracking Mobile Apps in the Post-Roe v. Wade Era. In: *Proceedings of the 37th IEEE/ACM International Conference on automated Software Engineering*. ASE '22. Association for Computing Machinery; 2023:1-6. <a href="https://doi.org/10.1145/3551349.3561343">https://doi.org/10.1145/3551349.3561343</a>

qualitative research is inherently personal and involves sharing personal experiences and perspectives.<sup>7</sup>

The evidence demonstrates participant trust in research can erode when participants have limited control over how and with whom their personal health data is shared.¹ Exposure of these data types and other sensitive data not explicitly mentioned through public repositories may inadvertently discourage patients from participating in medical research. Patients may also hesitate to seek care at academic medical centers where their medical records may be used for research. In particular, this may deter participation among racial and ethnic minority groups whose trust in medical research may already be limited.<sup>8</sup>

Additionally, data-sharing policies are much stricter in Canada, the European Union, and many other countries globally where U.S. researchers conduct NIH-funded research. These differences across countries complicate matters for researchers and grant administrator teams attempting to comply with differing and sometimes competing data-sharing policies between the U.S. and other countries.

Therefore, while PIDs combined with metadata can promote transparency, increased scientific integrity, and public trust in research, we suggest increased guidance and clarity on specific justifications for limiting data sharing and to address researchers' uncertainty about the appropriateness of particular research contexts that justify withholding.

Thank you for the opportunity to comment on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research.

Sincerely,

Susan Hull, MSN, RN-BC, NEA-BC, FAMIA

Supor C. Hull

ANI Co-chair

Nancy Beale, Ph.D., RN-BC

Tong ghal

ANI Co-chair

The <u>Alliance for Nursing Informatics</u> (ANI), co-sponsored by AMIA and HIMSS, advances nursing informatics leadership, practice, education, policy, and research through a unified voice of nursing informatics organizations. We transform health and healthcare through nursing informatics and

<sup>7</sup> Saunders B, Kitzinger J, Kitzinger C. Anonymising interview data: challenges and compromise in practice. *Qual Res*. 2015;15(5):616-632. https://doi.org/10.1177/1468794114550439

<sup>8</sup> Milani SA, Swain M, Otufowora A, Cottler LB, Striley CW. Willingness to Participate in Health Research Among Community-Dwelling Middle-Aged and Older Adults: Does Race/Ethnicity Matter? *J Racial Ethn Health Disparities*. 2021;8(3):773-782. <a href="https://doi.org/10.1007/s40615-020-00839-y">https://doi.org/10.1007/s40615-020-00839-y</a>

g Rehm HL, Page AJH, Smith L, Adams JB, Alterovitz G, Babb LJ, Barkley MP, Baudis M, Beauvais MJS, Beck T, Beckmann JS, Beltran S, Bernick D, Bernier A, Bonfield JK, Boughtwood TF, Bourque G, Bowers SR, Brookes, AJ . . . & Birney E. GA4GH: International policies and standards for data sharing across genomic research and healthcare. *Cell Genom.* 2021 Nov 10;1(2):100029. doi:10.1016/j.xgen.2021.100029

innovation. ANI is a collaboration of organizations representing more than 25,000 nurse informaticists and bringing together 29 distinct nursing informatics groups globally. ANI crosses academia, practice, industry, and nursing specialty boundaries and collaborates with the more than 4 million nurses in practice today. Contact ANI.

**Submit date: 4/24/2023** 

I am responding to this RFI: On behalf of an organization

Name: Lizbet Boroughs, MSPH Associate Director of Federal Relations and Kate Hudson, JD, Associate

Vice President and Counsel

Name of Organization: Association of American Universities

Type of Organization: Professional org association

Role: Member of the public

# 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

The Association of American Universities (AAU) thanks the National Institutes of Health for the opportunity to comment on NOT-OD-23-091, the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research. Founded in 1900, AAU is composed of America's leading research universities. AAU's 65 research universities transform lives through education, research, and innovation.

AAU strongly agrees with NIH's statements that "increasing access to publications and data resulting from federally funded research offers many benefits to the scientific community and the public," and that access "can accelerate research, generate higher quality scientific results, encourage greater scientific integrity, and enable future inquiry, discovery, and translation for NIH-supported research." Indeed, in 2021, AAU and its sister organization, the Association of Public & Land-grant Universities (APLU), published a joint Guide to Accelerate Public Access to Research Data to help inform our respective member institutions' activities on accessible research data. Leading up to the publication of this document, with funding from NIH and the National Science Foundation (NSF#1837847 and #1939279), AAU and APLU held a series of workshops and conferences with researchers, senior research officers, librarians, chief information officers, and organizations in support of increasing public access to research.

Given our past work and strong interest in public access, AAU is carefully monitoring various federal research agencies' implementation of the August 2022 guidance released by the Office of Science and Technology Policy (OSTP). Our joint response with APLU in January 2020 to NOT-OD-20-013 highlighted that additional specific clarification, outside the scope of the RFI, would enable robust participation and engagement by researchers and universities with NIH's Data Management and Sharing Policy. AAU's comments on NOT-OD-23-091 are informed by our collaborations and discussions with our members, APLU, the Association of American Medical Colleges (AAMC), the Federation of American Societies for Experimental Biology (FASEB), and the Council on Governmental Relations (COGR).

AAU appreciates that NIH is engaged in clarifying reasonable costs for publications that can be charged directly by individual PIs to grants. This approach should also encompass cost considerations at the broader University level. Preparations for publications are not only supported by direct costs but also pooled mechanisms such as facilities and administrative costs, library subscriptions, and additional University support from Other available revenue sources. Indeed, oversight of Data Management and Sharing (DMS) is a collaborative process and not solely the researcher's responsibility during an award's arc. Data curation; compliance with federal, state, and tribal laws; metadata requirements related to fields of study; and proper data storage are tasks that require resources and an integrated approach

well beyond the individual researcher's scope of direct costs. Universities with robust financial resources, data infrastructure, and library and faculty support may have the capacity to leverage these resources to respond to the added costs involved in ensuring that the new public access requirements are met, however, many institutions and their faculty may struggle to support these additional costs.

AAU suggests that NIH could ensure data access and help minimize costs by creating and supporting one agency-wide data repository, similar to the creation of PubMed Central, to serve this purpose for publications. This would be particularly useful for areas where no current NIH-supported disciplinary repository exists. AAU also suggests that agencies create overarching disciplinary-specific repositories to ensure that universities do not create a myriad of different repositories, which will diffuse the accessibility of data access overall.

Additionally, we urge the NIH to explore ways to ensure that faculty and institutions have the means to receive support for publication and data storage costs well beyond the length of an individual grant. Without financial support after the terms of a grant, researchers and universities will be unable to comply with open access and data management standards for NIH without incurring the costs themselves, which will undoubtedly have a more significant and inequitable impact on researchers and institutions without robust research infrastructure funding.

AAU appreciates NIH's continued engagement with the community on the unanticipated costs of its DMS policy.

## 2. Steps for improving equity in access and accessibility of publications.

Ultimately, data is limited in its utility if research data stewardship is not fundamental to the research endeavor. Conceptualizing and planning for data access and interoperability is a continually iterative process involving researchers, funders, institutions, health professionals, and the public. Data technology and analysis are not stagnant, and their evolution will require flexibility within NIH's public access guidance and continual training for program officers at the individual NIH institutes.

AAU is, therefore, supportive of NIH's collaborations with scientific societies, such as FASEB's "DataWorks! Help Desk," to improve data management at the individual researcher level. AAU also strongly supports the creation of disciplinary based data repositories to improve and ensure access to federally funded research results and believes that it is important for NIH to support and facilitate the creation of such repositories. As previously stated, we also recommend the creation of one overall NIH-supported data repository for areas where disciplinary repositories do not currently exist or are not feasible.

## 3. Methods for monitoring evolving costs and impacts on affected communities.

NIH proposes to actively monitor trends in publication fees and policies to ensure that they remain reasonable and equitable. This monitoring will be very important as we are concerned that the impact of the new public access policy could result in increasing publication fees in the form of Article Processing Charges (APCs), making the affordability of the costs of publishing significantly more challenging for some researchers and institutions. NIH's evidence of trends should also encompass not only fees and policies, but also monitor which institutions, disciplines, and labs have decreasing appearances in the most accessed journals to provide a more accurate picture of this effect.

Additionally, AAU emphasizes that publication fees are only one narrow measure to determine evolving costs and impacts of the NIH public access policy, and that simply monitoring trends in publication costs will not fully encapsulate this impact. We echo our colleagues at FASEB who stated in their response to NOT-OD-23-91 that the scientific peer review process required to ensure the highest standard of scientific integrity is not adequately reflected in publication fees. The human effort of oversight and compliance, long-term data access, and impacts on society journals must be considered, too.

## 4. Early input on considerations to increase findability and transparency of research.

AAU supports NIH's efforts to provide near term data points for utilizing Persistent Identifiers (PIDs) for different research products and metadata. We remain concerned, however, that without clear standards on PIDs and metadata, different approaches will inadvertently hamper accessibility and reproducibility. As NIH refines its recommendations regarding certain PID platforms and metadata storage, consistency across federal agencies will be key to effectuating more robust adoption; we applaud NIH's continued collaboration with the National Institute of Standards and Technology (NIST) and their efforts to develop a Research Data Framework. Developing and adopting standard metadata approaches could help facilitate the use of metadata across different datasets and disciplines, reducing barriers to sharing and reusing data.

The Association of Research Libraries, the California Digital Library, APLU, and AAU released a report, Implementing Effective Data Practices: Stakeholder Recommendations for Collaborative Research Support, in 2020 with recommendations for data practices supporting an open research ecosystem. AAU stands by the 2020 recommendations. The report identified five core PIDs that are fundamental and foundational to an open data ecosystem. Using these PIDs will ensure that basic metadata about research is standardized, networked, and discoverable in scholarly infrastructure:

- 1. Digital object identifiers (DOIs) to identify research data, as well as publications and
- Other outputs
- 2. Open Researcher and Contributor (ORCID) IDs to identify researchers
- 3. Research Organization Registry (ROR) IDs to identify research organization affiliations
- 4. Crossref Funder Registry IDs to identify research funders
- 5. Crossref Grant IDs to identify grants and Other types of research awards

We encourage NIH's efforts to identify and pilot a DOI system that would overlay existing NIH grant identifiers to allow for greater interoperability. NIH's current award identifiers have extremely limited utility outside of NIH. Such a DOI system should be further coordinated with Other federal agencies and affected research stakeholders. Further, the use of services and tools such as DataCite, ORCID, Crossref, figshare, and Others should be allowed as a direct cost in the grant proposal. Many of these tools require membership fees or charge fees for additional services. These entities are critical to local data management on University campuses and may require significant campus investment through direct fees or human capital.

Conclusion

AAU commends NIH's outreach and engagement with the scientific community to inform refinements to its DMS policy. A collaborative approach with stakeholders is imperative to ensure public access to federally funded research outputs. AAU strongly urges NIH to consider the creation and maintenance of discipline-specific repositories and to address the need for financial support following the end of a grant in order to allow for greater compliance with open access and data management obligations.

In addition to the specific areas delineated within NOT-OD-23-091, AAU suggests Other areas for further engagement in NIH's DMS policy: (1) longer-term costs of data to researchers and universities, (2) data interoperability challenges, (3) more clarity on researcher compliance guidance, and (4) the broad definition of "scientific data." AAU looks forward to additional opportunities for discussion.

# **Uploaded File:**

AAU-comments-NIH-RFI-Public-Access-to-Research-April-2023.pdf

**Description:** 

Email: <a href="mailto:lizbet.boroughs@aau.edu">lizbet.boroughs@aau.edu</a>



Boston University
Brandeis University
Brown University
California Institute of Technology
Carnegie Mellon University
Case Western Reserve University
Columbia University
Cornell University

Dartmouth College
Duke University
Emory University
Georgia Institute of Technology

Harvard University

Indiana University
The Johns Hopkins University
Massachusetts Institute of Technology
McGill University
Michigan State University
New York University
Northwestern University
The Ohio State University

The Pennsylvania State University
Princeton University
Purdue University
Rice University
Rutgers University – New Brunswick
Stanford University
Stony Brook University
State University of New York
Texas A&M University
Tufts University
Tufts University
Tulane University

University at Buffalo —
State University of New York
The University of Arizona
University of California, Berkeley
University of California, Irvine
University of California, Irvine
University of California, Los Angeles
University of California, San Diego
University of California, Santa Barbara
University of California, Santa Barbara
University of California, Santa Brobara
University of Colorado Boulder
University of Florida
University of Illinois, Urbana-Champaign
The University of Iowa
The University of Kansas

University of Missouri, Columbia
The University of North Carolina at
Chapel Hill
University of Oregon
University of Pennsylvania
University of Pittsburgh
University of Rochester
University of Southern California
The University of Texas at Austin
University of Toronto
The University of Utah
University of Virginia
University of Washington

The University of Wisconsin - Madison

Washington University in St. Louis

Yale University

University of Minnesota, Twin Cities

University of Michigan

To: Lyric Jorgenson, PhD

Acting Director, Office of Science Policy and

Acting NIH (National Institutes of Health) Associate Director for Science Policy

National Institutes of Health

From: Lizbet Boroughs, MSPH, Associate Director of Federal Relations

Kate Hudson, JD, Associate Vice President and Counsel

Association of American Universities

Date: April 24, 2023

RE: Comments in Response to NOT-OD-23-091, Request for Information on the NIH Plan

to Enhance Public Access to the Results of NIH-Supported Research

The Association of American Universities (AAU) thanks the National Institutes of Health for the opportunity to comment on NOT-OD-23-091, the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research. Founded in 1900, AAU is composed of America's leading research universities. AAU's 65 research universities transform lives through education, research, and innovation.

AAU strongly agrees with NIH's statements that "increasing access to publications and data resulting from federally funded research offers many benefits to the scientific community and the public," and that access "can accelerate research, generate higher quality scientific results, encourage greater scientific integrity, and enable future inquiry, discovery, and translation for NIH-supported research." Indeed, in 2021, AAU and its sister organization, the Association of Public & Land-grant Universities (APLU), published a joint <u>Guide to Accelerate Public Access to Research Data</u> to help inform our respective member institutions' activities on accessible research data. Leading up to the publication of this document, with funding from NIH and the National Science Foundation (NSF#1837847 and #1939279), AAU and APLU held a series of workshops and conferences with researchers, senior research officers, librarians, chief information officers, and organizations in support of increasing public access to research.

Given our past work and strong interest in public access, AAU is carefully monitoring various federal research agencies' implementation of the <u>August 2022 guidance</u> released by the Office of Science and Technology Policy (OSTP). Our <u>joint response</u> with APLU in January 2020 to NOT-OD-20-013 highlighted that additional specific clarification, outside the scope of the RFI, would enable robust participation and engagement by researchers and universities with <u>NIH's Data Management and Sharing Policy</u>. AAU's comments on NOT-OD-23-091 are informed by our collaborations and discussions with our members, APLU, the Association of American Medical Colleges (AAMC), the Federation of American Societies for Experimental Biology (FASEB), and the Council on Governmental Relations (COGR).

# Inequities in publishing opportunities

AAU appreciates that NIH is engaged in clarifying reasonable costs for publications that can be charged directly by individual PIs to grants. This approach should also encompass cost considerations at the broader university level. Preparations for publications are not only supported by direct costs but also pooled mechanisms such as facilities and administrative

costs, library subscriptions, and additional university support from other available revenue sources. Indeed, oversight of Data Management and Sharing (DMS) is a collaborative process and not solely the researcher's responsibility during an award's arc. Data curation; compliance with federal, state, and tribal laws; metadata requirements related to fields of study; and proper data storage are tasks that require resources and an integrated approach well beyond the individual researcher's scope of direct costs. Universities with robust financial resources, data infrastructure, and library and faculty support may have the capacity to leverage these resources to respond to the added costs involved in ensuring that the new public access requirements are met, however, many institutions and their faculty may struggle to support these additional costs.

AAU suggests that NIH could ensure data access and help minimize costs by creating and supporting one agency-wide data repository, similar to the creation of PubMed Central, to serve this purpose for publications. This would be particularly useful for areas where no current NIH-supported disciplinary repository exists. AAU also suggests that agencies create overarching disciplinary-specific repositories to ensure that universities do not create a myriad of different repositories, which will diffuse the accessibility of data access overall.

Additionally, we urge the NIH to explore ways to ensure that faculty and institutions have the means to receive support for publication and data storage costs well beyond the length of an individual grant. Without financial support after the terms of a grant, researchers and universities will be unable to comply with open access and data management standards for NIH without incurring the costs themselves, which will undoubtedly have a more significant and inequitable impact on researchers and institutions without robust research infrastructure funding.

AAU appreciates NIH's continued engagement with the community on the unanticipated costs of its DMS policy.

Improving equity in access and accessibility of publications to diverse communities and end users Ultimately, data is limited in its utility if research data stewardship is not fundamental to the research endeavor. Conceptualizing and planning for data access and interoperability is a continually iterative process involving researchers, funders, institutions, health professionals, and the public. Data technology and analysis are not stagnant, and their evolution will require flexibility within NIH's public access guidance and continual training for program officers at the individual NIH institutes.

AAU is, therefore, supportive of NIH's collaborations with scientific societies, such as FASEB's "DataWorks! Help Desk," to improve data management at the individual researcher level. AAU also strongly supports the creation of disciplinary based data repositories to improve and ensure access to federally funded research results and believes that it is important for NIH to support and facilitate the creation of such repositories. As previously stated, we also recommend the creation of one overall NIH-supported data repository for areas where disciplinary repositories do not currently exist or are not feasible.

## Methods for monitoring evolving costs and impacts on affected communities

NIH proposes to actively monitor trends in publication fees and policies to ensure that they remain reasonable and equitable. This monitoring will be very important as we are concerned that the impact of the new public access policy could result in increasing publication fees in the form of Article Processing Charges (APCs), making the affordability of the costs of publishing significantly more challenging for some researchers and institutions. NIH's evidence of trends should also encompass not only fees and policies, but also monitor which institutions, disciplines, and labs have decreasing appearances in the most accessed journals to provide a more accurate picture of this effect.

Additionally, AAU emphasizes that publication fees are only one narrow measure to determine evolving costs and impacts of the NIH public access policy, and that simply monitoring trends in publication costs will not fully encapsulate this impact. We echo our colleagues at FASEB who stated in their response to

NOT-OD-23-91 that the scientific peer review process required to ensure the highest standard of scientific integrity is not adequately reflected in publication fees. The human effort of oversight and compliance, long-term data access, and impacts on society journals must be considered, too.

# Input on considerations to increasing findability and transparency of research and effort to improve the use of PIDs and metadata

AAU supports NIH's efforts to provide near term data points for utilizing Persistent Identifiers (PIDs) for different research products and metadata. We remain concerned, however, that without clear standards on PIDs and metadata, different approaches will inadvertently hamper accessibility and reproducibility. As NIH refines its recommendations regarding certain PID platforms and metadata storage, consistency across federal agencies will be key to effectuating more robust adoption; we applaud NIH's continued collaboration with the National Institute of Standards and Technology (NIST) and their efforts to develop a Research Data Framework. Developing and adopting standard metadata approaches could help facilitate the use of metadata across different datasets and disciplines, reducing barriers to sharing and reusing data.

The Association of Research Libraries, the California Digital Library, APLU, and AAU released a report, Implementing Effective Data Practices: Stakeholder Recommendations for Collaborative Research Support, in 2020 with recommendations for data practices supporting an open research ecosystem. AAU stands by the 2020 recommendations. The report identified five core PIDs that are fundamental and foundational to an open data ecosystem. Using these PIDs will ensure that basic metadata about research is standardized, networked, and discoverable in scholarly infrastructure:

- 1. Digital object identifiers (DOIs) to identify research data, as well as publications and other outputs
- 2. Open Researcher and Contributor (ORCID) IDs to identify researchers
- 3. Research Organization Registry (ROR) IDs to identify research organization affiliations
- 4. Crossref Funder Registry IDs to identify research funders
- 5. Crossref Grant IDs to identify grants and other types of research awards

We encourage NIH's efforts to identify and pilot a DOI system that would overlay existing NIH grant identifiers to allow for greater interoperability. NIH's current award identifiers have extremely limited utility outside of NIH. Such a DOI system should be further coordinated with other federal agencies and affected research stakeholders. Further, the use of services and tools such as DataCite, ORCID, Crossref, figshare, and others should be allowed as a direct cost in the grant proposal. Many of these tools require membership fees or charge fees for additional services. These entities are critical to local data management on university campuses and may require significant campus investment through direct fees or human capital.

#### **Conclusion**

AAU commends NIH's outreach and engagement with the scientific community to inform refinements to its DMS policy. A collaborative approach with stakeholders is imperative to ensure public access to federally funded research outputs. AAU strongly urges NIH to consider the creation and maintenance of discipline-specific repositories and to address the need for financial support following the end of a grant in order to allow for greater compliance with open access and data management obligations. In addition to the specific areas delineated within NOT-OD-23-091, AAU suggests other areas for further engagement in NIH's DMS policy: (1) longer-term costs of data to researchers and universities, (2) data interoperability challenges, (3) more clarity on researcher compliance guidance, and (4) the broad definition of "scientific data." AAU looks forward to additional opportunities for discussion.

**Submit date: 4/24/2023** 

I am responding to this RFI: On behalf of an organization

Name: Beth Mathews-Bradshaw

Name of Organization: The Alliance for Aging Research

Type of Organization: Patient advocacy organization

Role: Patient advocate

1. How to best ensure equity in publication opportunities for NIH-supported investigators.

2. Steps for improving equity in access and accessibility of publications.

The Alliance for Aging Research agrees with removal of the 12-month embargo period. Accessibility to articles when first published is vitally important to patients affected by disease, particularly those relying on new research for effective therapeutics. It is also important to note that if the data is being referenced in publicly available news articles, e.g. The New York Times, patients affected by that disease should be able to access data that is the result of trials funded with taxpayer dollars. The Alliance would also like to see greater use by publishers, with access through PubMed, of allowing a free copy to patients for articles still under embargo, such as Elsevier does with its Patient Access program.

- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

The Alliance for Aging Research notes that the PubMed website is not easily used by the layperson. The FAQs and user guide sections are extremely long. The section on MeSH Terms is incredibly dense. That said, it is not hard to get search results; it is harder to know that you are getting the best results. The Alliance believes a less technical user guide specifically for the layperson would be helpful. Examples of searches illustrating how to focus results would also be beneficial.

Email: bmbradshaw@agingresearch.org

**Submit date:** 4/24/2023

I am responding to this RFI: On behalf of an organization

Name: Carter Alleman

Name of Organization: American Society for Pharmacology and Experimental Therapeutics

Type of Organization: Professional org association

Role: Member of the public

# 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Item III.D.1 notes "NIH intends to develop supplemental information that elaborates on and clarifies allowable costs for publication, consistent with these conditions." ASPET encourages NIH to include that such supplemental information covers all allowable paths for charging publishing costs, including from indirect costs and Other University general or restricted funds. ASPET also encourages NIH to include in this guidance coverage for all costs, such as open access fees, page charges, and submission fees among Other costs.

Inequities in the publishing world already exist, with those researchers at larger universities having the benefit of administrative support and scale in terms of libraries, while those in underserved areas and populations do not have the level of support at institutes to assist them with publishing. NIH should allow all avenues to be available for publication and should not limit how a grant is to be used for publications. Whether this will require an increase in the grant amount, or NIH including publication costs within the grant, is a matter for future study by NIH. However, if NIH has the goal to increase publications from these communities, NIH needs to make all efforts available and provide maximum flexibility.

NIH should also allow for flexibility and choice for both the authors and publishers in publishing research so that the appropriate reuse of articles can be determined by the author and publisher. ASPET encourages NIH to permit CC BY-NC license options that allow for the free reuse of content by the public (in line with the goals of NIH) but not for commercial purposes.

# 2. Steps for improving equity in access and accessibility of publications.

Scholarly societies, such as ASPET, are a unique partner in this area of improving equity in access and accessibility of publications. Operating simultaneously in the scientific enterprise, in education, and in business, societies can pull best practices and implement them across multiple sectors at once. However, financial support for equity efforts is lacking. With proper funding, scholarly societies would be ideal partners to improve equity in access and accessibility. Examples of practical steps that could be taken more broadly include plain language summaries, alt text for images, creating more videos, working with media on news stories, and engaging through social media. Societies are also well-situated to develop educational materials and facilitate training to support researchersand the broader diverse community on improving communication around the scientific process and a specific field of science. To facilitate this, resources from NIH could be specifically allocated to address the financial need for domain-specific experts, including scholarly societies.

## 3. Methods for monitoring evolving costs and impacts on affected communities.

ASPET recommends that NIH not monitor publication fees, which could lead to a system that favors quantity over quality. Any "one-size fits all" pricing structure which is the logical result of this type of monitoring does not enhance the publication's quality; it just streamlines the bookkeeping.

While there might be an interest in monitoring whether funded researchers are requesting more total resources in the direct versus indirect portion of the grant and resultant changes in awarded amounts over time, this would be challenging to monitor without an effective baseline. The determination of the baseline will shift as this Policy is implemented as there should be more articles published and discoveries occurring with more public access. While there are also the dangers, such as AI produced manuscripts and paper mills, that will need to be guarded against, that will also shift future baselines. Ultimately, NIH should allow the marketplace and competition between publishers to determine the reasonable publication costs.

If NIH feels there needs to be more publication avenues, there could be further discussion. However, if NIH's goal is to increase those affected communities' publication rates, NIH should work with its scientific societies to improve resources and education to allow those impacted to publish in existing journals.

Monitoring equity in funded grants will be important, as is understanding where and how the system is developing and evolving. To obtain a snapshot of the current environment and assess impact of policy changes, NIH could compare the total, median, and mean number of publication fees in the direct portion of grants for different stakeholder groups over time and as a percentage of total published articles funded by the agency.

# 4. Early input on considerations to increase findability and transparency of research.

ASPET supports NIH's commitment to engage withexisting identifier infrastructure and standards already in use across many scholarly societies. Requiring ORCID (Open Researcher and Contributor ID) for the corresponding and/or submitting author has been seamless for integration into societies' manuscript submission, peer review, and publication systems; requiring ORCID for all co-authors has posed more challenging but is improving with time. ASPET supportsNIH adoption of a DOI (Digital Object Identifier) overlay on existing grants; this activity could foster a more connected ecosystem of grants, publications, and data.

#### **Uploaded File:**

ASPET-NIH-Public-Access-Plan-RFI-Response.pdf

**Description:** ASPET Official Comment

Email: calleman@aspet.org

#### Council

#### Michael F. Jarvis

President
University of Illinois-Chicago

#### Namandjé N. Bumpus

President-Elect
U.S. Food and Drug Administration

#### Margaret E. Gnegy

Past President
University of Michigan Medical
School

### Kathryn A. Cunningham

Secretary/Treasurer
University of Texas Medical Branch

#### **Xinxin Ding**

Secretary/Treasurer-Elect University of Arizona College of Pharmacy

#### Carol L. Beck

Past Secretary/Treasurer
Thomas Jefferson University

#### Randy A. Hall

Councilor Emory University School of Medicine

#### John R. Traynor

Councilor University of Michigan

#### Nina Isoherranen

Councilor University of Washington

#### **Kenneth Tew**

Chair, Publications Committee Medical University of South Carolina

## Catherine M. Davis

FASEB Board Representative Uniformed Services University of the Health Sciences

#### Michael W. Wood

Chair, Program Committee Neupharm LLC

#### Martha I. Dávila-García

Chair, DEI Task Force Howard University College of Medicine

#### Dianicha Santana

Chair, Young Scientists Committee University of Illinois

## **David Jackson**

Executive Officer



Submitted online via RFI website on April 24, 2023.

# RE: Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research

On behalf of the members of the American Society for Pharmacology and Experimental Therapeutics (ASPET), we appreciate the opportunity to submit comments on the National Institutes of Health (NIH) Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research (NOT-OD-23-091).

ASPET is 4,000-member scientific society whose members conduct basic and clinical pharmacological research and work in academia, government, industry, and non-profit organizations. ASPET members conduct research leading to the development of new medicines and therapeutic agents to fight existing and emerging diseases. ASPET is a global pharmacology community that advances the science of drugs and therapeutics to accelerate the discovery of cures for disease. We are in constant pursuit of our Mission through research, education, innovation, and advocacy.

ASPET owns and self-publishes four journals covering a wide range of pharmacological topics. They are *The Journal of Pharmacology and Experimental Therapeutics (JPET), Pharmacological Reviews (Pharm Rev), Molecular Pharmacology (Mol Pharm),* and *Drug Metabolism and Disposition (DMD)*. ASPET co-publishes a fifth journal, *Pharmacology Research & Perspectives (PR&P)*, with the British Pharmacological Society and Wiley. All four of ASPET's wholly-owned journals continuously accept articles and have published continuously accepted articles and posted fully formatted versions as soon as they were ready. Formatted articles are freely available during a rolling five-year window, where articles are freely accessible for five years, starting 12 months after publication. After those five years are up, the articles go back to being under access control. NIH-funded articles are deposited in PubMed Central on behalf of authors by ASPET. All ASPET journals are Plan S compliant and meet NIH Data Availability requirements. Journal authors range from undergraduates, and postdoctoral students, to PhD scientists at universities, government agencies, and in industry.

NIH's proposal is commendable in its goal to allow for instantaneous access to NIH supported research publications. ASPET agrees with NIH that all involved in the scientific research enterprise are trustees of the public's funds and thus the public should have access to the research results. ASPET is concerned, though, that the NIH's goal does not account for the reality of the level of administrative burden that will occur, and the cost borne to the entire research enterprise, to achieve this goal.

Scholarly scientific societies were founded to convene researchers in a field and advance a particular branch of science. ASPET has long accomplished this goal through various means, including establishing best practices and standards, policy feedback, workforce and career development, awards and recognition, advocacy, education, and communicating advances in science through publications, conferences, and other means. We are led by and represent many of the same scientists who conduct research funded by NIH. As a nonprofit, revenues we collect are reinvested in advancing science and supporting the research community.

Before addressing NIH's interest in Section III of the NIH Plan to Enhance Public Access, ASPET would like to provide the following comments on Section II.

# **II. Scientific Data**

II.C. NIH will rely on the approaches and timelines for data sharing specified in the NIH DMS Policy. The NIH DMS Policy indicates that scientific data that are *not* associated with peer-reviewed scholarly publications should be made accessible as soon as possible, and no later than the end of the performance period for the research award.

The peer reviewed scholarly publication is the hallmark of NIH funded research. The publication has a structure that with known production values, processes, and locations. Including additional scientific data that is not associated with the peer reviewed scholarly publication for the sake of inclusion creates more administrative burden. There may be value in scientific data underlying null and negative findings, or other data that was tangential that is included in preprints, conference proceedings, or book chapters, but those decisions should be left to the professional judgment of the author as to whether they rise to the level of scientific data sufficient to be included in a peer reviewed scholarly publication. There is a high bar to meet with data that is included in the peer reviewed publication. NIH should keep its focus on peer reviewed publications and not the inclusion of all data that may exist.

The other issue that arises with the inclusion of the entire universe of data is that it becomes administratively burdensome. The current process for disseminating scholarly research is through journals with their own set of criteria for reviewing and validating data. If NIH expects that conference proceedings, book chapters, and preprints are to be included in its Public Access Plan, by which method would it like to see the material submitted? Does NIH expect its partner scientific associations, such as ASPET, to record all conference sessions and digitize all related documents and submit to NIH in a Dropbox file or is there more that needs to be done administratively to meet NIH's plan? NIH has not shown there is a need to have these additional proceedings included in the public record and thus this takes away from its intended goal of allowing public access to scholarly peer reviewed publications.

How to best ensure equity in publication opportunities for NIH-supported investigators. NIH policy already allows supported researchers to charge reasonable publishing costs - NIH seeks information on additional steps it might consider taking to ensure that proposed changes to implementation of the NIH Public Access Policy do not create new inequities in publishing opportunities or reinforce existing ones.

Item III.D.1 notes "NIH intends to develop supplemental information that elaborates on and clarifies allowable costs for publication, consistent with these conditions." ASPET encourages NIH to include that such supplemental information covers all allowable paths for charging publishing costs, including from indirect costs and other university general or restricted funds. ASPET also encourages NIH to include in this guidance coverage for all costs, such as open access fees, page charges, and submission fees among other costs.

Inequities in the publishing world already exist, with those researchers at larger universities having the benefit of administrative support and scale in terms of libraries, while those in

underserved areas and populations do not have the level of support at institutes to assist them with publishing. NIH should allow all avenues to be available for publication and should not limit how a grant is to be used for publications. Whether this will require an increase in the grant amount, or NIH including publication costs within the grant, is a matter for future study by NIH. However, if NIH has the goal to increase publications from these communities, NIH needs to make all efforts available and provide maximum flexibility.

NIH should also allow for flexibility and choice for both the authors and publishers in publishing research so that the appropriate reuse of articles can be determined by the author and publisher. ASPET encourages NIH to permit CC BY-NC license options that allow for the free reuse of content by the public (in line with the goals of NIH) but not for commercial purposes.

Steps for improving equity in access and accessibility of publications. NIH welcomes input on other steps that could be taken to improve equity in access to publications by diverse communities of users, including researchers, clinicians and public health officials, students and educators, and other members of the public.

Scholarly societies, such as ASPET, are a unique partner in this area of improving equity in access and accessibility of publications. Operating simultaneously in the scientific enterprise, in education, and in business, societies can pull best practices and implement them across multiple sectors at once. However, financial support for equity efforts is lacking. With proper funding, scholarly societies would be ideal partners to improve equity in access and accessibility. Examples of practical steps that could be taken more broadly include plain language summaries, alt text for images, creating more videos, working with media on news stories, and engaging through social media. Societies are also well-situated to develop educational materials and facilitate training to support researchers and the broader diverse community on improving communication around the scientific process and a specific field of science. To facilitate this, resources from NIH could be specifically allocated to address the financial need for domain-specific experts, including scholarly societies.

Methods for monitoring evolving costs and impacts on affected communities. NIH seeks information on effective approaches for monitoring trends in publication fees and equity in publication opportunities.

ASPET recommends that NIH not monitor publication fees, which could lead to a system that favors quantity over quality. Any "one-size fits all" pricing structure which is the logical result of this type of monitoring does not enhance the publication's quality; it just streamlines the bookkeeping.

While there might be an interest in monitoring whether funded researchers are requesting more total resources in the direct versus indirect portion of the grant and resultant changes in awarded amounts over time, this would be challenging to monitor without an effective baseline. The determination of the baseline will shift as this Policy is implemented as there should be more articles published and discoveries occurring with more public access. While there are also the dangers, such as AI produced manuscripts and paper mills, that will need to be guarded against, that will also shift future baselines. Ultimately, NIH should allow the marketplace and competition between publishers to determine the reasonable publication costs.

If NIH feels there needs to be more publication avenues, there could be further discussion. However, if NIH's goal is to increase those affected communities' publication rates, NIH should work with its scientific societies to improve resources and education to allow those impacted to publish in existing journals.

Monitoring equity in funded grants will be important, as is understanding where and how the system is developing and evolving. To obtain a snapshot of the current environment and assess impact of policy changes, NIH could compare the total, median, and mean number of publication fees in the direct portion of grants for different stakeholder groups over time and as a percentage of total published articles funded by the agency.

Early input on considerations to increase findability and transparency of research. NIH seeks suggestions on any specific issues that be considered in efforts to improve use of PIDs and metadata, including information about experiences institutions and researchers have had with adoption of different identifiers.

ASPET supports NIH's commitment to engage with existing identifier infrastructure and standards already in use across many scholarly societies. Requiring ORCID (Open Researcher and Contributor ID) for the corresponding and/or submitting author has been seamless for integration into societies' manuscript submission, peer review, and publication systems; requiring ORCID for all co-authors has posed more challenging but is improving with time. ASPET supports NIH adoption of a DOI (Digital Object Identifier) overlay on existing grants; this activity could foster a more connected ecosystem of grants, publications, and data.

#### Conclusion

ASPET commends NIH for engaging to improve the plan for public access and to develop a policy that allows researchers to comply more readily. We hope to continue the discussion and offer to work with NIH as it moves forward with its plan.

**Transforming Discoveries into Therapies** 

**Submit date: 4/24/2023** 

I am responding to this RFI: On behalf of an organization

Name: Kelsey Badger

Name of Organization: The Ohio State University

Type of Organization: University

Role: Institutional official

# 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

While we support the prioritization of public access that is inherent in the removal of embargo periods, we encourage the NIH to evaluate the increase in administrative burden this shortened timeline will place on institutions and their researchers when reporting article compliance in a timely fashion. We recommend that the NIH proactively establish additional agreements with publishers that will streamline the automatic deposit to PubMed Central.

Inequities in publishing opportunity can apply to both scholarly publication and research data. While we applaud the NIH's support of established data repositories under the DMS policy, we are concerned that gaps in existing repository infrastructure create inequities in some researchers' ability to comply with this expectation. In particular, the repository options for sharing sensitive human subjects research data are limited and often require a substantially higher cost than the options available for data that can be made openly available. The NIH has previously demonstrated leadership in developing the capacity of research data repositories, most notably through its support of the Generalist Repository Ecosystem Initiative. We encourage the NIH to continue this leadership by evaluating infrastructure gaps for the sharing of human subjects data and exploring opportunities to fund the development of the needed repositories.

We also recommend that the NIH clarify acceptable adjustments to data sharing timelines in the case of pending intellectual property claims. The lack of clear guidance on this issue places an undue burden on researchers who are making a good faith effort to comply with the DMS Policy.

# 2. Steps for improving equity in access and accessibility of publications.

The influx of publicly available research data under the NIH DMS Policy will create new opportunities for the development of interactive tools, lesson plans, and Other educational scaffolding that make data more accessible to the general public. We recommend the NIH collaborate with Other federal agencies to fully explore these opportunities for enhancing scientific and data literacies. Instructive examples include My NASA Data and the USGS Youth and Education in Science (YES) office.

### 3. Methods for monitoring evolving costs and impacts on affected communities.

We support the monitoring of article publication charges (APC) that are passed on to authors and propose that the NIH also consider extending this monitoring effort to the deposit fees that are assessed by research data repositories.

In both cases, we encourage monitoring efforts that do not rely exclusively on budget data from awarded studies. Because additional funding has not been allocated for managing data and sharing

research outputs, researchers may prioritize the use of data and article repositories that do not assess fees. As a result, the budget information from awarded studies may provide the misleading impression that preservation and public access do not require additional funding.

We strongly recommend increasing funding thresholds to account for the added costs of high-quality data management and sharing. When researchers are incentivized to find the lowest cost option for sharing data, they may not consider Other factors that are important in selecting a repository, such as whether data is curated to enhance discovery and potential for reuse.

This is a particularly critical issue when the data includes human subjects. Sufficient funding must be available for any and all necessary protections, including expert support for de-identification and the fees associated with repositories that offer disclosure risk assessment, mediated/controlled access, and the processing of legal documents such as data use agreements. It compromises the privacy and confidentiality of research participants to expect data sharing without fully funding these costs.

## 4. Early input on considerations to increase findability and transparency of research.

We applaud the supplemental funding the NIH awarded in 2022 to help existing NIH-funded data repositories increase their alignment with the OSTP Desirable Characteristics for Data Repositories. We recommend that additional cycles be considered. Moreover, we encourage the NIH to undertake an agency-wide audit of the current compliance of NIH-funded repositories with these characteristics. It is especially important to evaluate the extent to which NIH-funded repositories are currently using DOIs or comparable PIDs and to accelerate the adoption of this practice, which is essential to data discovery.

We encourage the NIH to collaborate with Other federal agencies in exploring the use of Machine Actionable Data Management Plans (maDMPs) as a strategy for increasing the findability and transparency of research outputs. Public presentations by the Research Data Alliance (RDA) have demonstrated how maDMPs can serve as important linking agents between existing systems for persistent identification of publications, datasets, authors, and institutions.

**Submit date: 4/24/2023** 

I am responding to this RFI: On behalf of an organization

Name: J. Carl Maxwell

Name of Organization: Association of American Publishers

Type of Organization: Professional org association

Role: Member of the public

1. How to best ensure equity in publication opportunities for NIH-supported investigators.

- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

# **Uploaded File:**

AAP\_Response\_NIH\_RFI\_NOT-OD-23-091\_04242023.pdf

**Description:** PDF of Association of American Publishers Response to NOT-OD-23-091, Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research

Email: cmaxwell@publishers.org

The Association of American Publishers (AAP) welcomes this opportunity to comment on the Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research (NOT-OD-23-091). AAP represents over 80 Professional and Scholarly Publishers, including dozens of scholarly societies representing over 750,000 doctors, scientists, researchers, and other members of the academy. AAP commercial publishers also publish journals on behalf of hundreds of non-profit scientific societies as well. A full list of AAP members may be found on our website: publisher.org

Scientific publishing has been a critical part of the scholarly communication ecosystem for centuries. AAP members take deep pride in the goods and services they provide that contribute to advancing science, economic prosperity, and public welfare. In addition, we are often the first line of defense in protecting scientific integrity and ensuring the public can have faith in science. Many of the advancements enabling open science are a direct result of our hard work and investment in a free, competitive marketplace, including online publication, pre-print servers, archiving, persistent identifiers, and metadata. The current discussion of open access was enabled by publisher's rapid movement online 25 years ago. These enhancements were the result of our ongoing passion to innovate in the development, discovery and dissemination of high quality, trusted reports about research for a global audience. We believe the marketplace in which we freely compete is critical for authors, the scientific community, and the quality of scholarly communication.

We appreciate that the National Institutes of Health recognizes the importance of publishers' investments in adding value to scholarly articles by seeking post publisher peer review article versions to satisfy open access requirements, as opposed to the author's original manuscript. However, for reasons we will discuss later, we suggest that it may be time for NIH to reconsider the article version necessary to satisfy statutory language and agency policy. As background to our suggestion, we would like to note the following key points about the current scholarly communication environment:

<u>First</u>, the essential services from publishers that NIH values so highly have a cost, and that cost must be <u>funded</u>. Currently, the cost of publisher services is recouped via a broad array of business models, most prominently by readers all over the world through subscriptions. The 2008 NIH open access memo and 2013 OSTP memos acknowledged the importance of this model by providing a one-year period for subscriptions before free online access occurred.

While we acknowledge OSTP's actions to adjust public access rules, and that publishers will work to address the new requirements, we nonetheless remain concerned about aspects of the memo echoed in NIH's draft policy. One area of particular concern is its call for immediate and free access to subscription manuscripts as an option to satisfy the government's public access requirements. Immediate Green open access, whereby a researcher deposits the Accepted Manuscript (AM) to a repository for free public availability, is especially concerning because free undermines the subscription, read-and-publish, subscribe to open, or similar licensing agreements on which scholarly societies and publishers depend for financial support and to produce the underlying article. Unfortunately, there is ample evidence of subscription cancellation domestically and abroad using programs such as Unsub.org and Unpaywall.org, as well as blogs tracking publisher-library negotiations, which provide subscribers with resources to evaluate and cancel subscriptions in favor of accessing open access alternatives. Assertions that trustworthy, high quality open access publications can be accomplished for free to readers, funders, and authors are simply not plausible.

Second, NIH open science policies should center and empower the researcher, enabling them to pursue their passion and knowledge, and to publish in the journal of their choice for maximum impact, without unfunded mandates and burdensome compliance regimes. AAP believes that in the pursuit of instant universal reader open access, policy makers risk losing focus on the needs of researchers, and the broader scientific community. Researchers should be able to decide how, when, and where they communicate their findings and interact with the academy and the broader public. We support broad instant public access as an important goal, but it should be achieved in a way that places less, not more, restrictions on researchers – including their freedom of choice in publication outlets and the licenses that apply to their work.

1. How to best ensure equity in publication opportunities for NIH-supported investigators. The NIH Public Access Plan aims to maintain the existing broad discretion for researchers and authors to choose how and where to publish their results. Consistent with current practice, the NIH Public Access Plan allows the submission of final published articles to PMC (in cases where a formal agreement is in place) to minimize the compliance burden on NIH-supported researchers and also maintains the flexibility of NIH-supported researchers to submit the final peer-reviewed manuscript. These submission routes are allowed regardless of whether or not the journal uses an open access model, a subscription model of publishing, or other publication model. This flexibility aims to protect against concerns that have been raised about certain publishing models potentially disadvantaging early career researchers and researchers from limited-resourced institutions or under-represented groups. NIH policy already allows supported researchers to charge reasonable publishing costs against their awards. NIH seeks information on additional steps it might consider taking to ensure that proposed changes to implementation of the NIH Public Access Policy do not create new inequities in publishing opportunities or reinforce existing ones.

AAP believes the best method for addressing issues of equity is through a vibrant, competitive, and diverse marketplace with a broad array of publishers and options for authors, including non-profit scholarly societies, university press, and dedicated open science publishers. We believe that helping researchers understand and budget for costs, combined with NIH seeking robust and sustainable funding from agency leaders and Congress, is the best way to ensure authors have a wide array of options to communicate their reports about research findings. Consistent with our vision for open science, centering on the researcher will help enhance scientific publications and ensure NIH Public Access requirements do not become a burdensome unfunded mandate.

Underfunding publication or imposing price caps could easily drive authors to predatory or other publishers who do not invest adequately in robust selection, peer-review, editing, and other enhancements. This would harm the quality and integrity of scholarly publications and potentially devastate the many U.S. scientific societies that play a critical role in publishing and supporting their scientific communities. Scientific societies rely on publishing revenues to support their communities and fund operations such as publisher owned pre-print services like SocArXiv and ChemRxiv, especially since dues-based models have struggled to keep pace with costs in recent decades. Agency imposed price controls also reduce publishers' incentive to innovate, further damaging the quality of scientific discourse.

As NIH explores costs and equity issues around open science, we suggest NIH also consider the viability of the author's original manuscript, or a pre-print version, as one path to achieving NIH's open science

goals. Pre-prints do not involve publisher additive value or intellectual property and most publishers allow and often encourage researchers to share their preprint immediately. This could be an option for NIH to pursue if budgets are constrained, and as part of the broader goal of open science. As mentioned, many publishers host and manage pre-print servers, encouraging discourse within the scientific community. We would welcome discussion with NIH about this and other options to meet its open science goals.

Where researchers are required to make the peer-reviewed version with value-add from publishers available, sustainable publication models will still be necessary to enable publishers to continue providing editorial and peer review, integrity, and quality checks, and as well as dissemination and preservation services in the long term.

2. Steps for improving equity in access and accessibility of publications. Removal of the currently allowable 12-month embargo period for NIH-supported publications will improve access to these research products for all. As noted in the NIH Public Access Plan, NIH also plans to continue making articles available in human and machine-readable forms to support automated text processing. NIH will also seek ways to improve the accessibility of publications via assistive devices. NIH welcomes input on other steps that could be taken to improve equity in access to publications by diverse communities of users, including researchers, clinicians and public health officials, students and educators, and other members of the public.

A financially robust scholarly publishing enterprise is well positioned to boost reader access. Assistive and interactive technologies can be brought to bear to empower scientists and researchers with disabilities, and AAP members are interested in partnering with NIH to explore the many ways we can lift the STEM community and achieve broad equity. Publishers already fund resources in this area, like the <a href="Access Text Network">Access Text Network</a>, and a sustainable scholarly publishing system will be able to build upon these efforts and expand accessibility.

As part of efforts to boost reader equity, it is important the agency center and empower researchers by allowing them to choose the license which best meets their needs. Broad open licenses may make sense for some researchers, while others may be concerned about undue modification, misinterpretation, or commercialization of their work. Researchers need the ability to choose the best license for their publication, including non-commercial, non-derivative versions (e.g., CC-BY-NC-ND) and we note that an open license is not necessary to use the ideas presented in scholarly communication. Early open access adopters, such as <a href="Harvard">Harvard</a> and <a href="MIT">MIT</a>, default to non-commercial licenses for their researchers and allow researchers to opt-out of funder licensing requirements.

Agency requirements restricting authors' ability to license their rights, for example through a rights retention/restriction strategy that mandates immediate Green open access, could significantly limit authors' options to bring their work to the scientific community. This is because immediate Green open access has significant potential to reduce the number of publication outlets available to authors by undermining the subscription, read-and-publish, or similar sources of funding on which they financially depend. Once that funding disappears, the publication outlet is no longer sustainable, and researchers lose an important vehicle to communicate their findings to the communities best placed to build upon them. This will only exacerbate the issues related to equity of access.

A reduction in publication outlets also decreases equity in publication opportunities and increases costs for the remaining journals. Moreover, rights restriction strategies undermine academic freedom, by allowing the government to assert broad control over the author's reports about research findings, even when these descriptions and reports are mostly funded through private sector investments and not through grants. AAP opposes policies that would grant agencies inappropriate rights in downstream copyrighted works generated through private sector investment in the peer review, editing, and publication process. We believe authors should have the freedom to decide how they assign their copyright, free from political interference. Rights restriction mandates will not eliminate the cost of publication but instead jeopardize the quality and integrity of peer reviewed publications.

3. **Methods for monitoring evolving costs and impacts on affected communities.** NIH proposes to actively monitor trends in publication fees and policies to ensure that they remain reasonable and equitable. NIH seeks information on effective approaches for monitoring trends in publication fees and equity in publication opportunities.

As competitors in an open market, publishers price fairly as they seek to publish the most informative and innovative science. Authors should have the ability to choose the way they communicate their research, free from interference, and they can balance the variety, cost, and impact of individual journals and publishers. AAP notes most publishers publicly list Article Publishing Charges on their websites as a matter of transparency, including adjustments for the type of license under which the article is published. NIH has ample resources to examine price of publication, budget, and assist researchers in developing cost by lines as part of grant proposals. See for example:

- American Chemical Society
- American Psychological Association
- Elsevier
- Wolters Kluwer
- The Public Library of Science (PLoS)
- Wiley

Alongside commitments to transparency, publishers also make a range of commitments and statements to price fairly and provide waiver programs. Taking some of the above publishers, for instance, ACS operates a Country Discount Policy; Elsevier provides waivers and information about its commitments to pricing in relation to quality, competitiveness and model uptake; and PLoS operates a range of models for open access funding support.

AAP urges NIH to collaborate with researchers to include publishing costs as part of grant applications to ensure authors have the resources to make the most authoritative version available to the widest possible audience and provide transparency for agency budgeting. We recommend that NIH develop programs to specifically fund traditionally marginalized communities and early career researchers to ensure they can bring their unique and important voices into the scholarly discourse. Where subscription-based business models are replaced by others, we recommend NIH take steps to ensure no one is left behind by boosting investment in scholarly communication.

4. **Early input on considerations to increase findability and transparency of research.** Section IV of the NIH Public Access Plan is a first step in developing the NIH's updated plan for PIDs and

metadata, which will be submitted to OSTP by December 31, 2024. NIH seeks suggestions on any specific issues that should be considered in efforts to improve use of PIDs and metadata, including information about experiences institutions and researchers have had with adoption of different identifiers.

AAP members have long been champions of PIDS and metadata as part of the overall publication process. Ultimately, creating a seamless user experience for the reader will allow access to the published article, as well as associated data, metadata, and other material. Linking to the VoR on the publisher website as part of any PMC manuscript will direct readers to the authoritative version and affiliated material and avoid disassociating the article from other material such as data and context.

**Submit date: 4/24/2023** 

I am responding to this RFI: On behalf of an organization

Name: Mary M. Langman

Name of Organization: Medical Library Association & Association of Academic Health Sciences Libraries

Type of Organization: Professional orgassociation

Role: Institutional official

## 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

The complexity of the current process for depositing publications requires significant infrastructure, training, and time that often falls on the lowest paid employees at major research institutions, especially administrative assistants, grant coordinators, and librarians. While this is already an undesirable effect of the policy on staff members at research institutions, the basic existence of these support positions privileges investigators at large research institutions over smaller institutions that primarily serve rural populations and communities of color. This is harmful to the research landscape as these constraints make it even harder to perform research that meets the needs of vulnerable populations. It is in the best interest of the scientific community and the NIH to limit the complexity of processes that fall to investigators and their support staff, but instead leverage or mandate the resources of publishers. Managing this complexity should be of primary concern when executing Section III.A.3.b.

Also see attached file.

#### 2. Steps for improving equity in access and accessibility of publications.

The NIH can improve equity in access and accessibility of publications by mandating that NIH-funded research be openly licensed for reuse, through a license such as CC-BY (Creative Commons Attribution), which unambiguously enables a variety of re-use possibilities and allows authors to retain rights to their published work. This would concretely clarify issues raised in section III.C.1, while relying on already existing legal infrastructure. Language surrounding this issue should be clear, so as to prevent publishers from taking advantage of CC-BY licenses by requiring authors to transfer their copyright to the publisher prior to assigning a CC-BY license, which is currently the practice for many publishers.

Also see attached file.

## 3. Methods for monitoring evolving costs and impacts on affected communities.

There also seems to be significant confusion about the difference between Public Access and Open Access and the ultimate goal of the NIH Public Access Policy. MLA and AAHSL recommend that the NIH clarify that while article processing charges are allowable costs for NIH awards, there are multiple options for meeting public access requirements, and gold open access publishing is only one of them. MLA and AAHSL agree with several elements of the Ivy Plus Libraries Confederation's comments (<a href="https://ivpluslibraries.org/2023/03/iplc-letter-to-the-office-of-science-technology-policy/">https://ivpluslibraries.org/2023/03/iplc-letter-to-the-office-of-science-technology-policy/</a>) on the Office of Science and Technology Policy's 2022 Memorandum, in particular their points about avoiding over-reliance on article processing charges and the importance of establishing a research dissemination infrastructure that is not the product of commercial publishing interests.

Also see attached file.

## 4. Early input on considerations to increase findability and transparency of research.

MLA and AAHSL call for the development of a robust infrastructure to ensure that NIH supports the findability of research data, potentially separate from PubMed Central, in particular the development of a single search tools to find datasets across multiple repositories. This single search could take advantage of the "Associated Data" field currently available for articles, while allowing data to be searched for directly, rather than publications that have associated data. bioCaddie's dataMED (<a href="https://datamed.org/about.php">https://datamed.org/about.php</a>) is an example of the kind of NIH-supported search interface that would be especially useful, or potentially expanding the scope of the new NCBI Datasets interface (<a href="https://www.ncbi.nlm.nih.gov/datasets/">https://www.ncbi.nlm.nih.gov/datasets/</a>) beyond genetic data.

MLA and AAHSL agree with and affirm the need for a DOI-equivalent for data, and for an infrastructure that easily links datasets to published articles consistently. NLM already has an ecosystem for linking citations (PubMed) to full text (PubMed Central); we recommend that NLM add anOther layer to link both of those to deposited data. Key to these efforts is the consistent use of standard identifiers across research disciplines, and the establishment of standard methods for citing datasets.

Also see attached file.

Also see attached file.

## **Uploaded File:**

2023\_MLA-AAHSL\_comments\_NIH-pub-access-plan.pdf

**Description:** Medical Library Association and Association of Academic Health Sciences Libraries complete set of comments Re: Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research

Medical Library Association & Association of Academic Health Sciences Libraries Comments on the NIH Public Access Plan April 24, 2023





Re: Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research

# **Notice Number:**

NOT-OD-23-091

# Data Management/Sharing

While the Request For Information specifically welcomed input on Section III ("Scholarly Publications") of the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research and requested answers to four specific questions, the Medical Library Association (MLA) and the Association of Academic Health Sciences Libraries (AAHSL) have specific input to share on Section II ("Scientific Data"), the majority of which relates to requesting additional clarity on the application of the NIH Data Management and Sharing Policy:

- Clarity and additional guidance is needed about the types and level of data included under any
  future public access policy, particularly for qualitative and image data: for example, does a
  researcher need to digitize and share images of all microscopy slides, or just the slides used to
  support the data in the publication?
- Any future public access policy that includes data must reaffirm the importance of allowing for multiple data repositories, particularly for mixed-methods research with multiple data types.
- Regarding data versioning, the concept of a "final version" of data is somewhat outdated, and
  there is a need to allow for multiple "versions of record" that can be preserved, while at the
  same time requiring metadata that clarifies which set of data was used to support any given
  publication.
- Section II.B needs clarification about how researchers should address data that might be patentable.

# Scholarly Publications

# Selection of Publication Venues

MLA and AAHSL agree with the importance of ensuring that authors retain the right to select appropriate dissemination venues for their grant-funded research. The associations also recognize that creating and maintaining an updated list of unethical journal/publishers to avoid is not viable. Likewise, maintaining a list of approved journals/publishers is restrictive and unnecessarily limiting. However, MLA and AAHSL believe that providing funded-researchers with support for identifying appropriate journals/publishers that engage in ethical publishing practices is a key role to ensure that federally-funded research has the widest reach and greatest impact.

MLA and AAHSL recommend that funders:

Medical Library Association & Association of Academic Health Sciences Libraries Comments on the NIH Public Access Plan April 24, 2023

- Establish standards for journals' publishing practices that are based not on impact metrics or reputation, but instead on ethical practices related to peer-review, access to research data, and disclosure of fees/publishing costs;
- Provide guidance to investigators about how to select journals and publishers that follow ethical publishing practices.

# PubMed Central Submission Process & Publisher Embargo Periods

MLA and AAHSL have specific concerns related to Section III.A.3.b, in particular the requirements for the NIH-funded investigator to make their own arrangements with the journal to deposit the published article to PubMed Central without an embargo:

- An increasing number of publishers no longer submit NIH-supported accepted manuscripts to PubMed Central on behalf of authors, increasing the compliance burden on individual investigators.
- This burden will be especially problematic for very large grants, such as Clinical and Translational Science Awards or training awards, potentially shifting the cost to research offices and libraries.
   The compliance burden will also be felt by individual researchers in institutions lacking a robust research administration infrastructure.

MLA and AAHSL recommend that NIH and other federal funders require that all publishers take on the responsibility of submitting manuscripts to PubMed Central on behalf of authors in order to avoid the complications and challenges listed above.

There also seems to be significant confusion about the difference between Public Access and Open Access and the ultimate goal of the NIH Public Access Policy. MLA and AAHSL recommend that the NIH clarify that while article processing charges are allowable costs for NIH awards, there are multiple options for meeting public access requirements, and gold open access publishing is only one of them. MLA and AAHSL agree with several elements of the Ivy Plus Libraries Confederation's comments (<a href="https://ivpluslibraries.org/2023/03/iplc-letter-to-the-office-of-science-technology-policy/">https://ivpluslibraries.org/2023/03/iplc-letter-to-the-office-of-science-technology-policy/</a>) on the Office of Science and Technology Policy's 2022 Memorandum, in particular their points about avoiding over-reliance on article processing charges and the importance of establishing a research dissemination infrastructure that is not the product of commercial publishing interests.

MLA and AAHSL are also concerned that one of the major changes to the current NIH Public Access Policy mandated by the OSTP, ie, the removal of the embargo period, has not been addressed in this Plan, which has already created confusion among researchers and the librarians who support them alike.

# PubMed Central Features/Functionality

Related to the organization and usability of the PubMed Central interface, MLA and AAHSL recommend that:

• Usage (reading/downloading/citation rates, etc.) of NIH-funded research outputs be displayed in PubMed Central to track their impact.

Medical Library Association & Association of Academic Health Sciences Libraries Comments on the NIH Public Access Plan April 24, 2023

- PubMed Central's interface should include the option for making the PDF-version of a manuscript the default view, as this is common on publisher websites and is generally considered more readable.
- PubMed Central should prioritize making data/text mining of its contents easily accessible to researchers, facilitating sharing and supporting reuse.

# Copyright/Creative Commons Licenses

MLA and AAHSL advocate for federal adoption of Creative Commons Licenses as an option for maintaining authorship rights. This license is already widely used, has legal backing in place, and would require low effort for journals to implement. It would also further NIH's goals for increasing equity.

# **Preprints**

MLA and AAHSL acknowledge the increasingly crucial role of preprints in accelerating scientific discovery, and applauded the NIH Preprint Pilot as part of NIH's efforts to combat the COVID-19 pandemic. MLA and AAHSL recommend that preprints be considered in any future public access policies as a legitimate research product that is a viable interim means of compliance with public access policy mandates.

# Mechanism to Increase Findability and Transparency of Research

MLA and AAHSL call for the development of a robust infrastructure to ensure that NIH supports the findability of research data, potentially separate from PubMed Central, in particular the development of a single search tool to find datasets across multiple repositories. This single search could take advantage of the "Associated Data" field currently available for articles, while allowing data to be searched for directly, rather than publications that have associated data. bioCaddie's dataMED (<a href="https://datamed.org/about.php">https://datamed.org/about.php</a>) is an example of the kind of NIH-supported search interface that would be especially useful, or potentially expanding the scope of the new NCBI Datasets interface (<a href="https://www.ncbi.nlm.nih.gov/datasets/">https://www.ncbi.nlm.nih.gov/datasets/</a>) beyond genetic data.

MLA and AAHSL agree with and affirm the need for a DOI-equivalent for data, and for an infrastructure that easily links datasets to published articles consistently. NLM already has an ecosystem for linking citations (PubMed) to full text (PubMed Central); we recommend that NLM add another layer to link both of those to deposited data. Key to these efforts is the consistent use of standard identifiers across research disciplines, and the establishment of standard methods for citing datasets.

**Submit date: 4/24/2023** 

I am responding to this RFI: On behalf of an organization

Name: Andrew Bostjancic

Name of Organization: Taylor and Francis Group

Type of Organization: Other

Type of Organization-Other: Academic Publisher

Role: Member of the public

## 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

T&F is committed to delivering a range of publishing options and content types that are inclusive, holistic and provide opportunities for researchers working across career stages and disciplines. We are keen to continually develop approaches to ensure equity and diversity in publication opportunities and we know that this requires input and collaboration between multiple stakeholders from across the scholarly ecosystem. Specifically, it requires publishers to help researchers and more marginalized communities across career stage by providing training to navigate the publishing landscape - and understand the options available. It requires funders to investigate their processes for grant selections, so that grant opportunities are not exclusively awarded to the same highly resourced researchers and institutions. University efforts to expand opportunities through institutional grants can help to reduce inequities and provide a diversity of voices. Collaborative commitment to tackle the challenge of increasing equity from diverse stakeholders ensures that all knowledge makers are given the opportunity to contribute, irrespective of race, ethnicity, gender, geography, language, discipline, or funding source. NIH can be a leader in convening these stakeholders to help discuss ways to broaden equity.

T&F is committed to equity in publication opportunities and has taken the following steps to answer this call to action.

T&F offers over 300 dedicated OA journals, and more than 95% of our venues offer an OA pathway. We work with authors to find the best home for their work. Across our portfolio we also offer an increasing number of tailored fully open access publishing venues which increase the opportunities for researchers to publish research outside of more selective venues, and ensures that regardless of the results (e.g. negative, null, incremental research), there is an outlet for researchers to make their findings discoverable and accessible to all. This includes our 'open research' publishing venues provided by F1000. The F1000 publishing model combines the speed of preprints with the benefits of full publication. This includes functionality that ensures the robustness, quality, and transparency of research using rigorous editorial checks, open data, and invited open peer review. Authors are given autonomy throughout the entire publishing process.

Publishing venues that operate on this basis help to remove the barriers to publication that many researchers face, particularly early career researchers, and are entirely aligned with the DORA principles. T&F is signed up and committed to the DORA declaration, the Managing Director of our imprint F1000, is a member of DORA's Board of Advisers - and through this we are developing ways to support

researchers across all career stages and disciplines to share the outputs of their research in more transparent and accessible ways.

In addition to providing more trusted and reputable routes for researchers to publish their work, our role as a publisher is to support initiatives to build capacity and skills to help deliver trust and value in the research we receive and publish. An example of this is when in 2019, T&F launched the 'Excellence in Peer Review: Taylor & Francis Reviewer Training Network'. This aims to provide clear practical advice to researchers to improve the quality of their reviews and introduce the key principles to early career researchers and researchers from under-represented groups. This initiative encourages greater inclusion and participation in peer review.

We support the initiative for Transforming Institutions by Gendering contents and Gaining Equality in Research (TRIGGER). This aims to understand and address the causes behind under-representation of women in Science, Technology, Engineering, Mathematics, and Medicine (STEMM) subjects.

We were also the 2021 publisher winner of ABC International Excellence Award for Accessible Publishing, recognized by the Award's judges for an "innovative approach to alternative text for images, graphs, and diagrams."

T&F believes in the importance of public access to amplify and communicate research that delivers change and improves lives. We would like to encourage the NIH to collaborate actively with publishers to ensure we are positioned to provide the services that are needed to drive equity and access to research.

#### **Question 1 Recommendations**

- 1. Convene a cross-stakeholder discussion/s to refine NIH's requirements and ensure implementation of the plan in the most optimum way to deliver equity.
- 2. Continue active collaboration with the academic publishing community to elicit feedback on the implementation of the plan and provide a route for us to share the global and disciplinary specific feedback we receive around access and equity issues.

## 2. Steps for improving equity in access and accessibility of publications.

Ensuring all functionality and content is accessible to all people is a laudable ambition. Developing clear guidelines for formatting with a focus on accessibility will improve access for everyone. One of the primary roles of publishers is to transform content from authors into a final product through typesetting and copyediting. This labor-intensive effort alongside the creation and sharing of article metadata is critical for making content machine readable and discoverable.

Across the company, T&F is developing new formats and tailored views of research that are designed to support access, use, and reuse of research. One of the emerging tools is the implementation of Plain Language Summaries (PLSs). These additional abstracts allow us to succinctly summarize the key points from a piece of scientific research to a non-technical audience. Creating PLSs tailored views of content is an important way to increases access, engagement in research content and findings to the various communities and stakeholders who are the ultimate users of research, including policymakers, students, educators, and the public.

Through our society partners, funders and Other expert community links, we have a wealth of experience in developing research access and dissemination strategies and solutions. By collaborating alongside knowledge creators and federal agencies, publishers can create models and formats that are designed to deliver the requirements of our stakeholders. Emerging scientific innovations require training for authors to remain at the forefront of their fields. T&F works alongside our expert academic editors and societies and we have a depth of experience in providing research communication, sharing, and dissemination training to researchers across the career stages and across disciplines e.g. How to manage and share data; How to publish for reach and impact; How to peer review effectively. We are willing and able to support the NIH in providing training to its various cohorts of grantees.

We provide guidance and best practice to our authors and editorial boards to ensure that content is published with adherence to various accessibility standards. For example, we have in-house experts who can provide authors with a guide to alternative text so that they can provide the best descriptions. We also provide content in a variety of formats including PDF, ePub2, ePub3, and HTML formats to expand equity and accessibility. T&F has adopted this practice and works to provide a suitable format - we provide these formats on request from individuals and institutions.

In 2022, T&F brought on our first Accessibility Officer to provide oversight, coordination, guidance, and leadership to the organization's Accessibility Working Group. This addition has already provided the organization with a more effective and efficient accessibility strategy. If not done so already, the NIH could consider appointing staff resources with specific remit and responsibility for ensuring accessibility.

#### Question 2 Recommendations

- 1. Provide training to grantees on key aspects of how best to communicate and disseminate research in ways that ensure compliance of NIH requirements. Ensure awareness of best practice and standards to support discoverability and access.
- 2. Collaborate with publishers to develop more tailored research use-focused findings and output to maximize the potential for research to reach its target audience/s.
- 3. Create guides encouraging the use of alternative text for visual or print impaired individuals.
- 4. Appoint staff resources to support NIH Accessibility requirements.

## 3. Methods for monitoring evolving costs and impacts on affected communities.

T&F acts as a responsible steward with the funds we receive from researchers in return for the publishing services that we provide. Our role is to preserve academic freedom and provide routes to share, disseminate, and deliver impact from research. We provide options for researchers working across all career stages and disciplines to reach their intended audiences and their communities of interest and help build careers and research capacity. We do not support blunt measures and restrictions on where researchers can publish - instead preferring to develop solutions collaboratively to deliver sustainable publishing solutions that preserve academic freedom and choice, while maximizing the reach, access and potential impact of research.

When calculating prices for APCs, T&F aims to be transparent with our costs and mitigate inequities with our stakeholders. We continue to balance this transparency with market considerations and remain compliant with U.S. antitrust price fixing laws. List price APCs across T&F journals range from US \$600 to

US \$4,800. The list price APC is reviewed at least annually across journals and varies across several factors, including:

Funding available for the journal: this varies by discipline. Additionally, some journals are supported through grants, typically from their owning society, meaning charges are subsidized.

Impact: highly selective journals typically charge higher APCs. The APC on the accepted article also covers the work and analysis put into rejected content.

Discipline: we set APCs based on funding patterns within the field, as well as benchmarking against APCs on related journals to ensure that rates are realistic and equitable among communities.

Demographics of submissions / publications: considering the geography of submissions allows us to price fairly to market.

The type of research output: shorter article types and non-traditional formats typically incur lower APCs.

It should be noted that many customers do not pay the list price APC, benefitting from flexible funding options including:

- Discounts of up to 100% where a professional member association or learned society provide additional support.
- Discounts due to their organization's participation in a membership scheme or transformative agreement, which usually allow researchers to submit without any individual payment on their part.

T&F is committed to cost transparency and providing our published authors with world class services so that their work can have the greatest impact on society.

#### **Recommendations Question 3**

- 1. Empower authors to make the decisions for disseminating their research.
- 2. Provide training materials for authors and grant managers to collaborate on finding the best route to publish.

## 4. Early input on considerations to increase findability and transparency of research.

We are entirely aligned to support any push that the NIH has in promoting the use and integration of persistent identifiers (PIDs), research descriptors, and metadata into grant and publishing workflows.

PIDs and associated metadata are the essential foundation blocks to enable the discoverability and access of research and its findings. Like many publishers, T&F is a member of Crossref and ensures high quality metadata around all the research it publishes; we are also building our capabilities for inclusion of funding and grant data associated with articles by utilizing the Crossref Funding Registry.

Several funding agencies are also now members of Crossref (e.g. Wellcome) and register DOIs for all their awarded grants. By assigning a PID (e.g. a DOI) to its grants, the NIH would provide an identifier that can be captured by publishers in the article submission workflow and thereby allow grant output connections to be made and greatly simplify impact-related (and ROI) tracking for the NIH.

Adding grant IDs would add new information into this network of PIDs and provide increased transparency and create the possibility for robust ROI calculations for funders. This wider network of PIDS would include:

- Researcher IDs e.g ORCiD
- Institution IDs e.g. ROR or Ringgold
- Funder IDs e.g. FundRef
- Project IDs e.g. RAID
- Research object IDs e.g. DOIs for publications, data, preprints, code and Other outputs

Adding all (or a selection) of these PIDs into the metadata of research articles and objects stored in Other online locations (e.g. data repositories) will ensure progress to a more machine-readable ecosystem to enable analysis and ROI for funders. Most of the PID issuing agencies - ORCID, Crossref, Datacite, RRIDs - operate on a not-for-profit basis and are the commonly used standards across the research system. To support the simple capture of relevant research and researcher meta-data in its grant workflows, we recommend the NIH consider:

- Providing integrated links
- Drop-down lists
- APIs to Other websites

#### **Recommendations Question 4**

- 1. Align with Other funders to assign common PIDs for NIH grants consider using the established framework provided by the Crossref Funding Registry.
- 2. Utilize current and prevalent PID infrastructure where possible to avoid creating additional learnings (and need for interoperability building) for researchers.
- 3. Adopt researcher-centered practices to capture key descriptive information using auto-complete/ integrated links, drop-down lists, and APIs to Other websites to keep simple, avoid manual entry, and ensure accurate completion; include PIDs assignment for grant-related information in existing NIH systems/those used by its researchers where possible.
- 4. Monitor and adopt industry and global standards and best practices where applicable.

#### **Uploaded File:**

TF-NIH-Public-Access-Plan-RFI-Comments.pdf

Description: Full comments plus additional compliance clarification request

Email: andrew.bostjancic@taylorandfrancis.com



April 24th, 2023

National Institutes of Health Office of the Director 9000 Rockville Pike Bethesda, Maryland 20892

**RE:** Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research

**Notice Number: NOT-OD-23-091** 

Taylor & Francis is a leading global research publisher, focused on advancing science and fostering human progress through knowledge – something we've been doing since 1798. Across the organization, we provide a wide range of publishing outlets for scholarly research, including books, eBooks, journals, and open research publishing venues. We are committed to expanding the range of fully open access publishing options across our portfolio. We partner with over 150 US-based learned societies and expert associations to make research available to the communities they serve.

We appreciate the opportunity to provide comment on NIH's Public Access Plan and offer the following recommendations:

# **Question 1 Recommendations**

- 1. Convene a cross-stakeholder discussion/s to refine NIH's requirements and ensure implementation of the plan in the most optimum way to deliver equity.
- 2. Continue active collaboration with the academic publishing community to elicit feedback on the implementation of the plan and provide a route for us to share the global and disciplinary specific feedback we receive around access and equity issues.

## **Question 2 Recommendations**

- 1. Provide training to grantees on key aspects of how best to communicate and disseminate research in ways that ensure compliance of NIH requirements. Ensure awareness of best practice and standards to support discoverability and access.
- 2. Collaborate with publishers to develop more tailored research and more focused findings to maximize the potential for research to reach its target audience/s.
- 3. Create guides encouraging the use of alternative text for visual or print impaired individuals.
- **4.** Appoint staff resources to support NIH Accessibility requirements

#### **Question 3 Recommendations**

- 1. Empower authors to make the decisions for disseminating their research.
- **2.** Provide training materials for authors and grant managers to collaborate on finding the best route to publish.

# **Question 4 Recommendations**

- **1.** Align with other funders to assign common PIDs for NIH grants consider using the established framework provided by the Crossref Funding Registry.
- **2.** Utilize current and prevalent PID infrastructure where possible to avoid creating additional learnings (and need for interoperability building) for researchers.
- **3.** Adopt researcher-centered practices to capture key descriptive information using autocomplete/ integrated links, drop-down lists, and APIs to other websites to keep simple, avoid manual entry, and ensure accurate completion; include PIDs assignment for grant-related information in existing NIH systems/those used by its researchers where possible.
- **4.** Monitor and adopt industry and global standards and best practices where applicable.

#### **Full Comments**

1. How to best ensure equity in publication opportunities for NIH-supported investigators.

T&F is committed to delivering a range of publishing options and content types that are inclusive, holistic and provide opportunities for researchers working across career stages and disciplines. We are keen to continually develop approaches to ensure equity and diversity in publication opportunities and we know that this requires input and collaboration between multiple stakeholders from across the scholarly ecosystem. Specifically, it requires publishers to help researchers and more marginalized communities across career stage by providing training to navigate the publishing landscape – and understand the options available. It requires funders to investigate their processes for grant selections, so that grant opportunities are not exclusively awarded to the same highly resourced researchers and institutions. University efforts to expand opportunities through institutional grants can help to reduce inequities and provide a diversity of voices. Collaborative commitment to tackle the challenge of increasing equity from diverse stakeholders ensures that all knowledge makers are given the opportunity to contribute, irrespective of race, ethnicity, gender, geography, language, discipline, or funding source. NIH can be a leader in convening these stakeholders to help discuss ways to broaden equity.

T&F is committed to equity in publication opportunities and has taken the following steps to answer this call to action.

T&F offers over 300 dedicated OA journals, and more than 95% of our venues offer an OA pathway. We work with authors to find the best home for their work. Across our portfolio we also offer an increasing number of tailored fully open access publishing venues which increase the opportunities for researchers to publish research outside of more selective venues, and ensures that regardless of the results (e.g. negative, null, incremental research), there is an outlet

for researchers to make their findings discoverable and accessible to all. This includes our 'open research' publishing venues provided by F1000. The F1000 publishing model combines the speed of preprints with the benefits of full publication. This includes functionality that ensures the robustness, quality, and transparency of research using rigorous editorial checks, open data, and invited open peer review. Authors are given autonomy throughout the entire publishing process.

Publishing venues that operate on this basis help to remove the barriers to publication that many researchers face, particularly early career researchers, and are entirely aligned with the DORA principles. T&F is signed up and committed to the DORA declaration, the Managing Director of our imprint F1000, is a member of DORA's Board of Advisers – and through this we are developing ways to support researchers across all career stages and disciplines to share the outputs of their research in more transparent and accessible ways.

In addition to providing more trusted and reputable routes for researchers to publish their work, our role as a publisher is to support initiatives to build capacity and skills to help deliver trust and value in the research we receive and publish. An example of this is when in 2019, T&F launched the 'Excellence in Peer Review: Taylor & Francis Reviewer Training Network'. This aims to provide clear practical advice to researchers to improve the quality of their reviews and introduce the key principles to early career researchers and researchers from under-represented groups. This initiative encourages greater inclusion and participation in peer review.

We support the initiative for Transforming Institutions by Gendering contents and Gaining Equality in Research (TRIGGER). This aims to understand and address the causes behind underrepresentation of women in Science, Technology, Engineering, Mathematics, and Medicine (STEMM) subjects.

We were also the 2021 publisher winner of ABC International Excellence Award for Accessible Publishing, recognized by the Award's judges for an "innovative approach to alternative text for images, graphs, and diagrams."

T&F believes in the importance of public access to amplify and communicate research that delivers change and improves lives. We would like to encourage the NIH to collaborate actively with publishers to ensure we are positioned to provide the services that are needed to drive equity and access to research.

# **Question 1 Recommendations**

- 1. Convene a cross-stakeholder discussion/s to refine NIH's requirements and ensure implementation of the plan in the most optimum way to deliver equity.
- 2. Continue active collaboration with the academic publishing community to elicit feedback on the implementation of the plan and provide a route for us to share the global and disciplinary specific feedback we receive around access and equity issues.

2. Steps for improving equity in access and accessibility of publications. NIH welcomes input on other steps that could be taken to improve equity in access to publications by diverse communities of users, including researchers, clinicians and public health officials, students and educators, and other members of the public.

Ensuring all functionality and content is accessible to all people is a laudable ambition. Developing clear guidelines for formatting with a focus on accessibility will improve access for everyone. One of the primary roles of publishers is to transform content from authors into a final product through typesetting and copyediting. This labor-intensive effort alongside the creation and sharing of article metadata is critical for making content machine readable and discoverable.

Across the company, T&F is developing new formats and tailored views of research that are designed to support access, use, and reuse of research. One of the emerging tools is the implementation of Plain Language Summaries (PLSs). These additional abstracts allow us to succinctly summarize the key points from a piece of scientific research to a non-technical audience. Creating PLSs tailored views of content is an important way to increases access, engagement in research content and findings to the various communities and stakeholders who are the ultimate users of research, including policymakers, students, educators, and the public.

Through our society partners, funders and other expert community links, we have a wealth of experience in developing research access and dissemination strategies and solutions. By collaborating alongside knowledge creators and federal agencies, publishers can create models and formats that are designed to deliver the requirements of our stakeholders. Emerging scientific innovations require training for authors to remain at the forefront of their fields. T&F works alongside our expert academic editors and societies and we have a depth of experience in providing research communication, sharing, and dissemination training to researchers across the career stages and across disciplines e.g. How to manage and share data; How to publish for reach and impact; How to peer review effectively. We are willing and able to support the NIH in providing training to its various cohorts of grantees.

We provide guidance and best practice to our authors and editorial boards to ensure that content is published with adherence to various accessibility standards. For example, we have in-house experts who can provide authors with a guide to alternative text so that they can provide the best descriptions. We also provide content in a variety of formats including PDF, ePub2, ePub3, and HTML formats to expand equity and accessibility. T&F has adopted this practice and works to provide a suitable format – we provide these formats on request from individuals and institutions.

In 2022, T&F brought on our first Accessibility Officer to provide oversight, coordination, guidance, and leadership to the organization's Accessibility Working Group. This addition has already provided the organization with a more effective and efficient accessibility strategy. If not done so already, the NIH could consider appointing staff resources with specific remit and responsibility for ensuring accessibility.

# **Question 2 Recommendations**

- 1. Provide training to grantees on key aspects of how best to communicate and disseminate research in ways that ensure compliance of NIH requirements. Ensure awareness of best practice and standards to support discoverability and access.
- **2.** Collaborate with publishers to develop more tailored research use-focused findings and output to maximize the potential for research to reach its target audience/s.
- 3. Create guides encouraging the use of alternative text for visual or print impaired individuals.
- **4.** Appoint staff resources to support NIH Accessibility requirements.
- 3. Methods for monitoring evolving costs and impacts on affected communities. NIH proposes to actively monitor trends in publication fees and policies to ensure that they remain reasonable and equitable. NIH seeks information on effective approaches for monitoring trends in publication fees and equity in publication opportunities.

T&F acts as a responsible steward with the funds we receive from researchers in return for the publishing services that we provide. Our role is to preserve academic freedom and provide routes to share, disseminate, and deliver impact from research. We provide options for researchers working across all career stages and disciplines to reach their intended audiences and their communities of interest and help build careers and research capacity. We do not support blunt measures and restrictions on where researchers can publish – instead preferring to develop solutions collaboratively to deliver sustainable publishing solutions that preserve academic freedom and choice, while maximizing the reach, access and potential impact of research.

When calculating prices for APCs, T&F aims to be transparent with our costs and mitigate inequities with our stakeholders. We continue to balance this transparency with market considerations and remain compliant with U.S. antitrust price fixing laws. List price APCs across T&F journals range from US \$600 to US \$4,800. The list price APC is reviewed at least annually across journals and varies across several factors, including:

**Funding available for the journal:** this varies by discipline. Additionally, some journals are supported through grants, typically from their owning society, meaning charges are subsidized. **Impact:** highly selective journals typically charge higher APCs. The APC on the accepted article also covers the work and analysis put into rejected content.

**Discipline:** we set APCs based on funding patterns within the field, as well as benchmarking against APCs on related journals to ensure that rates are realistic and equitable among communities.

**Demographics of submissions / publications**: considering the geography of submissions allows us to price fairly to market.

**The type of research output:** shorter article types and non-traditional formats typically incur lower APCs.

It should be noted that many customers do not pay the list price APC, benefitting from flexible funding options including:

• Discounts of up to 100% where a professional member association or learned society provide additional support.

 Discounts due to their organization's participation in a membership scheme or transformative agreement, which usually allow researchers to submit without any individual payment on their part.

T&F is committed to cost transparency and providing our published authors with world class services so that their work can have the greatest impact on society.

# **Recommendations Question 3**

- 1. Empower authors to make the decisions for disseminating their research.
- **2.** Provide training materials for authors and grant managers to collaborate on finding the best route to publish.
- 4. Early input on considerations to increase findability and transparency of research. Section IV of the NIH Public Access Plan is a first step in developing the NIH's updated plan for PIDs and metadata, which will be submitted to OSTP by December 31, 2024. NIH seeks suggestions on any specific issues that should be considered in efforts to improve use of PIDs and metadata, including information about experiences that institutions and researchers have had with adoption of different identifiers.

We are entirely aligned to support any push that the NIH has in promoting the use and integration of persistent identifiers (PIDs), research descriptors, and metadata into grant and publishing workflows.

PIDs and associated metadata are the essential foundation blocks to enable the discoverability and access of research and its findings. Like many publishers, T&F is a member of Crossref and ensures high quality metadata around all the research it publishes; we are also building our capabilities for inclusion of funding and grant data associated with articles by utilizing the Crossref Funding Registry.

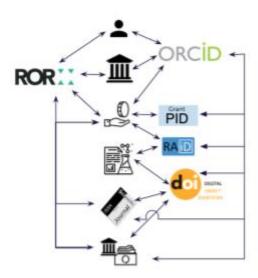
Several funding agencies are also now members of Crossref (e.g. Wellcome) and register DOIs for all their awarded grants. By assigning a PID (e.g. a DOI) to its grants, the NIH would provide an identifier that can be captured by publishers in the article submission workflow and thereby allow grant output connections to be made and greatly simplify impact-related (and ROI) tracking for the NIH.

Adding grant IDs would add new information into this network of PIDs and provide increased transparency and create the possibility for robust ROI calculations for funders. This wider network of PIDS would include:

- Researcher IDs e.g ORCiD
- Institution IDs e.g. ROR or Ringgold
- Funder IDs e.g. FundRef
- Project IDs e.g. RAID
- Research object IDs e.g. DOIs for publications, data, preprints, code and other outputs

Adding all (or a selection) of these PIDs into the metadata of research articles and objects stored in other online locations (e.g. data repositories) will ensure progress to a more machine-readable ecosystem to enable analysis and ROI for funders. Most of the PID issuing agencies – ORCID, Crossref, Datacite, RRIDs – operate on a not-for-profit basis and are the commonly used standards across the research system. To support the simple capture of relevant research and researcher meta-data in its grant workflows, we recommend the NIH consider:

- Providing integrated links
- Drop-down lists
- APIs to other websites



An approximate representation of a PID-enabled research information workflow. This image is based on the workflows described in Developing a persistent identifier roadmap for open access to UK research <a href="http://repository.jisc.ac.uk/id/eprint/7840">http://repository.jisc.ac.uk/id/eprint/7840</a>

# **Recommendations Question 4**

- 1. Align with other funders to assign common PIDs for NIH grants consider using the established framework provided by the Crossref Funding Registry.
- **2.** Utilize current and prevalent PID infrastructure where possible to avoid creating additional learnings (and need for interoperability building) for researchers.
- **3.** Adopt researcher-centered practices to capture key descriptive information using autocomplete/ integrated links, drop-down lists, and APIs to other websites to keep simple, avoid manual entry, and ensure accurate completion; include PIDs assignment for grant-related information in existing NIH systems/those used by its researchers where possible.
- **4.** Monitor and adopt industry and global standards and best practices where applicable.

## **Additional Public Access Plan Feedback**

III.A.3.b. Final published article submission: the NIH-supported investigator will be expected to either ensure the final peer-reviewed manuscript is submitted to PMC upon acceptance for publication, to be made publicly available at the time of publication or arrange with the journal to deposit the individual published article to PMC without a post-publication embargo.

T&F requests clarification of this guidance to ensure compliance with the F1000 publishing model. According to the guidance, "the final peer-reviewed manuscript is submitted to PMC upon acceptance for publication."

F1000 operates an open research, post-publication peer review model across all its publishing venues (and provides publishing services for a number of funding agencies including the Bill and Melinda Gates Foundation, European Commission and Wellcome). Like more traditional journals, once content is approved through the F1000 peer review process, a 'final peer reviewed manuscript' is created and indexed broadly, including in PubMed and PMC. Our F1000 publishing model was designed to be entirely compliant with OA and open data requirements and mandates of organizations focused on driving open access and delivering research more broadly (including alignment with NIH policies).

We request that the NIH Public Access Plan does not inadvertently create ambiguity or exclude content from reputable and recognized publishers, such as F1000, who are operating a different model to deliver fully OA, peer reviewed and validated content to PMC.

Thank you for considering our recommendations. We are committed to working with NIH to produce and implement a Public Access Plan that brings benefits to all stakeholders in the research system.

We would very much welcome the opportunity to discuss our findings and recommendations with you at your earliest convenience.

Kind regards,

Andrew Bostjancic

andrew Bostjancie

Open Research Policy and External Affairs Manager

Taylor & Francis

**Submit date:** 4/24/2023

I am responding to this RFI: On behalf of an organization

Name: Holly Sue Zullo

Name of Organization: Huntsman Cancer Institute

Type of Organization: University

Role: Institutional official

- 1. How to best ensure equity in publication opportunities for NIH-supported investigators.
- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

# **Uploaded File:**

HCI-Response-to-NOT-OD-23-091.docx

Email: holly.zullo@hci.utah.edu

# Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research

April 24, 2023

Re: NOT-OD-23-091

Dear NIH Program Officials,

Thank you for the opportunity to comment on the NIH plan to enhance public access to the results of NIH-supported research. Below we list the particular areas in which input was sought, along with our responses.

# 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Please see some comments below.

# 2. Steps for improving equity in access and accessibility of publications.

• We feel that PMCentral is generally working well. However, it is really critical to minimize the work that PIs have to do to submit or make these accessible.

# 3. Methods for monitoring evolving costs and impacts on affected communities.

- We've heard concerns from investigators about the requirements to publish open access, or
  the likely impact of a requirement of making publications immediately accessible. There is
  concern that journals will automatically increase publication fees if they are required to make
  their publications immediately available (because their business model is disrupted if they don't
  have the ability to charge for publications, their likely (perhaps only?) response will be to
  increase publication fees). These fees are often already in the \$3,000-\$12,000 range per
  publication.
- Labs now have to make difficult decisions regarding publications, e.g., whether to publish a
  finding immediately, or combine it with other findings often up to a year later just to afford
  the publication fees. This is particularly problematic when we try to support our trainees we
  often have undergraduate students publish manuscripts as first authors, but this will just
  become too expensive. Overall, there is concern that immediate publication will change the
  market, resulting in increases in publication fees, fewer publications, and less timely
  publications and certainly fewer publications of trainees.
- This may also apply specifically to US investigators (e.g., if journals decide to charge more for investigators who have the 'immediate accessible' requirement, then it will limit their abilities to publish, compared to, for example, Chinese groups)

# Some additional comments below:

- One of our investigators reported that her last two open access papers (Moonshot projects required to be published without embargo, as the new policy now requires for all NIH-funded studies) cost \$9,500 and \$11,500, respectively, to be published with immediate open access. Although one can budget for these astronomical expenses in grants, this just adds to the everrising costs of research supplies and personnel with no concomitant increase in the budget limit for research project grants from NIH. People who can't afford this will be forced to publish in less well-respected journals, which may affect their careers. This could disproportionally affect URM or more junior investigators.
- Requiring release of unpublished data by the end of a grant period is inequitable to smaller
  labs or those with fewer resources who may not have the staff or other resources needed to
  finish analysis by the end of the grant. Data are often being collected right up until the end of
  the grant period, and then still need to be analyzed. Forcing these data to be posted
  immediately, whether published or not, will allow 'big fish' (or anyone) to swarm in, take the
  data, and scoop the younger/less advanced investigator who did all the work generating the

- data. It would be more reasonable to allow a period of time for ongoing analysis (2-3 years), if applicable, before requiring the data be posted.
- Overall, the release of unpublished data immediately at the end of a grant period is unrealistic, considering that there is often a time lag in publication, especially if postdocs or trainees are involved. This may scoop some of our trainees who often have to write a thesis before they can publish.
- We have concerns about preliminary or less-than-fully-analyzed data being posted and incorrect conclusions drawn because the data have not been peer-reviewed.
- 4. Early input on considerations to increase findability and transparency of research.
  - No comments.

**Submit date:** 4/24/2023

I am responding to this RFI: On behalf of an organization

Name: Jennifer Regala

Name of Organization: American Urological

Type of Organization: Professional org association

Role: Institutional official

## 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Responses to this question, as given by the collective letters the American Urological Association (AUA) signed as detailed in our introduction letter (attached as a PDF to the RFI), state our position well and in detail. The AUA does want to emphasize these very important points:

- Peer review adds value to our high-impact research. Our flagship journal, The Journal of Urology®, is more than 100 years old, and is highly regarded in urology and in medicine at large. This peer review comes with a high financial and intellectual cost, and we ask that the NIH consider this considerable investment by the AUA when studying financial realities of implementing Open Access policies.
- We widely and generously distribute our research despite subscription paywalls. Non-subscribers around the world have a plethora of ways to absorb our research in a multitude of valuable formats, from podcasts to social media to author-written insights of articles.

CMSS has asked, and we echo this request, for a 2-year delay to the mandate so all stakeholders can work together to develop sustainable policies focused on reliable, equitable, high-quality scientific content. We agree with and want to reinforce an important point in the CMSS letter:

"Policies that restrict publishers' abilities to collaborate with authors to realize their protection of rights under United States copyright law would further limit revenue streams on which we depend, including royalties, licensing, reprints, and advertising. We urge the NIH not to include rights retention language or license requirements in the final policy Other than the grantee's right to deposit the manuscript. Preserving a Green OA route presents a sustainable business model that should be embraced. Expanding rights retention policies beyond the deposition of the manuscript would also erode the publisher's ability to monitor usage of the content in support of the author's intellectual property."

#### 2. Steps for improving equity in access and accessibility of publications.

- We agree with the fundamental importance of accessibility to premiere urological AUA research publications. Green Open Access remains a viable solution to this question of access.
- We do innovate in deliverability and accessibility of our content as a main goal of each of our publications.
- We strive to deliver our content in a way that eliminates or at least avoids dissemination and promotion of misinformation, which we believe will be diminished if "open" research is posted without context and/or curation.

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

It is undeniable that the AUA works tirelessly to publish the most rigorously peer-reviewed, impactful urological research. We also make every effort to provide the ability to reproduce the outputs we publish.

We highlight another crucial point from the CMSS letter we signed: "It is also worth noting that requiring all publishers to supply financial information in pursuit of fixed pricing conflicts with fair trade." The AUA's position is that any monitoring should be done by market because of the variances.

#### 4. Early input on considerations to increase findability and transparency of research.

In principle, the AUA agrees with the importance of discoverability and transparency of research. In practice, though, we need to work with our extensive scholarly publishing community - researchers, librarians, vendors, society publishers, commercial publishers, government representatives, and beyond - to build the infrastructure that will support all affected entities. The future of research depends on slow, deliberative collaboration to adopt the changes that will advance science in the United States and across the world.

To emphasize the CMSS letter: "As the medical and research community collectively work to increase the public's trust in health and science, these proposed changes could unintentionally foster misinformation."

#### **Uploaded File:**

NIH-OA-RFI\_AUA-Response\_April-24-2023\_FINAL.pdf

**Description:** The attached letter is the American Urological Association's official response to the NIH RFI. Each RFI question is also answered individually in the above response boxes.

Email: jregala@auanet.org



# Advancing Urology<sup>TM</sup>

American Urological Association Response to National Institutes of Health (NIH) Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research

# Introduction

Thank you for the opportunity to provide our response to the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research. Although we are signatories to group letters from the Council of Medical Specialty Societies (CMSS), Wolters Kluwer Health Publishing, and an independent group of medical publishers that includes ASCO, the New England Journal of Medicine, and others, the American Urological Association (AUA) and its executive and editorial leadership believe it is important to record our own response to this request for information. We agree with the concept of making federally funded research accessible to all, but we believe the unintended consequences of the mandates suggested by our government will create barriers to publication for many.

Additionally, the good works of societies and associations will be severely impacted if these sweeping changes are made without thoughtful input from groups like ours. As a global membership organization with approximately 25,000 members, revenue dollars from subscriptions to our scholarly journals are vital to our mission. We reinvest these earnings into our scholarly publications, staffing of the editorial office, our peer review systems, and content platforms. These revenues also allow for the AUA's investment in creation of guidelines that impact urological patient care across the world, funding research that ultimately becomes part of the research publishing ecosystem, educating urological providers in all stages of their careers, and so much more. Careful consideration for this paradigm shift across the society/association landscape must be given, or the advantages of open access will be erased by the perils of the many lost opportunities our research community will suffer. We appreciate the accessibility you are trying to offer to global readers, patients, and taxpayers; however, we already put a tremendous amount of work into that effort now. At the AUA, we make our research accessible in a number of ways: posting summaries of all research articles in our subscription journals on our 100% freely accessible platform, www.AUANews.net; creating publicly accessible video and visual abstracts; sharing each published article to multiple social media channels with a shareable link that eliminates the subscription paywall; disseminating journal content to the research community via email; and other opportunities such as podcasts and webinars.

We look forward to collaborating with the NIH to continue to enhance accessibility, but we ask that in addition to listening opportunities that you engage with the AUA and other like-minded



# Advancing Urology<sup>TM</sup>

organizations to understand what will be sacrificed for an open access ideal that is commendable in principle but will be destructive if actually implemented. We believe that common ground can be attained through open dialogue and shared decision-making.

# Question 1: How to best ensure equity in publication opportunities for NIH-supported investigators.

The NIH Public Access Plan aims to maintain the existing broad discretion for researchers and authors to choose how and where to publish their results. Consistent with current practice, the NIH Public Access Plan allows the submission of final published articles to PubMed Central (PMC) (in cases where a formal agreement is in place) to minimize the compliance burden on NIH-supported researchers and also maintains the flexibility of NIH-supported researchers to submit the final peer-reviewed manuscript. NIH seeks information on additional steps it might consider taking to ensure that proposed changes to implementation of the NIH Public Access Policy do not create new inequities in publishing opportunities or reinforce existing ones.

Responses to this question, as given by the collective letters the AUA signed as detailed in our introduction, state our position well and in detail. The AUA does want to emphasize these very important points:

- Peer review adds value to our high-impact research. Our flagship journal, The Journal of Urology®, is more than 100 years old, and is highly regarded in urology and in medicine at large. This peer review comes with a high financial and intellectual cost, and we ask that the NIH consider this considerable investment by the AUA when studying financial realities of implementing Open Access policies.
- We widely and generously distribute our research despite subscription paywalls. Nonsubscribers around the world have a plethora of ways to absorb our research in a multitude of valuable formats, from podcasts to social media to author-written insights of articles.

CMSS has asked, and we echo this request, for a 2-year delay to the mandate so all stakeholders can work together to develop sustainable policies focused on reliable, equitable, high-quality scientific content. We agree with and want to reinforce an important point in the CMSS letter:

"Policies that restrict publishers' abilities to collaborate with authors to realize their protection of rights under United States copyright law would further limit revenue streams on which we depend, including royalties, licensing, reprints, and advertising. We urge the NIH not to include rights retention language or license requirements in the



# Advancing Urology<sup>™</sup>

final policy other than the grantee's right to deposit the manuscript. Preserving a Green OA route presents a sustainable business model that should be embraced. Expanding rights retention policies beyond the deposition of the manuscript would also erode the publisher's ability to monitor usage of the content in support of the author's intellectual property."

# Question 2: Steps for improving equity in access and accessibility of publications.

Removal of the currently allowable 12-month embargo period for NIH-supported publications will improve access to these research products for all. As noted in the NIH Public Access Plan, NIH also plans to continue making articles available in human and machine-readable forms to support automated text processing. NIH will also seek ways to improve the accessibility of publications via assistive devices. NIH welcomes input on other steps that could be taken to improve equity in access to publications by diverse communities of users, including researchers, clinicians and public health officials, students and educators, and other members of the public.

- We agree with the fundamental importance of accessibility to premiere urological AUA research publications. Green Open Access remains a viable solution to this question of access.
- We do innovate in deliverability and accessibility of our content as a main goal of each of our publications.
- We strive to deliver our content in a way that eliminates or at least avoids dissemination and promotion of misinformation, which we believe will be diminished if "open" research is posted without context and/or curation.

Question 3: Methods for monitoring evolving costs and impacts on affected communities.

NIH proposes to actively monitor trends in publication fees and policies to ensure that they remain reasonable and equitable. NIH seeks information on effective approaches for monitoring trends in publication fees and equity in publication opportunities.

It is undeniable that the AUA works tirelessly to publish the most rigorously peer-reviewed, impactful urological research. We also make every effort to provide the ability to reproduce the outputs we publish.

We highlight another crucial point from the CMSS letter we signed: "It is also worth noting that requiring all publishers to supply financial information in pursuit of fixed pricing conflicts with fair trade." The AUA's position is that any monitoring should be done by market because of the variances.

Question 4: Early input on considerations to increase findability and transparency of research.



# Advancing Urology<sup>TM</sup>

Section IV of the NIH Public Access Plan is a first step in developing the NIH's updated plan for persistent identifiers (PIDs) and metadata, which will be submitted to OSTP by December 31, 2024. NIH seeks suggestions on any specific issues that should be considered in efforts to improve use of PIDs and metadata, including information about experiences institutions and researchers have had with adoption of different identifiers.

In principle, the AUA agrees with the importance of discoverability and transparency of research. In practice, though, we need to work with our extensive scholarly publishing community – researchers, librarians, vendors, society publishers, commercial publishers, government representatives, and beyond – to build the infrastructure that will support all affected entities. The future of research depends on slow, deliberative collaboration to adopt the changes that will advance science in the United States and across the world.

To emphasize the CMSS letter: "As the medical and research community collectively work to increase the public's trust in health and science, these proposed changes could unintentionally foster misinformation."

**Submit date: 4/24/2023** 

I am responding to this RFI: On behalf of an organization

Name: Jennifer Griffiths

Name of Organization: Springer Nature

Type of Organization: Other

Type of Organization-Other: Publisher

## 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Summary: To ensure equity in implementation of both publishing and open data aspects of the NIH Public Access Plan, NIH-supported investigators need the resources to support and enable their choice of compliance route.

Ensuring equity in publication opportunities for NIH-supported investigators means ensuring that every investigator, regardless of field, career stage, grant size, gender, ethnicity and institutional affiliation, has the resources available to them to choose where to publish and the route for compliance that enables that choice.

The plan in its current form allows compliance through either deposition of the "final peer-reviewed manuscript upon acceptance" (III.A.3.a. - i.e. "Accepted Manuscript" submission without any embargo / zero embargo green OA) or final published article submission (III.A.3.a. - ie. submission of the "Version of Record" / gold OA). Most journals in which NIH investigators currently choose to publish - including our own - support only one of these two routes: Gold Open Access - where the Version of Record is made freely available at publication.

By contrast, most journals and their publishers do not support the zero embargo green OA route - where an unfinished Accepted Manuscript is made openly available at the same time that the Version of Record is published. Such a model is simply not sustainable: it undermines the subscription model that supports it and slows progress towards the sustainable and scalable options for public access that gold OA enables. Gold OA is the only sustainable model for trusted open access. So, to best ensure equity in publication opportunities NIH must make sure the funding is in place to support any NIH investigator that might wish to publish in journals which only support the Gold OA route.

Our work has also shown that authors complying through the Gold OA route are likely to achieve greater reach and impact for their papers than if they had elected for compliance via the Accepted Manuscript route. This dichotomy has the potential to exacerbate existing inequities between NIH-fundees and/or create new ones. Researchers that are less well-funded (which is more common for early career researchers, those in fields with small grant sizes, and those at minority-serving institutions or HBCUs) can be further disadvantaged because they are more likely to have to comply via zero-embargo Green, missing out on the impact and reach of gold OA .

The zero-embargo Green access approach is also unsustainable since it prevents maintenance of subscription income to pay for the costs and work of publishing. So ultimately it is essential that sufficient funding is made available to pay for reasonable APCs for Gold OA publication. The calls on NIH funding can be minimized where such funding is pooled with University library budgets via

Transformative Agreements (TAs). Ultimately to achieve a full transition to sustainable open access there needs to be a way to align and maximize use of available funds to spread the load. TAs don't solve all sustainability and equity issues but, by combining funder and library funds, they are a strong step in the right direction ... one that has proven to be a scalable solution that substantially reduces the administrative burden on researchers. Regardless of whether NIH grant funds are used to contribute to centralized TAs or to support author-mediated payments to enable Gold OA, the NIH needs to budget for, and monitor, such costs.

Specifically, we recommend that to avoid creating new inequalities or exacerbating existing ones during this transitional phase NIH should ask grantees to include an estimate of reasonable publishing costs for articles arising from the grant as a standard budget line item.

This approach will ensure that authors that are planning to comply via Gold OA will have requested sufficient funds to cover reasonable APCs. It will also enable NIH to better monitor and track potential inequities arising from, or being exacerbated by, differences in impact between the two different compliance routes.

We are aligned with STM's recent position statement regarding zero embargo Green OA / "Rights Retention Strategies" and their response to NIH's RFI. In particular we support the argument that many journals need exclusive publishing rights to support sustainable business models and continued investment. Our longstanding position on this topic is clear: mandatory obligations being placed on grant fundees (already overburdened with compliance obligations), to openly license unfinished versions of their papers put them in a difficult position, undermine progress towards full sustainable public access for research papers and force publishers to maintain paywalls and defend subscription revenue.

To demonstrate their commitment to maintaining researchers' free choice about where to publish, as well as the integrity and independence of the QA processes that publishers implement, NIH should not place any such burden upon the researchers it funds.

The scientific data requirements of the NIH Public Access Plan also put a substantial compliance burden on NIH-supported investigators. Publishers are ideally placed to support requirements to make scientific data "freely available and publicly accessible by default at the time of publication" through policy and infrastructural support for integration of machine readable persistent identifiers (PIDs). However, as for achieving equity in publishing opportunities, to achieve full open data compliance will require sufficient support to be put in place for every investigator, regardless of field, career stage, grant size, gender, ethnicity and institutional affiliation.

#### 2. Steps for improving equity in access and accessibility of publications.

Summary: To improve equity in access and accessibility of publications NIH needs to monitor and maximize the proportion of NIH-supported publications complying through Gold OA.

Gold OA maximizes access not only by enabling free online access to humans and machines but also by enabling re-use, re-formatting, aggregation, and Other procedures to make the content discoverable, accessible and usable by diverse communities according to their specific needs. The Version of Record, which Gold OA makes accessible, is the complete, authoritative and up-to-date version of the paper,

curated and maintained by publishers and editors. Our work shows that researchers prefer the VoR over the unfinished Accepted Manuscript, both as readers and authors.

So there are significant disadvantages for those that do not have access to the VoR. Therefore to maximize the equity benefits as the NIH Public Access Plan is implemented it is important that the proportion of compliance through Gold OA is maximized and monitored. The full equity benefits of the NIH Public Access Plan can only be realized when there are no paywalls around any NIH-supported VoRs. Until then less-well resourced researchers and, more importantly, a large proportion of the US public, including many clinicians, public health officials, students and educators, will only have access to unfinished inferior versions of any papers that have complied with the plan via the zero embargo Green route.

Given this: we recommend that NIH should include an explicit preference / encouragement for compliance via Gold OA in its guidance for researchers, as for example included In the FAQs for the NASA policy for the Science Mission Directorate

## 3. Methods for monitoring evolving costs and impacts on affected communities.

Summary: To monitor costs and impacts of the NIH Public Access Plan, the NIH should, where possible, work with institutions and their libraries to leverage Transformative Agreements and Other equivalent centralized payment arrangements. Differences in impact between green and Gold OA compliance paths and their knock-on effect on equity should be monitored.

The only sustainable publishing model requires payment of publication fees (APCs) so there should be guidance to grantees that these need to be estimated and included in their applications. The funding burden on NIH for these can be minimized if grant money is pooled with University library money and this is best achieved via Transformative Agreements (TAs). These TAs can then be used to monitor and report on these costs to universities and funders like the NIH.

TAs don't solve all sustainability and equity issues but, by combining funder and library funds, they are a strong step in the right direction that has proven to be a scalable solution that substantially reduces the administrative burden on researchers. Regardless of whether NIH grants are used to contribute to centralized TAs or to support author-mediated payments to enable Gold OA, the NIH needs to budget for, and monitor, such costs.

Our work has shown that authors complying through the Gold OA route are likely to achieve greater reach and impact for their papers than if they had elected for compliance via the Accepted Manuscript route. This dichotomy has the potential to exacerbate existing inequities between NIH-fundees and/or create new ones. Researchers that are less well-funded (which is more common for early career researchers, those in fields with small grant sizes, and those at minority-serving institutions or HBCUs) can be further disadvantaged because they are more likely to have to comply via zero-embargo Green, missing out on the impact and reach of Gold OA .

Therefore we recommend that differences in impact between green and gold OA compliance paths and their knock-on impact on potentially disadvantaged NIH-investigators should be quantified and regularly reported.

#### 4. Early input on considerations to increase findability and transparency of research.

Summary: Publishers are key partners in deploying and integrating metadata and PIDs to enable a more efficient, transparent and impactful open science ecosystem

Publishers are ideally placed to support increasing findability and transparency of research through policy and infrastructural support for integration of machine readable persistent identifiers (PIDs).

We would welcome the chance to work through with NIH the most beneficial PIDs and metadata and their use cases. These are some of the PIDs and metadata we are already including in our publications:

DOI (Digital Object Identifier) for outputs/publications, i.e. eBooks, ejournals, journal articles and chapters

ORCID (Open Researcher and Contributor iD) for persons, i.e. authors and editors

Crossref Funder ID for grant-giving organizations

GRID ID (Global Research Identifier Database iD) and ISNI ID (International Standard Name Identifier) for research organizations/affiliations.

Grant Numbers: we collect "Grant Numbers" and incorporate them in our metadata that is also deposited at Crossref

Conference Series ID

Clinical Trial ID

Article, Issue Copyright Holder

Article, Issue Copyright Year

**Keywords** 

Registration, Received, Accepted, Issue Online Dates

Article Citation ID

We also actively contribute in multiple ways to cross-industry efforts in this area through STM, Crossref, ORCID, CHORUS (for example our participation in the CHORUS/CSIRO pilot on research resources and facilities) and Others.

We recommend that NIH works closely with publishers in general, and particularly these pre-existing cross-industry organizations, to maximize the impact of the revised plan for PIDs and metadata.

#### **Uploaded File:**

Springer-Nature-NIH-RFI-response-FINAL.pdf

**Description:** Fulltext with links and additional points

Email: jennifer.griffiths@us.nature.com



# Springer Nature NIH public access plan RFI response

# 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

The NIH Public Access Plan aims to maintain the existing broad discretion for researchers and authors to choose how and where to publish their results. Consistent with current practice, the NIH Public Access Plan allows the submission of final published articles to PubMed Central (PMC) (in cases where a formal agreement is in place) to minimize the compliance burden on NIH-supported researchers and also maintains the flexibility of NIH-supported researchers to submit the final peer-reviewed manuscript. NIH seeks information on additional steps it might consider taking to ensure that proposed changes to implementation of the NIH Public Access Policy do not create new inequities in publishing opportunities or reinforce existing ones.

# **SN Response:**

Summary: To ensure equity in implementation of both publishing and open data aspects of the NIH Public Access Plan, NIH-supported investigators need the resources to support and enable their choice of compliance route.

Ensuring equity in publication opportunities for NIH-supported investigators means ensuring that every investigator, regardless of field, career stage, grant size, gender, ethnicity and institutional affiliation, has the resources available to them to choose where to publish and the route for compliance that enables that choice.

The plan in its current form allows compliance through either deposition of the "final peer-reviewed manuscript upon acceptance" (III.A.3.a. - i.e. "Accepted Manuscript" submission without any embargo / zero embargo green OA) or final published article submission (III.A.3.a. - ie. submission of the "Version of Record" / gold OA). Most journals in which NIH investigators currently choose to publish - including our own - support **only one** of these two routes: Gold Open Access — where the Version of Record is made freely available at publication.

By contrast, most journals and their publishers <u>do not support</u> the zero embargo green OA route - where an unfinished Accepted Manuscript is made openly available at the same time that the Version of Record is published. Such a model is simply not sustainable: it undermines the subscription model that supports it and slows progress towards the sustainable and scalable options for public access



that gold OA enables. Gold OA is the only sustainable model for trusted open access. So, to best ensure equity in publication opportunities NIH must make sure the funding is in place to support <u>any</u> NIH investigator that might wish to publish in journals which only support the Gold OA route.

Our work has also shown that authors complying through the Gold OA route are likely to achieve greater reach and impact for their papers than if they had elected for compliance via the Accepted Manuscript route. This dichotomy has the potential to exacerbate existing inequities between NIH-fundees and/or create new ones. Researchers that are less well-funded (which is more common for early career researchers, those in fields with small grant sizes, and those at minority-serving institutions or HBCUs) can be further disadvantaged because they are more likely to have to comply via zero-embargo Green, missing out on the impact and reach of gold OA.

The zero-embargo Green access approach is also unsustainable since it prevents maintenance of subscription income to pay for the costs and work of publishing. So ultimately it is essential that sufficient funding is made available to pay for reasonable APCs for Gold OA publication. The calls on NIH funding can be minimized where such funding is pooled with university library budgets via Transformative Agreements (TAs). Ultimately to achieve a full transition to sustainable open access there needs to be a way to align and maximize use of available funds to spread the load. TAs don't solve all sustainability and equity issues but, by combining funder and library funds, they are a strong step in the right direction ... one that has proven to be a scalable solution that substantially reduces the administrative burden on researchers. Regardless of whether NIH grant funds are used to contribute to centralized TAs or to support author-mediated payments to enable Gold OA, the NIH needs to budget for, and monitor, such costs.

Specifically, we recommend that to avoid creating new inequalities or exacerbating existing ones during this transitional phase NIH should ask grantees to include an estimate of reasonable publishing costs for articles arising from the grant as a standard budget line item.

This approach will ensure that authors that are planning to comply via Gold OA will have requested sufficient funds to cover reasonable APCs. It will also enable NIH to better monitor and track potential inequities arising from, or being exacerbated by, differences in impact between the two different compliance routes.

We are aligned with STM's recent <u>position statement</u> regarding zero embargo Green OA / "Rights Retention Strategies" and their response to NIH's Rfl. In particular we support the argument that many journals need exclusive publishing rights to support sustainable business models and continued investment. <u>Our longstanding position</u> on this topic is clear: mandatory obligations being placed on grant fundees (already <u>overburdened with compliance obligations</u>), to openly license unfinished versions of their papers put them in a difficult position, undermine progress towards full sustainable public access for research papers and force publishers to maintain paywalls and defend subscription revenue.



To demonstrate their commitment to maintaining researchers' free choice about where to publish, as well as the integrity and independence of the QA processes that publishers implement, NIH should not place any such burden upon the researchers it funds.

The scientific data requirements of the NIH Public Access Plan also put a substantial compliance burden on NIH-supported investigators. Publishers are ideally placed to support requirements to make scientific data "freely available and publicly accessible by default at the time of publication" through policy and infrastructural support for integration of machine readable persistent identifiers (PIDs). However, as for achieving equity in publishing opportunities, to achieve full open data compliance will require sufficient support to be put in place for every investigator, regardless of field, career stage, grant size, gender, ethnicity and institutional affiliation.

# 2. Steps for improving equity in access and accessibility of publications.

Removal of the currently allowable 12-month embargo period for NIH-supported publications will improve access to these research products for all. As noted in the NIH Public Access Plan, NIH also plans to continue making articles available in human and machine-readable forms to support automated text processing. NIH will also seek ways to improve the accessibility of publications via assistive devices. NIH welcomes input on other steps that could be taken to improve equity in access to publications by diverse communities of users, including researchers, clinicians and public health officials, students and educators, and other members of the public.

# **SN Response:**

Summary: To improve equity in access and accessibility of publications NIH needs to monitor and maximize the proportion of NIH-supported publications complying through Gold OA.

Gold OA maximizes access not only by enabling free online access to humans and machines but also by enabling re-use, re-formatting, aggregation, and other procedures to make the content discoverable, accessible and usable by diverse communities according to their specific needs. The Version of Record, which Gold OA makes accessible, is the complete, authoritative and up-to-date version of the paper, curated and maintained by publishers and editors. <u>Our work</u> shows that researchers prefer the VoR over the unfinished Accepted Manuscript, both as readers and authors.

So there are significant disadvantages for those that do not have access to the VoR. Therefore to maximize the equity benefits as the NIH Public Access Plan is implemented it is important that the proportion of compliance through Gold OA is maximized and monitored. The full equity benefits of the NIH Public Access Plan can only be realized when there are no paywalls around any



NIH-supported VoRs. Until then less-well resourced researchers and, more importantly, <u>a large</u> <u>proportion of the US public, including many clinicians, public health officials, students and <u>educators</u>, will only have access to unfinished inferior versions of any papers that have complied with the plan via the zero embargo Green route.</u>

Given this: we recommend that NIH should include an explicit preference / encouragement for compliance via Gold OA in its guidance for researchers, as for example included In the FAQs for the NASA policy for the Science Mission Directorate

# 3. Methods for monitoring evolving costs and impacts on affected communities.

NIH proposes to actively monitor trends in publication fees and policies to ensure that they remain reasonable and equitable. NIH seeks information on effective approaches for monitoring trends in publication fees and equity in publication opportunities.

# **SN** Response:

Summary: To monitor costs and impacts of the NIH Public Access Plan, the NIH should, where possible, work with institutions and their libraries to leverage Transformative Agreements and other equivalent centralized payment arrangements. Differences in impact between green and Gold OA compliance paths and their knock-on effect on equity should be monitored.

The only sustainable publishing model requires payment of publication fees (APCs) so there should be guidance to grantees that these need to be estimated and included in their applications. The funding burden on NIH for these can be minimized if grant money is pooled with university library money and this is best achieved via Transformative Agreements (TAs). These TAs can then be used to monitor and report on these costs to universities and funders like the NIH.

TAs don't solve all sustainability and equity issues but, by combining funder and library funds, they are a strong step in the right direction that has proven to be a scalable solution that substantially reduces the administrative burden on researchers. Regardless of whether NIH grants are used to contribute to centralized TAs or to support author-mediated payments to enable Gold OA, the NIH needs to budget for, and monitor, such costs.

Our work has shown that authors complying through the Gold OA route are likely to achieve greater reach and impact for their papers than if they had elected for compliance via the Accepted Manuscript route. This dichotomy has the potential to exacerbate existing inequities between NIH-fundees and/or create new ones. Researchers that are less well-funded (which is more common



for early career researchers, those in fields with small grant sizes, and those at minority-serving institutions or HBCUs) can be further disadvantaged because they are more likely to have to comply via zero-embargo Green, missing out on the impact and reach of Gold OA .

Therefore we recommend that differences in impact between green and gold OA compliance paths and their knock-on impact on potentially disadvantaged NIH-investigators should be quantified and regularly reported.

# 4. Early input on considerations to increase findability and transparency of research.

Section IV of the NIH Public Access Plan is a first step in developing the NIH's updated plan for persistent identifiers (PIDs) and metadata, which will be submitted to OSTP by December 31, 2024. NIH seeks suggestions on any specific issues that should be considered in efforts to improve use of PIDs and metadata, including information about experiences institutions and researchers have had with adoption of different identifiers.

# **SN Response:**

Summary: Publishers are key partners in deploying and integrating metadata and PIDs to enable a more efficient, transparent and impactful open science ecosystem

Publishers are ideally placed to support increasing findability and transparency of research through policy and infrastructural support for integration of machine readable persistent identifiers (PIDs). We would welcome the chance to work through with NIH the most beneficial PIDs and metadata and their use cases. These are *some* of the PIDs and metadata we are already including in our publications:

- **DOI** (Digital Object Identifier) for outputs/publications, i.e. eBooks, ejournals, journal articles and chapters
- ORCID (Open Researcher and Contributor iD) for persons, i.e. authors and editors
- Crossref Funder ID for grant-giving organizations
- **GRID ID** (Global Research Identifier Database iD) and **ISNI ID** (International Standard Name Identifier) for research organizations/affiliations.
- Grant Numbers: we collect "Grant Numbers" and incorporate them in our metadata that is also deposited at Crossref
- Conference Series ID
- Clinical Trial ID
- Article, Issue Copyright Holder
- Article, Issue Copyright Year



- Keywords
- Registration, Received, Accepted, Issue Online Dates
- Article Citation ID

We also actively contribute in multiple ways to cross-industry efforts in this area through STM, Crossref, ORCID, CHORUS (for example our participation in the <a href="CHORUS/CSIRO pilot on research resources">CHORUS/CSIRO pilot on research resources and facilities</a>) and others.

We recommend that NIH works closely with publishers in general, and particularly these pre-existing cross-industry organizations, to maximize the impact of the revised plan for PIDs and metadata.

# **Springer Nature NIH Response: Additional Points**

Further to the direct responses we want to raise several specific additional points regarding the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research

- 1. The plan outlines an expectation that the deposition of the "final peer-reviewed manuscript" (the Accepted Manuscript) to the NIHMS system upon acceptance. As authors in journals offering both OA and subscription publishing options choose their preferred option subsequent to acceptance it would be helpful to make it clear that the deposition can occur as soon as possible after this decision has been made.
- 2. We would also seek clarity on the timeline for this plan to become implemented as policy: exactly what is meant by "an effective date no later than December 31, 2025" e.g. would the policy apply to all papers from that date that (a) arise from new grant calls, (b) arise from new grants awarded, (c) are submitted to a journal or (d) are published? We recommend option (a) since that would allow all stakeholders the maximum amount of time to adapt to this new policy.
- 3. The plan states that the Gold OA option involves, "publishing in a journal with a formal agreement with NLM to submit "final published articles" (the Version of Record) to be available in PMC on publication."
  - a. What if a journal is best suited to the research to be published but does not have an agreement?
  - b. Can authors deposit the VoR in PMC themselves? If not, what is the rationale for prohibiting this?



- 4. The plan implies authors are free to choose where to publish but it also implies restrictions to what funding will be approved. As per our main responses we recommend that NIH should monitor and maximize the proportion of NIH-supported publications complying through Gold OA.and ensure there is sufficient funding to support gold OA for all papers that NIH-supported investigators choose to publish in journals that only support that route. However <u>if</u> funding restrictions are to be applied:
  - a. How will these be communicated? How much funding is available for Gold OA?
  - b. We need to work together on education/signposting for researchers on how they should budget for publication fees.
  - c. NIH should create a mechanism for authors to fund publishing charges after the grant has closed
- 5. The plan states that a requirement is that "Costs are charged consistently regardless of the source of support". Most reputable publishers, including Springer Nature, grant full or partial waivers for APCs for authors without access to sufficient funding. For this reason we recommend that the wording is clarified to indicate that the intent of this requirement is that NIH-supported researchers should not be charged at a higher level compared to other authors, rather than ruling out variation in APC pricing to take account of financial need.

**Submit date: 4/24/2023** 

I am responding to this RFI: On behalf of an organization

Name: Roy Kaufman

Name of Organization: Copyright Clearance Center

Type of Organization: Other

Type of Organization-Other: CCC is a not-for-profit organization founded in 1977 at the suggestion of

Congress to facilitate collective copyright licensing for the text sector.

Role: Member of the public

1. How to best ensure equity in publication opportunities for NIH-supported investigators.

- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

The following text, with attachments and links, has been uploaded in PDF format. For convenience, text is pasted herein as well. We recommend using the PDF version.

Response of Copyright Clearance Center (CCC) to Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research (RFI)

Notice Number:

NOT-OD-23-091

CCC welcomes the opportunity to submit this response to Question 4 of the NIH's Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research. More importantly, we welcome NIH's interest in the use of PIDs and metadata to increase findability and transparency of scientific research.

Background on CCC.

CCC is a not-for-profit organization founded in 1977 at the suggestion of Congress to facilitate collective copyright licensing for the text sector. Presently, among Other lines of business, CCC provides licenses to content from over 10,000 rightsholders for whom we serve as an agent. We provide these licenses to more than 35,000 business organizations (Business Users) around the world. CCC is a supplier of knowledge management software called RightFind®, which is used by a subset of these Business Users to manage and access content. We also provide (1) Other software services, (2) library staffing, (3) content enrichment, data and metadata services, and (4) content delivery. On October 19, 2021, U.S. Secretary of Commerce Gina Raimondo announced that we were awarded a Market Development Cooperator Program grant, administered by the Commerce Department's International Trade Administration, to support our work with standards development organizations.

Our fastest growing business is managing the agreement- and fee-administration process on behalf of publishers who collect fees or Otherwise track usage from authors, institutions, consortia, government and Other funding bodies for immediate open access (OA). We do this primarily through our RightsLink®

for Scientific Communications software platform (RLSC). RLSC is by far the market leader in managing open access agreements and payments, doing so for many of the top publishers of NIH-funded research.

PIDs and Metadata.

Through both our knowledge management work with Business Users and our work on behalf of publishers, CCC experiences firsthand the promise of persistent identifiers (PIDs) when applied early, consistently and persistently. We are also painfully aware of the problems related to the entropy that results from lack of early, consistent, and persistent application thereof.

A healthy research and publishing ecosystem requires PIDs and robust, rich, quality metadata to make connections among people, organizations, places, and digital objects. For example, in RLSC alone, we depend on dozens of author, institution, and manuscript metadata elements to apply the appropriate business logic and workflows necessary to automate and scale OA on the path toward open science.

Also, even within a seemingly unified sector such as scientific publishing, it is sometimes necessary to accommodate multiple PIDs serving the same purpose, such as organizational identifiers. While in some ways accommodating multiple PIDs increases work and decreases interoperability, PIDs have different scope, attributes, and audiences. Some users prefer PIDs with ISO certification, while Others prefer PIDs with established business models to ensure sustainability and maintenance, while Others focus on ability to use without cost to access PIDs. When one PID has been selected for use by a stakeholder as part of master data management, being forced to accommodate a different PID can have significant costs and introduce unnecessary friction. Accordingly at CCC, we accommodate a variety of organizational IDs in RLSC and have long preferred the features of Ringgold for our primary use.

Review of data quality of bibliographic records from the MEDLINE database

In 2022, three CCC colleagues reviewed the data quality of bibliographic records in the Medline database. A paper detailing the results of their research have been posted on bioRxiv and is attached to this document (Bramley, R, Howe, S, Marmanis, H 2022, Notes on the data quality of bibliographic records from the MEDLINE database, doi: <a href="https://doi.org/10.1101/2022.09.30.510312">https://doi.org/10.1101/2022.09.30.510312</a>; hereafter, "Bramley, et al"). As noted in the paper:

[T]he PubMed database, which contains over 33.8 million records collected over many decades, suffers from several data quality issues. These issues relate to, in part, character encodings, the absence of persistent identifiers, differences in human languages, and schema changes. These shortcomings should not be surprising since PubMed aggregates information produced by different publishers and XML providers, a fact that leads naturally to the presence of "multi-source" problems.

Among the conclusions of the paper are (1) "[g]iven the incompleteness and uniqueness of identifying fields, the disambiguation of author names remains a significant problem for PubMed, particularly for records dating before 2014, and (2) [o]verall, there is an improvement in the use of identifiers; in particular, records created since 2015 exhibit an increase in external identifiers. However, the data quality for institutional identifiers is poor and their use has been diminishing over time."

Mapping metadata management across the research lifecycle.

In late 2022, CCC and Media Growth Strategies undertook a thorough examination of metadata management across the research lifecycle. This review builds on an existing body of work to uncover

multiple system complexities and breakages, which - separately and together - create missed opportunities for the communities for whom OA and open science models are designed to serve.

CCC has made this information publicly available in interactive infographic form at <a href="https://www.copyright.com/stateofmetadata/">https://www.copyright.com/stateofmetadata/</a>, and we have attached a chart summarizing where metadata breakages occur throughout the research lifecycle and how they impact various stakeholder groups.. Drawn directly from research interviews, the infographic depicts the significant economic impact that a fragmented metadata supply chain is having today on researchers, institutions, funders, and publishers. Researchers in particular shoulder a significant administrative burden that ultimately disrupts and delays the process of scientific discovery.

The infographic is a living document which will be updated and modified based on ongoing community feedback.

As the scholarly communications community continues its shift to OA and open science, stakeholders require a robust network of interoperable systems for making critical and necessary improvements, and much progress is underway. In that environment, a dedication to data stewardship across each stakeholder group, and the service providers supporting them, will lead to greater data sharing; reliable, trustworthy metrics on research impact; and a responsive, equitable rewards system. NIH can lead the way.

Question 4 of the RFI states: "NIH seeks suggestions on any specific issues that should be considered in efforts to improve use of PIDs and metadata, including information about experiences institutions and researchers have had with adoption of different identifiers."

First, we recommend that NIH review the research, findings and recommendations set forth in Bramley, et al.

Second, NIH, as the premier funder of biomedical research in the US, is well positioned to help research and lead by example by requiring PIDs at appropriate points. As can be seen in the above-referenced infographic, grant application is one of the first organized parts of the lifecycle where PIDs can be effectively mandated. Once mandated and used, PIDs can flow throughout the lifecycle to improve everything from grant management to expression in PubMed. We urge NIH to review the infographic, sign up for updates, and provide feedback should NIH believe there are amendments and changes needed.

We have three specific recommendations with respect to mandated use of PIDs.

- 1. NIH should mandate that grant applications include organizations IDs for the institutions(s) affiliated with each researcher listed on the grant application, and Funder Registry IDs for the distinct funders of the grant. The requirement should insist that grant applications include at least one of the following organizational identifiers used in the scholarly publishing ecosystem and NIH should make metadata fields available for all four:
- A. Ringgold- a proprietary global organization identifier system owned by CCC with over 600,000 unique records and rich hierarchical metadata used today by (1) most large and mid-sized commercial and non-commercial publishers, and (2) a range of critical infrastructure providers in the publishing ecosystem. For publishers, Ringgold often is part of a master data management strategy. Ringgold is also used by

some funders, academic institutions, and consortia. Ringgold maps one-to-one with ISNI and the Funder Registry.

- B. ISNI- ISO standard name identifier system with 1,697,000 unique organizational records of which a minimum of 500,000 are relevant to the research sector. ISNI is free to use and has been adopted by many national libraries. It lacks the hierarchical metadata of Ringgold but enjoys the rigor and authority of ISO accreditation. The relevant organization records in ISNI map one-to-one with Ringgold.
- C. ROR- Research Organization Registry (ROR) is a global, community-led registry of open persistent identifiers for research organizations. ROR is free to use and has been adopted by some publishers, institutions, and overseas funders. It contains 104,000 unique identifiers and some hierarchical metadata. It can map to ISNI and the Funder Registry, but not on a one-to-one basis.
- D. Funder Registry (formerly known as FundRef) -Funder Registry is an open registry of grant-giving organization names and identifiers, with 32,000 unique identifiers for funders. It is donated by Elsevier to CrossRef and is updated approximately every 4-6 weeks. The Funder Registry ID can be used for author affiliations where the funder and affiliation are one and the same.
- 2. NIH should mandate that grant applicants include one or both of the following individual identifiers for all researchers in grant applications, and NIH should make metadata fields available for both.
- a. ORCID- ORCID, which stands for Open Researcher and Contributor ID, is a global, not-for-profit organization sustained by fees from member organizations. ORCID is the most broadly adopted identifier system for individuals in scientific publishing.
- b. ISNI- While not as well adopted as ORCID in research and science, ISNI has been broadly adopted in adjacent and non-adjacent fields.
- 3. NIH should mandate that appropriate PIDs be used at each stage reporting, while remaining flexible as to which PIDs it mandates, and should reevaluate its mandated PIDs on an ongoing basis. New PIDs such as RAiD (Research Activity Identifier) and DataCite (DOI-based system for research outputs) are being developed regularly and can help connect people, places and research. Likewise, Other existing PIDs such as, e.g., Scopus Affiliation ID (AF-ID) and Author ID (AU-ID) are currently used in certain relevant applications. Appropriate PIDs should be mandated at each stage of the workflow, while recognizing that the needs of researchers and the availability of PIDs change over time.

As a final recommendation, we suggest that NIH follow the lead of Wellcome Trust and the Bill and Melinda Gates Foundation, among Others, in registering grants for DOIs. This will help enable connectivity of PIDs and the discoverability of the grants, maximizing return to US taxpayers.

Respectfully submitted for Copyright Clearance Center by,

Roy S Kaufman

#### **Uploaded File:**

NIH-RFI-with-attachments.pdf

Description: The attached PDF contains our full response with attavhments and links.



Response of Copyright Clearance Center (CCC) to Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research (RFI)

Notice Number:

NOT-OD-23-091

CCC welcomes the opportunity to submit this response to Question 4 of the NIH's <u>Request</u> for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported <u>Research</u>. More importantly, we welcome NIH's interest in the use of PIDs and metadata to increase findability and transparency of scientific research.

# Background on CCC.

CCC is a not-for-profit organization founded in 1977 at the suggestion of Congress to facilitate collective copyright licensing for the text sector. Presently, among other lines of business, CCC provides licenses to content from over 10,000 rightsholders for whom we serve as an agent. We provide these licenses to more than 35,000 business organizations (Business Users) around the world. CCC is a supplier of knowledge management software called RightFind®, which is used by a subset of these Business Users to manage and access content. We also provide (1) other software services, (2) library staffing, (3) content enrichment, data and metadata services, and (4) content delivery. On October 19, 2021, <u>U.S. Secretary of Commerce Gina Raimondo announced</u> that we were awarded a Market Development Cooperator Program grant, administered by the Commerce Department's International Trade Administration, to support our work with standards development organizations.

Our fastest growing business is managing the agreement- and fee-administration process on behalf of publishers who collect fees or otherwise track usage from authors, institutions, consortia, government and other funding bodies for immediate open access (OA). We do this primarily through our RightsLink® for Scientific Communications software platform (RLSC). RLSC is by far the market leader in managing open access agreements and payments, doing so for many of the top publishers of NIH-funded research.

#### PIDs and Metadata.

Through both our knowledge management work with Business Users and our work on behalf of publishers, CCC experiences firsthand the promise of persistent identifiers (PIDs) when applied early, consistently and persistently. We are also painfully aware of the problems

+1.978.750.8400

copyright.com

info@copyright.com

Phone

Fmail

Web

related to the entropy that results from lack of early, consistent, and persistent application thereof.

A healthy research and publishing ecosystem requires PIDs and robust, rich, quality metadata to make connections among people, organizations, places, and digital objects. For example, in RLSC alone, we depend on dozens of author, institution, and manuscript metadata elements to apply the appropriate business logic and workflows necessary to automate and scale OA on the path toward open science.

Even within a seemingly unified sector such as scientific communications, it is sometimes necessary to accommodate multiple PIDs serving the same purpose, such as organizational identifiers. While in some ways accommodating multiple PIDs increases work and decreases interoperability, PIDs have different scope, attributes, and audiences. Some users prefer PIDs with ISO certification, while others prefer PIDs with established business models to ensure sustainability and maintenance, while others focus on ability to use without cost to access PIDs. When one PID has been selected for use by a stakeholder as part of master data management, being forced to accommodate a different PID can have significant costs and introduce unnecessary friction. Accordingly at CCC, we accommodate a variety of organizational IDs in RLSC and have long preferred the features of Ringgold for our primary use.<sup>1</sup>

# Review of data quality of bibliographic records from the MEDLINE database

In 2022, three CCC colleagues reviewed the data quality of bibliographic records in the Medline database. A paper detailing the results of their research have been posted on bioRxiv and is attached to this document (Bramley, R, Howe, S, Marmanis, H 2022, Notes on the data quality of bibliographic records from the MEDLINE database, doi: <a href="https://doi.org/10.1101/2022.09.30.510312">https://doi.org/10.1101/2022.09.30.510312</a>; hereafter, "Bramley, et al"). As noted in the paper:

[T]he PubMed database, which contains over 33.8 million records collected over many decades, suffers from several data quality issues. These issues relate to, in part, character encodings, the absence of persistent identifiers, differences in human languages, and schema changes. These shortcomings should not be surprising since

222 Rosewood Drive Danvers, MA 01923 USA Phone +1.978.750.8400 Email info@copyright.com

Web copyright.com

<sup>&</sup>lt;sup>1</sup> CCC adopted Ringgold as its preferred organizational PID approximately 8 years ago. CCC acquired Ringgold in 2022 so that we could ensure its continued viability given its importance to ourselves and our clients.

PubMed aggregates information produced by different publishers and XML providers, a fact that leads naturally to the presence of "multi-source" problems.

Among the conclusions of the paper are (1) "[g]iven the incompleteness and uniqueness of identifying fields, the disambiguation of author names remains a significant problem for PubMed, particularly for records dating before 2014, and (2) [o]verall, there is an improvement in the use of identifiers; in particular, records created since 2015 exhibit an increase in external identifiers. However, the data quality for institutional identifiers is poor and their use has been diminishing over time."

# Mapping metadata management across the research lifecycle.

In late 2022, CCC and Media Growth Strategies undertook a thorough examination of metadata management across the research lifecycle. This review builds on an existing body of work to uncover multiple system complexities and breakages, which — separately and together — create missed opportunities for the communities for whom OA and open science models are designed to serve.

CCC has made this information publicly available in interactive infographic form at <a href="https://www.copyright.com/stateofmetadata/">https://www.copyright.com/stateofmetadata/</a>, and we have attached a chart summarizing where metadata breakages occur throughout the research lifecycle and how they impact various stakeholder groups. Drawn directly from research interviews, the infographic depicts the significant economic impact that a fragmented metadata supply chain is having today on researchers, institutions, funders, and publishers. Researchers in particular shoulder a significant administrative burden that ultimately disrupts and delays the process of scientific discovery.

The infographic is a living document which will be updated and modified based on ongoing community feedback.

As the scholarly communications community continues its shift to OA and open science, stakeholders require a robust network of interoperable systems for making critical and necessary improvements, and much progress is underway. In that environment, a dedication to data stewardship across each stakeholder group, and the service providers supporting them, will lead to greater data sharing; reliable, trustworthy metrics on research impact; and a responsive, equitable rewards system. NIH can lead the way.

+1.978.750.8400

copyright.com

info@copyright.com

Phone

Email

Web

Question 4 of the RFI states: "NIH seeks suggestions on any specific issues that should be considered in efforts to improve use of PIDs and metadata, including information about experiences institutions and researchers have had with adoption of different identifiers."

First, we recommend that NIH review the research, findings and recommendations set forth in Bramley, et al.

**Second**, NIH, as the premier funder of biomedical research in the US, is well positioned to help research and lead by example by requiring PIDs at appropriate points. As can be seen in the above-referenced infographic, grant application is one of the first organized parts of the lifecycle where PIDs can be effectively mandated. Once mandated and used, PIDs can flow throughout the lifecycle to improve everything from grant management to expression in PubMed. We urge NIH to review the infographic, sign up for updates, and provide feedback should NIH believe there are amendments and changes needed.

We have three specific recommendations with respect to mandated use of PIDs:

1. NIH should mandate that grant applications include organizations IDs for the institutions(s) affiliated with each researcher listed on the grant application, and Funder Registry IDs for the distinct funders of the grant. The requirement should insist that grant applications include at least one of the following organizational identifiers used in the scholarly publishing ecosystem and NIH should make metadata fields available for all four:

A. Ringgold- a proprietary global organization identifier system owned by CCC with over 600,000 unique records and rich hierarchical metadata used today by (1) most large and mid-sized commercial and non-commercial publishers, and (2) a range of critical infrastructure providers in the publishing ecosystem. For publishers, Ringgold often is part of a master data management strategy. Ringgold is also used by some funders, academic institutions, and consortia. Ringgold maps one-to-one with ISNI and the Funder Registry.

B. ISNI- ISO standard name identifier system with 1,697,000 unique organizational records of which a minimum of 500,000 are relevant to the research sector. ISNI is free to use and has been adopted by many national libraries. It lacks the hierarchical metadata of Ringgold but enjoys the rigor and authority of ISO accreditation. The

relevant organization records in ISNI map one-to-one with Ringgold.

- C. ROR- Research Organization Registry (ROR) is a global, community-led registry of open persistent identifiers for research organizations. ROR is free to use and has been adopted by some publishers, institutions, and overseas funders. It contains 104,000 unique identifiers and some hierarchical metadata. It can map to ISNI and the Funder Registry, but not on a one-to-one basis.
- D. Funder Registry (formerly known as FundRef) —Funder Registry is an open registry of grant-giving organization names and identifiers, with 32,000 unique identifiers for funders. It is donated by Elsevier to CrossRef and is updated approximately every 4-6 weeks. The Funder Registry ID can be used for author affiliations where the funder and affiliation are one and the same.
- 2. NIH should mandate that grant applicants include one or both of the following individual identifiers for all researchers in grant applications, and NIH should make metadata fields available for both.
  - a. ORCID- ORCID, which stands for Open Researcher and Contributor ID, is a global, not-for-profit organization sustained by fees from member organizations. ORCID is the most broadly adopted identifier system for individuals in scientific publishing.
  - b. ISNI- While not as well adopted as ORCID in research and science, ISNI has been broadly adopted in adjacent and non-adjacent fields.
- 3. NIH should mandate that appropriate PIDs be used at each stage reporting, while remaining flexible as to which PIDs it mandates, and should reevaluate its mandated PIDs on an ongoing basis. New PIDs such as RAiD (Research Activity Identifier) and DataCite (DOI-based system for research outputs) are being developed regularly and can help connect people, places and research. Likewise, other existing PIDs such as, e.g., Scopus Affiliation ID (AF-ID) and Author ID (AU-ID) are currently used in certain relevant applications. Appropriate PIDs should be mandated at each stage of the workflow, while recognizing that the needs of researchers and the availability of PIDs change over time.

+1.978.750.8400

copyright.com

info@copyright.com

Phone

Email

Web

As a final recommendation, we suggest that NIH follow the lead of Wellcome Trust and the Bill and Melinda Gates Foundation, among others, in registering grants for DOIs. This will help enable connectivity of PIDs and the discoverability of the grants, maximizing return to US taxpayers.

+1.978.750.8400

copyright.com

info@copyright.com

Phone

Email

Web

Respectfully submitted for Copyright Clearance Center by,

Roy S Kaufman

# Notes on the data quality of bibliographic records from the MEDLINE database

Robin Bramley\* Stephen Howe<sup>†</sup> Haralambos Marmanis<sup>†</sup>
August 17, 2022

#### Abstract

The US National Library of Medicine has created and maintains the PubMed® database, a collection of over 33.8 million records that contain citations and abstracts from the biomedical and life sciences literature. That database is an important resource for researchers and information service providers alike. As part of our work related to the creation of an author graph for coronaviruses, we encountered several data quality issues with records from a curated subset of the PubMed database called MEDLINE. We provide a data quality assessment for records selected from the MEDLINE database and report on several issues ranging from parsing issues (e.g., character encodings and schema definition weaknesses) to low scores against several data quality metrics (e.g., identifier completeness, validity, and uniqueness).

# 1 Introduction

PubMed is an enormously valuable resource for the biomedical and health fields. The PubMed database is a voluminous collection of medical literature citations that is free, easily accessible, and has been a data source for many works in the information retrieval and life sciences communities. As machine learning becomes more prevalent in various branches of the life sciences, the number of works that rely on the PubMed database increases. Many papers that cited PubMed have appeared within the proceedings of The International Conference on Data and Text Mining in Biomedicine series e.g., DTMBIO '10 [1]. In ACM's Digital Library[2], the year 2021 was a new high point at 235 for computing research articles that mentioned PubMed in the full-text collection, up from 1 in 1998 and 115 in 2010. Many information providers utilize the PubMed database, and there are a variety of machine learning models trained on PubMed[3]. It should be no surprise that, during the COVID-19 pandemic,

<sup>\*</sup>Copyright Clearance Center Limited, London, United Kingdom

<sup>&</sup>lt;sup>†</sup>Copyright Clearance Center Inc., Danvers, Massachusetts, United States of America

the PubMed database has been crucial in providing timely and frictionless access to the scientific literature [4].

However, the PubMed database, which contains over 33.8 million records [5] collected over many decades, suffers from several data quality issues. These issues relate to, in part, character encodings, the absence of persistent identifiers, differences in human languages, and schema changes. These shortcomings should not be surprising since PubMed aggregates information produced by different publishers and XML providers, a fact that leads naturally to the presence of "multi-source problems" [6].

MEDLINE is a curated subset of PubMed, its records are indexed with a controlled vocabulary called MeSH [7] and include information regarding funding, genetic, chemical, and other metadata. Articles in MEDLINE predominantly come from a set of indexed journals and a reference data file of these journals is available separately [8]. MEDLINE was made available online, through PubMed, in 1997.

In this article, we will provide an account of our experience in working with the curated MEDLINE records and report on the data quality issues that we encountered. We will describe, at length, the problem of Author Name Disambiguation, which is widely acknowledged as a source of errors when processing bibliographic databases in general, due to the challenges of synonyms (e.g., "John Doe", "John T Doe", and "JT Doe" referring to the same individual) and homonyms (i.e., two different people who share the same name such as "John Smith") [9]. Other problem areas that we will discuss include issues with character encodings, date related issues, the presence of persistent identifiers (and lack thereof), affiliation disambiguation, language related data issues, and schema data quality issues. Knowing how to address these challenges is valuable for practitioners who need to work with MEDLINE (or databases like MEDLINE) and process its records so that they can be used in their information systems.

#### 1.1 PubMed data

The PubMed database is available as XML, based on a DTD (currently the 2019 version) [10]. The compressed files are made available via an FTP server (they are also accessible by HTTPS) and each one of them contains up to 30,000 citation records. Every year, in mid- December, the data are consolidated and an annual baseline is produced. This is followed by incremental daily update files that include deletions.

A PubMed XML file has a root element of PubmedArticleSet that contains 1, or more, PubmedArticle or PubmedBookArticle children. The DTD also permits 0 or 1 DeleteCitation elements, and these can be seen in the update files. The elements of the PubmedArticle are divided into the MedlineCitation and the optional PubmedData - we have colloquially referred to these as the "front" and "back" matter respectively.

The description of the XML elements [11], also outlines potential discrepancies caused by schema changes, or policy changes to the collected data. For

example, records created before 2002 only contained author initials instead of full, first or middle, names; moreover, records between 1988 and 2013 only included the affiliation for the first author.

# 1.1.1 Known DTD shortcomings

There are two known problems with the DTD that have not yet been addressed. The first known problem is that authors cannot be linked to their CollectiveName. Some publishers have tried to work around this by interspersing CollectiveName elements and Author elements. In a wheat genome sequencing consortium paper (PMID 30115783), one of the contributors was a member of 12 groups, so that person appears as an Author record 12 times. This multiplicity complicates the author name disambiguation, as it may be impossible to distinguish a duplicate author entry from a valid homonym.

The second problem is related to a shortcoming in the 2019 DTD. Specifically, the back matter PubmedData element may contain a ReferenceList with many Reference elements, but it doesn't prevent the presence of many ReferenceList elements each with one Reference. Consequently, extraction must be able to handle both because both have been observed in the records. Furthermore, the ReferenceList definition permits deeply nested ReferenceList elements, as shown below:

<!ELEMENT ReferenceList (Title?, Reference\*, ReferenceList\*) >

#### 1.1.2 Escape characters

Escape sequence characters may appear within text fields such as the article title or abstract text. For example, if you wanted to represent a record in JSON, then you would have to beware of trailing backslashes and double quotes. Backslashes can also be problematic for the language used to parse the record. Furthermore, it may be necessary to remove other special characters such as new line characters (e.g., carriage return, line feed), tabs, and so on.

#### 1.1.3 Extended characters

PubMed encompasses articles published in many different languages, sometimes multiple languages. Consequently, fields such as the affiliation string, or parts of the author's name, may contain extended characters. This is an important consideration for the disambiguation of author names.

## 1.2 Open Source libraries

Since PubMed has been a canonical source of biomedical citations, there are open source libraries to assist with parsing the records. Whilst none of these libraries were appropriate for our needs, they are included here for completeness.

For Python, pubmed\_parser [12] is an active project, but only handles a constrained field list. The pymed [13] project, which is now archived, only

parsed and cleansed a limited subset of the fields. It also seems that the design was intended to wrap the API.

For Java, there is pubmed-parser [14], which is based around the Java Architecture for XML Binding (JAXB). This project only had a short flurry of commits over 6 days in April 2021, consequently it is unclear whether this is actively maintained.

# 2 Materials and Methods

This work will identify challenges that can be faced when working with the MEDLINE data and categorize them along several dimensions of data quality [15].

# 2.1 Data acquisition

The PubMed baseline files were downloaded from their respective NLM FTP folders [16][17] and uploaded to separate folders on an S3 bucket.

## 2.2 Data processing

Figure 1 illustrates our data processing approach. The PubMed gzipped XML files were processed using Apache Spark 3.1.1 on Amazon EMR 6.3.1. The initial ingestion process extracted a few key properties, such as the PMID and DOI (from the PubmedData if present), before splitting the XML into two fragments representing the front matter (bibliographic metadata) and the back matter (references).

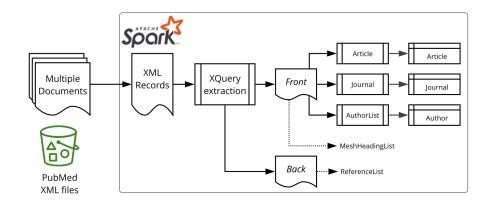


Figure 1: Data processing overview.

The baseline files were ingested first, then the update files were subsequently processed to apply updates, inserts and deletions. Record updates were applied

by sorting the records by their PMID in conjunction with the DateRevised property; only the newest records were retained. Note that the PMID Version attribute is not suitable for this purpose as it is only used by Public Library of Science (PLOS) records [11].

Spark SQL [18] is designed for tabular data, with the key construct being the DataFrame. Whereas XML documents are represented using a hierarchical structure that allows for repeating elements (a one-to-many relationship). This leads to an inherent mismatch between the two data formats that requires data transformation.

There is a spark-xml module [19], but we discovered during our initial experiments that the PubMed XML was too complex for spark-xml, as it resulted in heavily nested DataFrames, and incorrect query results. Consequently, we solved the XML to DataFrame impedance mismatch by performing an XQuery [20] operation per target entity type (e.g. Article, Author, etc.) as shown on the right-hand side of Figure 1.

The spark-xml XmlInputFormat class was retained for loading the XML files into Spark, with the ingestion and extraction utilizing XQuery queries to extract properties, via the Saxon-HE [21] library as provided by the Elsevier Labs spark-xml-utils [22] module.

To ease maintenance of the complex XQuery queries, we adopted a pattern whereby the XQuery output produces a JSON document. This makes the target property for a particular XPath or XQuery expression transparent (Figure 2) and inserting new elements does not break downstream code because it does not rely on positional information. The last part of that transformation phase is to leverage the read method of the SparkSession object which parses the JSON documents to DataFrame records. Note that Figure 2 also represents the handling of escape characters using the XQuery replace function.

```
'"forename": "', replace(replace($x/ForeName, '\\', '\\\\'), '"', '\\"'), '", ',
'"initials": "', replace(replace($x/Initials, '\\', '\\\\'), '"', '\\"'), '", ',
'"lastname": "', replace(replace($x/LastName, '\\', '\\\\'), '"', '\\"'), '", ',
'"suffix": "', normalize-space(replace(replace($x/Suffix, '^[,\\, |+', .''), '[,\\, '], \\]
```

Figure 2: JSON representation within XQuery.

## 2.3 Data analysis

The resulting DataFrames were analyzed using Spark SQL in Apache Zeppelin [23]. For string fields, we consider the length in characters and in words (by splitting on spaces). Metrics were rounded to 3 decimal places (or less).

The plots were produced in R, with the box plots using log-scale for the y-axis.

## 2.4 Definitions

- $\mathcal{N} = \text{number of records}$
- $\mathcal{M}$  = number of records missing a value for the target property
- $\mathcal{D}$  = distinct values of those present (excludes null / blank)
- V defined by count of records matching a regex for identifiers (Table 1)
- $\mathcal{P} = \text{present} = \mathcal{N} \mathcal{M}$
- Completeness metric =  $\mathcal{P} / \mathcal{N}$
- $Validity metric = \mathcal{V} / \mathcal{P}$
- $Uniqueness \text{ metric} = \mathcal{D} / \mathcal{P}$

Identifier	Regular expression
DOI [24]	"^10.\d{4,9}/[;()/:a-zA-Z0-9]+\$" <sup>1</sup>
ORCID [25]	$\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ $
ISNI [26] (presentation)	"[0-9]{4} [0-9]{4} [0-9]{4} [0-9]{3}[0-9X]"
ISNI (compact)	"[0-9]{15}[0-9X]"
GRID [27]	"grid $\.\d{4,6}\.$ [0-9a-f] ${1,2}$ "

Table 1: Regular expressions for identifier validation

## 2.5 Limitations of the study

The source dataset comprises the PubMed 2022 baseline plus daily update files to 1252 (30th March 2022).

It should be noted that our study includes only the PubmedArticle records, not the PubmedBookArticle records. The PubmedArticle records are only those from the MEDLINE subset (based on the Status attribute), and further excludes news articles, and those articles without a title; this gives a total of 28,986,590 article records. News articles were excluded from extraction because journalists, anecdotally those from the British Medical Journal, skew attempts to identify prolific authors through aggregation.

Other applied constraints are as follows:

- Only Author records with the ValidYN attribute of Y have been extracted, not Investigator records. For these 120,191,520 authors, only the first Affiliation element is considered.
- The DataBank element provides links to external datasets such as clinical trials. These identifiers were not investigated as part of the reported study.

<sup>&</sup>lt;sup>1</sup>Adapted from https://www.crossref.org/blog/dois-and-matching-regular-expressions/

- For alternative article identifiers, we did not extract the ELocationID element nor Publisher Item Identifiers (PII) from the PubmedData.
- For Journals, ISSNs were not analyzed.

## 2.5.1 Approximation

Five number summary information is produced using Spark's DataFrameStatFunctions approxQuantiles method with an error margin of 0.0001, an example is shown below:

articleDF.stat.approxQuantile("doi\_len", Array(0.0,0.25,0.5,0.75,1.0),
0.0001)

However, the distinct counts do not leverage the Spark SQL approx\_count\_distinct function, rather the dataframe.select("column").distinct.count approach was used.

# 3 Results and discussion

In this section, we'll present our results related to data quality for the entities and fields shown in Figure 3. The PubMed XML data model is article-centric, but we will work our way from left to right.

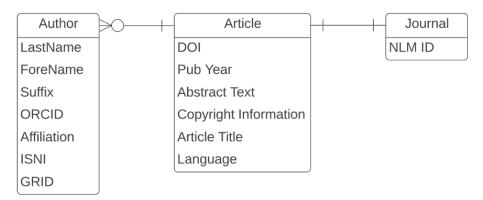


Figure 3: Entity Relationship Diagram for a subset of PubMed.

#### 3.1 Data quality issues related to author names

One of the important considerations regarding author records is that PubMed has not always recorded all the authors of a paper. The number of authors was limited to 10 between the years 1984 and 1995, and to 25 between the years 1996 and 1999 [11].

The most common last names in MEDLINE are Romanized Chinese names (Table 2), which can be very challenging to disambiguate. Looking at the length

characteristics (Figure 4), there are a few obvious problems, namely pollution of the author elements by incorrectly entered collective names (Table 3), and single character last names potentially caused by name transposition errors (Table 4).

LastName	Occurrences
Wang	1,086,073
Li	895,976
Zhang	878,544
Chen	722,753
Liu	703,743
Lee	547,636
Kim	523,687
Yang	433,439
Wu	360,532
Huang	309,375

Table 2: Top 10 LastName values.

LastName	Length	
Endocrinology Genetics And Metabolism Group Pediatric Branch Of Chi-		
nese Medical Association Neonatal Screening Group Specialist Committee For		
Prevention And Control Of Birth Defects Chinese Association Of Preven-		
tive Medicine Prevention And Control Committee Of Birth Defects Pediatric		
Branch Of Chinese Medical Association		
The Group Of Minimally Invasive Spinal Surgery And Enhanced Recovery	211	
Professional Committee Of Orthopedic Surgery And Enhanced Recovery As-		
sociation Of China Rehabilitation Technology Transformation And Promotion		
Genetic Disease Society Guangdong Precision Medicine Application Associ-	209	
ation Prenatal Diagnosis Group Maternal And Child Health Care Society		
Guangdong Medical Association Expert Committee Of Prenatal Diagnosis		
Arir Associazione Riabilitatori dell'Insufficienza Respiratoria Sip Società Ital-	201	
iana di Pneumologia Aifi Associazione Italiana Fisioterapisti And Sifir Società		
Italiana di Fisioterapia E Riabilitazione		
This Paper Is A Co-Publication Between European Journal Of Preventive Car-	176	
diology European Heart Journal Acute Cardiovascular Care And European		
Journal Of Cardiovascular Nursing		
Committee For Birth Defect Prevention And Control Chinese Association Of	174	
Preventive Medicine Genetic Testing And Precision Medicine Branch Chinese		
Association Of Birth Health		
Consensus Group Of Experts On Application Of Metagenomic Next Genera-	152	
tion Sequencing In The Pathogen Diagnosis In Clinical Moderate And Severe		
Infections		
Expert Committee Of The Inter-Laboratory Quality Assessment Of Prenatal	150	
Screening And Diagnosis Clinical Test Center Of The National Health Com-		
mission		
For The Antimalarial Therapeutic Efficacy Monitoring Group National Malaria	142	
Elimination Programme The Federal Ministry Of Health Abuja Nigeria		
On Behalf Of The Association Of Rural Surgeons Of India-Lancet Commission	142	
On Global Surgery Consensus Committee Arsi-LCoGS Consensus Committee		

Table 3: Ten longest Last Name values.

LastName	Occurrences
S	756
A	704
Е	636
M	592
О	563
K	497
R	453
Р	363
G	306
V	279

Table 4: Top 10 shortest LastName values.

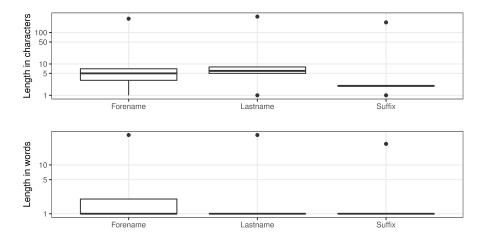


Figure 4: Author name character / word distributions.

The author forename field is 99.913% complete. Regarding the length, before 1945, the longest value in the forename field was 3 characters long, which reflects the policy to only hold author initials. The distributions, in Figure 4, clearly show that there are outliers. As shown in Table 5, these are primarily for working groups (a validity error), but the first row represents a different form of data preparation error where the affiliation has been concatenated with the forename.

PMID	LastName value	ForeName value	Length
34313229	Choi	Moon Hyung Department Of Ra-	276
		diology Eunpyeong St Mary's	
		Hospital College Of Medicine	
		The Catholic University Of Ko-	
		rea Seoul Republic Of Korea	
		Catholic Smart Imaging Cen-	
		ter Eunpyeong St Mary's Hos-	
		pital College Of Medicine The	
		Catholic University Of Korea	
		Seoul Republic Of Korea	
33145749	En Representación Del Grupo de	En Representación Del Grupo de	123
	Trastornos de la Conducta Y Del	Trastornos de la Conducta Y Del	
	Movimiento Durante El Sueño de	Movimiento Durante El Sueño de	
	la Sociedad Española de Sueño	la Sociedad Española de Sueño	
32329046	En Representación Del Grupo	En Representación Del Grupo	106
	de Estudio de Enfermedades	de Estudio de Enfermedades	
	Desmielinizantes de la Comu-	Desmielinizantes de la Comu-	
	nidad Autónoma de Madrid	nidad Autónoma de Madrid	
32433836	Pharmakopsychiatrie	The Therapeutic Drug Moni-	102
		toring Task Force Of The Ar-	
		beitsgemeinschaft Für Neuropsy-	
		chopharmakologie Und	

Table 5: ForeName values over 100 characters.

Completeness does not apply to author suffixes since not everyone has a suffix to their name. In terms of uniqueness there are 823 distinct values across 483,541 entries. There are also consistency issues, examples of which can be observed in Table 6 (e.g., Jr, Junior, Júnior). Figure 4 shows the range of suffix lengths and clearly indicates that there is something wrong with at least some records. When we look at the longest values for author suffixes (Table 7) and the most common single character values (Table 8), it becomes clear that there are multiple data issues related to the author suffix field; the general theme of misplaced values, or value "pollution", occurs across fields and is a major data quality weakness for the MEDLINE records.

Suffix value	Occurrences
Jr	374,510
3rd	74,260
2nd	20,364
4th	5,828
Sr	4,075
Junior	535
Júnior	380
Filho	241
PhD	238
5th	204
Neto	200
III	199
Dr	146
6th	129
MD	99

Table 6: Top 15 suffixes.

Suffix value	Length	
Brian Buckley Caitlin Cornell Alyssa Fuller Eric Hojnowski Ryan LaFollette Yelena		
Livshits Todd Michaelis Claire Motyl Tarakad Ramachandran Devan Rahmachan-		
drin Sofia Seckler Evaline Tso And Kate Zmijewski-Mekeem		
European Society Of Clinical Microbiology And Infectious Diseases Escmid Vaccine	98	
Study Group Evasg		
(Conceptualization; Review and editing; Read and approved final version of		
manuscript)		
Faculty of Bioscience and Bioindustry, Tokushima University, Tokushima, Japan		
BA, MBBS (Hons), FRANZCP, PhD, Dip Psychodynamic Psychotherapy, Cert ATP	72	
on behalf of the Portuguese visual impairment study group (PORVIS-group)		
(Writing original draft; Read and approved final version of manuscript)		
RN, Cert Psych Nurs, BA (Hons), Dip Ed, B Ed, M Ed, PhD, FACMHN		
DVM, PhD, Diplomate ABVP (Dairy Practice), SFHEA, NVS, MRCVS		
B Phil (Hons), B Soc & Soc & Community Development)	60	

Table 7: Ten longest suffixes.

Suffix value	Occurrences
*	32
S	12
K	11
W	11
J	8
F	8
†	8
A	7
P	7
M	5

Table 8: Top 10 shortest suffixes.

The PubMed DTD does not have a dedicated field for an email address. From 1996, NLM included "the first author's electronic mail (e-mail) address at the end of <Affiliation>, if present in the journal. Furthermore, as of October 1, 2013, NLM no longer edits affiliation data to add e-mail address" [11]

A word of caution about relying on email addresses as a discriminator for author name disambiguation; the most common email address is user@example.com which occurred 2023 times in the MEDLINE dataset of this study. Additionally, there are other non-specific email addresses such as journal editorial mailboxes.

Since 2010, the PubMed DTD has included an Identifier element, which has been used from 2013 [11]. However, it has less than 3% completeness (Table 9) and it is worth noting that there are occurrences where the same ORCID identifier has been incorrectly allocated to multiple authors within a paper.

Identifier	Completeness	Validity	Uniqueness
ORCID	2.820%	99.915%	40.921%

Table 9: Author ORCID measures.

## 3.2 Data quality issues related to affiliation names

An author's institutional affiliation is a very important information field, but the completeness is only around 42%. We have not derived a validity score, but there are quality problems within that set that are obvious from the length distributions (Figure 5). As previously mentioned, this field may contain values that aren't written in English as well as non-ASCII characters.

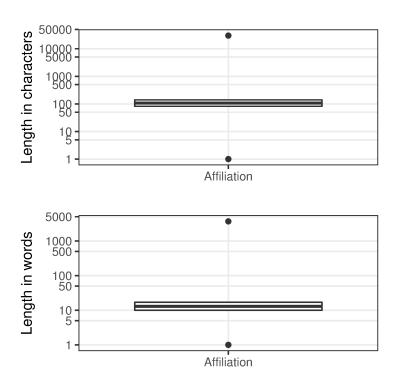


Figure 5: Affiliation character / word distributions.

In Figure 5, the outliers at the top of the range, which we have termed "narrative affiliations", typically describe the affiliations for many, if not all, of the contributors to the paper (e.g., see Figure 6 where we show the entry from the article with PMID 32308221). These narrative affiliations may also be repeated for all the author entries within the author list. At the other end of the range, there are many incomplete, or indistinguishable entries (Table 10).

#### **Affiliations**

- 1 Amy Meyer, is at the University of Missouri School of Medicine, Columbia, Missouri. Hariharan Regunath, MD, MSMA member since 2019, is in the Department of Medicine, Division of Pulmonary, Critical Care and Environmental Medicine, and Division of Infectious Diseases, University of Missouri, Columbia, Missouri.
- 2 Christian Rojas-Moreno, MD, William Salzer, MD, and Gordon Christensen, MD is in the Department of Medicine, Division of Infectious Diseases, University of Missouri, Columbia, Missouri.

Figure 6: An example of narrative affiliations.

Affiliation string	Occurrences
	5,761
,.	2,463
London, UK.	601
Editor-in-Chief.	468
London.	405
Pathology.	360
GSK, Siena, Italy.	342
Duke University.	341
Harvard University.	332
McGill University.	329
Paris, France.	323
School of Medicine.	303
Yale University.	301
Editor.	295
Radiology.	262

Table 10: Top 15 affiliations under 20 characters long.

Our parsing has not included any special case exclusions. We note that pubmed\_parser [12] excludes "For a full list of the authors' affiliations please see the Acknowledgements section." - though this exact string only occurs once within our selected dataset of over 51 million affiliation strings! It should also be noted that "as of October 1, 2013, NLM no longer performs quality control of the affiliation data" [11].

Whilst multiple affiliations were possible from the 2015 DTD [11], this is a good place to mention how some data providers concatenate multiple affiliations for an author in a single element. Here is an example for Yong-Beom Park (PMID 29465366):

Division of Rheumatology, Department of Internal Medicine, Yonsei University College of Medicine, Seoul; and Institute for Immunology and Immunological Diseases, Yonsei University College of Medicine, Seoul, Republic of Korea.

Affiliation identifiers, such as ISNI and GRID, were possible from the 2015 DTD [11]. We've captured values for those too in Table 11.

Identifier	Completeness	Validity	Uniqueness
ISNI	0.002%	99.965%	22.803%
GRID	0.003%	100.000%	23.752%
Affiliation	42.526%	N/A	45.979%

Table 11: Key measures for Affiliations / Affiliation identifiers.

# 3.3 Data quality issues related to articles

#### 3.3.1 Article persistent identifiers

As can be seen in Table 12, the application of digital object identifiers (DOI), although not perfect, reaches a respectable score in terms of uniqueness but there are issues with validity of those identifiers and a significantly low score in terms of completeness; we'll examine the impact that earlier publications have on DOI completeness.

Identifier	Completeness	Validity	Uniqueness
DOI	71.373%	99.377%	99.949%

Table 12: MEDLINE article identifiers.

## 3.3.2 Publication year

In the full PubMed database, there are over 100,000 records with a publication year earlier than 1900. In our selected data set from MEDLINE, there are only 3 that are clearly wrong (Table 13). In the first two examples, the publication year has the upper value from the journal pagination range. These erroneous publication years caused Parquet compatibility problems with Spark 3 (see issue SPARK-31404: https://issues.apache.org/jira/browse/SPARK-31404) when constructing a Date column, as they pre-date the introduction of the Gregorian calendar in 1582 and Spark implements a Proleptic Gregorian calendar as of version 3.

PMID	Publication Year
11662976	1132
11665278	1041
32422596	1

Table 13: Example of erroneous publication year values.

Figure 7 illustrates the volume of citation records with a valid DOI per publication year with 2022 in progress. Note that as of Q1 2022 there are not yet articles scheduled for publication in subsequent years.

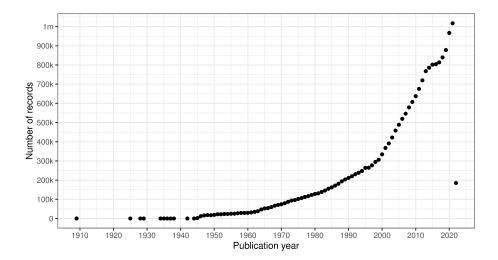


Figure 7: Count of citation records with a valid DOI per publication year (excluding erroneous years).

## 3.3.3 Abstract

The abstract field was added to the PubMed record in 1975 [11]. The abstract text, which may be subject to copyright restrictions, is a prime candidate for text mining. Consequently, for the two-thirds of records with an abstract, it's useful to understand their length distribution (Figure 8) and the erroneous values that they contain. Whilst the uniqueness is 99.942%, there is still a significant number (over 11 thousand abstracts) with non-unique abstract values. From the length information, we can infer that there are clearly meaningless abstract entries towards the lower end of these ranges, as seen in Table 14.

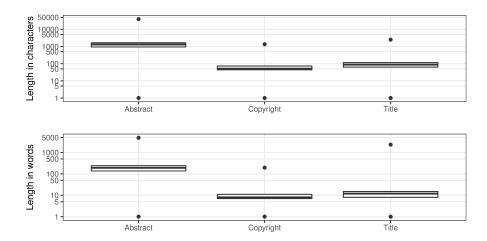


Figure 8: Article character / word distributions.

Abstract text	Occurrences
[Figure: see text].	579
	182
Not available.	106
N/A.	51
n/a.	50
no summary.	48
Null.	41
NA.	29
No Abstract.	22
<p></p> .	20
Editorial.	17
EDITORIAL.	16
	13
No abstract.	13
None.	10

Table 14: Top 15 abstracts under 20 characters long.

# 3.3.4 Copyright

An important consideration when mining MEDLINE should be whether copyrighted material is being used. The NLM terms and conditions clearly state that they do not provide legal advice [28]. The copyright information field was introduced in 1999 [11], with a completeness measure of almost 22% of the records that have an abstract. From Table 15, it is evident that Elsevier is

most consistent in supplying copyright statements although there is some lack of consistency regarding the actual values. Figure 8 shows the distributions of character length and word tokens, it should be clear that at the low end of the range there must be some invalid values (Table 16).

Copyright information	Occurrences
Copyright © 2020 Elsevier Inc. All rights reserved.	40,773
© 2021. The Author(s).	39,577
Copyright © 2020 Elsevier B.V. All rights reserved.	39,221
Copyright © 2020 Elsevier Ltd. All rights reserved.	39,220
Copyright © 2016 Elsevier Inc. All rights reserved.	38,600
Copyright © 2019 Elsevier Inc. All rights reserved.	37,672
Copyright © 2018 Elsevier Inc. All rights reserved.	37,414
Copyright © 2017 Elsevier Ltd. All rights reserved.	36,833
Copyright © 2017 Elsevier Inc. All rights reserved.	36,817
Copyright © 2018 Elsevier Ltd. All rights reserved.	36,766

Table 15: Top 10 copyright statements.

Copyright information	Occurrences
© 2013.	6,941
excerpt	4,996
© The author(s).	3,193
© FASEB.	1,444
full text	1,238
©2011 AACR.	1,159
©2013 AACR.	1,145
©2012 AACR.	958
Celsius.	956
© 2017 The Authors.	925

Table 16: Top 10 short copyright statements.

#### 3.3.5 Title

MEDLINE has just over 7,500 records without an ArticleTitle element, leading to a completeness value of 99.974%. The uniqueness of the title field is approaching 98%. Like our observations for the abstracts, there are standard article titles that relate to the publication type towards the lower end of the character length and number of word token ranges (Figure 8; see also Table 17).

Article title	Occurrences
[Not Available].	13,440
Reply.	1,972
Invited commentary.	1,896
Editorial comment.	1,676
Editorial.	1,465
Response.	1,312
Discussion.	1,052
Editorial Comment.	1,051
Preface.	974
The authors reply.	768
In reply.	714
Introduction.	585
In Reply.	519
Authors' response.	469
Foreword.	428

Table 17: Top 15 article titles under 20 characters long.

#### 3.3.6 Language

Another important consideration for text mining is the language, or languages, that the article is published in. It should be noted that PubMed includes translated titles, in square brackets, where appropriate. The language element contains language codes from the US Library of Congress MARC [29] schema, such as "chi" for Chinese. The language code table [30] includes "und" for undetermined and "mul" for multiple languages. However, language codes can also be concatenated together; for example, "fregerita" means the article was published in French, German, and Italian.

The language field is complete for the entirety of the MEDLINE records, but if we treat a solitary value of "und" or "mul" (238,470 and 1,399 occurrences, respectively) as invalid then the validity of this field is 99.55%. This excludes cases where they are present with other values too. From a recency perspective, "und" last occurred in 2002, and that is the only occurrence since 1985; "mul" occurred once in both 2016 and 2015, but before that it was last seen in 2011.

The maximum number of languages specified for a record is 6, but the 75th percentile is 1. Considering the values individually by splitting the strings and exploding the resulting array, allows us to produce the top 10 languages (Table 18). Note that almost 84% of records within the MEDLINE sample are published in English. The next most common language, German, only accounts for about 3% of articles.

Language code	Occurrences
eng	24,290,379
ger	861,109
fre	744,111
rus	697,806
jpn	429,283
spa	364,920
chi	329,153
ita	305,526
und	239,588
pol	172,956

Table 18: Top 10 languages.

### 3.4 Data quality issues related to journals

The key identifier provided in MEDLINE for a journal is the US National Library of Medicine (NLM) identity. When compared to the J\_MEDLINE reference data set of MEDLINE indexed journals [8], the NLM identifiers have a referential integrity [15] measurement of 99.989%. There were 146 NLM identifiers that were not included within the J\_MEDLINE dataset, affecting 3,045 articles. When considering a graph representation of the dataset, this would result in dangling edges that may not be permitted by some graph storage engines, such as Neo4j.

### 3.5 Data quality issues related to time evolution

In this section we consider the change over time for some of the key identifiers. Are there any obvious trends in whether identifiers are becoming more pervasive or prevalent in newer citation records? Here are some general observations: DOIs are almost ubiquitous for new articles (Figure 9), ORCIDs have been on the rise to just under 17% of authors per year (Figure 10), but GRID and ISNI usage peaked in 2017, having first appeared in 2015 (Figure 11). That leaves us with the tedious task of disambiguating the affiliation of the authors in the records. As can be seen in Figure 12, the vast majority of recent records contain an affiliation string for all authors; this is due to a policy change in 2014 to collect affiliations for all contributors [11].

### 4 Conclusions

PubMed is an enormously valuable resource for the biomedical sciences and healthcare, yet, those attempting to identify authors and affiliations, or otherwise use the records from that database, need to be aware of the quality issues within the dataset. This article has highlighted some of those data quality concerns.

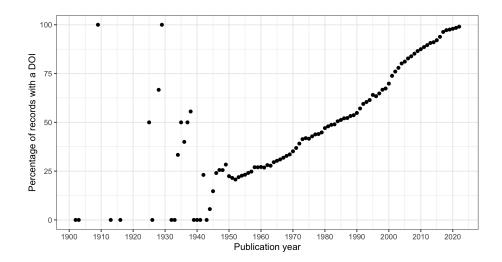


Figure 9: DOI percentage of articles per publication year.

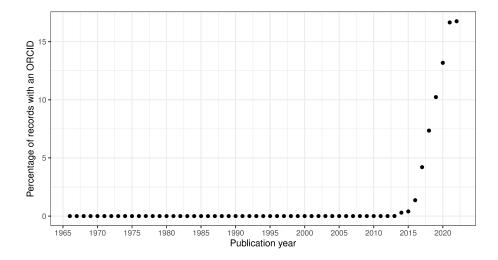


Figure 10: ORCID percentage of authors per publication year.

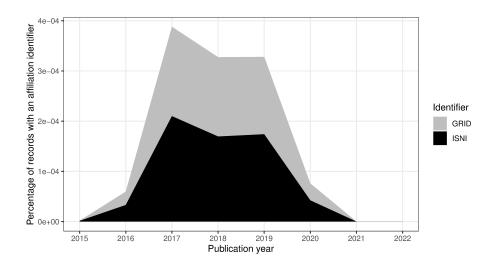


Figure 11: ISNI & GRID percentage of authors per publication year.

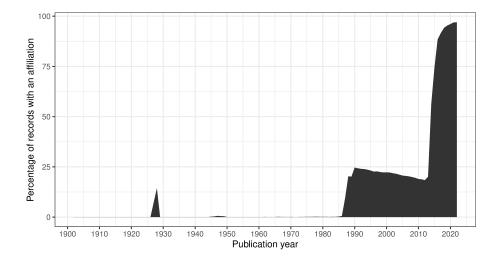


Figure 12: Percentage of authors per publication year with an affiliation string.

The data are subject to many human errors, such as typographical errors, and system related errors such as inconsistent representations of author names (leading to the synonym problem) and affiliations. There is a lack of author identifiers (contributing to the homonym problem) and a significant lack of affiliation identifiers. Being an aggregated source, the PubMed database suffers from multi-source problems such as inconsistent representations from the upstream XML providers that result in a high degree of lexicographic entropy.

In summary, our work supports the following conclusions:

- Given the incompleteness and uniqueness of identifying fields, the disambiguation of author names remains a significant problem for PubMed, particularly for records dating before 2014.
- PubMed has excellent integrity for NLM-internal identifiers (e.g., MeSH), though there is the noted exception around the J\_MEDLINE dataset. Beyond the NLM database, the majority of articles are labelled with a DOI, and the DTD provides support for identifiers for authors, institutions, both of which are far from complete. The DTD also caters for grant information, and auxiliary data through the DataBank elements, though these were beyond the scope of our work.
- Overall, there is an improvement in the use of identifiers; in particular, records created since 2015 exhibit an increase in external identifiers. However, the data quality for institutional identifiers is poor and their use has been diminishing over time.

Unless the data quality issues are addressed retroactively, they will weaken (if not entirely distort) any subsequent data analysis. Perhaps, an intervention in current publishing systems, to prevent the data sources of PubMed from manifesting the data quality issues mentioned herein, is the best one can hope for the future. Much like the application of machine learning has been applied within the NLM for indexing (e.g., with the MTI tooling [31]), the NLM could enhance their process with systems that possess a learning architecture to improve and accelerate the curation of the PubMed records. It is also possible that another information provider will provide an open data repository containing cleansed PubMed data, although a proprietary offering is more likely.

Another possibility for better use of the PubMed treasure trove is the creation of an open source library for cleansing the data, or at least properly identify the data quality issues, and optimize the amount of information that one can obtain from processing the PubMed records. Once this is accomplished with one programming language the open source community can augment the library and expand its adoption in other programming languages, for example by porting the library.

Lastly, the community would benefit from the availability of open source libraries that can accurately perform author name disambiguation, or a substantial set of "gold data" that can be used for training and validation; that dataset, however, should be orders of magnitude larger than the ones that are currently available (e.g., the 'amorgani/AND' dataset [32] [33]).

### References

- [1] DTMBIO '10: Proceedings of the ACM fourth international workshop on Data and text mining in biomedical informatics. URL: https://dl.acm.org/doi/proceedings/10.1145/1871871.
- [2] ACM Digital Library. URL: https://dl.acm.org.
- [3] Lee J et al. "BioBERT: a pre-trained biomedical language representation model for biomedical text mining." In: *Bioinformatics* 36.4 (2020), pp. 1234–1240. DOI: 10.1093/bioinformatics/btz682.
- [4] LitCovid, NLM. URL: https://www.ncbi.nlm.nih.gov/research/coronavirus/.
- [5] About PubMed, NLM. URL: https://pubmed.ncbi.nlm.nih.gov/about/.
- [6] Rahm and Do. "Data Cleaning: Problems and Current Approaches". In: Bulletin of the IEEE Computer Society Technical Committee on Data Engineering (2000).
- [7] MeSH: Medical Subject Headings. URL: https://www.nlm.nih.gov/mesh/meshhome.html.
- [8] Dataset of MEDLINE indexed journals. URL: https://ftp.ncbi.nlm.nih.gov/pubmed/J\_Medline.txt.
- [9] Sanyal DK, Bhowmick PK, and Das PP. "A review of author name disambiguation techniques for the PubMed bibliographic database". In: *Journal of Information Science* 47.2 (2021), pp. 227–254. DOI: 10.1177/0165551519888605.
- [10] PubMed 2019 DTD. URL: http://dtd.nlm.nih.gov/ncbi/pubmed/out/pubmed\_190101.dtd.
- [11] MEDLINE PubMed XML Element Descriptions and their Attributes. URL: https://www.nlm.nih.gov/bsd/licensee/elements\_descriptions.html.
- [12] Titipat Achakulvisut, Daniel E. Acuna, and Konrad Kording. "Pubmed Parser: A Python Parser for PubMed Open-Access XML Subset and MEDLINE XML Dataset XML Dataset". In: Journal of Open Source Software 5.46 (2020), p. 1979. DOI: 10.21105/joss.01979. URL: https://doi.org/10.21105/joss.01979.
- [13] pymed. URL: https://github.com/gijswobben/pymed.
- [14] pubmed-parser. URL: https://github.com/thecloudcircle/pubmed-parser.

- [15] DAMA International. DAMA DMBOK Data Management Body of Knowledge. 2nd Edition. New Jersey, USA.: Technics Publications, 2017.
- [16] PubMed baseline download files. URL: https://ftp.ncbi.nlm.nih.gov/pubmed/baseline/.
- [17] PubMed daily update files. URL: https://ftp.ncbi.nlm.nih.gov/pubmed/updatefiles/.
- [18] Michael Armbrust et al. "Spark SQL: Relational Data Processing in Spark". In: Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data. SIGMOD '15. Melbourne, Victoria, Australia: Association for Computing Machinery, 2015, pp. 1383–1394. ISBN: 9781450327589. DOI: 10.1145/2723372.2742797. URL: https://doi.org/10.1145/ 2723372.2742797.
- [19] The spark-xml project. URL: https://github.com/databricks/spark-xml.
- [20] XQuery 3.0 standard. URL: https://www.w3.org/TR/2014/REC-xquery-30-20140408/.
- [21] Saxon-HE library, Saxonica. URL: http://www.saxonica.com/.
- [22] The spark-xml-utils project. URL: https://github.com/elsevierlabs-os/spark-xml-utils.
- [23] Apache Zeppelin. URL: https://zeppelin.apache.org.
- [24] Digital Object Identifier (DOI). URL: https://www.doi.org.
- [25] Open Researcher and Contributor ID (ORCID). URL: https://orcid.org.
- [26] ISO 27729, International Standard Name Identifier (ISNI). URL: https://isni.org.
- [27] Global Research Identifier Database (GRID), Digital Science. URL: https://www.grid.ac.
- [28] NLM terms and conditions. URL: https://www.nlm.nih.gov/databases/download/terms\_and\_conditions.html.
- [29] MARC definition, US Library of Congress. URL: https://www.loc.gov/marc/faq.html#definition.
- [30] MEDLINE/PubMed Language Table. URL: https://www.nlm.nih.gov/bsd/language\_table.html.
- [31] NLM Medical Text Indexer (MTI). URL: https://lhncbc.nlm.nih.gov/ii/tools/MTI.html.
- [32] Vishnyakova D, Rodriguez-Esteban R, and Rinaldi F. "A new approach and gold standard toward author disambiguation in MEDLINE." In: *J Am Med Inform Assn* 26 (2019), pp. 1037–1045. DOI: 10.1093/jamia/ocz028.
- [33] Vishnyakova D et al., AND Author Name Disambiguation corpus. URL: https://github.com/amorgani/AND/.

# The State of Scholarly Metadata: 2023



In late 2022, CCC and Media Growth Strategies undertook a thorough examination of metadata management across the research lifecycle.

This in-depth review builds on an existing body of work to uncover multiple policy and system complexities and breakages, which – separately and together – create missed opportunities for the communities for whom Open Access (OA) and Open Science models are designed to serve.

CCC is sharing this analysis with the scholarly communications community to spark dialogue and to drive action. Drawn directly from our research interviews, this living infographic depicts the significant economic and social impact that a fragmented metadata supply chain has today on researchers, institutions, funders, and publishers. Researchers, in particular, shoulder a significant administrative burden that ultimately disrupts and delays the process of scientific discovery.

As the scholarly communications community continues its shift to full OA, stakeholders recognize that new strategies, inclusive policies, and a robust network of interoperable data and systems are essential for making critical infrastructure improvements, and much progress is underway. In that environment, a dedication to data stewardship across each stakeholder group, and the service providers supporting them, will lead not only to a smoother OA transition, but also to greater research integrity; data sharing; reliable, trustworthy metrics on research impact; and a responsive, equitable rewards and recognition system.

### Research stage

# **Idea Development**



# RESEARCHER

Researcher seeks collaborators; meets with colleagues and library / research office staff

### **CHALLENGES**

### **Underutilization of ORCID**

Some institutions don't require researchers to use ORCID; records can be outdated if authors don't consistently update; ORCID may not be accessible to authors in some geographies.

### **→** IMPACT

If authors can't be identified with a standard ID, they may not be able to authenticate to content, get credited appropriately for their work, secure OA funding, or complete downstream processes without unnecessary manual effort. Costly manual effort is also required of publishers, institutions, and funders to disambiguate authors retrospectively.

### Research stage

# **Proposal Submission**



### RESEARCHER

Researcher submits application for funding



Funder selects reviewers and begins application review



Funder logs funding terms in grant management system

### **CHALLENGES**

### **Inconsistent Metadata Capture**

Variability across grant application process/systems results in possible loss of metadata necessary to determine OA funding entitlements at a later stage, e.g., institutional affiliations.

# **CHALLENGES**

# **Legacy System Limitations**

Low adoption of standardized PIDs (FundRef, RAiD, Ringgold, ISNI, ROR) due to limitations of legacy systems and/or lack of awareness.

# **CHALLENGES**

# **Low-Quality Data**

Free text fields are great for gathering feedback; they're not designed to capture granular data like an organizational identifier. Researchers often confuse proposal numbers with grant IDs later in the publication process--they need structure to improve the accuracy of data capture.

# **→** IMPACT

Without disambiguated grant and funder details, grants may not be effectively utilized in later publication stages, leaving OA funding unclaimed and shifting coverage to research institutions. In an ecosystem that values a sustainable OA shift, this impacts everyone.

# **→** IMPACT

Hindered conflict of interest management among peer reviewers threatens research integrity, and low-quality data results in low accuracy of later-stage funding identification, tracking, and analysis of research output.

# **→** IMPACT

Lack of registered grant DOIs makes it difficult and costly to link funding to particular research outputs, resulting in missed OA opportunities as well as incomplete analysis to inform future funding investments.

# Research stage

# Research & Authoring



# RESEARCHER

Researcher conducts literature review



# RESEARCHER

Researcher posts pre-print / shares early outputs



# RESEARCHER

Researcher selects publication for submission

# **CHALLENGES**

# **Researcher Inequities & Research Barriers**

- Valid research coming from under-represented researchers is hard to find due to lack of metadata, including DOIs.
- Search and discovery are difficult due to inconsistency in identifying the user and enabling appropriate access to research.
- Authors from under-represented areas may not have equitable access to search and discovery services or equitable opportunities for publication.

# Global inequities hinder scientific progress.

# **CHALLENGES**

# **Poor Connections Across Research Outputs**

Lack of persistent identifiers (PIDs) and inconsistent application of PIDs across research outputs e.g., data sets, equipment, setting(s), samples, software

# **CHALLENGES**

# Risk of OA non-compliance

Metadata lost upstream makes managing funding compliance onerous.

# **→** IMPACT

Inability to easily find, verify, and reuse the data and artifacts underlying research, making it difficult to accurately interpret, cite and reproduce research findings.

# **→** IMPACT

Lack of available information about both corresponding author and all co-authors leads to manual input to identify funder and institutional mandates at best and missed funding requirements at worst.

# Research stage Publication



RESEARCHER

Researcher submits article



NSTITUTION

Institution funds OA publication



PUBLISHE

Publisher indexes metadata to enable search & discovery

### CHALLENGES

### **Missed Funding Opportunities**

- Under-utilization of metadata validation services
- If the researcher has submitted before, outdated information from their existing profile can be pulled into the submission
- Inconsistency between journal policies and metadata procedures
- Lack of funding information captured at submission and validated at acceptance
- Demand for increased interoperability between IDs

### **CHALLENGES**

# Missed Funding Opportunities & Costly Billing Complications

If funder/institution information manually input by the author does not use a standardized name or PID (e.g., abbreviations, nicknames), this can interfere with matching to the correct OA funding source.

### **CHALLENGES**

### **Unnecessary Manual Intervention**

Publishers are sometimes manually entering PIDs prior to registering DOIs for a more complete publication record.

# **→** IMPACT

Without granular, accurate organizational affiliation identifiers for a manuscript, coupled with incomplete funding details, authors may miss the opportunity to get OA funding or miss the chance to opt into OA due to affordability concerns. OA initiatives driven by institutions and funders may lack uptake as a result. Publishers are also unable to automate processes that reduce the cost of business model transformation. Manual effort is required to retrospectively cover the publication with proper funding sources, driving up the cost of publishing. No one benefits in this scenario.

# **→** IMPACT

Publishers and institutions take on the time and expense of manually finding the papers that should have matched to an agreement and collaborating on a resolution. Funding decisions cannot be based on abbreviations or free-form data.

# **→ IMPACT**

This is a laborious practice with high economic and opportunity costs that could be reduced with earlier, automated PID assertion and/or validation.

## Research stage

### Reuse & Measurement



RESEARCHER

Researcher evaluates research impact



### MOTITITION

Institution assesses historical subscription & publication data to inform institutional deals



FUNDER

Funder evaluates research impact



**PUBLISHER** 

Publisher assesses historical subscription and publication to inform institutional deals

### **CHALLENGES**

### **Problematic Research Impact Measurement**

Difficult to track research/researcher impact due to lack of adoption of metadata standards.

### **CHALLENGES**

### **Problematic Deal Modeling**

- Lack of consistent affiliation and funding data makes modeling future agreements hard for institutions.
- Data is not standardized across publisher platforms, creating unnecessary manual work to gather and normalize data for analysis.

# CHALLENGES

# **Problematic Research Impact Measurement**

Difficult to track funder impact due to lack of adoption of metadata standards.

# **CHALLENGES**

# **Problematic Deal Modeling**

Lack of consistent affiliation and funding data makes modelling future agreements difficult for publishers and institutions.

### **→** IMPACT

Researcher rewards and recognition decisions, or future opportunities for funding, may be based on incomplete or inaccurate data, affecting reputation and career advancement.

### **→** IMPACT

The transition to modern models of OA publication is delayed, prolonging a mixed-model landscape and the availability of open outputs to advance science.

# **★ IMPACT**

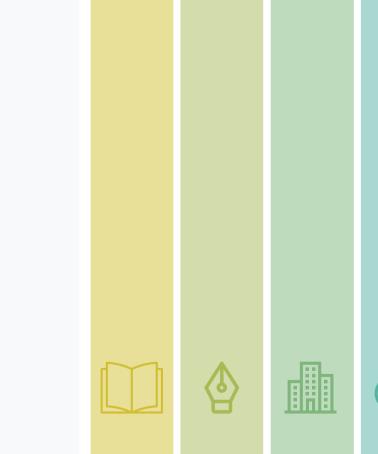
Incomplete analysis to support future funding investments and to report activities to the public.

# **★ IMPACT**

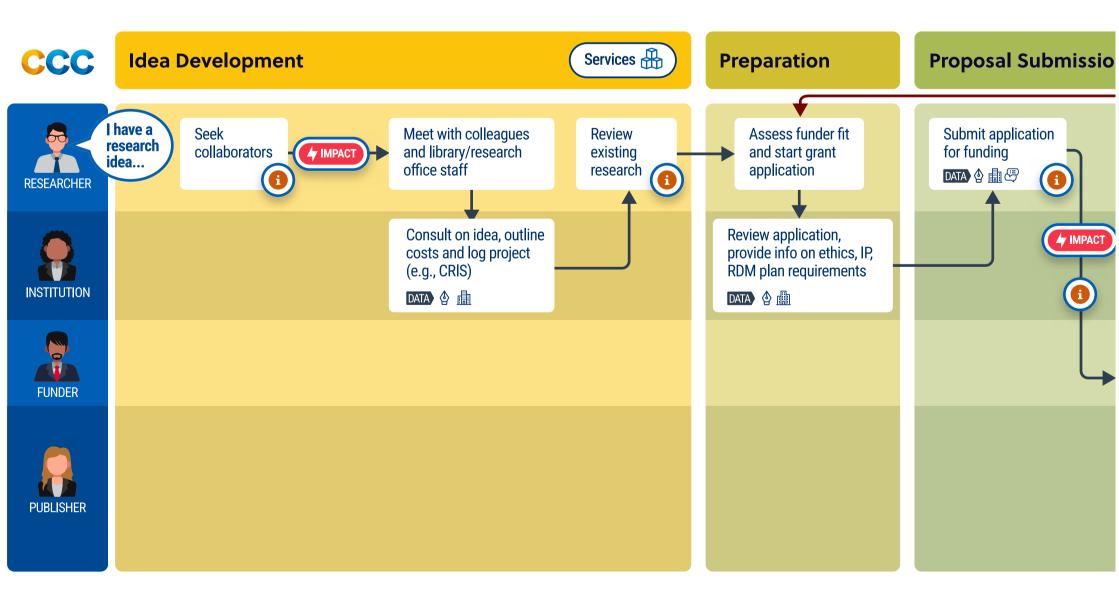
The transition to OA is delayed, putting some publishers at risk of losing authors to funding mandates and losing revenue that is necessary to sustain operations.

To view the interactive map, visit

# stateofmetadata.com







# **About CCC**

A pioneer in voluntary collective licensing, CCC (Copyright Clearance Center) helps organizations integrate, access, and share information through licensing, content, software, and professional services. With expertise in copyright, information management, artificial intelligence, and machine learning, CCC and its subsidiary RightsDirect collaborate with stakeholders to design and deliver innovative information solutions that power decision-making by harnessing information from a wide variety of data sources and content assets.



Email: <a href="mailto:rkaufman@copyright.com">rkaufman@copyright.com</a>

**Submit date: 4/24/2023** 

I am responding to this RFI: On behalf of an organization

Name: Lisa Braverman

Name of Organization: American Society for Radiation Oncology

Type of Organization: Professional org association

Role: Member of the public

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

At the American Society for Radiation Oncology (ASTRO), we recognize author groups are diverse and have obtained varying levels of federal funding. Requiring zero-embargo Open Access favors researchers who have sufficient funding to pay Article Publication Charges (APCs). Such a policy disadvantages early career researchers and research teams that, for reasons including family/medical leave, have received lower levels of funding. To combat this significant disparity, free Green OA routes should be publicized as the primary method of compliance with the OSTP mandate.

ASTRO strongly encourages the NIH to request a two-year extension from OSTP for mandate compliance, to January 1, 2028. The additional time will allow for greater author education and will help minimize disparities created by this policy.

#### 2. Steps for improving equity in access and accessibility of publications.

ASTRO supports a requirement that all publicly available versions of an article be linked to the publisher/professional society-supplied version of record. A two-year delay of required compliance with the OSTP mandate will enable a more robust, automated technical system to be implemented that will enable linking to occur by default. While research accessibility is critically important, confusion about medical article versioning is a danger to public health and must be avoided.

- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

DOIs have been successful and should be preserved. DOIs should be used to denote article versions of record. Preprints, Green OA depositions, research data, and any related materials should link back to a single DOI of the version of record. To enable this process and reduce confusion and public health threats regarding article versioning, federally funded research made publicly accessible within one week of article publication should be considered in compliance with the OSTP policy.

Email: lisa.braverman@astro.org

**Submit date: 4/24/2023** 

I am responding to this RFI: On behalf of myself

Name: Roy Kaufman

Name of Organization: Copyright Clearance Center

Type of Organization: Other

Type of Organization-Other: CCC is a not-for-profit organization founded in 1977 at the suggestion of

Congress to facilitate collective copyright licensing for the text sector.

**Role:** Member of the public

1. How to best ensure equity in publication opportunities for NIH-supported investigators.

- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

Please see the attached PDF, which contains our entire submission on these questions, with recommendations, attachments and links.

### **Uploaded File:**

NIH-RFI-with-attachments-1.pdf

**Description:** CCC's submission with recommendations, attachments and links.

Email: rkaufman@copyright.com



Response of Copyright Clearance Center (CCC) to Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research (RFI)

Notice Number:

NOT-OD-23-091

CCC welcomes the opportunity to submit this response to Question 4 of the NIH's <u>Request</u> for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported <u>Research</u>. More importantly, we welcome NIH's interest in the use of PIDs and metadata to increase findability and transparency of scientific research.

### Background on CCC.

CCC is a not-for-profit organization founded in 1977 at the suggestion of Congress to facilitate collective copyright licensing for the text sector. Presently, among other lines of business, CCC provides licenses to content from over 10,000 rightsholders for whom we serve as an agent. We provide these licenses to more than 35,000 business organizations (Business Users) around the world. CCC is a supplier of knowledge management software called RightFind®, which is used by a subset of these Business Users to manage and access content. We also provide (1) other software services, (2) library staffing, (3) content enrichment, data and metadata services, and (4) content delivery. On October 19, 2021, <u>U.S. Secretary of Commerce Gina Raimondo announced</u> that we were awarded a Market Development Cooperator Program grant, administered by the Commerce Department's International Trade Administration, to support our work with standards development organizations.

Our fastest growing business is managing the agreement- and fee-administration process on behalf of publishers who collect fees or otherwise track usage from authors, institutions, consortia, government and other funding bodies for immediate open access (OA). We do this primarily through our RightsLink® for Scientific Communications software platform (RLSC). RLSC is by far the market leader in managing open access agreements and payments, doing so for many of the top publishers of NIH-funded research.

#### PIDs and Metadata.

Through both our knowledge management work with Business Users and our work on behalf of publishers, CCC experiences firsthand the promise of persistent identifiers (PIDs) when applied early, consistently and persistently. We are also painfully aware of the problems

+1.978.750.8400

copyright.com

info@copyright.com

Phone

Fmail

Web

related to the entropy that results from lack of early, consistent, and persistent application thereof.

A healthy research and publishing ecosystem requires PIDs and robust, rich, quality metadata to make connections among people, organizations, places, and digital objects. For example, in RLSC alone, we depend on dozens of author, institution, and manuscript metadata elements to apply the appropriate business logic and workflows necessary to automate and scale OA on the path toward open science.

Even within a seemingly unified sector such as scientific communications, it is sometimes necessary to accommodate multiple PIDs serving the same purpose, such as organizational identifiers. While in some ways accommodating multiple PIDs increases work and decreases interoperability, PIDs have different scope, attributes, and audiences. Some users prefer PIDs with ISO certification, while others prefer PIDs with established business models to ensure sustainability and maintenance, while others focus on ability to use without cost to access PIDs. When one PID has been selected for use by a stakeholder as part of master data management, being forced to accommodate a different PID can have significant costs and introduce unnecessary friction. Accordingly at CCC, we accommodate a variety of organizational IDs in RLSC and have long preferred the features of Ringgold for our primary use.<sup>1</sup>

### Review of data quality of bibliographic records from the MEDLINE database

In 2022, three CCC colleagues reviewed the data quality of bibliographic records in the Medline database. A paper detailing the results of their research have been posted on bioRxiv and is attached to this document (Bramley, R, Howe, S, Marmanis, H 2022, Notes on the data quality of bibliographic records from the MEDLINE database, doi: <a href="https://doi.org/10.1101/2022.09.30.510312">https://doi.org/10.1101/2022.09.30.510312</a>; hereafter, "Bramley, et al"). As noted in the paper:

[T]he PubMed database, which contains over 33.8 million records collected over many decades, suffers from several data quality issues. These issues relate to, in part, character encodings, the absence of persistent identifiers, differences in human languages, and schema changes. These shortcomings should not be surprising since

222 Rosewood Drive Danvers, MA 01923 USA Phone +1.978.750.8400 Email info@copyright.com

Web copyright.com

<sup>&</sup>lt;sup>1</sup> CCC adopted Ringgold as its preferred organizational PID approximately 8 years ago. CCC acquired Ringgold in 2022 so that we could ensure its continued viability given its importance to ourselves and our clients.

PubMed aggregates information produced by different publishers and XML providers, a fact that leads naturally to the presence of "multi-source" problems.

Among the conclusions of the paper are (1) "[g]iven the incompleteness and uniqueness of identifying fields, the disambiguation of author names remains a significant problem for PubMed, particularly for records dating before 2014, and (2) [o]verall, there is an improvement in the use of identifiers; in particular, records created since 2015 exhibit an increase in external identifiers. However, the data quality for institutional identifiers is poor and their use has been diminishing over time."

### Mapping metadata management across the research lifecycle.

In late 2022, CCC and Media Growth Strategies undertook a thorough examination of metadata management across the research lifecycle. This review builds on an existing body of work to uncover multiple system complexities and breakages, which — separately and together — create missed opportunities for the communities for whom OA and open science models are designed to serve.

CCC has made this information publicly available in interactive infographic form at <a href="https://www.copyright.com/stateofmetadata/">https://www.copyright.com/stateofmetadata/</a>, and we have attached a chart summarizing where metadata breakages occur throughout the research lifecycle and how they impact various stakeholder groups. Drawn directly from research interviews, the infographic depicts the significant economic impact that a fragmented metadata supply chain is having today on researchers, institutions, funders, and publishers. Researchers in particular shoulder a significant administrative burden that ultimately disrupts and delays the process of scientific discovery.

The infographic is a living document which will be updated and modified based on ongoing community feedback.

As the scholarly communications community continues its shift to OA and open science, stakeholders require a robust network of interoperable systems for making critical and necessary improvements, and much progress is underway. In that environment, a dedication to data stewardship across each stakeholder group, and the service providers supporting them, will lead to greater data sharing; reliable, trustworthy metrics on research impact; and a responsive, equitable rewards system. NIH can lead the way.

+1.978.750.8400

copyright.com

info@copyright.com

Phone

Email

Web

Question 4 of the RFI states: "NIH seeks suggestions on any specific issues that should be considered in efforts to improve use of PIDs and metadata, including information about experiences institutions and researchers have had with adoption of different identifiers."

First, we recommend that NIH review the research, findings and recommendations set forth in Bramley, et al.

**Second**, NIH, as the premier funder of biomedical research in the US, is well positioned to help research and lead by example by requiring PIDs at appropriate points. As can be seen in the above-referenced infographic, grant application is one of the first organized parts of the lifecycle where PIDs can be effectively mandated. Once mandated and used, PIDs can flow throughout the lifecycle to improve everything from grant management to expression in PubMed. We urge NIH to review the infographic, sign up for updates, and provide feedback should NIH believe there are amendments and changes needed.

We have three specific recommendations with respect to mandated use of PIDs:

1. NIH should mandate that grant applications include organizations IDs for the institutions(s) affiliated with each researcher listed on the grant application, and Funder Registry IDs for the distinct funders of the grant. The requirement should insist that grant applications include at least one of the following organizational identifiers used in the scholarly publishing ecosystem and NIH should make metadata fields available for all four:

A. Ringgold- a proprietary global organization identifier system owned by CCC with over 600,000 unique records and rich hierarchical metadata used today by (1) most large and mid-sized commercial and non-commercial publishers, and (2) a range of critical infrastructure providers in the publishing ecosystem. For publishers, Ringgold often is part of a master data management strategy. Ringgold is also used by some funders, academic institutions, and consortia. Ringgold maps one-to-one with ISNI and the Funder Registry.

B. ISNI- ISO standard name identifier system with 1,697,000 unique organizational records of which a minimum of 500,000 are relevant to the research sector. ISNI is free to use and has been adopted by many national libraries. It lacks the hierarchical metadata of Ringgold but enjoys the rigor and authority of ISO accreditation. The

relevant organization records in ISNI map one-to-one with Ringgold.

- C. ROR- Research Organization Registry (ROR) is a global, community-led registry of open persistent identifiers for research organizations. ROR is free to use and has been adopted by some publishers, institutions, and overseas funders. It contains 104,000 unique identifiers and some hierarchical metadata. It can map to ISNI and the Funder Registry, but not on a one-to-one basis.
- D. Funder Registry (formerly known as FundRef) —Funder Registry is an open registry of grant-giving organization names and identifiers, with 32,000 unique identifiers for funders. It is donated by Elsevier to CrossRef and is updated approximately every 4-6 weeks. The Funder Registry ID can be used for author affiliations where the funder and affiliation are one and the same.
- 2. NIH should mandate that grant applicants include one or both of the following individual identifiers for all researchers in grant applications, and NIH should make metadata fields available for both.
  - a. ORCID- ORCID, which stands for Open Researcher and Contributor ID, is a global, not-for-profit organization sustained by fees from member organizations. ORCID is the most broadly adopted identifier system for individuals in scientific publishing.
  - b. ISNI- While not as well adopted as ORCID in research and science, ISNI has been broadly adopted in adjacent and non-adjacent fields.
- 3. NIH should mandate that appropriate PIDs be used at each stage reporting, while remaining flexible as to which PIDs it mandates, and should reevaluate its mandated PIDs on an ongoing basis. New PIDs such as RAiD (Research Activity Identifier) and DataCite (DOI-based system for research outputs) are being developed regularly and can help connect people, places and research. Likewise, other existing PIDs such as, e.g., Scopus Affiliation ID (AF-ID) and Author ID (AU-ID) are currently used in certain relevant applications. Appropriate PIDs should be mandated at each stage of the workflow, while recognizing that the needs of researchers and the availability of PIDs change over time.

+1.978.750.8400

copyright.com

info@copyright.com

Phone

Email

Web

As a final recommendation, we suggest that NIH follow the lead of Wellcome Trust and the Bill and Melinda Gates Foundation, among others, in registering grants for DOIs. This will help enable connectivity of PIDs and the discoverability of the grants, maximizing return to US taxpayers.

+1.978.750.8400

copyright.com

info@copyright.com

Phone

Email

Web

Respectfully submitted for Copyright Clearance Center by,

Roy S Kaufman

# Notes on the data quality of bibliographic records from the MEDLINE database

Robin Bramley\* Stephen Howe<sup>†</sup> Haralambos Marmanis<sup>†</sup>
August 17, 2022

#### Abstract

The US National Library of Medicine has created and maintains the PubMed® database, a collection of over 33.8 million records that contain citations and abstracts from the biomedical and life sciences literature. That database is an important resource for researchers and information service providers alike. As part of our work related to the creation of an author graph for coronaviruses, we encountered several data quality issues with records from a curated subset of the PubMed database called MEDLINE. We provide a data quality assessment for records selected from the MEDLINE database and report on several issues ranging from parsing issues (e.g., character encodings and schema definition weaknesses) to low scores against several data quality metrics (e.g., identifier completeness, validity, and uniqueness).

### 1 Introduction

PubMed is an enormously valuable resource for the biomedical and health fields. The PubMed database is a voluminous collection of medical literature citations that is free, easily accessible, and has been a data source for many works in the information retrieval and life sciences communities. As machine learning becomes more prevalent in various branches of the life sciences, the number of works that rely on the PubMed database increases. Many papers that cited PubMed have appeared within the proceedings of The International Conference on Data and Text Mining in Biomedicine series e.g., DTMBIO '10 [1]. In ACM's Digital Library[2], the year 2021 was a new high point at 235 for computing research articles that mentioned PubMed in the full-text collection, up from 1 in 1998 and 115 in 2010. Many information providers utilize the PubMed database, and there are a variety of machine learning models trained on PubMed[3]. It should be no surprise that, during the COVID-19 pandemic,

<sup>\*</sup>Copyright Clearance Center Limited, London, United Kingdom

<sup>&</sup>lt;sup>†</sup>Copyright Clearance Center Inc., Danvers, Massachusetts, United States of America

the PubMed database has been crucial in providing timely and frictionless access to the scientific literature [4].

However, the PubMed database, which contains over 33.8 million records [5] collected over many decades, suffers from several data quality issues. These issues relate to, in part, character encodings, the absence of persistent identifiers, differences in human languages, and schema changes. These shortcomings should not be surprising since PubMed aggregates information produced by different publishers and XML providers, a fact that leads naturally to the presence of "multi-source problems" [6].

MEDLINE is a curated subset of PubMed, its records are indexed with a controlled vocabulary called MeSH [7] and include information regarding funding, genetic, chemical, and other metadata. Articles in MEDLINE predominantly come from a set of indexed journals and a reference data file of these journals is available separately [8]. MEDLINE was made available online, through PubMed, in 1997.

In this article, we will provide an account of our experience in working with the curated MEDLINE records and report on the data quality issues that we encountered. We will describe, at length, the problem of Author Name Disambiguation, which is widely acknowledged as a source of errors when processing bibliographic databases in general, due to the challenges of synonyms (e.g., "John Doe", "John T Doe", and "JT Doe" referring to the same individual) and homonyms (i.e., two different people who share the same name such as "John Smith") [9]. Other problem areas that we will discuss include issues with character encodings, date related issues, the presence of persistent identifiers (and lack thereof), affiliation disambiguation, language related data issues, and schema data quality issues. Knowing how to address these challenges is valuable for practitioners who need to work with MEDLINE (or databases like MEDLINE) and process its records so that they can be used in their information systems.

#### 1.1 PubMed data

The PubMed database is available as XML, based on a DTD (currently the 2019 version) [10]. The compressed files are made available via an FTP server (they are also accessible by HTTPS) and each one of them contains up to 30,000 citation records. Every year, in mid- December, the data are consolidated and an annual baseline is produced. This is followed by incremental daily update files that include deletions.

A PubMed XML file has a root element of PubmedArticleSet that contains 1, or more, PubmedArticle or PubmedBookArticle children. The DTD also permits 0 or 1 DeleteCitation elements, and these can be seen in the update files. The elements of the PubmedArticle are divided into the MedlineCitation and the optional PubmedData - we have colloquially referred to these as the "front" and "back" matter respectively.

The description of the XML elements [11], also outlines potential discrepancies caused by schema changes, or policy changes to the collected data. For

example, records created before 2002 only contained author initials instead of full, first or middle, names; moreover, records between 1988 and 2013 only included the affiliation for the first author.

### 1.1.1 Known DTD shortcomings

There are two known problems with the DTD that have not yet been addressed. The first known problem is that authors cannot be linked to their CollectiveName. Some publishers have tried to work around this by interspersing CollectiveName elements and Author elements. In a wheat genome sequencing consortium paper (PMID 30115783), one of the contributors was a member of 12 groups, so that person appears as an Author record 12 times. This multiplicity complicates the author name disambiguation, as it may be impossible to distinguish a duplicate author entry from a valid homonym.

The second problem is related to a shortcoming in the 2019 DTD. Specifically, the back matter PubmedData element may contain a ReferenceList with many Reference elements, but it doesn't prevent the presence of many ReferenceList elements each with one Reference. Consequently, extraction must be able to handle both because both have been observed in the records. Furthermore, the ReferenceList definition permits deeply nested ReferenceList elements, as shown below:

<!ELEMENT ReferenceList (Title?, Reference\*, ReferenceList\*) >

#### 1.1.2 Escape characters

Escape sequence characters may appear within text fields such as the article title or abstract text. For example, if you wanted to represent a record in JSON, then you would have to beware of trailing backslashes and double quotes. Backslashes can also be problematic for the language used to parse the record. Furthermore, it may be necessary to remove other special characters such as new line characters (e.g., carriage return, line feed), tabs, and so on.

#### 1.1.3 Extended characters

PubMed encompasses articles published in many different languages, sometimes multiple languages. Consequently, fields such as the affiliation string, or parts of the author's name, may contain extended characters. This is an important consideration for the disambiguation of author names.

#### 1.2 Open Source libraries

Since PubMed has been a canonical source of biomedical citations, there are open source libraries to assist with parsing the records. Whilst none of these libraries were appropriate for our needs, they are included here for completeness.

For Python, pubmed\_parser [12] is an active project, but only handles a constrained field list. The pymed [13] project, which is now archived, only

parsed and cleansed a limited subset of the fields. It also seems that the design was intended to wrap the API.

For Java, there is pubmed-parser [14], which is based around the Java Architecture for XML Binding (JAXB). This project only had a short flurry of commits over 6 days in April 2021, consequently it is unclear whether this is actively maintained.

### 2 Materials and Methods

This work will identify challenges that can be faced when working with the MEDLINE data and categorize them along several dimensions of data quality [15].

### 2.1 Data acquisition

The PubMed baseline files were downloaded from their respective NLM FTP folders [16][17] and uploaded to separate folders on an S3 bucket.

### 2.2 Data processing

Figure 1 illustrates our data processing approach. The PubMed gzipped XML files were processed using Apache Spark 3.1.1 on Amazon EMR 6.3.1. The initial ingestion process extracted a few key properties, such as the PMID and DOI (from the PubmedData if present), before splitting the XML into two fragments representing the front matter (bibliographic metadata) and the back matter (references).

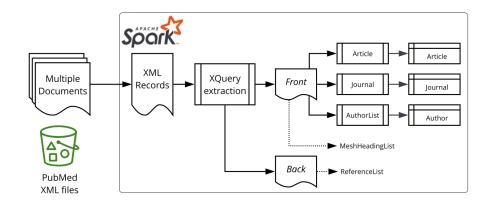


Figure 1: Data processing overview.

The baseline files were ingested first, then the update files were subsequently processed to apply updates, inserts and deletions. Record updates were applied

by sorting the records by their PMID in conjunction with the DateRevised property; only the newest records were retained. Note that the PMID Version attribute is not suitable for this purpose as it is only used by Public Library of Science (PLOS) records [11].

Spark SQL [18] is designed for tabular data, with the key construct being the DataFrame. Whereas XML documents are represented using a hierarchical structure that allows for repeating elements (a one-to-many relationship). This leads to an inherent mismatch between the two data formats that requires data transformation.

There is a spark-xml module [19], but we discovered during our initial experiments that the PubMed XML was too complex for spark-xml, as it resulted in heavily nested DataFrames, and incorrect query results. Consequently, we solved the XML to DataFrame impedance mismatch by performing an XQuery [20] operation per target entity type (e.g. Article, Author, etc.) as shown on the right-hand side of Figure 1.

The spark-xml XmlInputFormat class was retained for loading the XML files into Spark, with the ingestion and extraction utilizing XQuery queries to extract properties, via the Saxon-HE [21] library as provided by the Elsevier Labs spark-xml-utils [22] module.

To ease maintenance of the complex XQuery queries, we adopted a pattern whereby the XQuery output produces a JSON document. This makes the target property for a particular XPath or XQuery expression transparent (Figure 2) and inserting new elements does not break downstream code because it does not rely on positional information. The last part of that transformation phase is to leverage the read method of the SparkSession object which parses the JSON documents to DataFrame records. Note that Figure 2 also represents the handling of escape characters using the XQuery replace function.

```
'"forename": "', replace(replace($x/ForeName, '\\', '\\\\'), '"', '\\"'), '", ',
'"initials": "', replace(replace($x/Initials, '\\', '\\\\'), '"', '\\"'), '", ',
'"lastname": "', replace(replace($x/LastName, '\\', '\\\\'), '"', '\\"'), '", ',
'"suffix": "', normalize-space(replace(replace($x/Suffix, '^[,\\, |+', .''), '[,\\, '], \\]
```

Figure 2: JSON representation within XQuery.

### 2.3 Data analysis

The resulting DataFrames were analyzed using Spark SQL in Apache Zeppelin [23]. For string fields, we consider the length in characters and in words (by splitting on spaces). Metrics were rounded to 3 decimal places (or less).

The plots were produced in R, with the box plots using log-scale for the v-axis.

#### 2.4 Definitions

- $\mathcal{N} = \text{number of records}$
- $\mathcal{M}$  = number of records missing a value for the target property
- $\mathcal{D}$  = distinct values of those present (excludes null / blank)
- V defined by count of records matching a regex for identifiers (Table 1)
- $\mathcal{P} = \text{present} = \mathcal{N} \mathcal{M}$
- Completeness metric =  $\mathcal{P} / \mathcal{N}$
- $Validity metric = \mathcal{V} / \mathcal{P}$
- $Uniqueness metric = \mathcal{D} / \mathcal{P}$

Identifier	Regular expression
DOI [24]	"^10.\d{4,9}/[;()/:a-zA-Z0-9]+\$" <sup>1</sup>
ORCID [25]	$\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ $
ISNI [26] (presentation)	$"[0-9]{4} [0-9]{4} [0-9]{4} [0-9]{3}[0-9X]"$
ISNI (compact)	"[0-9]{15}[0-9X]"
GRID [27]	$"grid \. d{4,6} \. [0-9a-f]{1,2}"$

Table 1: Regular expressions for identifier validation

#### 2.5 Limitations of the study

The source dataset comprises the PubMed 2022 baseline plus daily update files to 1252 (30th March 2022).

It should be noted that our study includes only the PubmedArticle records, not the PubmedBookArticle records. The PubmedArticle records are only those from the MEDLINE subset (based on the Status attribute), and further excludes news articles, and those articles without a title; this gives a total of 28,986,590 article records. News articles were excluded from extraction because journalists, anecdotally those from the British Medical Journal, skew attempts to identify prolific authors through aggregation.

Other applied constraints are as follows:

- Only Author records with the ValidYN attribute of Y have been extracted, not Investigator records. For these 120,191,520 authors, only the first Affiliation element is considered.
- The DataBank element provides links to external datasets such as clinical trials. These identifiers were not investigated as part of the reported study.

<sup>&</sup>lt;sup>1</sup>Adapted from https://www.crossref.org/blog/dois-and-matching-regular-expressions/

- For alternative article identifiers, we did not extract the ELocationID element nor Publisher Item Identifiers (PII) from the PubmedData.
- For Journals, ISSNs were not analyzed.

#### 2.5.1 Approximation

Five number summary information is produced using Spark's DataFrameStatFunctions approxQuantiles method with an error margin of 0.0001, an example is shown below:

articleDF.stat.approxQuantile("doi\_len", Array(0.0,0.25,0.5,0.75,1.0),
0.0001)

However, the distinct counts do not leverage the Spark SQL approx\_count\_distinct function, rather the dataframe.select("column").distinct.count approach was used.

### 3 Results and discussion

In this section, we'll present our results related to data quality for the entities and fields shown in Figure 3. The PubMed XML data model is article-centric, but we will work our way from left to right.

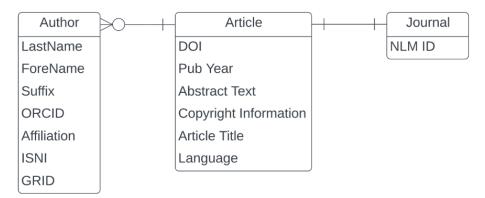


Figure 3: Entity Relationship Diagram for a subset of PubMed.

#### 3.1 Data quality issues related to author names

One of the important considerations regarding author records is that PubMed has not always recorded all the authors of a paper. The number of authors was limited to 10 between the years 1984 and 1995, and to 25 between the years 1996 and 1999 [11].

The most common last names in MEDLINE are Romanized Chinese names (Table 2), which can be very challenging to disambiguate. Looking at the length

characteristics (Figure 4), there are a few obvious problems, namely pollution of the author elements by incorrectly entered collective names (Table 3), and single character last names potentially caused by name transposition errors (Table 4).

LastName	Occurrences
Wang	1,086,073
Li	895,976
Zhang	878,544
Chen	722,753
Liu	703,743
Lee	547,636
Kim	523,687
Yang	433,439
Wu	360,532
Huang	309,375

Table 2: Top 10 LastName values.

LastName	Length
Endocrinology Genetics And Metabolism Group Pediatric Branch Of Chi-	322
nese Medical Association Neonatal Screening Group Specialist Committee For	
Prevention And Control Of Birth Defects Chinese Association Of Preven-	
tive Medicine Prevention And Control Committee Of Birth Defects Pediatric	
Branch Of Chinese Medical Association	
The Group Of Minimally Invasive Spinal Surgery And Enhanced Recovery	211
Professional Committee Of Orthopedic Surgery And Enhanced Recovery As-	
sociation Of China Rehabilitation Technology Transformation And Promotion	
Genetic Disease Society Guangdong Precision Medicine Application Associ-	209
ation Prenatal Diagnosis Group Maternal And Child Health Care Society	
Guangdong Medical Association Expert Committee Of Prenatal Diagnosis	
Arir Associazione Riabilitatori dell'Insufficienza Respiratoria Sip Società Ital-	201
iana di Pneumologia Aifi Associazione Italiana Fisioterapisti And Sifir Società	
Italiana di Fisioterapia E Riabilitazione	
This Paper Is A Co-Publication Between European Journal Of Preventive Car-	176
diology European Heart Journal Acute Cardiovascular Care And European	
Journal Of Cardiovascular Nursing	
Committee For Birth Defect Prevention And Control Chinese Association Of	174
Preventive Medicine Genetic Testing And Precision Medicine Branch Chinese	
Association Of Birth Health	
Consensus Group Of Experts On Application Of Metagenomic Next Genera-	152
tion Sequencing In The Pathogen Diagnosis In Clinical Moderate And Severe	
Infections	
Expert Committee Of The Inter-Laboratory Quality Assessment Of Prenatal	150
Screening And Diagnosis Clinical Test Center Of The National Health Com-	
mission	
For The Antimalarial Therapeutic Efficacy Monitoring Group National Malaria	142
Elimination Programme The Federal Ministry Of Health Abuja Nigeria	
On Behalf Of The Association Of Rural Surgeons Of India-Lancet Commission	142
On Global Surgery Consensus Committee Arsi-LCoGS Consensus Committee	

Table 3: Ten longest Last Name values.

LastName	Occurrences
S	756
A	704
Е	636
M	592
О	563
K	497
R	453
P	363
G	306
V	279

Table 4: Top 10 shortest LastName values.

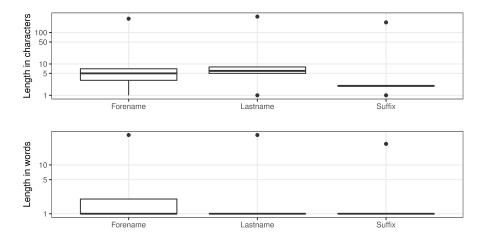


Figure 4: Author name character / word distributions.

The author forename field is 99.913% complete. Regarding the length, before 1945, the longest value in the forename field was 3 characters long, which reflects the policy to only hold author initials. The distributions, in Figure 4, clearly show that there are outliers. As shown in Table 5, these are primarily for working groups (a validity error), but the first row represents a different form of data preparation error where the affiliation has been concatenated with the forename.

PMID	LastName value	ForeName value	Length
34313229	Choi	Moon Hyung Department Of Ra-	276
		diology Eunpyeong St Mary's	
		Hospital College Of Medicine	
		The Catholic University Of Ko-	
		rea Seoul Republic Of Korea	
		Catholic Smart Imaging Cen-	
		ter Eunpyeong St Mary's Hos-	
		pital College Of Medicine The	
		Catholic University Of Korea	
		Seoul Republic Of Korea	
33145749	En Representación Del Grupo de	En Representación Del Grupo de	123
	Trastornos de la Conducta Y Del	Trastornos de la Conducta Y Del	
	Movimiento Durante El Sueño de	Movimiento Durante El Sueño de	
	la Sociedad Española de Sueño	la Sociedad Española de Sueño	
32329046	En Representación Del Grupo	En Representación Del Grupo	106
	de Estudio de Enfermedades	de Estudio de Enfermedades	
	Desmielinizantes de la Comu-	Desmielinizantes de la Comu-	
	nidad Autónoma de Madrid	nidad Autónoma de Madrid	
32433836	Pharmakopsychiatrie	The Therapeutic Drug Moni-	102
		toring Task Force Of The Ar-	
		beitsgemeinschaft Für Neuropsy-	
		chopharmakologie Und	

Table 5: ForeName values over 100 characters.

Completeness does not apply to author suffixes since not everyone has a suffix to their name. In terms of uniqueness there are 823 distinct values across 483,541 entries. There are also consistency issues, examples of which can be observed in Table 6 (e.g., Jr, Junior, Júnior). Figure 4 shows the range of suffix lengths and clearly indicates that there is something wrong with at least some records. When we look at the longest values for author suffixes (Table 7) and the most common single character values (Table 8), it becomes clear that there are multiple data issues related to the author suffix field; the general theme of misplaced values, or value "pollution", occurs across fields and is a major data quality weakness for the MEDLINE records.

Suffix value	Occurrences
Jr	374,510
3rd	74,260
2nd	20,364
4th	5,828
Sr	4,075
Junior	535
Júnior	380
Filho	241
PhD	238
5th	204
Neto	200
III	199
Dr	146
6th	129
MD	99

Table 6: Top 15 suffixes.

Suffix value	Length
Brian Buckley Caitlin Cornell Alyssa Fuller Eric Hojnowski Ryan LaFollette Yelena	
Livshits Todd Michaelis Claire Motyl Tarakad Ramachandran Devan Rahmachan-	
drin Sofia Seckler Evaline Tso And Kate Zmijewski-Mekeem	
European Society Of Clinical Microbiology And Infectious Diseases Escmid Vaccine	98
Study Group Evasg	
(Conceptualization; Review and editing; Read and approved final version of	86
manuscript)	
Faculty of Bioscience and Bioindustry, Tokushima University, Tokushima, Japan	
BA, MBBS (Hons), FRANZCP, PhD, Dip Psychodynamic Psychotherapy, Cert ATP	
on behalf of the Portuguese visual impairment study group (PORVIS-group)	
(Writing original draft; Read and approved final version of manuscript)	
RN, Cert Psych Nurs, BA (Hons), Dip Ed, B Ed, M Ed, PhD, FACMHN	
DVM, PhD, Diplomate ABVP (Dairy Practice), SFHEA, NVS, MRCVS	
B Phil (Hons), B Soc & Soc & Community Development)	

Table 7: Ten longest suffixes.

Suffix value	Occurrences
*	32
S	12
K	11
W	11
J	8
F	8
†	8
A	7
P	7
M	5

Table 8: Top 10 shortest suffixes.

The PubMed DTD does not have a dedicated field for an email address. From 1996, NLM included "the first author's electronic mail (e-mail) address at the end of <Affiliation>, if present in the journal. Furthermore, as of October 1, 2013, NLM no longer edits affiliation data to add e-mail address" [11]

A word of caution about relying on email addresses as a discriminator for author name disambiguation; the most common email address is user@example.com which occurred 2023 times in the MEDLINE dataset of this study. Additionally, there are other non-specific email addresses such as journal editorial mailboxes.

Since 2010, the PubMed DTD has included an Identifier element, which has been used from 2013 [11]. However, it has less than 3% completeness (Table 9) and it is worth noting that there are occurrences where the same ORCID identifier has been incorrectly allocated to multiple authors within a paper.

Identifier	Completeness	Validity	Uniqueness
ORCID	2.820%	99.915%	40.921%

Table 9: Author ORCID measures.

#### 3.2 Data quality issues related to affiliation names

An author's institutional affiliation is a very important information field, but the completeness is only around 42%. We have not derived a validity score, but there are quality problems within that set that are obvious from the length distributions (Figure 5). As previously mentioned, this field may contain values that aren't written in English as well as non-ASCII characters.

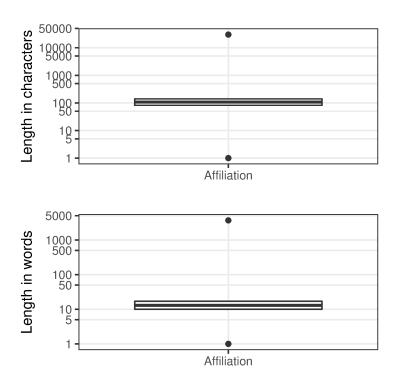


Figure 5: Affiliation character / word distributions.

In Figure 5, the outliers at the top of the range, which we have termed "narrative affiliations", typically describe the affiliations for many, if not all, of the contributors to the paper (e.g., see Figure 6 where we show the entry from the article with PMID 32308221). These narrative affiliations may also be repeated for all the author entries within the author list. At the other end of the range, there are many incomplete, or indistinguishable entries (Table 10).

#### **Affiliations**

- 1 Amy Meyer, is at the University of Missouri School of Medicine, Columbia, Missouri. Hariharan Regunath, MD, MSMA member since 2019, is in the Department of Medicine, Division of Pulmonary, Critical Care and Environmental Medicine, and Division of Infectious Diseases, University of Missouri, Columbia, Missouri.
- 2 Christian Rojas-Moreno, MD, William Salzer, MD, and Gordon Christensen, MD is in the Department of Medicine, Division of Infectious Diseases, University of Missouri, Columbia, Missouri.

Figure 6: An example of narrative affiliations.

Affiliation string	Occurrences	
	5,761	
,.	2,463	
London, UK.	601	
Editor-in-Chief.	468	
London.	405	
Pathology.	360	
GSK, Siena, Italy.	342	
Duke University.	341	
Harvard University.	332	
McGill University.	329	
Paris, France.	323	
School of Medicine.	303	
Yale University.	301	
Editor.	295	
Radiology.	262	

Table 10: Top 15 affiliations under 20 characters long.

Our parsing has not included any special case exclusions. We note that pubmed\_parser [12] excludes "For a full list of the authors' affiliations please see the Acknowledgements section." - though this exact string only occurs once within our selected dataset of over 51 million affiliation strings! It should also be noted that "as of October 1, 2013, NLM no longer performs quality control of the affiliation data" [11].

Whilst multiple affiliations were possible from the 2015 DTD [11], this is a good place to mention how some data providers concatenate multiple affiliations for an author in a single element. Here is an example for Yong-Beom Park (PMID 29465366):

Division of Rheumatology, Department of Internal Medicine, Yonsei University College of Medicine, Seoul; and Institute for Immunology and Immunological Diseases, Yonsei University College of Medicine, Seoul, Republic of Korea.

Affiliation identifiers, such as ISNI and GRID, were possible from the 2015 DTD [11]. We've captured values for those too in Table 11.

Identifier	Completeness	Validity	Uniqueness
ISNI	0.002%	99.965%	22.803%
GRID	0.003%	100.000%	23.752%
Affiliation	42.526%	N/A	45.979%

Table 11: Key measures for Affiliations / Affiliation identifiers.

### 3.3 Data quality issues related to articles

#### 3.3.1 Article persistent identifiers

As can be seen in Table 12, the application of digital object identifiers (DOI), although not perfect, reaches a respectable score in terms of uniqueness but there are issues with validity of those identifiers and a significantly low score in terms of completeness; we'll examine the impact that earlier publications have on DOI completeness.

Γ.	Identifier	Completeness	Validity	Uniqueness
	DOI	71.373%	99.377%	99.949%

Table 12: MEDLINE article identifiers.

#### 3.3.2 Publication year

In the full PubMed database, there are over 100,000 records with a publication year earlier than 1900. In our selected data set from MEDLINE, there are only 3 that are clearly wrong (Table 13). In the first two examples, the publication year has the upper value from the journal pagination range. These erroneous publication years caused Parquet compatibility problems with Spark 3 (see issue SPARK-31404: https://issues.apache.org/jira/browse/SPARK-31404) when constructing a Date column, as they pre-date the introduction of the Gregorian calendar in 1582 and Spark implements a Proleptic Gregorian calendar as of version 3.

PMID	Publication Year
11662976	1132
11665278	1041
32422596	1

Table 13: Example of erroneous publication year values.

Figure 7 illustrates the volume of citation records with a valid DOI per publication year with 2022 in progress. Note that as of Q1 2022 there are not yet articles scheduled for publication in subsequent years.

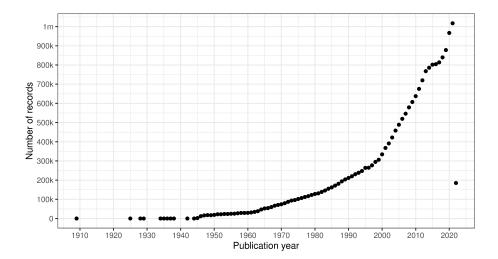


Figure 7: Count of citation records with a valid DOI per publication year (excluding erroneous years).

#### 3.3.3 Abstract

The abstract field was added to the PubMed record in 1975 [11]. The abstract text, which may be subject to copyright restrictions, is a prime candidate for text mining. Consequently, for the two-thirds of records with an abstract, it's useful to understand their length distribution (Figure 8) and the erroneous values that they contain. Whilst the uniqueness is 99.942%, there is still a significant number (over 11 thousand abstracts) with non-unique abstract values. From the length information, we can infer that there are clearly meaningless abstract entries towards the lower end of these ranges, as seen in Table 14.

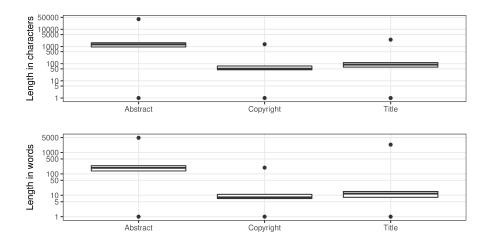


Figure 8: Article character / word distributions.

Abstract text	Occurrences			
[Figure: see text].	579			
	182			
Not available.	106			
N/A.	51			
n/a.	50			
no summary.	48			
Null.	41			
NA.	29			
No Abstract.	22			
<p></p> .	20			
Editorial.	17			
EDITORIAL.	16			
	13			
No abstract.	13			
None.	10			

Table 14: Top 15 abstracts under 20 characters long.

#### 3.3.4 Copyright

An important consideration when mining MEDLINE should be whether copyrighted material is being used. The NLM terms and conditions clearly state that they do not provide legal advice [28]. The copyright information field was introduced in 1999 [11], with a completeness measure of almost 22% of the records that have an abstract. From Table 15, it is evident that Elsevier is

most consistent in supplying copyright statements although there is some lack of consistency regarding the actual values. Figure 8 shows the distributions of character length and word tokens, it should be clear that at the low end of the range there must be some invalid values (Table 16).

Copyright information	Occurrences
Copyright © 2020 Elsevier Inc. All rights reserved.	40,773
© 2021. The Author(s).	39,577
Copyright © 2020 Elsevier B.V. All rights reserved.	39,221
Copyright © 2020 Elsevier Ltd. All rights reserved.	39,220
Copyright © 2016 Elsevier Inc. All rights reserved.	38,600
Copyright © 2019 Elsevier Inc. All rights reserved.	37,672
Copyright © 2018 Elsevier Inc. All rights reserved.	37,414
Copyright © 2017 Elsevier Ltd. All rights reserved.	36,833
Copyright © 2017 Elsevier Inc. All rights reserved.	36,817
Copyright © 2018 Elsevier Ltd. All rights reserved.	36,766

Table 15: Top 10 copyright statements.

Copyright information	Occurrences			
© 2013.	6,941			
excerpt	4,996			
© The author(s).	3,193			
© FASEB.	1,444			
full text	1,238			
©2011 AACR.	1,159			
©2013 AACR.	1,145			
©2012 AACR.	958			
Celsius.	956			
© 2017 The Authors.	925			

Table 16: Top 10 short copyright statements.

#### 3.3.5 Title

MEDLINE has just over 7,500 records without an ArticleTitle element, leading to a completeness value of 99.974%. The uniqueness of the title field is approaching 98%. Like our observations for the abstracts, there are standard article titles that relate to the publication type towards the lower end of the character length and number of word token ranges (Figure 8; see also Table 17).

Article title	Occurrences		
[Not Available].	13,440		
Reply.	1,972		
Invited commentary.	1,896		
Editorial comment.	1,676		
Editorial.	1,465		
Response.	1,312		
Discussion.	1,052		
Editorial Comment.	1,051		
Preface.	974		
The authors reply.	768		
In reply.	714		
Introduction.	585		
In Reply.	519		
Authors' response.	469		
Foreword.	428		

Table 17: Top 15 article titles under 20 characters long.

#### 3.3.6 Language

Another important consideration for text mining is the language, or languages, that the article is published in. It should be noted that PubMed includes translated titles, in square brackets, where appropriate. The language element contains language codes from the US Library of Congress MARC [29] schema, such as "chi" for Chinese. The language code table [30] includes "und" for undetermined and "mul" for multiple languages. However, language codes can also be concatenated together; for example, "fregerita" means the article was published in French, German, and Italian.

The language field is complete for the entirety of the MEDLINE records, but if we treat a solitary value of "und" or "mul" (238,470 and 1,399 occurrences, respectively) as invalid then the validity of this field is 99.55%. This excludes cases where they are present with other values too. From a recency perspective, "und" last occurred in 2002, and that is the only occurrence since 1985; "mul" occurred once in both 2016 and 2015, but before that it was last seen in 2011.

The maximum number of languages specified for a record is 6, but the 75th percentile is 1. Considering the values individually by splitting the strings and exploding the resulting array, allows us to produce the top 10 languages (Table 18). Note that almost 84% of records within the MEDLINE sample are published in English. The next most common language, German, only accounts for about 3% of articles.

Language code	Occurrences
eng	24,290,379
ger	861,109
fre	744,111
rus	697,806
jpn	429,283
spa	364,920
chi	329,153
ita	305,526
und	239,588
pol	172,956

Table 18: Top 10 languages.

#### 3.4 Data quality issues related to journals

The key identifier provided in MEDLINE for a journal is the US National Library of Medicine (NLM) identity. When compared to the J\_MEDLINE reference data set of MEDLINE indexed journals [8], the NLM identifiers have a referential integrity [15] measurement of 99.989%. There were 146 NLM identifiers that were not included within the J\_MEDLINE dataset, affecting 3,045 articles. When considering a graph representation of the dataset, this would result in dangling edges that may not be permitted by some graph storage engines, such as Neo4j.

#### 3.5 Data quality issues related to time evolution

In this section we consider the change over time for some of the key identifiers. Are there any obvious trends in whether identifiers are becoming more pervasive or prevalent in newer citation records? Here are some general observations: DOIs are almost ubiquitous for new articles (Figure 9), ORCIDs have been on the rise to just under 17% of authors per year (Figure 10), but GRID and ISNI usage peaked in 2017, having first appeared in 2015 (Figure 11). That leaves us with the tedious task of disambiguating the affiliation of the authors in the records. As can be seen in Figure 12, the vast majority of recent records contain an affiliation string for all authors; this is due to a policy change in 2014 to collect affiliations for all contributors [11].

#### 4 Conclusions

PubMed is an enormously valuable resource for the biomedical sciences and healthcare, yet, those attempting to identify authors and affiliations, or otherwise use the records from that database, need to be aware of the quality issues within the dataset. This article has highlighted some of those data quality concerns.

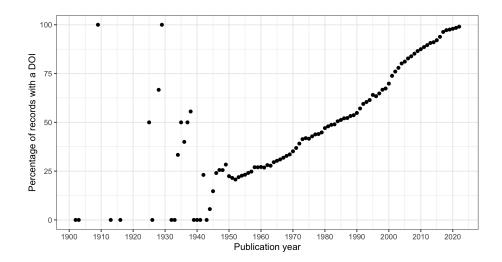


Figure 9: DOI percentage of articles per publication year.

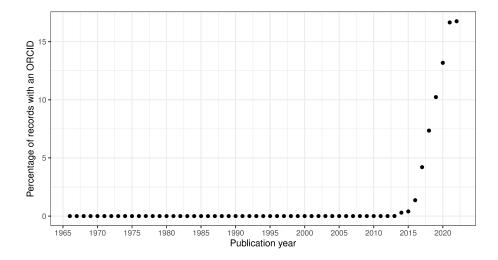


Figure 10: ORCID percentage of authors per publication year.

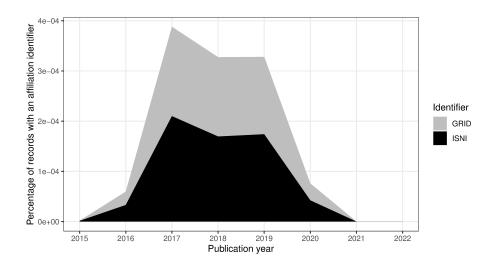


Figure 11: ISNI & GRID percentage of authors per publication year.

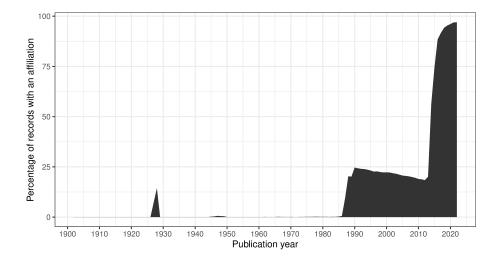


Figure 12: Percentage of authors per publication year with an affiliation string.

The data are subject to many human errors, such as typographical errors, and system related errors such as inconsistent representations of author names (leading to the synonym problem) and affiliations. There is a lack of author identifiers (contributing to the homonym problem) and a significant lack of affiliation identifiers. Being an aggregated source, the PubMed database suffers from multi-source problems such as inconsistent representations from the upstream XML providers that result in a high degree of lexicographic entropy.

In summary, our work supports the following conclusions:

- Given the incompleteness and uniqueness of identifying fields, the disambiguation of author names remains a significant problem for PubMed, particularly for records dating before 2014.
- PubMed has excellent integrity for NLM-internal identifiers (e.g., MeSH), though there is the noted exception around the J\_MEDLINE dataset. Beyond the NLM database, the majority of articles are labelled with a DOI, and the DTD provides support for identifiers for authors, institutions, both of which are far from complete. The DTD also caters for grant information, and auxiliary data through the DataBank elements, though these were beyond the scope of our work.
- Overall, there is an improvement in the use of identifiers; in particular, records created since 2015 exhibit an increase in external identifiers. However, the data quality for institutional identifiers is poor and their use has been diminishing over time.

Unless the data quality issues are addressed retroactively, they will weaken (if not entirely distort) any subsequent data analysis. Perhaps, an intervention in current publishing systems, to prevent the data sources of PubMed from manifesting the data quality issues mentioned herein, is the best one can hope for the future. Much like the application of machine learning has been applied within the NLM for indexing (e.g., with the MTI tooling [31]), the NLM could enhance their process with systems that possess a learning architecture to improve and accelerate the curation of the PubMed records. It is also possible that another information provider will provide an open data repository containing cleansed PubMed data, although a proprietary offering is more likely.

Another possibility for better use of the PubMed treasure trove is the creation of an open source library for cleansing the data, or at least properly identify the data quality issues, and optimize the amount of information that one can obtain from processing the PubMed records. Once this is accomplished with one programming language the open source community can augment the library and expand its adoption in other programming languages, for example by porting the library.

Lastly, the community would benefit from the availability of open source libraries that can accurately perform author name disambiguation, or a substantial set of "gold data" that can be used for training and validation; that dataset, however, should be orders of magnitude larger than the ones that are currently available (e.g., the 'amorgani/AND' dataset [32] [33]).

### References

- [1] DTMBIO '10: Proceedings of the ACM fourth international workshop on Data and text mining in biomedical informatics. URL: https://dl.acm.org/doi/proceedings/10.1145/1871871.
- [2] ACM Digital Library. URL: https://dl.acm.org.
- [3] Lee J et al. "BioBERT: a pre-trained biomedical language representation model for biomedical text mining." In: *Bioinformatics* 36.4 (2020), pp. 1234–1240. DOI: 10.1093/bioinformatics/btz682.
- [4] LitCovid, NLM. URL: https://www.ncbi.nlm.nih.gov/research/coronavirus/.
- [5] About PubMed, NLM. URL: https://pubmed.ncbi.nlm.nih.gov/about/.
- [6] Rahm and Do. "Data Cleaning: Problems and Current Approaches". In: Bulletin of the IEEE Computer Society Technical Committee on Data Engineering (2000).
- [7] MeSH: Medical Subject Headings. URL: https://www.nlm.nih.gov/mesh/meshhome.html.
- [8] Dataset of MEDLINE indexed journals. URL: https://ftp.ncbi.nlm.nih.gov/pubmed/J\_Medline.txt.
- [9] Sanyal DK, Bhowmick PK, and Das PP. "A review of author name disambiguation techniques for the PubMed bibliographic database". In: *Journal of Information Science* 47.2 (2021), pp. 227–254. DOI: 10.1177/0165551519888605.
- [10] PubMed 2019 DTD. URL: http://dtd.nlm.nih.gov/ncbi/pubmed/out/pubmed\_190101.dtd.
- [11] MEDLINE PubMed XML Element Descriptions and their Attributes. URL: https://www.nlm.nih.gov/bsd/licensee/elements\_descriptions.html.
- [12] Titipat Achakulvisut, Daniel E. Acuna, and Konrad Kording. "Pubmed Parser: A Python Parser for PubMed Open-Access XML Subset and MEDLINE XML Dataset XML Dataset". In: Journal of Open Source Software 5.46 (2020), p. 1979. DOI: 10.21105/joss.01979. URL: https://doi.org/10.21105/joss.01979.
- [13] pymed. URL: https://github.com/gijswobben/pymed.
- [14] pubmed-parser. URL: https://github.com/thecloudcircle/pubmed-parser.

- [15] DAMA International. DAMA DMBOK Data Management Body of Knowledge. 2nd Edition. New Jersey, USA.: Technics Publications, 2017.
- [16] PubMed baseline download files. URL: https://ftp.ncbi.nlm.nih.gov/pubmed/baseline/.
- [17] PubMed daily update files. URL: https://ftp.ncbi.nlm.nih.gov/pubmed/updatefiles/.
- [18] Michael Armbrust et al. "Spark SQL: Relational Data Processing in Spark". In: Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data. SIGMOD '15. Melbourne, Victoria, Australia: Association for Computing Machinery, 2015, pp. 1383–1394. ISBN: 9781450327589. DOI: 10.1145/2723372.2742797. URL: https://doi.org/10.1145/ 2723372.2742797.
- [19] The spark-xml project. URL: https://github.com/databricks/spark-xml.
- [20] XQuery 3.0 standard. URL: https://www.w3.org/TR/2014/REC-xquery-30-20140408/.
- [21] Saxon-HE library, Saxonica. URL: http://www.saxonica.com/.
- [22] The spark-xml-utils project. URL: https://github.com/elsevierlabs-os/spark-xml-utils.
- [23] Apache Zeppelin. URL: https://zeppelin.apache.org.
- [24] Digital Object Identifier (DOI). URL: https://www.doi.org.
- [25] Open Researcher and Contributor ID (ORCID). URL: https://orcid.org.
- [26] ISO 27729, International Standard Name Identifier (ISNI). URL: https://isni.org.
- [27] Global Research Identifier Database (GRID), Digital Science. URL: https://www.grid.ac.
- [28] NLM terms and conditions. URL: https://www.nlm.nih.gov/databases/download/terms\_and\_conditions.html.
- [29] MARC definition, US Library of Congress. URL: https://www.loc.gov/marc/faq.html#definition.
- [30] MEDLINE/PubMed Language Table. URL: https://www.nlm.nih.gov/bsd/language\_table.html.
- [31] NLM Medical Text Indexer (MTI). URL: https://lhncbc.nlm.nih.gov/ii/tools/MTI.html.
- [32] Vishnyakova D, Rodriguez-Esteban R, and Rinaldi F. "A new approach and gold standard toward author disambiguation in MEDLINE." In: *J Am Med Inform Assn* 26 (2019), pp. 1037–1045. DOI: 10.1093/jamia/ocz028.
- [33] Vishnyakova D et al., AND Author Name Disambiguation corpus. URL: https://github.com/amorgani/AND/.

# The State of Scholarly Metadata: 2023



In late 2022, CCC and Media Growth Strategies undertook a thorough examination of metadata management across the research lifecycle.

This in-depth review builds on an existing body of work to uncover multiple policy and system complexities and breakages, which – separately and together – create missed opportunities for the communities for whom Open Access (OA) and Open Science models are designed to serve.

CCC is sharing this analysis with the scholarly communications community to spark dialogue and to drive action. Drawn directly from our research interviews, this living infographic depicts the significant economic and social impact that a fragmented metadata supply chain has today on researchers, institutions, funders, and publishers. Researchers, in particular, shoulder a significant administrative burden that ultimately disrupts and delays the process of scientific discovery.

As the scholarly communications community continues its shift to full OA, stakeholders recognize that new strategies, inclusive policies, and a robust network of interoperable data and systems are essential for making critical infrastructure improvements, and much progress is underway. In that environment, a dedication to data stewardship across each stakeholder group, and the service providers supporting them, will lead not only to a smoother OA transition, but also to greater research integrity; data sharing; reliable, trustworthy metrics on research impact; and a responsive, equitable rewards and recognition system.

### Research stage

# **Idea Development**



# RESEARCHER

Researcher seeks collaborators; meets with colleagues and library / research office staff

### **CHALLENGES**

### **Underutilization of ORCID**

Some institutions don't require researchers to use ORCID; records can be outdated if authors don't consistently update; ORCID may not be accessible to authors in some geographies.

### **→** IMPACT

If authors can't be identified with a standard ID, they may not be able to authenticate to content, get credited appropriately for their work, secure OA funding, or complete downstream processes without unnecessary manual effort. Costly manual effort is also required of publishers, institutions, and funders to disambiguate authors retrospectively.

### Research stage

### **Proposal Submission**



### RESEARCHER

Researcher submits application for funding



Funder selects reviewers and begins application review



Funder logs funding terms in grant management system

### **CHALLENGES**

### **Inconsistent Metadata Capture**

Variability across grant application process/systems results in possible loss of metadata necessary to determine OA funding entitlements at a later stage, e.g., institutional affiliations.

## **CHALLENGES**

### **Legacy System Limitations**

Low adoption of standardized PIDs (FundRef, RAiD, Ringgold, ISNI, ROR) due to limitations of legacy systems and/or lack of awareness.

# **CHALLENGES**

## **Low-Quality Data**

Free text fields are great for gathering feedback; they're not designed to capture granular data like an organizational identifier. Researchers often confuse proposal numbers with grant IDs later in the publication process--they need structure to improve the accuracy of data capture.

### **→** IMPACT

Without disambiguated grant and funder details, grants may not be effectively utilized in later publication stages, leaving OA funding unclaimed and shifting coverage to research institutions. In an ecosystem that values a sustainable OA shift, this impacts everyone.

### **→** IMPACT

Hindered conflict of interest management among peer reviewers threatens research integrity, and low-quality data results in low accuracy of later-stage funding identification, tracking, and analysis of research output.

### **→** IMPACT

Lack of registered grant DOIs makes it difficult and costly to link funding to particular research outputs, resulting in missed OA opportunities as well as incomplete analysis to inform future funding investments.

### Research stage

### Research & Authoring



### RESEARCHER

Researcher conducts literature review



### RESEARCHER

Researcher posts pre-print / shares early outputs



### RESEARCHER

Researcher selects publication for submission

## **CHALLENGES**

### **Researcher Inequities & Research Barriers**

- Valid research coming from under-represented researchers is hard to find due to lack of metadata, including DOIs.
- Search and discovery are difficult due to inconsistency in identifying the user and enabling appropriate access to research.
- Authors from under-represented areas may not have equitable access to search and discovery services or equitable opportunities for publication.

### Global inequities hinder scientific progress.

### **CHALLENGES**

### **Poor Connections Across Research Outputs**

Lack of persistent identifiers (PIDs) and inconsistent application of PIDs across research outputs e.g., data sets, equipment, setting(s), samples, software

## **CHALLENGES**

### Risk of OA non-compliance

Metadata lost upstream makes managing funding compliance onerous.

### **→** IMPACT

Inability to easily find, verify, and reuse the data and artifacts underlying research, making it difficult to accurately interpret, cite and reproduce research findings.

### **→** IMPACT

Lack of available information about both corresponding author and all co-authors leads to manual input to identify funder and institutional mandates at best and missed funding requirements at worst.

# Research stage Publication



RESEARCHER

Researcher submits article



NSTITUTION

Institution funds OA publication



PUBLISHE

Publisher indexes metadata to enable search & discovery

### CHALLENGES

### **Missed Funding Opportunities**

- Under-utilization of metadata validation services
- If the researcher has submitted before, outdated information from their existing profile can be pulled into the submission
- Inconsistency between journal policies and metadata procedures
- Lack of funding information captured at submission and validated at acceptance
- Demand for increased interoperability between IDs

### **CHALLENGES**

# Missed Funding Opportunities & Costly Billing Complications

If funder/institution information manually input by the author does not use a standardized name or PID (e.g., abbreviations, nicknames), this can interfere with matching to the correct OA funding source.

### **CHALLENGES**

### **Unnecessary Manual Intervention**

Publishers are sometimes manually entering PIDs prior to registering DOIs for a more complete publication record.

### **→** IMPACT

Without granular, accurate organizational affiliation identifiers for a manuscript, coupled with incomplete funding details, authors may miss the opportunity to get OA funding or miss the chance to opt into OA due to affordability concerns. OA initiatives driven by institutions and funders may lack uptake as a result. Publishers are also unable to automate processes that reduce the cost of business model transformation. Manual effort is required to retrospectively cover the publication with proper funding sources, driving up the cost of publishing. No one benefits in this scenario.

### **→** IMPACT

Publishers and institutions take on the time and expense of manually finding the papers that should have matched to an agreement and collaborating on a resolution. Funding decisions cannot be based on abbreviations or free-form data.

### **→ IMPACT**

This is a laborious practice with high economic and opportunity costs that could be reduced with earlier, automated PID assertion and/or validation.

### Research stage

### Reuse & Measurement



RESEARCHER

Researcher evaluates research impact



### MOTITITION

Institution assesses historical subscription & publication data to inform institutional deals



FUNDER

Funder evaluates research impact



**PUBLISHER** 

Publisher assesses historical subscription and publication to inform institutional deals

### **CHALLENGES**

### **Problematic Research Impact Measurement**

Difficult to track research/researcher impact due to lack of adoption of metadata standards.

### **CHALLENGES**

### **Problematic Deal Modeling**

- Lack of consistent affiliation and funding data makes modeling future agreements hard for institutions.
- Data is not standardized across publisher platforms, creating unnecessary manual work to gather and normalize data for analysis.

# CHALLENGES

### **Problematic Research Impact Measurement**

Difficult to track funder impact due to lack of adoption of metadata standards.

### **CHALLENGES**

### **Problematic Deal Modeling**

Lack of consistent affiliation and funding data makes modelling future agreements difficult for publishers and institutions.

### **→** IMPACT

Researcher rewards and recognition decisions, or future opportunities for funding, may be based on incomplete or inaccurate data, affecting reputation and career advancement.

### **→ IMPACT**

The transition to modern models of OA publication is delayed, prolonging a mixed-model landscape and the availability of open outputs to advance science.

### **★ IMPACT**

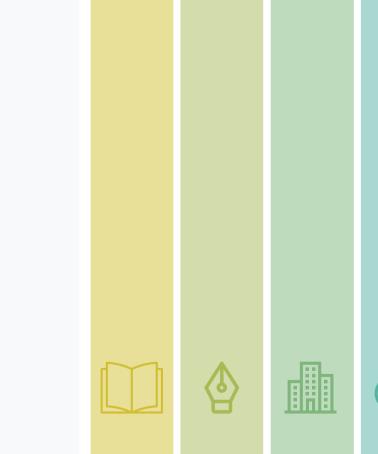
Incomplete analysis to support future funding investments and to report activities to the public.

### **★ IMPACT**

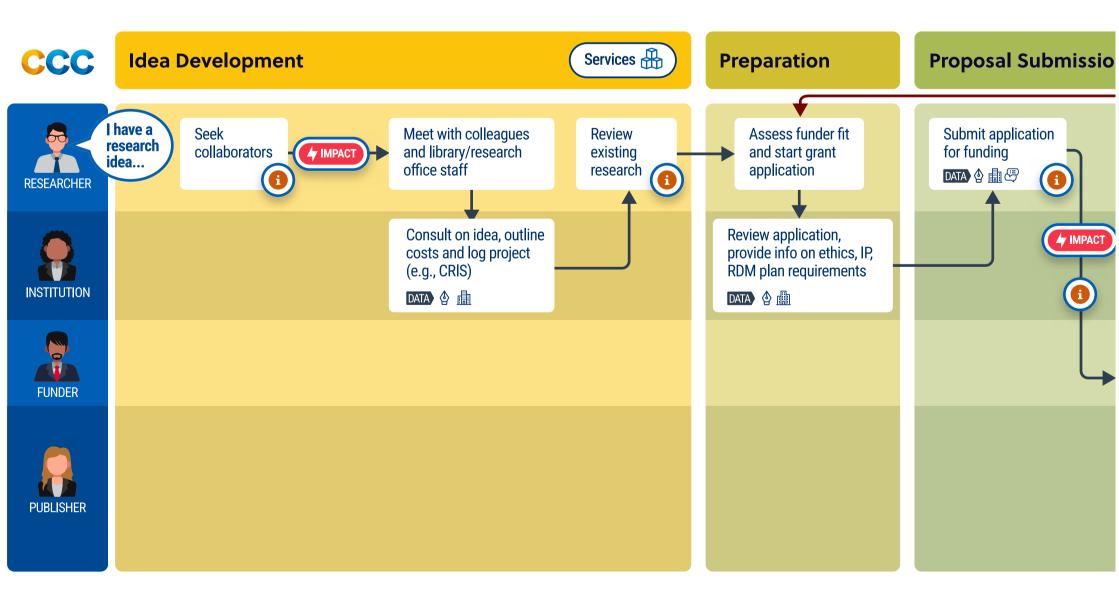
The transition to OA is delayed, putting some publishers at risk of losing authors to funding mandates and losing revenue that is necessary to sustain operations.

To view the interactive map, visit

# stateofmetadata.com







### **About CCC**

A pioneer in voluntary collective licensing, CCC (Copyright Clearance Center) helps organizations integrate, access, and share information through licensing, content, software, and professional services. With expertise in copyright, information management, artificial intelligence, and machine learning, CCC and its subsidiary RightsDirect collaborate with stakeholders to design and deliver innovative information solutions that power decision-making by harnessing information from a wide variety of data sources and content assets.



**Submit date: 4/24/2023** 

I am responding to this RFI: On behalf of an organization

Name: Abigail Goben

Name of Organization: Research Data Access and Preservation Association

Type of Organization: Professional org association

- 1. How to best ensure equity in publication opportunities for NIH-supported investigators.
- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

#### **Uploaded File:**

RDAP\_NIHRFIResponse\_PublicAccess\_NOT-OD-23-091.pdf

**Description:** A response related to the NIH RFI addressing concerns about data sharing.

Email: agoben@uic.edu

The Research Data Access & Preservation (RDAP) Association thanks the NIH for the opportunity to respond to this Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research.

We appreciate the NIH's consideration of challenges facing researchers in moving towards compliance with the recommendations from the OSTP memo "Ensuring Free, Immediate, and Equitable Access to Federally Funded Research" and the opportunity to provide feedback on the NIH plans for implementation. However, while the NIH has asked primarily for feedback relating to manuscript sharing, we believe this request for information should consider all NIH-sponsored research outputs -- specifically the research datasets underlying peer-reviewed scholarly publications.

As data professionals, we frequently collaborate with researchers in developing and implementing robust data management and sharing plans, and provide education to assist researchers in understanding and adopting good practices throughout the data lifecycle. The NIH DMSP policy and related guidance documents are a significant improvement in policy and expectations for data sharing across a broad array of research disciplines. However, there remain significant challenges in aligning outcomes with the OSTP recommendations for preserving and sharing data underlying publications.

We believe there are a number of steps that NIH can take to improve researchers' abilities to comply with the emerging requirements in response to the OSTP memo while maximizing data sharing from NIH-funded research in alignment with the goals of the NIH DMSP. These opportunities will also advance the development of a national infrastructure that reduces the burden on individual institutions and researchers.

- NIH should review the federal repositories currently available and enhance their capabilities to accept big data, and sensitive/controlled data. This includes a review of NIH-supported repositories and how they align with the National Science and Technology Council "Desirable Characteristics of Data Repositories for Federally Funded Research.
- NIH should develop mechanisms for rewarding the sharing of high-quality datasets as
  well as high-impact datasets. While researchers are currently rewarded for publications
  in various venues, there is less in place for recognition of datasets which can be reused
  across projects and disciplines. The significant labor to collect, curate, and manage this
  data should be recognized.
- NIH should engage in a collaborative partnership with a stakeholder group similar to the FDP program. Data managers, data librarians, and data curators, such as many of the members of RDAP, are well positioned to understand the challenges and opportunities in expanding data sharing and preservation, and are engaged in developing efficient data management, sharing and curation workflows.
- We encourage NIH to work with the US Repositories Network (USRN). The USRN's goal
  is to improve collaboration and cohesion across all open repositories, including those
  that host peer-reviewed publications and scientific data.

- We encourage NIH to bolster data management expertise for grant managers and we offer to engage RDAP members to support collaboration on this continuing education.
- We encourage the NIH to provide expanded clarification and guidance for researchers and institutional offices of technology transfer around the challenges of managing data sharing compliance while pursuing protection of intellectual property.

The NIH Data Management and Sharing Policy (DMSP) as currently written relies upon individual researchers and their home institutions to provide the majority of the investment, administrative burden, and resources to preserve and share data. This disproportionately affects researchers at less advantaged and smaller institutions, who are not likely to have local resources, thereby increasing inequities. We suggest that NIH should put greater emphasis on the development of national infrastructure, to reduce the burden on individual institutions and researchers. Scalable innovations in infrastructure can improve data sharing capacity and access to all types of data regardless of size or format. There are good examples of this approach, such as GenBank. Our primary recommendation is to take the initial goals of the DMSP and the emerging requirements from the OSTP, which are both aligned with investment, and to begin to invest in national data sharing infrastructure. Such infrastructure would better equalize the opportunities for researchers nationwide to maximally share their data.

Further, while the NIH enables researchers to incorporate the costs of data management and sharing into research budgets, there was no concurrent increase in award budgets. Taken as an either-or option, it may follow that researchers will prioritize staff, equipment, and other items in their budgets directly attributable to research outcomes as opposed to data sharing. This choice may be exacerbated if a budget reduction is required during the application and award process. Thus at the end of a research project the resources available for data management and sharing activities are likely to be inadequate. This impact will be compounded for lessfunded areas, such as women's and minority health research. Poorly-resourced data sharing has the potential to have a negative compounding effect on access and reuse of data in these domains. Several publications have attempted to address the underlying costs of data management and sharing, including the NASEM report Life-Cycle Decisions for Biomedical Data: The Challenge of Forecasting Costs. We encourage the review and integration of their findings in evaluation of raising budget caps, especially for R01, R03, and R21 grants.

The reliance on the NIH DMSP for sharing the data underlying the publications continues to fail to address the underlying resource disparities between institutions and the increase of risk introduced by mandated data sharing. Many research institutions do not have generalist institutional repositories that are available for any sort of data sharing and requiring their reliance on the repositories in the GREI initiative creates a disparity for those researchers and institutions and their ability to maintain oversight of the data generated from their awards. Additionally, even most large research institutions, including many "R1" universities, do not have data repositories or the necessary staff managing them to facilitate curation, preservation and sharing for sensitive and controlled data. Nor are there enough disciplinary or association-based repositories who can absorb the more sensitive data. This introduces significant increased

challenges for researchers seeking to appropriately share data who are likely to not have adequate university or disciplinary resources, and will likely harm researchers at smaller and less well-resourced universities more.

The difficulty of navigating the current landscape of repositories is compounded by the challenge that while NIH hosts a number of repositories, these are limited in the types and sizes of data they will ingest. The scope of data that each NIH-supported repository will ingest and steward is also not clear. While we strongly support the current repositories, several of these NIH-supported repositories have actively limited their intake in the past year or told researchers planning Data Management and Sharing Plans that they will not serve as the final storage and sharing place for data.

Over all, the current policy and repository options places the burden of preservation and sharing to the individual or to the institution and will create a two-tier system of researchers -- those wholly reliant on public repositories such as figshare and Dryad or on the minimal amount of storage tied to manuscripts in PubMedCentral (currently at 2GB) where they can store publicly available data but who are then giving up control of their research -- and researchers at larger institutions who will have more resources but likely still inadequately prepared for the complexity and scale of managing and sharing data. Additionally, the OSTP Nelson memorandum indicated that the federal funding agencies want a more comprehensive and immediate solution. It accelerates and complicates the current workflows.

As researchers rush to meet the immediate data sharing requirements of the NIH DMSP and the OSTP memo, this is likely to drive them to very limited data sharing in order to connect their data in PubMedCentral. As PubMedCentral is not a data repository and does not provide any additional metadata or guide to the data, this will not greatly enhance data discoverability and reuse. Instead, we recommend the creation of a PubMedCentral equivalent for datasets. This would be possible using underlying Datacite resources but should be cautiously approached to prevent the intervention and blockade to access by publishers seeking maximum profits.

We support open systems that enable broad and equitable access to both datasets themselves and their accompanying metadata materials. It is incumbent on NIH to monitor data sharing expenditures and the choices of repositories to ensure that data sharing infrastructure does not become susceptible to the influence of commoditization and for-profit motives.

Understandably, the NIH policy is written to accommodate a broad range of disciplinary communities and their evolving data sharing practices and related infrastructure. This could be improved by working with these communities to solidify conceptual details, such as defining the discrete unit of data necessary to validate the research results reported in a paper. For example, is computational reproducibility a minimum standard? Because there is a lack of clarity, it can be difficult for us to adequately advise researchers in our institutions. Lack of clarity can also lead researchers to employ minimalist solutions, rather than maximize data sharing. While societies and publishers have some influence on these issues, NIH has the capacity to bring faster and more uniform change to the practices around data and publications.

It is also imperative to address concerns around public data sharing requirements and how they interact with intellectual property and patent concerns. The current policy provides exceptions only for SBIR and STTR grants. The only guidance from the DMSP implementation team, provided in a Q&A webinar by Director Paine in September 2022, has been that general intellectual property protection isn't in alignment with the data sharing requirements to maximally share data. However, neither the policy nor the supporting materials provided at sharing.nih.gov guide researchers or their institutions in how to address this conflict. More robust and uniform guidance from NIH is needed to assist researchers and other stakeholders in addressing reproducibility and data sharing goals, while also addressing valid intellectual property interests.

The gap between the policy and researchers implementation presents many challenges, particularly data with controlled and sensitive topics. Lack of guidance opens researchers and institutions up to risk of data breaches and data loss as they strive to meet the spirit and requirements of the NIH Policy and Nelson Memo in maximizing data sharing. The general guidance provided to date does not adequately address the very real and significant concerns of researchers and their institutions. RDAP looks forward to the opportunity to talk further with NIH representatives in improving this guidance for the benefit of all involved.

#### **About RDAP**

The RDAP community is an engaged community of information professionals committed to creating, maintaining, advancing, and teaching best practices for research data, access, and preservation. Many RDAP community members lead research data management services and activities at numerous US academic institutions and are responsible for helping their researchers meet both the spirit and requirements of the NIH Public Access Plan and NIH Data Management and Sharing Policy. Learn more about RDAP at <a href="https://rdapassociation.org/">https://rdapassociation.org/</a>.

**Submit date: 4/24/2023** 

I am responding to this RFI: On behalf of an organization

Name: Janine Chiappa McKenna

Name of Organization: American Gastroenterological Association

Type of Organization: Professional org association

Role: Institutional official

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

The NIH-proposed plan mandates zero embargo, which is intended to be more equitable for readers by making research more rapidly and freely available, and researcher-authors by allowing them to self-archive their manuscript in a public repository (i.e., "green" open access) without having to pay article processing charges (APCs, i.e., "gold" open access). However, the increase in free content will lead to difficult decisions for publishers that could result in greater inequities for researchers, as outlined below.

- 1. PUBLISHING OPTIONS FOR RESEARCHER-AUTHORS MAY BECOME LIMITED TO FEE-BASED OPEN-ACCESS JOURNALS. Currently, the publishing ecosystem includes journals that are fully open access, exclusively subscription-based, or "hybrid," meaning they offer both open access content as well as content behind a subscription paywall. This gives researcher-authors diversity in choice depending on their preferred publication method. Further, the existence of hybrid journals allows "green" open access as a cost-effective publishing option. Biomedical and clinical journals, such as those published by AGA, publish high volumes of manuscripts resulting from NIH-funded research. With the increase in free content inevitably leading to a decline in individual and library subscription revenue, hybrid journals will likely convert to a fully online, open access model meaning that researcher-authors will be increasingly limited to journals requiring article processing charges (APCs) and "green" open access will no longer be available as a cost-effective publishing option.
- 2. PUBLISHING MAY BECOME RESTRICTED TO ONLY RESEARCHER-AUTHORS WITH SIGNIFICANT GRANT FUNDING. The work of publishers and the services they provide will not decrease because a journal converts to fully open-access, meaning that expenses will not change and existing revenue coming from subscriptions will need to be covered by raising APCs for researcher-authors. Further, the NIH Public Access Policy will apply to all NIH-funded authors regardless of their total funding or how small a role they play in a research project or manuscript. Therefore, the policy may create inequities in that only well-funded investigators or those at institutions with additional resources will be able to afford these fees, or authors must reallocate grant funds from research expenses to publication costs. Early-career researchers in particular may be penalized.
- 3. CLARITY FOR RESEARCHER-AUTHORS ON WHEN THE PUBLIC ACCESS MANDATE APPLIES WOULD LIKELY REDUCE THEIR BURDEN. We suggest that the NIH indicate a minimum threshold of funding and/or level of participation by researcher-authors at which the immediate public access mandate would apply to a particular manuscript. This is particularly important as science increasingly moves to a "team science" model with large, collaborative research teams developing manuscripts that can have tens or even hundreds of authors who are not contributing equally. Minimal contributions to studies or

use of funded shared resources by NIH-funded researchers should not qualify a paper for the proposed mandate.

4. PROVIDE PUBLISHERS THE ABILITY TO MAKE THEIR OWN DECISIONS REGARDING RIGHTS RETENTION. As NIH seeks to make peer-reviewed content accessible without an embargo, AGA requests that the NIH refrain from requiring reuse rights under licenses that restrict our ability to establish copyright. Instead, AGA should retain the rights associated with the final version of record, both as a resource for the association as well as to ensure an author's research isn't misappropriated and turned into derivative works that could lack integrity or worse, cause patient harm. Under copyright provisions, we guard against misuse of author content by requiring third parties to follow our policies regarding appropriate use of published content.

#### 2. Steps for improving equity in access and accessibility of publications.

We have outlined several suggestions below regarding guidelines and procedures that may help improve equity in access and accessibility of publications resulting from NIH-funded research.

- 1. EDUCATE AUTHORS ON APPROPRIATE REPORTING OF FEDERAL FUNDING IN MANUSCRIPTS AND WHEN THE PUBLIC ACCESS MANDATE APPLIES. Overreporting is commonplace and even incentivized as researcher-authors attempt to demonstrate significant progress on their funded research through the volume of publications. However, we are aware that grantees, or work done on their behalf from Other institutions, have inappropriately deposited articles in PubMed Central because NIH funding was acknowledged in a manuscript that was loosely related to but not a direct result of the funded research. Therefore, we urge NIH to provide clear conditions under which authors should acknowledge NIH funding in their manuscripts and adhere to the public access mandate. Consistent communication and education to the research community regarding these conditions will also be essential.
- 2. SOLICIT FEEDBACK FROM THE RESEARCHER-AUTHOR COMMUNITY ON NECESSARY PUBLICATION COSTS AND PROVIDE CLEAR GUIDANCE ON BUDGETING PUBLICATION FEES. Although NIH states publication fees are an allowable expense, AGA members have shared experiences of publication budgets in their grants being reduced from their original proposal. As a result, there is not enough funding in their grants to cover publication fees for the multiple papers that will typically result from a single grant. Based on public comments thus far from NIH, it seems unlikely that there will be increases in agency funding to sufficiently cover researcher-authors' publication costs. We urge NIH to continue open dialogue with the researcher community to understand their challenges toward developing potential solutions.
- 3. REDUCE DUPLICATION OF WORK BY PUBLISHERS AND THE NATIONAL LIBRARY OF MEDICINE THROUGH PARTNERSHIPS THAT STREAMLINE THE PROCESS OF DEPOSITING MANUSCRIPTS. Currently, the National Library of Medicine (NLM) duplicates the work of publishers by preparing text files for online publication in PubMed Central. We urge NLM to consider seeking licensing agreements with publishers in which publishers would provide high-quality machine-readable, highly tagged extensible markup language (XML) in exchange for a fee rather than both parties doing similar work. This licensing arrangement would also ensure compliance of deposits into PubMed Central.

Alternatively, PubMed Central could become a centralized bibliographic database that links to journal websites rather than separately hosting its own full-text journal articles. This would be an innovative

approach that would also incentivize publishers to develop more ancillary content and enhance user features.

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

Budgets submitted by grant applicants could be an informative tool for monitoring publication costs, if there was a system by which this data could be pooled across NIH institutes and centers and categorized by different types of research. We also welcome ongoing dialogue with NIH regarding equity in publication opportunities as this is a priority area for AGA's publications under our AGA Equity Project, an organization-wide initiative prioritizing diversity, equity and inclusion in our policies, processes, and programs. For example, AGA participates in Research4Life and offers fee waivers for researcher-authors who require financial assistance, such as early-career researchers or researchers from under-resourced regions or institutions who may lack sufficient funds to cover our journals' publication fees.

#### 4. Early input on considerations to increase findability and transparency of research.

We urge NIH to engage with publishers and the PID community to use or adapt what has already been created, rather than creating its own system. For example, PubMed currently replaces publisher DOIs in the references of papers in its repository; by removing publisher DOIs from reference links or choosing to include links to the PubMed Central version instead of the version of record (VOR), the NLM is depriving the user of access to associated editorials, letters to the editor, infographics, and Other ancillary materials that may provide additional value and context to the reader.

For researchers, we suggest that NIH employ DOIs for grants and require them for datasets published. By adopting PIDs already in use in scholarly publishing, journals can include persistent links to critical pieces of research for users to access.

Email: jmckenna@gastro.org

**Submit date: 4/24/2023** 

I am responding to this RFI: On behalf of an organization

Name: Gwen Twillman

Name of Organization: American Society for Nutrition

Type of Organization: Professional org association

Role: Institutional official

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

The American Society for Nutrition (ASN) broadly supports the efforts of NIH to develop educational materials and standards to improve article accessibility and PubMed Central procedures for processing.

Increasing diversity, equity and inclusion in nutrition and related sciences is a strategic priority for ASN. ASN allocates publication waivers for underrepresented and early career scientists. NIH could further its goals by dedicating publication resources for underrepresented scientists and the scientific societies that support them.

ASN also encourages NIH to follow a model like that of the Bill and Melinda Gates Foundation for coverage of publishing fees. The Gates Foundation uses a central budget to pay for article processing charges and publisher fees. Grantees do not have to use funds out of their research budget or seek reimbursement from the Foundation. Instead, invoices are directed to the Foundation for payment from the central budget directly to the publisher or service provider. The Gates Foundation supports publication of research funded, in part or whole, by the Foundation and only requires a valid Gates grant number. The central budget covers open access publishing fees and additional publishing costs such as page charges. The grantee is responsible for managing any publisher agreements and covering any additional costs beyond these fees. It is ASN's understanding that fees are paid at any point in time from this central budget, even after the end of a grant funding period.

#### 2. Steps for improving equity in access and accessibility of publications.

ASN is committed to the translation of science to a variety of audiences - researchers, clinicians, policymakers, public health professionals and lay audiences. Sample tactics include blog posts, statements of significance, press releases and outreach to traditional and social media, as well as journal features such as Great Debates in Nutrition, Nutrition for the Clinician, and AJCN in Press podcast. Training researchers to properly communicate their science also is an ASN priority.

Financial support for activities to translate science and improve equity in access is lacking. NIH funding to help scientific societies continue and increase these efforts will help ensure their sustainability.

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

Scientific research societies that publish scholarly journals, such as ASN, invest the income from their journals back into the scientific research community by supporting professional development and educational opportunities, including training for the next generation of scientists, meetings, and awards. ASN encourages NIH to consider additional support for the scientific research community in the form of professional development activities, particularly those helping early-career and underrepresented

researchers prepare and support scholarly publications, such as training young professionals to serve as peer reviewers.

Income from publications also funds editorial expenses that ensure a rigorous and fair peer review process, foster scientific integrity and trust in science, and furthers science advancement.

Activities that monitor evolving publication costs must also consider and evaluate any negative consequences on organizations that prioritize rigor and reproducibility of science over publication volumes. For example, the American Society for Nutrition funds the following activities to ensure a peer review robust process: compliance with industry and ethical standards in the conduct and reporting of research; compensation of editors to oversee peer review, a Statistical Review Board to confirm data analysis, and senior-level staff to monitor trends; plagiarism screening; management of ethical investigations; and Other best practices.

4. Early input on considerations to increase findability and transparency of research.

Email: gtwillman@nutrition.org

**Submit date: 4/24/2023** 

I am responding to this RFI: On behalf of an organization

Name: Edward Pentz

Name of Organization: Crossref

Type of Organization: Other

Type of Organization-Other: Non-profit, open scholarly infrastructure provider

Role: Institutional official

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Researchers should be free to publish their manuscripts in the most appropriate journal that meets the NIH Public Access Plan requirements. By registering its grants with Crossref and getting Crossref Grant DOIs, NIH can ensure that published outputs from NIH-supported researchers are easily connected to the related grant without any additional burden on the researchers. With over 18,000 members from 150 countries and over 100,000 journals, Crossref metadata and DOIs will support connecting the publications of NIH-supported researchers to the global research discovery ecosystem wherever they publish. Our growing membership includes many new formats and models for publishing, with incentives in place such as our new GEM Program (Global Equitable Membership) which enables zero-fee participation in the system by members in the least economically-advantaged parts of the world. Crossref also encourages critical metadata that are used for downstream analysis, such as references, data citation, and increasingly important for assessment - abstracts.

#### 2. Steps for improving equity in access and accessibility of publications.

Open persistent identifiers and metadata are essential to providing equitable access to publications. Crossref encourages NIH to register Crossref grant DOIs and metadata, including ORCID IDs and ROR IDs. Our open and robust API is open to everyone, used by tens of thousands of systems across the research ecosystem, and is heavily relied upon for text-mining and Other machine uses.

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

Open persistent identifiers and metadata are essential to monitoring trends with publication fees and where research outputs from NIH-supported researchers are made available. To enable this, Crossref encourages NIH to register Crossref grant DOIs and metadata, including ORCID IDs and ROR IDs.

NIH could encourage its grantees to publish in outlets that provide the richest possible metadata and therefore increased evidence and accessibility for the community.

### 4. Early input on considerations to increase findability and transparency of research.

In looking at identifiers and metadata and how to improve their use, we encourage NIH to focus on a number of critical questions: How open are they? How are they funded and how sustainable are they? How are they integrated with the global scholarly research ecosystem? How broadly are they used? What services are available to register, resolve and disseminate the persistent identifiers and metadata? Are there complementary services available that support Other goals such as research integrity? How

and by whom are they governed? How global/wide-reaching are they? The answers to all of these questions will also answer how truly persistent and trustworthy the operation and services are.

Crossref would be happy to collaborate with the NIH on connecting NIH grants with the wider open scholarly infrastructure that Crossref provides. As the leading Registration Agency providing DOI services, we represent by far the largest community of stakeholders involved in documenting the progress of science, so updates and future enhancements can be developed and—crucially—adopted at scale.

The Grant DOI program is unique to Crossref and has been ramping up for the last couple of years. We currently have over 76,000 registered grants, including 8,700 from the US Department of Energy's Office of Scientific and Technical Information (DOE-OSTI), with Other US federal agencies actively exploring membership and grant registration.. Crossref is ready to fully support NIH registering its grants with us so they too can connect with the global network of research metadata.

We look forward to working with the NIH alongside our work with Other agencies on meeting the shared goal of "a robust, connected ecosystem where institutions, researchers, research outputs, and funding sources are linked consistent with FAIR principles". In Other words: the Research Nexus.

Ensuring free, immediate, and equitable access to metadata that captures the scholarly record is an essential part of meeting the goals of the NIH Public Access policy and the OSTP memo and supporting Open Science globally.

#### **Uploaded File:**

Crossref-NIH-Public-Access-RFI-1.pdf

**Description:** Letter from Crossref with detailed feedback on the Public Access Plan

Email: epentz@crossref.org

#### Submitted through:

https://osp.od.nih.gov/nih-plan-to-enhance-public-access-to-the-results-of-nih-supported-research/

April 24, 2023

Office of The Director National Institutes of Health 9000 Rockville Pike Bethesda, Maryland 20892

RE: Crossref Comments in Response to NOT-OD-23-091, "Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research"

To Whom It May Concern:

I'm writing on behalf of Crossref as its Executive Director, in response to the Request for Information (RFI) on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research issued on February 21, 2023.

Crossref is a not-for-profit organization that sits at the heart of the global exchange of research information with a mission to make it possible to find, cite, link, assess, and reuse research objects. We do this by developing and maintaining open scholarly infrastructure, following the Principles of Open Scholarly Infrastructure (POSI) and the FAIR principles (Findability, Accessibility, Interoperability, and Reuse), that support open research. The service we are best known for is enabling the registration and dissemination of open metadata and persistent identifiers (DOIs - Digital Object Identifiers) for many objects and resources related to scholarly research: journal articles, books, book chapters, preprints, datasets, standards, grants, and many other artifacts. We have over 18,000 members (including universities, libraries, government agencies, government and private funders, museums, scientific societies, and publishers) from 150 countries worldwide, who have so far created metadata for over 143 million scholarly research objects, and these Crossref DOIs are resolved (clicked and followed) over 1 billion times per month. We provide additional services that enable the community to make connections between objects or to assess their trustworthiness, and our open metadata and API enable anyone interested in research to incorporate it into their own systems. We also maintain dedicated feeds to key partners such as ORCID, with over 3 million authors having now granted us permission to programmatically add works information to their ORCID records. Crossref contributed to the creation of ORCID, as we have with ROR.

Crossref welcomes NIH's intent to incorporate guidance on the uses of persistent identifiers and metadata in its Public Access Plan. Our specific comments are focused on issue number 4:

**4. Early input on considerations to increase findability and transparency of research.**Section IV of the NIH Public Access Plan is a first step in developing the NIH's updated plan for persistent identifiers (PIDs) and metadata, which will be submitted to OSTP by December 31, 2024.
NIH seeks suggestions on any specific issues that should be considered in efforts to improve use of PIDs and metadata, including information about experiences institutions and researchers have had with adoption of different identifiers.

In looking at identifiers and metadata and how to improve their use, we encourage NIH to focus on a number of critical questions: How open are they? How are they funded and how sustainable are they? How are they integrated with the global scholarly research ecosystem? How broadly are they used? What services are available to register, resolve and disseminate the persistent identifiers and metadata? Are there complementary services available that support other goals such as research integrity? How and by whom are they governed? How global/wide-reaching are they? The answers to all of these questions will also answer how truly persistent and trustworthy the operation and services are.

Crossref metadata is made openly available without any reuse restrictions via a public REST API and is integrated into thousands of scholarly information systems and services, including Pubmed and Pubmed Central. Crossref's open metadata includes basic bibliographic metadata, DOIs, abstracts, references, funding and licensing information, corrections and retractions and other open identifiers such as ORCID IDs for researchers, DataCite DOIs for research data and ROR IDs for organizations. This enables a connected, discoverable scholarly record - what we call the <u>research nexus</u> - and it's important that NIH grants are connected to this open scholarly infrastructure and the open research ecosystem that it enables across the world.

As the RFI correctly notes of the current situation with NIH grants and their identifiers: "they are not registered or indexed to ensure uniqueness beyond NIH and they are not retrievable using a standardized communications protocol that would allow for interoperability".

We are pleased to note that "NIH is exploring use of the digital object identifier (DOI) system that would overlay existing NIH grant identifiers to resolve these issues". Crossref would be happy to collaborate with the NIH on connecting NIH grants with the wider open scholarly infrastructure that Crossref provides. As the leading Registration Agency providing DOI services, we represent by far the largest community of stakeholders involved in documenting the progress of science, so updates and future enhancements can be developed and—crucially—adopted at scale.

The Grant DOI program is unique to Crossref and has been ramping up for the last couple of years. We currently have over 76,000 registered grants, including 8,700 from the US Department of Energy's Office of Scientific and Technical Information (DOE-OSTI), with other US federal agencies actively exploring membership and grant registration.. Crossref is ready to fully support NIH registering its grants with us so they too can connect with the global network of research metadata.

We are encouraged by this statement: "NIH will coordinate this exploration with efforts of other Federal agencies and relevant external/internal impacted communities to assess how to best develop a robust, connected ecosystem where institutions, researchers, research outputs, and funding sources are linked consistent with FAIR principles".

In response to the OSTP memo in November 2022, Crossref <u>outlined in a detailed</u> how funding agencies can meet OSTP (and Open Science) guidance using existing open infrastructure, which includes Crossref, and also ORCID, ROR, and DataCite.

We look forward to working with the NIH alongside our work with other agencies on meeting the shared goal of "a robust, connected ecosystem where institutions, researchers, research outputs, and funding sources are linked consistent with FAIR principles". In other words: the Research Nexus.

Ensuring free, immediate, and equitable access to metadata that captures the scholarly record is an essential part of meeting the goals of the NIH Public Access policy and the OSTP memo and supporting Open Science globally.

Your sincerely

Ed Pentz

**Executive Director** 

Crossref

**Submit date: 4/24/2023** 

Name: Simon Bacon

Name of Organization: Behavioral Medicine Research Council

Type of Organization: Nonprofit research organization

**Role:** Scientific researcher

1. How to best ensure equity in publication opportunities for NIH-supported investigators.

See attached paper

2. Steps for improving equity in access and accessibility of publications.

See attached paper

3. Methods for monitoring evolving costs and impacts on affected communities.

See attached paper

4. Early input on considerations to increase findability and transparency of research.

See attached paper

**Uploaded File:** 

Segerstrom-et-al.-2023.pdf

Description: This is our recent BMRC position statement on open science which covers a number of

elements from the aspects detailed above

Email: simon.bacon@concordia.ca



# Open Science in Health Psychology and Behavioral Medicine: A Statement From the Behavioral Medicine Research Council

Suzanne C. Segerstrom, PhD, MPH<sup>1</sup>, Michael A. Diefenbach, PhD<sup>2</sup>, Kyra Hamilton, PhD<sup>3</sup>, Daryl B. O'Connor, PhD<sup>5</sup>, A. Janet Tomiyama, PhD<sup>6</sup>, with the Behavioral Medicine Research Council

<sup>1</sup>Department of Psychology, University of Kentucky, 125 Kastle Hall, Lexington, KY 40506-0044, USA

Institute of Health System Science, Feinstein Institutes of Medical Sciences, Northwell Health, Manhasset, NY, USA

<sup>3</sup>Griffith University, School of Applied Psychology, Mt. Gravatt, QLD, Australia

<sup>4</sup>Health Sciences Research Institute, University of California – Merced, Merced, CA, USA

<sup>5</sup>School of Psychology, University of Leeds, Leeds, UK

<sup>6</sup>Department of Psychology, University of California – Merced, Merced, CA, USA

Suzanne C. Segerstrom segerstrom@uky.edu

#### **Abstract**

Open Science practices include some combination of registering and publishing study protocols (including hypotheses, primary and secondary outcome variables, and analysis plans) and making available preprints of manuscripts, study materials, de-identified data sets, and analytic codes. This statement from the Behavioral Medicine Research Council (BMRC) provides an overview of these methods, including preregistration; registered reports; preprints; and open research. We focus on rationales for engaging in Open Science and how to address shortcomings and possible objections. Additional resources for researchers are provided. Research on Open Science largely supports positive consequences for the reproducibility and reliability of empirical science. There is no solution that will encompass all Open Science needs in health psychology and behavioral medicine's diverse research products and outlets, but the BMRC supports increased use of Open Science practices where possible.

Keywords Reproducibility · Methodology · Privacy · Publication bias

#### Summary of Recommendations of the Behavioral Medicine Research Council

#### Preregistration

The BMRC strongly recommends the practice of preregistration when engaging in hypothesis-driven research, with transparent reporting of deviations from preregistered plans. The BMRC further encourages the inclusion of sample diversity considerations in preregistration.

#### Registered Reports

The BMRC recognizes the value of journals in the area of health psychology and behavioral medicine to introduce Registered Reports as a new article format.

#### **Preprints and Postprints**

The BMRC views peer-reviewed, accepted science as the best form of evidence and recommends a close evaluation of the role of preprints for health psychology and behavioral medicine research and to compare this role with the use of preprints among physicists and economists.

#### Open Research

The BMRC encourages open research practices at a minimum as required by funding entities and publications. In practice, research materials should be as open as possible and as closed as necessary, respecting privacy, laws, and cultural knowledge.

#### Civility, Collegiality, and Collaboration

The BMRC urges researchers to be tolerant and to work together in a collaborative, collegial, and civil manner, acknowledging scientific and methodological differences and similarities.

#### Equity

The BMRC recognizes the advantages and disadvantages of Open Science in achieving equity in health psychology and behavioral medicine. A more equitable research environment is needed to advance equitable open science. Open access publication cost and institutional recognition of open science practices may inadvertently disadvantage underrepresented scientists.

#### **Overview**

The present article resulted from a dialogue among representatives of the Behavioral Medicine Research Council (BMRC; representing four large international organizations in behavioral medicine and health psychology), focusing on the need to communicate our science openly and equitably while maintaining rigorous research standards. The need for this dialogue arose from multiple developments that happened over the past decade: First, legislative actions require data generated through federal funding to be made available if requested by other researchers. Second, the scientific field was confronted with high-profile incidents in which studies could not be replicated, including cases in which the original data had been fabricated or falsified [1–3]. Third, questions of equity in data quality and data access have become increasingly prominent. Fourth, the introduction of new and innovative publishing recommendations and formats (e.g., preregistration and registered reports) has prompted the need for greater transparency. The aim of the present BMRC statement on Open Science is threefold: (a) to provide a snapshot of Open Science practices in three of the most prominent journals in our field; (b) to critically evaluate the most common Open Science practices for our field; and (c) to provide recommendations for the adoption of such practices, including preregistration, registered reports, preprints and postprints, and open research.

#### Relevance

As members of the research community, we accept the need to publish the results of our research efforts, and we are often reminded that if it is not published, "it has not happened." Yet, the traditional publication system has been criticized for not providing equitable access to publicly funded research results [4]. Instead, journals tend to favor positive findings over null or contradictory results (see the well-known "file-drawer problem") [5]. Additionally, non-registered research is open to post-hoc analytic reports by researchers and may contribute to the reproducibility problem through so-called "questionable research practices" (see below). For example, one study found that 57% of studies published prior to 2000 (when registration for large clinical trials was introduced) reported beneficial intervention effects on the primary outcome compared to only 8% of trials published after 2000 [6].

Since the publication of the Open Science Collaboration's 2015 paper [7] estimating the reproducibility of psychological science, there have been many important developments to address these issues. The research community has suggested several practices, together known as "Open Science." Open Science includes some combination of registering and publishing study protocols (including hypotheses, primary and secondary outcome variables, and analysis plans) and making available preprints of manuscripts, study materials, de-identified data sets, and analytic codes. Open Science is important for health psychology and behavioral medicine. Research in this field has the potential to profoundly impact individual, community, and population health and well-being, as well as healthcare practices and policies. The potential societal impact of our work underscores the importance of ensuring experimental rigor, transparency, reproducibility, and equitable access to advance our science.

Uptake of Open Science practices has been steady and there is clear evidence of a steep upward trajectory [8].

Progress has accelerated since leading funders signed on to improving reproducibility [9] and journals and publishers started to embrace the Transparency and Openness Promotion (TOP) guidelines (see Box 1), preregistration, and new article formats such as registered reports. For example, in 2012, registered reports were first proposed by the journals Cortex and Perspectives on Psychological Science and then launched in these journals (along with in Social Psychology) in 2013 [10]. Over 300 journals now offer the registered reports format across a large number of disciplines including psychology and medicine. Despite these numerous developments and advances, there remains much room for improvement.

### Frequency of Open Science Practices in Annals of Behavioral Medicine, Health Psychology, and Psychosomatic Medicine, 2018–2020

As a starting point, we examined Open Science practices in the primary journals of the BMRC's constituent organizations and how patterns and trends in transparency and openness have changed (data and code available at https://osf.io/ wytz3/). In an analysis of Open Science practices in Annals of Behavioral Medicine, Health Psychology, and Psychosomatic Medicine, coders indicated for each empirical study or review published in 2018, 2019, and 2020 whether it was preregistered (the study protocol was predefined in its entirety or in part); was a Registered Report (acceptance in principle was based on the review of the introduction and methods only, before data collection and/or analysis); made a statement on protocol sharing, data sharing, or material sharing; or whether it was gold open access (for further definitions, see the Open Research Coding Checklist in the Supplemental Materials) [11]. We sampled for 3 years to ensure a sufficient sampling time frame to provide a good overview of the frequency of Open Science practices. Open Science practices overall were low (Table 1), except for the relatively high number of articles published as gold open access in Annals of Behavioral Medicine and Health Psychology (48.3% and 51.1%, respectively). This result is consistent with an analysis of reporting practices in 2018 in these three journals plus the American Journal of Preventive Medicine, in which there was low occurrence of elements such as explicit description of analyses as primary or secondary (16% of 162 sampled papers) and if and when studies were registered (13.6%) [12].

No clear pattern emerged from 2018 to 2020 (Table 1). If anything, there was evidence of reductions in some practices over time. It is difficult to reconcile these observations as journals and funders have become more stringent in their reporting requirements and need for registration. However, study registration did increase from 2008 to 2018 [12]. Annals of Behavioral Medicine, Health Psychology, and Translational Behavioral Medicine are signees to the TOP Guidelines [13-15] (see Box 1), which establish guidelines for data citation; data, materials, and code transparency; design and analysis; preregistration; and replication. Psychosomatic Medicine will become a signatory in 2023 [16]. Annals of Behavioral Medicine and Health Psychology's new instructions to authors emphasize open science practices in accordance with their TOP guidelines [14, 17]. Journals can customize whether TOP guidelines are required or optional, however, it is likely that increased adherence to TOP guidelines will be key to improving uptake of open science practices in the future.

#### Box 1. Open Science Resources for Researchers

#### Reporting Guidelines

American Psychological Association Reporting Guidelines:
Quantitative: https://apastyle.apa.org/jars/quant-table-1.pdf
Qualitative: https://apastyle.apa.org/jars/qual-table-1.pdf
Mixed methods: https://apastyle.apa.org/jars/mixed-table-1.pdf

EQUATOR Network: https://www.equator-network.org/

Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA): http://www.prismastatement.org/

Consolidated Standards of Reporting Trials (CONSORT): http://www.consort-statement.org/

Transparency and Openness Promotion (TOP) Guidelines: https://www.cos.io/initiatives/top-quidelines

See Current Signatories tab for participating journals

#### Preregistration

Systematic reviews and meta-analyses: https://www.crd.york.ac.uk/PROSPERO/

Clinical trials: https://clinicaltrials.gov/ https://www.isrctn.com/ Registered reports (after in principle acceptance): https://osf.io/rr/

See Participating Journals tab for journals offering this article format: https://www.cos.io/initiatives/registered-reports

Preregistration templates: http://dx.doi.org/10.23668/psycharchives.4584, https://osf.io/zab38/wiki/home/

#### **Preprints and Postprints**

Electronic preprints and postprints: https://www.eprints.org/uk/

Most journals' postprint policies: https://v2.sherpa.ac.uk/view/publisher\_list/1.html SPARC author addendum: https://sparcopen.org/our-work/author-rights/brochure-html/

American Psychological Association policy: https://www.apa.org/pubs/journals/resources/internetposting-guidelines

Australian Resource Council policy: https://www.arc.gov.au/policies-strategies/policy/arc-open-accesspolicy-version-20171

Australian National Health and Medical Research Council policy: https://www.nhmrc.gov.au/aboutus/resources/open-access-policy

#### Open Research

British Psychological Society policy: https://www.bps.org.uk/news-and-policy/open-data-positionstatement

National Institutes of Health Policy for Data Management and Sharing: https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-013.html

FAIR Principles: https://www.go-fair.org/fair-principles/

About Creative Commons licenses: https://creativecommons.org/about/cclicenses/

Generalist open science repositories (all assign DOIs) Harvard Dataverse: http://dataverse.harvard.edu/

Mendeley Data (also indexes 6 other repositories): https://data.mendeley.com/

Open Science Framework: http://osf.io/

Zenodo: https://zenodo.org/

Other repositories exist for specialty areas such as neuroscience. See Meyer, 2018.

Synthetic databases: https://www.synthpop.org.uk/get-started.html

#### Videos, Primers, and How-to Guides

OSF (preregistration): https://www.youtube.com/watch?v=8QK2-udwoK8

OSF (how-to guides): https://help.osf.io/hc/en-us

UK Reproducibility Network: https://www.ukrn.org/primers/

Synthetic databases: https://www.dsquintana.com/talk/riots\_synthetic/

Table 1 Open Science Practices in Behavioral Medicine Research Council Society Journals

	By journal			By year		
	ABM, %	HP, %	PM, %	2018, %	2019, %	2020, %
1. Does the article state whether or not the study (or some aspect of the study) was preregistered? (Yes)	23.2	10.4	14.4	18.7	11.4	16.9
2. Is the article a Registered Report? (Yes)	0	3.8	0	0	1.7	3.2
3. Does the article link to an accessible protocol? (Yes)	10.5	10.1	11.9	17.1	11.1	4.6
4. Does the article state whether or not data are available? (Yes)	15.4	6.8	5	2.0	14.8	9.9
5. Does the article state whether the study materials are available (on a free to access repository or similar) or make them available in the paper or supplementary materials section? (Yes)	28.9	21.9	11.5	15.8	28.6	19.4
6. Is the article gold open access? (Yes)	48.3	51.1	8.5	42.9	53.2	23.7

These findings mirror psychology at large and also echo a recent pulse survey conducted by the Society of Behavioral Medicine examining the work presented at the 2019 annual meeting of the society [15, 18]. Nearly three-quarters of all presentations (e.g., papers, posters, and symposia) did not report using any Open Science practice. Taken together, these findings should represent a call to action for health psychology and behavioral medicine researchers to integrate Open Science practices into research programs and investigate the barriers to uptake [19, 20].

Nevertheless, health psychology and behavioral medicine researchers have been early adopters of some key Open Science practices [21]. We have been exemplars in preregistering systematic reviews and meta-analyses and following the Preferred Reporting Items for Systematic Reviews and Meta-analyses and the Consolidated Standards of Reporting Trials guidelines [21]. Moreover, for many years, perhaps due to our close collaborative relationships with medicine or due to regulatory requirements, it has been standard practice for health psychology and behavioral medicine researchers to preregister randomized controlled trials in open-access trial repositories. As of April 2021, Translational Behavioral Medicine, published by the Society of Behavioral Medicine, has adopted the badge system for open data and open materials, thus providing an incentive for authors to make available their data and study materials to other researchers.

#### **Preregistration**

The number of published null results has increased over time in U.S. National Heart, Lung and Blood Institute (NHLBI) funded clinical trials, potentially as the result of introducing registration for large clinical trials on clinicaltrials.gov around the year 2000 [22]. Specifically, 57% of studies published before 2000 reported beneficial intervention effects on the primary outcome, compared with only 8% of trials published after 2000. Thus, the year 2000 marked the beginning of a natural experiment that resulted in greater constraints on reporting clinical trial results, which may have led to greater transparency in reporting standards.

When analyses are conducted transparently, questionable research practices are less likely. Questionable research practices are actions that may not constitute outright scientific fraud but threaten the validity of scientific conclusions [23]. They come in many forms but commonly arise from post hoc activities to produce a more easily publishable paper. One example is "p-hacking," which is the practice of taking actions such as removing observations or adding covariates solely to lower p values below .05 [24]. Another example is "HARKing," which stands for hypothesizing after results are known [25]. HARKing violates the fundamental tenet of formulating hypotheses a priori before an experiment is conducted. Yet another example is the overuse of "researcher degrees of freedom," wherein many statistical tests are run and only those that reach the threshold for statistical significance are reported.

There are numerous benefits of preregistration, not least that registering empirical work helps reduce the use of questionable research practices [26]. It is consistent with the requirements of truly confirmatory research, while not precluding the performance of exploratory research and data analysis [27]. Preregistration involves the precise specification and

documentation of all the main aspects of an empirical study and registering these in a repository in advance of conducting the work. As a result, researchers give careful and thorough consideration of the study hypotheses, design, data acquisition, and data analysis plans *a priori*, allowing time to finetune all aspects of the research process and ensuring that the research team has an agreed-upon, clear understanding of the proposed research. It also provides the researcher the opportunity to specify which hypotheses are confirmatory and which are exploratory. Presenting exploratory results as confirmatory misrepresents the scientific process and is another kind of questionable research practice [28].

One commonly raised objection is that preregistration is not possible in the case of secondary data analysis. Indeed, because the cost of collecting data is high, many of us engage in secondary data analysis of large data collection efforts, such as the Health and Retirement Survey or the Midlife in the United States Study. However, preregistration before *analysis* is possible, and thus Open Science is not at odds with secondary data analysis. Of course, whether or not a secondary data analysis is preregistered, manuscripts should be transparent about whether the research questions were formulated before the analyses were conducted and specifying which were exploratory.

The BMRC strongly recommends the practice of preregistration when engaging in hypothesis-driven research, with transparent reporting of deviations from preregistered plans.

#### **Registered Reports**

Null findings are more likely to remain in a researcher's file drawer and/or are less likely to be accepted for publication [29]. This science-wide problem is not limited to health psychology and behavioral medicine. However, as outlined earlier, the impact of publication bias is of greater consequence in our disciplines than many others, therefore making the introduction of Registered Reports a particularly important development for our field.

The Registered Report is relatively new type of article that aims to increase scientific transparency by implementing peer review before study results are obtained. Once the researcher has developed an idea and designed the study, including details of measures, sample size, inclusion/exclusion criteria, and data analysis plan, they submit a Stage 1 Registered Report (including Introduction and Method sections) for peer review. The key difference from the standard scientific process is that the researcher does not commence data collection until the Stage 1 Registered Report has received an In-Principle Acceptance. Once the data are collected and written up, the full Registered Report will be accepted for publication irrespective of the findings or their statistical significance, conditional on adherence to the Registered Report. Comparing 71 published Registered Reports in psychology with a random sample of 152 hypothesis-testing studies, 96% of standard reports had positive results compared with only 44% positive results in the Registered Reports [6]. Yet, the quality of Registered Reports has been shown to be higher than conventional publications [30]. At this time, Annals of Behavioral Medicine and Health Psychology do not offer registered reports. Psychosomatic *Medicine* is introducing the format in 2023.

The BMRC recognizes the value for journals in the area of behavioral medicine and health psychology to introduce

Registered Reports as a new article format. Over time, this change is likely to help encourage the uptake of this new approach to conducting science and improve the robustness of our evidence base [31].

#### **Preprints and Postprints**

A *preprint* is a version of a scholarly work, often a complete draft and after feedback from coauthors, uploaded to a public server without undergoing formal peer review. A postprint is a version of a scholarly piece of work that is uploaded to a public server after formal peer review [32]. The emphasis placed on preprints (and perhaps postprints) is often discipline-specific. For example, the preprint server arXiv.org has been essential for physics, mathematics, and computer sciences for almost three decades and EconStor has long been the norm as a disciplinary repository for economics and business. In contrast, the preprint server PsyArXiv.com was established in 2016 for the psychological sciences and is still in its infancy.

Preprints and postprints are important to Open Science as they provide open and rapid (in the case of preprints) access to scholarly work. This ensures the work is made publicly available to all interested parties, especially those in developing nations where institutional funds to publish, read, and subscribe to scientific journals are limited. Empirically, journal articles deposited on a preprint/postprint server have sizably higher citation and altmetric counts compared to non-deposited articles [33].

Given the momentum of Open Science and the unprecedented explosion of preprints in COVID-19 times, most psychology journals now permit the posting of preprints. However, most journals do not permit posting the publisher-prepared PDF but may allow posting the original author-formatted document. It is, therefore, important that authors check the journal policy on posting preprints and postprints (see Box 1). It is also possible for authors to negotiate for permission to post their preprints and postprints using tools such as the SPARC Author Addendum (see Box 1).

Preprints and "peer reviewed" published papers represent a continuum in the evolution of a body of work and can be formally linked, ensuring that the "peer reviewed" published paper supersedes the preprint as the version of record that should be cited [34]. Best practice is to update the preprint to the author-formatted document with each submission, ensuring that the available preprint is the final version submitted to the journal and providing a digital online identifier (DOI) for the published version of record. Some services will automatically link the preprint and published version-of-record DOIs. Conversely, a journal may require that the DOI for the preprint be provided in the version of record. For the member society journals, Annals of Behavioral Medicine and Psychosomatic Medicine have explicit preprint policies that allow for posting to non-commercial (NC) preprint servers and set forth DOI requirements. The American Psychological Association has a policy for its journals (including Health Psychology) that also allows posted preprints, with more stringent rules about copyright and warnings about "manuscripts that have garnered significant media attention as preprints" (see Box 1).

There are further advantages (and disadvantages) to posting preprints (see Table 1 in Ref [35]), and these can be considered from the perspective of the academic and early career

researcher (ECR), funding bodies, and journal publishers. From time to submission to paper publication, the publication process is unpredictable, variable, and often time-consuming—particularly problematic for ECRs who rely on the timely publication of their work to gain recognition for their efforts [36]. Depositing a scholarly piece of work in a preprint server ensures that the work is made publicly available almost immediately and to all, democratizing the flow of information. Authors can also receive feedback beyond a selected few who review the scholarly work during a formal peer-review process and make their judgments of appropriateness of and interest in the work. Moreover, preprints can be revised and updated far more efficiently than submitting corrections after publication. Further, a preprint documents the history of the ideas and thus becomes a timestamp establishing priority of scientific discovery and innovation, debunking the myth that preprints lead to scooping [34]. Posting preprints can also benefit academics, particularly ECRs, increasing visibility, facilitating networking, accelerating training time, optimizing research design and quality, and developing reviewer skills [36].

From the perspective of funding bodies and journals, there can be substantial benefits from the widespread adoption of preprints [34]. Although funders typically ask for "peer-reviewed publications" as demonstrated evidence of researchers' work in the field, they often allow the detail of "other scientific contributions". Such contributions could include preprints. Preprints provide tangible evidence of researchers' most recent work. Funding decisions should be based on the merit of the research, and preprints help to uphold this principle by allowing independent assessment of researchers' ideas rather than relying on journal names or impact factors as a proxy for quality [34]. Comments on preprints can also provide a more efficient formal review process, possibly improving the final manuscript.

Despite the many benefits, some concerns and challenges must be addressed, particularly concerning preprints (see Table 1 in Ref [35]). One concern with preprints is that servers will be flooded with weak papers only meant to assert priority. This can lead to misleading findings and confusion and distortion of study conclusions as well as premature media coverage, which is potentially dangerous given that preprints can shape scientific and global discourse [34], a phenomenon witnessed with the acceleration of preprints around COVID-19 [37, 38]. Given preprints have the potential, knowingly or not, to misrepresent knowledge, an important empirical question to be considered is: how can the scientific field ensure preprints positively and accurately shape knowledge? Also, how can the distinction between preprints and formal "peer-reviewed" papers be upheld, especially to lay readerships, and in all stages of the communication process (including conventional media, social media, and policy)? Should the notion be embraced that preprints and "peer reviewed" papers exist in parallel, synergizing and fulfilling complementary functions? Preprints facilitate rapid communication of scientific findings, whereas "peer reviewed" papers provide formal certification processes that promote reliability and reproducibility [34, 38].

Among 3,759 researchers across multiple disciplines, Open Science content and independent verification of author claims were essential for judging preprint credibility [39]. Peer reviews and author information were rated as less critical. Nevertheless, upholding fundamental principles and practices

of peer review should be maintained when assessing the quality of preprints, and papers should adhere to respected article reporting standards (see Box 1).

The BMRC recognizes the potential value of preprints as mechanisms to improving transparency and faster dissemination. However, the lack of regulation and potential to produce harm are significant concerns, and we view peer-reviewed, accepted science as the best form of evidence.

The BMRC recommends a close evaluation of the role of preprints for health psychology and behavioral medicine research and to compare this role with the use of preprints among physicists and economists.

#### **Engaging in Open Research**

Open research involves openly sharing one's research materials with others, including data, syntax, protocols, experimental stimuli, and so on [8, 28, 40-43]. One guideline for open data comes from the FAIR (findable, accessible, interoperable, and reusable) principles (see Box 1), which will be invoked below [44]. However, many researchers have reservations. They have proprietary feelings about data that took significant resources to collect, syntax that took significant expertise and time to write, and stimuli that took significant piloting to refine [20, 45-47]. Furthermore, making data, code, and other material shareable requires additional work (e.g., creating a codebook, cleaning data to ensure anonymity, labeling data, and commenting on code so it is interpretable) [45]. Promoting FAIR data will require planning for and budgeting money and time to prepare the data for open access. Researchers may also be concerned that their research will be "scooped" [20, 47, 48].

On the other hand, the resources involved in research materials and data are often taxpayer-funded and therefore arguably belong in the public domain. Delivering our findings transparently to the public is a first principle and an ethical obligation of the scientific community, ensuring quality and eschewing gatekeeping. In addition, open research benefits the entire field in that more resources are available to more researchers [20, 47]. Meta-analysis of individual participant data (sometimes called mega-analysis), facilitated by open research, is beginning to take over from meta-analysis of published results. Individual participant data meta-analyses are better powered and can better address moderators and confounding variables [49].

Less well-known are the benefits to the individual researcher. First, additional work to make data and syntax shareable is an academic work product. It is, therefore, possible to create a *curriculum vitae* (CV) line for publicly available datasets and syntax files, particularly when the data are extensive and extensively documented or when the syntax uses innovative and reusable approaches to problems. Many data repositories assign a DOI, making data *findable* and citable, and journals should mandate citation of data in papers using those data [29] (this mandate is part of the Open Science TOP Guidelines.) The license associated with the data (see below) can generate citations for the work. Furthermore, data and code sharing are associated with citation advantages for the publication itself [50].

Second, open research creates opportunities to find new collaborators and to publish research with other groups [47, 50]. Sharing data, for example, does not automatically mean

allowing others unfettered use of the data. Many different licenses can be applied to data, from CC0 (public domain) to CC BY (credit given to the creator, using the DOI) and additions including NC (non-commercial use only), SA (adaptations must be shared under the same terms), and ND (no derivatives or adaptations of the work permitted) (see Box 1). If a creator is interested in collaborating on shared data, a more restrictive license (e.g., CC-NC-ND) prevents new and different uses except when collaborating with the creator. Licenses are part of making data *reusable*. Simulated datasets (see below) are another method for finding new collaborators rather than sharing data in the public domain. Embargo periods are also possible [40].

Third, the process of making research materials shareable often reveals errors before sharing. One would typically want to make sure that a lab member or colleague can understand materials and reproduce results, that is, recreate the same results using the same data (or simulated data) and code. Unfortunately, errors are rife in the scientific literature. Too few research results are reproducible from the data (e.g., only 63% of meta-analyses were reproducible within 0.1 of the reported effect size) [51]. Typographical errors sneak in, perhaps contributing to many misreported p values [52]. The process of making data and code open is likely to reduce errors, corrections, and even retractions insofar as it motivate reproducibility checks before publication. Psychological Science articles with open data had only 5% major discrepancies on reproduction in measures of central tendency, variation, p values, effect sizes, test statistics, count/proportions, and degrees of freedom [53]. By contrast, articles in psychology published between 1985 and 2013 had 7%-15% major discrepancies in p values alone [52]. Open data and the researchers who publish them were perceived as more trustworthy [47].

Finally, open research is increasingly a requirement by funders and journals [20]. For example, the National Institutes of Health (NIH) requires a Data Management and Sharing Plan in grant applications (see Box 1) and will soon require that "researchers will maximize the appropriate sharing of scientific data, acknowledging certain factors (i.e., legal, ethical, or technical) that may affect the extent to which scientific data are preserved and shared." The policy defines data as: "The recorded factual material commonly accepted in the scientific community as of sufficient quality to validate and replicate research findings, regardless of whether the data are used to support scholarly publications. Scientific data do not include laboratory notebooks, preliminary analyses, completed case report forms, drafts of scientific papers, plans for future research, peer reviews, communications with colleagues, or physical objects, such as laboratory specimens" (emphasis added).

#### Making One's Research Open

Making one's research open is not difficult, although some elements are more difficult than others, and every step toward more open research is important (see resources in Box 1) [54]. Repositories exist for deposition of open research materials. Some journals and universities provide data repositories, and there are general and discipline-specific repositories (Box 1). Repositories are important for preventing broken or deleted links to an individual scientist or lab's web page and ameliorate low response rates when data are requested. Registration and indexing in a searchable resource such as a repository is

part of making data *findable*. Data may be shared as used in a particular publication (NIH will expect this step on publication) or as a complete study dataset (NIH will expect this step at the end of the funding period). The former is essential to assessing a study's reproducibility, and the latter avoids waste of resources associated with questions unasked of a particular dataset. It is important to share data in a form that will not become technologically inaccessible and is compatible with different software and therefore *interoperable*. For example, .csv files are more robust than .sav (SPSS) files.

Perhaps the most challenging issue in open data is privacy [55]. Many consent forms do not include language about data sharing but doing so is now best practice [56]. Consent rates were generally not affected by this language in psychological research, and the majority of consented participants in genomic research chose public release of anonymized data [57, 58]. Qualitatively, participant concerns about open data center around privacy invasion and release to irresponsible third parties [57, 59]; addressing these concerns during the informed consent process might improve consent rates. Local institutional review boards may also limit open data due to privacy concerns [20]. Finally, some data may preclude sharing because culture-specific knowledge is required to use them, or a cultural group does not permit it [48, 60]. Participants from underrepresented racial or ethnic groups may be less amenable to data sharing than White participants [48]. Industry funders and even academic institutions may prohibit open data or raise barriers to open data, such as complex approval processes. Sharing should be as open as possible and as closed as necessary to protect privacy and adhere to regulations (e.g., British Psychological Society [BPS] open data policy, see Box 1).

There are often federal guidelines regarding what is considered private health information and how de-identification is achieved (e.g., in the USA, the Safe Harbor method) [61]. However, a conservative rule of thumb is that if a person could definitively identify themselves in a dataset, then it is possible that others could also identify them and further measures may be necessary (see BPS open data policy, Box 1). Many data can be anonymized, but there are still options for open research where that is impossible [55, 62]. One solution for quantitative data is a synthetic dataset (see Box 1). Synthetic datasets preserve the variances and covariances of the original data but do not include any of the original data. A synthetic dataset will reproduce the original results given the same analysis. Furthermore, a synthetic dataset allows others to explore additional analyses or test other hypotheses and get the same results they would get with the actual data but precludes publication of those results. The original scientist(s) who obtained the original data must be included to create a publishable product. Simulated datasets can be quite large regarding the number of variables and number of observations and are easily generated using the R package synthpop [62]. Commercial solutions for electronic medical record data are also available [63].

Code associated with a particular publication should be shared alongside the data, whether real or synthetic. Both pieces are necessary to evaluate reproducibility, that is, the ability of an outside person to obtain the same results, given the same data and code. (Reproducibility is distinguished from replicability, which is the ability to obtain the same results given the same methods but new data.) Ideally, the code includes all the steps taken in cleaning, scoring, and

analyzing data—that is, a third party could take the raw data and the code and obtain the reported results. Comments detailing the purpose and rationale for each step should be included in the code [45].

The BMRC recognizes the value of open research to improve value, accuracy, and collaboration in health psychology and behavioral medicine research.

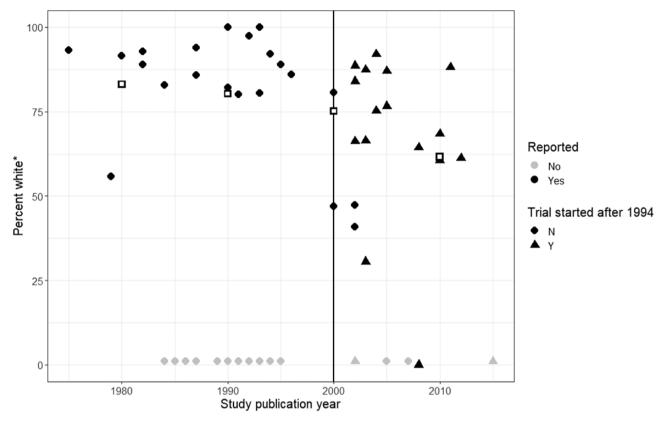
The BMRC encourages open research practices at a minimum as required by funding entities and publications. In practice, research materials should be as open as possible and as closed as necessary, respecting privacy, laws, and cultural knowledge.

#### **Open Science and Equity**

Open Science has the potential to both improve and obstruct equity for underrepresented groups in science [48]. On one hand, the availability of preprints/postprints (with their attendant benefits and drawbacks, see above) and open data may benefit scientists with fewer resources, who may not have subscription access to journals or the financial or logistical ability to collect large samples of participants [64]. Researchers from underrepresented groups highly endorsed open science values of rigor, reproducibility, and transparency and believed that research dissemination was an important equity issue [48]. Collaborations arising from open science may benefit researchers from underrepresented groups and generate adequately powered samples of underrepresented groups [48, 64]. Some practices (preprints and postprints) do not incur a significant burden, and others (preregistration) may save time in the long run [65].

On the other hand, researchers from underrepresented groups were also concerned that financial and time resources required for some open science practices [20, 45–47] would further disadvantage scholars from underrepresented groups [48]. In financial terms, open-access publication should be considered in an equity context; the cost to publish open access can be prohibitive even for well-resourced investigators (e.g., at the time of writing,  $\leq 9,500$  at Nature [66], or at the current exchange rate, US\$10,165). In time terms, scholars from underrepresented groups already bear an unequal burden in mentoring and service work (the "minority tax"). More recognition for open science practices in evaluation and promotion is not necessarily a cure: Groups who do not bear additional burdens might benefit disproportionately because they have more time to engage in open science practices. A more equitable research environment is needed to advance equitable open science, including decreasing the "minority tax" imposed on additional service contributions [67] and multilevel, multidimensional initiatives to increase individual and structural equity for female and underrepresented researchers [68].

Finally, preregistration might include attendant pressures to improve statistical power by relying on populations that are not hard to recruit and thereby decrease diversity. To probe this question, reported racial/ethnic diversity in the clinical trials included in Ref [22] was examined. Figure 1 shows the results (data and code available at https://osf.io/wytz3/). There is a clear trend toward more diversity after the preregistration requirement was put in place in 2000. However, this era also coincides with the March 1994 NIH requirement that grant applications include gender and ethnic diversity such that "for Phase III clinical trials... women



**Fig. 1.** Percent White participants in Ref. [22] as a function of study publication year and whether trial recruitment started after the publication of National Institutes of Health guidance in 1994. Reports in which racial/ethnic descriptions were not included are shown at the bottom of the graph. Census estimates for the USA are shown in open boxes. \* indicates where the sample was described only as percent of a nonwhite group, the remainder was assumed to be White for the purpose of this illustration.

and minorities and their subpopulations must be included such that valid analyses of differences in intervention effect can be accomplished" [69]. This requirement followed 1990 guidance on the "inclusion of women and members of minority groups in all NIH-supported biomedical and behavioral research involving humans subjects" [69]. A few conclusions may be drawn from these data: first, diversity increased following requirements rather than guidance; second, racial/ethnic qualities of the sample were more likely to be reported following the onset of requirements and preregistration; and third, before requirements and preregistration, the proportion of white participants usually exceeded Census estimates (open squares in Fig. 1); afterward, the proportion was closer to census estimates. The added requirement of preregistration did not appear to harm diversity in these clinical trials. However, preregistration does not typically require consideration of diversity as do NIH grant applications. Insofar as preregistration benefits researchers by requiring them to carefully consider how their study will be performed and why, the addition of diversity elements to preregistration would force researchers to address generalizability with regard to diversity and representation; oversampling may be necessary to appropriately characterize some groups [70]. The recruited sample could also be compared against the preregistration targets.

The BMRC recognizes the advantages and disadvantages of Open Science in achieving equity in health psychology and behavioral medicine. A more equitable research environment is needed to advance equitable open science. Open access publication cost and institutional recognition of Open Science practices may inadvertently disadvantage underrepresented scientists.

### The Need for Civility, Collegiality, and Collaboration

There have been numerous important and innovative developments in how scientific research is conducted. These changes have been described by some as a scientific revolution and there has been much talk of psychological science undergoing a renaissance [21]. However, there has also been discussion of the "tone debate" and concerns about the civility of the conduct of the scientific debate surrounding replication and reproducibility [71]. These concerns have centered around the need to be respectful and collegial in scientific discourse, to critique the science and not the scientist, and to recognize that there are different reactions to Open Science practices. In the latter case, for example, preregistration can be viewed by some as a commitment to do exactly what was proposed; however, it is also important to remember that preregistration is "a plan, not a prison" [72]. Deviations should be transparently reported but not demonized, allowing dispassionate and scientific scrutiny of the rationale and consequences of deviations. In the context of study replications more generally, the BMRC notes that failures of replication may reflect critical issues of context [73] and this failure to replicate and consequent drive to generate new hypotheses is part of the scientific method.

The BMRC urges researchers to be tolerant and to work together in a collaborative, collegial, and civil manner.

#### Conclusion

We have argued that Open Science in health psychology and behavioral medicine can potentially increase reproducibility, replication, openness, and transparency, which will improve our science's quality and reliability. There is no onesize-fits-all solution that will encompass all Open Science needs in health psychology and behavioral medicine's diverse research products and outlets: for example, qualitative science and community-based participatory research will require a different approach than quantitative science; clinical trials will require a different approach than observational studies. Different scientists and journals will have different research foci both in topic and approach and will adopt Open Science guidelines accordingly. When deciding to engage in or with Open Science practices and evaluations, researchers should include collegiality and equity in their priorities. However, there are sufficient resources and motivating data that health psychology and behavioral medicine research as a discipline should continue to move toward Open Science. This will ultimately improve the robustness of our evidence base in the longer term. As such, the BMRC recommends that health psychology and behavioral medicine adopt more Open Science practices such as preregistration, registered reports, and open research and that the field continue to monitor the viability of preprints as a method of scientific communication.

#### **Acknowledgments**

This statement was developed through a collaboration among the Society for Health Psychology, the Society of Behavioral Medicine, the American Psychosomatic Society, and the Academy of Behavioral Medicine Research and has been published jointly in *Health Psychology*, the *Annals of Behavioral Medicine*, and *Psychosomatic Medicine*.

The authors thank Ava Cazares, Joon Soo Kim, Thomas Mistretta, Halie Pfister, Cristina Pinheiro, Charis Stanek, Christy Wang, and Andrea Yacoub for their assistance in coding Open Science practices.

Behavioral Medicine Research Council members: Simon L. Bacon, Gary G. Bennett, Elizabeth Brondolo, Susan M. Czajkowski, Karina W. Davidson, Elissa S. Epel, Tracey A. Revenson, and John M. Ruiz.

#### **Compliance With Ethical Standards**

Authors' Statement of Conflict of Interest and Adherence to Ethical Standards Authors Suzanne C. Segerstrom, Michael A. Diefenbach, Kyra Hamilton, Daryl B. O'Connor, and A. Janet Tomiyama declare that they have no conflict of interest. All procedures, including the informed consent process, were conducted in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2000.

Contributions S.C.S.: Authors' Conceptualization, Investigation, Data Curation, Writing-Original Draft, Writing-Review and Editing, Visualization, Supervision. M.A.D.: Conceptualization, Investigation, Writing-Original Draft, Writing-Review and Editing, Kyra Hamilton: Conceptualization, Investigation, Writing-Original Draft, Writing-Review and Editing. Daryl B. O'Connor: Conceptualization, Investigation, Writing-Original Draft, Writing-Review and Editing. A. Janet Tomiyama: Conceptualization, Investigation, Writing-Original Draft, Writing-Review and Editing.

Disclaimer The content of this paper is solely the responsibility of the authors and does not necessarily represent the official views or policies of the US National Cancer Institute, National Institutes of Health, or Department of Health and Human Services.

Ethical Approval This review was not formally registered, and also there was no analytic plan in this review apart from descriptive statistics.

Note The mission of the Behavioral Medical Research Council (BMRC) is to identify strategic, high-priority research goals, and encourage multidisciplinary and multicenter research networks to pursue them. The BMRC consists of representatives of the following organizations: Academy of Behavioral Medicine Research; American Psychosomatic Society; Society for Health Psychology; and the Society of Behavioral Medicine. More information about the BMRC can be found at https://www.behavioralmedicineresearchcouncil.org/.

Data availability Data used to construct the table and figure presented in this review are available in a public archive: https://osf.io/wytz3/. (iv)Analytic code availability. Analytic code used to construct the table and figure presented in this review are available in a public archive: https://osf.io/wytz3/. (v) Materials availability. There were no materials in this review.

#### Supplementary Material

Supplementary material is available at *Annals of Behavioral Medicine* online.

#### References

- Yong E. Replication studies: Bad copy. Nature. 2012; 485:298– 300
- Camerer CF, Dreber A, Forsell E, et al. Evaluating replicability of laboratory experiments in economics. *Science*. 2016;351:1433– 1436.
- 3. Prinz F, Schlange T, Asadullah K. Believe it or not: How much can we rely on published data on potential drug targets? *Nat Rev Drug Discov*. 2011;10:712–712.
- Franco A, Malhotra N, Simonovits G. Underreporting in psychology experiments: Evidence from a study registry. Soc Psychol Personal Sci. 2016;7:8–12.
- Franco A, Malhotra N, Simonovits G. Publication bias in the social sciences: Unlocking the file drawer. Science. 2014;345:1502–1505.
- Scheel AM, Schijen MRMJ, Lakens D. An excess of positive results: Comparing the standard psychology literature with registered reports. Adv Methods Pract Psychol Sci. 2021;4.
- Open Science Collaboration. Estimating the reproducibility of psychological science. Science. 2015;349:aac4716.
- Nosek BA, Hardwicke TE, Moshontz H, et al. Replicability, robustness, and reproducibility in psychological science. 2021;73:719– 748.

- Collins FS, Tabak LA. Policy: NIH plans to enhance reproducibility. Nature. 2014;505:612–613.
- 10. Chambers CD, Tzavella L. The past, present and future of Registered Reports. *Nat Hum Behav*. 2022;6:29–42.
- Norris E, He Y, Loh R, West R, Michie S. Assessing markers of reproducibility and transparency in smoking behaviour change intervention evaluations. *J Smok Cessat*. 2021;2021;e6694386.
- McVay MA, Cooper KB, Carrera Seoane M, Donahue ML, Scherer LD. Transparent reporting of hypotheses and analyses in behavioral medicine research: An audit of publications in 2018 and 2008. Health Psychol Behav Med. 2021;9:285–297.
- 13. Nosek BA, Alter G, Banks GC, et al. Promoting an open research culture. *Science*. 2015;348:1422–1425.
- Freedland KE. Health psychology adopts transparency and openness promotion (TOP) guidelines. Health Psychol. 2021;40:227–229
- McVay MA, Conroy DE. Transparency and openness in behavioral medicine research. *Transl Behav Med*. 2021;11:287–290.
- Segerstrom SC. Psychosomatic Medicine: looking forward. Psychosom Med. 2022;84:265–266.
- Revenson TA, Zoccola PM. New instructions to authors emphasize open science, transparency, full reporting of sociodemographic characteristics of the sample, and avoidance of piecemeal publication. *Ann Behav Med.* 2022;56:415–417.
- Hardwicke TE, Thibault RT, Kosie JE, Wallach JD, Kidwell MC, Ioannidis JPA. Estimating the prevalence of transparency and reproducibility-related research practices in psychology (2014– 2017). Perspect Psychol Sci. 2022;17:239–251.
- Gagliardi D, Cox D, Li Y. Institutional inertia and barriers to the adoption of open science. In: *The Transformation of University Institutional and Organizational Boundaries*. Leiden, Netherlands: Brill;2015.
- Houtkoop BL, Chambers C, Macleod M, Bishop DVM, Nichols TE, Wagenmakers EJ. Data sharing in psychology: A survey on barriers and preconditions. Adv Methods Pract Psychol Sci. 2018;1:70–85.
- O'Connor DB. Leonardo da Vinci, preregistration and the architecture of science: Towards a more open and transparent research culture. Health Psychol Bull. 2021;5:39–45.
- Kaplan RM, Irvin VL. Likelihood of null effects of large NHLBI clinical trials has increased over time. PLOS One. 2015;10:e0132382.
- 23. John LK, Loewenstein G, Prelec D. Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychol Sci.* 2012;23:524–532.
- Winter JC de, Dodou D. A surge of p-values between 0.041 and 0.049 in recent decades (but negative results are increasing rapidly too). *PeerJ*. 2015;3:e733.
- 25. Kerr NL. HARKing: Hypothesizing after the results are known. Personal Soc Psychol Rev. 1998;2:196–217.
- Bosnjak M, Fiebach CJ, Mellor D, et al. A template for preregistration of quantitative research in psychology: Report of the joint psychological societies preregistration task force. *Am Psychol.* 2022;77:602–615.
- Fife DA, Rodgers JL. Understanding the exploratory/confirmatory data analysis continuum: Moving beyond the "Replication Crisis." Am Psychol. 2022;77:453–466.
- 28. Munafò MR, Nosek BA, Bishop DVM, et al. A manifesto for reproducible science. *Nat Hum Behav*. 2017;1:1–9.
- Simonsohn U, Nelson LD, Simmons JP. P-curve: A key to the filedrawer. J Exp Psychol Gen. 2014;143:534–547.
- 30. Soderberg CK, Errington TM, Schiavone SR, et al. Initial evidence of research quality of registered reports compared with the standard publishing model. *Nat Hum Behav*. 2021;5:990–997.
- 31. Norris E, O'Connor DB. Science as behaviour: Using a behaviour change approach to increase uptake of open science. *Psychol Health*. 2019;34:1397–1406.
- 32. Harnad S. Electronic preprints and postprints. In: *Encyclopedia of Library and Information Science*. New York: Marcel Dekker; 2003.
- 33. Serghiou S, Ioannidis JPA. Altmetric scores, citations, and publication of studies posted as preprints. *JAMA*. 2018;319:402.

- 34. Berg JM, Bhalla N, Bourne PE, et al. Preprints for the life sciences. *Science*. 2016;352:899–901.
- 35. Elmore SA. Preprints: What role do these have in communicating scientific results? *Toxicol Pathol*. 2018;46:364–365.
- Sarabipour S, Debat HJ, Emmott E, Burgess SJ, Schwessinger B, Hensel Z. On the value of preprints: An early career researcher perspective. *PLoS Biol.* 2019;17:e3000151.
- 37. Fraser N, Brierley L, Dey G, et al. The evolving role of preprints in the dissemination of COVID-19 research and their impact on the science communication landscape. *PLoS Biol.* 2021;19:e3000959.
- 38. Vlasschaert C, Topf JM, Hiremath S. Proliferation of papers and preprints during the coronavirus disease 2019 pandemic: Progress or problems with peer review? *Adv Chronic Kidney Dis*. 2020;27:418–426.
- Soderberg CK, Errington TM, Nosek BA. Credibility of preprints: An interdisciplinary survey of researchers. R Soc Open Sci. 2020;7:201520.
- 40. Martone ME, Garcia-Castro A, VandenBos GR. Data sharing in psychology. *Am Psychol.* 2018;73:111–125.
- 41. Hesse BW, Conroy DE, Kwaśnicka D, et al. We're all in this together: Recommendations from the Society of Behavioral Medicine's Open Science Working Group. *Transl Behav Med*. 2021;11:693–698.
- 42. Lindsay DS. Sharing data and materials in psychological science. *Psychol Sci.* 2017;28:699–702.
- Schönbrodt FD, Maier M, Heene M, Bühner M. Forschungstransparenz als hohes wissenschaftliches Gut stärken. *Psychol Rundsch*. 2018;69:37–44.
- 44. Wilkinson MD, Dumontier M, Aalbersberg IjJ, et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data*. 2016;3:160018.
- 45. Trisovic A, Lau MK, Pasquier T, Crosas M. A large-scale study on research code quality and execution. *arXiv:210312793 Cs.* 2021. Doi:10.48550/arXiv.2103.12793.
- Obels P, Lakens D, Coles NA, Gottfried J, Green SA. Analysis of open data and computational reproducibility in registered reports in psychology. Adv Methods Pract Psychol Sci. 2020;3:229–237.
- 47. Abele-Brehm AE, Gollwitzer M, Steinberg U, Schönbrodt FD. Attitudes toward open science and public data sharing: A survey among members of the German Psychological Society. *Soc Psychol*. 2019;50:252–260.
- 48. Lui PP, Gobrial S, Pham S, Giadolor W, Adams N, Rollock D. Open science and multicultural research: Some data, considerations, and recommendations. *Cultur Divers Ethnic Minor Psychol* (Online ahead of print April 11, 2022). Doi:10.1037/cdp0000541.
- Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: Rationale, conduct, and reporting. BMJ. 2010;340:c221.
- 50. McKiernan EC, Bourne PE, Brown CT, et al. How open science helps researchers succeed. *ELife*. 2016;5:e16800.
- Lakens D, Hilgard J, Staaks J. On the reproducibility of meta-analyses: Six practical recommendations. BMC Psychol. 2016;4:24.
- Nuijten MB, Hartgerink CHJ, van Assen MALM, Epskamp S, Wicherts JM. The prevalence of statistical reporting errors in psychology (1985–2013). Behav Res Methods. 2016;48:1205–1226.
- 53. Hardwicke TE, Bohn M, MacDonald K, et al. Analytic reproducibility in articles receiving open data badges at the journal *Psychological Science*: An observational study. *R Soc Open Sci.* 2021;8:201494.
- Kathawalla U-K, Silverstein P, Syed M. Easing into open science: A guide for graduate students and their advisors. *Collabra Psychol*. 2021;7:18684.
- 55. Walsh CG, Xia W, Li M, Denny JC, Harris PA, Malin BA. Enabling open-science initiatives in clinical psychology and psychiatry without sacrificing patients' privacy: Current practices and future challenges. Adv Methods Pract Psychol Sci. 2018;1:104–114.
- Meyer MN. Practical tips for ethical data sharing. Adv Methods Pract Psychol Sci. 2018;1:131–144.
- 57. Cummings JA, Zagrodney JM, Day TE. Impact of open data policies on consent to participate in human subjects research:

- Discrepancies between participant action and reported concerns. *PLoS One*, 2015;10:e0125208.
- 58. McGuire AL, Oliver JM, Slashinski MJ, et al. To share or not to share: A randomized trial of consent for data sharing in genome research. *Genet Med.* 2011;13:948–955.
- Trinidad SB, Fullerton SM, Bares JM, Jarvik GP, Larson EB, Burke W. Informed consent in genome-scale research: What do prospective participants think? AJOB Prim Res. 2012;3:3–11.
- 60. Lui PP, Skewes M, Gobrial S, Rollock D. Advancing transparency and impact of research: Initiating crosstalk between indigenous research and mainstream "Open Science." J Indig Res. 2021;9.
- Health and Human Services. Methods for De-identification of PHI. Washington, DC: HHS.gov;2015.
- 62. Quintana DS. A synthetic dataset primer for the biobehavioural sciences to promote reproducibility and hypothesis generation. *ELife*. 2020;9:e53275.
- 63. Reiner Benaim A, Almog R, Gorelik Y, et al. Analyzing medical research results based on synthetic data and their relation to real data results: Systematic comparison from five observational studies. *JMIR Med Inform.* 2020;8:e16492.
- 64. Syed M, Kathawalla UK. Cultural psychology, diversity, and representation in open science. In: *Cultural Methods in Psychology: Describing and Transforming Cultures*. New York: Oxford University Press;2022:427–454.

- Tackett JL, Brandes CM, Reardon KW. Leveraging the Open Science Framework in clinical psychological assessment research. *Psychol Assess*. 2019;31:1386–1394.
- 66. Seltzer R. Open Access Comes to Selective Journal. Washington, DC: Inside Higher Ed; 2020.
- 67. Williamson T, Goodwin CR, Ubel PA. Minority tax reform—avoiding overtaxing minorities when we need them most. *N Engl J Med*. 2021;384:1877–1879.
- 68. Bilimoria D, Singer LT. Institutions Developing Excellence in Academic Leadership (IDEAL): A partnership to advance gender equity, diversity, and inclusion in academic STEM. Equal Divers Incl Int J. 2019;38:362–381.
- NIH Guide. NIH Guidelines on the Inclusion of Women and Minorities as Subjects in Clinical Research;2000. https://grants.nih. gov/grants/funding/women\_min/guidelines\_amended\_10\_2001.htm.
- 70. Vaughan R. Oversampling in health surveys: Why, when, and how? *Am J Public Health*. 2017;107:1214–1215.
- 71. Derksen M, Field SM. The tone debate: Knowledge, self, and social order. *Rev Gen Psychol*. 2022;26:172–183.
- 72. DeHaven A. *Preregistration: A Plan, Not a Prison.* Charlottesville, VA: Center for Open Science; 2017.
- Nosek BA, Hardwicke TE, Moshontz H, et al. Replicability, robustness, and reproducibility in psychological science. *Annu Rev Psychol.* 2022;73:719–748.

I am responding to this RFI: On behalf of an organization

Name: Megan von Isenburg

Name of Organization: Data Discovery Collaboration

Type of Organization: Professional org association

Role: Institutional official

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

The Data Discovery Collaboration (DDC) is a multi-institutional consortium that works together in order to address concerns around data discovery through discussions of metadata, outreach, software development, and systems and metadata interoperability. In preparing this response, multiple members of the DDC came together to respond to this RFI from the lens of data discovery, based on the perspective that data deposit is a publication opportunity and that data publications are increasingly common.

Many journals require data publication with article publication that may exceed NIH Data Management and Sharing Policy requirements. Some researchers may have larger-than-ordinary data needs (e.g., working with human subjects data, working with complex imaging and 'omics data, working with very large datasets) that cannot be satisfied through PMC supplemental files size limits. These inequities in data publication are particularly stark for human subjects research, which may have costly and time-consuming requirements. Additionally, some researchers may come from institutions with limited financial and infrastructural resources or differential expertise in data sharing. Thus, the NIH should examine how to best support researchers across fields and across levels of institutional support and resources for data publication. If they do not do so, it is likely that some researchers will have access to publication opportunities that Others will not. This issue is particularly relevant as high profile journals like Nature tend to have these types of requirements.

To address these issues, we suggest increased monetary support, exploration into a PMC-style repository designed according to data standards, or a federated data repository or catalog interface for data associated with PMC articles.

#### 2. Steps for improving equity in access and accessibility of publications.

The DDC suggests two steps for improving equity in access and accessibility of publications. The first is workforce development training for licensing options and copyright. Authors do not always know what rights and licenses are available to them, such as Creative Commons licenses. Providing education could help even the playing field for writers as well as ensure broader access for readers.

Second, there should be standards set for Data Availability Statements (DAS) to allow for clearer and machine-readable information about when and how data associated with a publication is available. Currently, DASs are not standardised and the quality of DAS's can vary greatly across publications and articles.

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

It is important for the NIH to monitor trends in publication fees and policies, including those related to data deposit. There is some risk in researchers turning to external open access options with costly APCs if the NIH does not make enhancements to researcher ability to deposit data associated with a publication within PMC and Other NIH repositories.

#### 4. Early input on considerations to increase findability and transparency of research.

From the perspective of data as a publication and product of research, it is essential to enhance the findability of data sets resulting from NIH research regardless of which repository is used. We encourage development in federated search and data catalog options to increase findability.

In addition PIDs should be used to link data sets to their associated publications in PMC, journals, and Other systems. Data availability statements should be standardised and machine readable.

Ideally, no new PIDs should be created wherever industry standards exist, such as DOI, ORCID, and ROR. If industry standard PIDs are unable to be programmed into NIH systems or to be explicitly required by NIH, then cross-walking NIH PIDs with industry standard PIDs will be essential.

Data published as a supplementary file in PMC should be discoverable independently from their associated articles from multiple points, including topic, author, affiliation, etc. Supplementary data files are not adequately searchable at present. Ideally, published data and published articles stemming from the same research should be linked but independently discoverable.

**Description:** The Data Discovery Collaboration (DDC) is a multi-institutional consortium that works together in order to address concerns around data discovery through discussions of metadata, outreach, software development, and systems and metadata interoperability. I

Email: megan.vonisenburg@duke.edu

I am responding to this RFI: On behalf of an organization

Name of Organization: American Society of Hematology

Type of Organization: Other

Type of Organization-Other: Medical Specialty Society

**Role:** Member of the public

- 1. How to best ensure equity in publication opportunities for NIH-supported investigators.
- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

## **Uploaded File:**

ASH-Response-to-NIH-Guidelines-4.24.23.pdf

Email: sleous@hematology.org

# American Society of Hematology

AND STATE OF LAND OF OLY WHAT STATES OF OLY WHAT ST

2021 L Street, NW, Suite 900, Washington, DC 20036-4929 ph 202.776.0544 fax 202.776.0545 e-mail ASH@hematology.org

April 24, 2023

President
Robert Brodsky, MD
Johns Hopkins University
Ross Building, Room 1025
720 Rutland Avenue

Baltimore, MD 21205 Phone 410-502-2546

2023

#### **President-Elect**

Mohandas Narla, DSc New York Blood Center 310 E 67th Street New York, NY 10065 Phone 212-570-3056

#### Vice President

Belinda Avalos, MD Atrium Health 215 N. Pine Street, Unit 4703 Charlotte, NC 28202 Phone 980-442-2000

#### Secretary

Cynthia Dunbar, MD NHLBI/NIH Translational Stem Cell Biology Branch Building 10-CRC, Room 5E-3332 10 Center Drive Bethesda, MD 20892 Phone 301-402-1363

#### Treasurer

Joseph Mikhael, MD, FRCPC, MEd Translational Genomics Research Institute, City of Hope Cancer Center 445 N. Fifth Street Phoenix, AZ 85004 Phone 602-343-8445

#### Councillors

Christopher Flowers, MD, MS Alison Loren, MD, MS Bob Löwenberg, MD, PhD Charlotte Niemeyer, MD Sarah O'Brien, MD, MSc Betty Pace, MD Jamile Shammo, MD Wendy Stock, MD, MA

#### **Executive Director**

Martha Liggett, Esq.

National Institutes of Health Office of Science Policy 6705 Rockledge Drive Suite 630 Bethesda, MD 20892

Re: NOT-OD-23-091 Request for Information in the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research

To Whom it May Concern:

The American Society of Hematology (ASH) appreciates the opportunity to provide comments on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research (NIH Public Access Plan) and the 2022 White House Office of Science and Technology Policy (OSTP) memo on Ensuring Free, Immediate, and Equitable Access to Federally Funded Research.

ASH represents more than 18,000 clinicians and scientists worldwide who are committed to the study and treatment of blood and blood-related diseases. These disorders encompass malignant hematologic disorders such as leukemia, lymphoma, and multiple myeloma, as well as classical hematological conditions such as sickle cell anemia, thalassemia, bone marrow failure, venous thromboembolism, and hemophilia. As part of its mission to further the understanding, diagnosis, treatment, and prevention of disorders affecting the blood, ASH currently publishes two peer-reviewed journals, *Blood* and *Blood Advances*.

As a non-profit society publisher, ASH brings our best practices to the peer review of the articles and to wide dissemination of scholarly content in the field of hematology. The integrity of peer review is vital to sharing research findings in a way that assures accuracy, integrity, and the transmission of science that promotes new evidence vital to patient care. We are committed to public accessibility of scientific evidence as well as the need to preserve the US research enterprise as a source of high-quality scientific information. Our Society strongly recommends a two-year delay of the NIH Public Access Plan to adhere to the 2022 OSTP memo on Ensuring Free, Immediate, and Equitable Access to Federally Funded Research. This time would allow us to work with you to develop policies that sustain reliable, equitable, high quality scientific content.

ASH provides the following comments on NIH's Public Access Plan that focus on ensuring equity in publication opportunities for NIH-supported investigators, steps for improving equity in access and accessibility of publications, and early input on considerations to increase findability and transparency of research.

Ensuring equity in publication opportunities for NIH-supported investigators

Shifting of Revenue Streams

Opening papers prior to the current 12-month embargo will result in the loss of subscription revenue from institutions and individuals and, for many publishers, a corresponding decrease in advertising revenue. In order for publishers to provide the scientific community with the support it has become accustomed to, including, but not limited to, maintaining the integrity of the science, robust peer review, support for discoverability, reproducibility and dissemination of the science, the financial burden will shift to the authors. Diligent peer review, management and public disclosures of conflicts, and data and figure integrity checks are vital parts of a responsible publication process. Threats to the integrity of the content, such as plagiarism, paper mills, inappropriate AI generated content, and fraudulent data, are always present and require steady attention. While no system is perfect, peerreview increases the opportunity to mitigate these risks and protect the public from ensuing harms. Publishers also provide additional benefits to their communities by providing educational material, alternative metrics and enhanced metadata that may also suffer due to diminishing revenue. All of this requires resources that are likely to be endangered if publishers lose the revenue that currently sustains this work. Such losses could occur in the form of cancelled subscriptions, insufficient total article processing charge (APC) income, and lost licensing fees for approved reuse of content, among others.

Policies that restrict publishers' abilities to collaborate with authors to realize their protection of rights under United States copyright law would further limit revenue streams on which we depend, including royalties, licensing, reprints, and advertising. We urge the NIH not to include rights retention language or license requirements in the final policy other than the grantee's right to deposit the manuscript. Preserving a Green OA route presents a sustainable business model that should be embraced. Expanding rights retention policies beyond the deposition of the manuscript would also erode the publisher's ability to monitor usage of the content in support of the author's intellectual property.

## Access to funding

OSTP and NIH state that grants can be used to cover publication costs, which is a positive step; however, it is important that NIH increase the total amount of grant funding per award so that the additional Article Processing Charges, including potential fees to deposit papers into PubMed Central for example, will not reduce the funds available for research.

In addition, we are concerned that certain grants do not permit use of funds for publication fees. As such, ASH recommends that NIH exempt certain types of infrastructure-related grants (e.g., cancer center support grants, CTSAs, NCORPs) and teaching grants (K awards, T awards) from reporting funding to journals and thus requiring deposit.

The broad reach and impact of this proposed plan will be a challenge to implement and enforce if compliance is mandated for all NIH funded investigators regardless of how much funding they received or how small a role any given individual plays in a research project or manuscript. The NIH should instead apply a minimum threshold of funding and/or level of participation by authors and researchers before subjecting papers to the proposed mandate.

#### Copyright protection

Copyright protection is the first line of defense for any author against the misuse of their research, and publishers stand ready to defend investigators' intellectual property. Journals customarily allow authors to post their paper on their institutions' site, make use of their work at conferences, but this

policy needs to clearly state that making the content freely accessible does not give anyone the right to create derivative products without permission. Clarification that the rights remain with the copyright holder needs to be articulated. The final guidance should also clarify that authors are obligated to follow the NIH Guidelines only for the papers they author as a result of NIH funding.

## Definition of First Publication

There is confusion in the community concerning the definition of First Publication. The Society is interpreting NIH's draft language regarding first publication to mean that the manuscript uploaded to PubMed Central in compliance with this policy will be embargoed until the first appearance of the final typeset article. Are we also correct in understanding that the Pub Med Central first publication will include a link to the publisher's site? Clarification of this matter in the final policy is strongly recommended to avoid confusion in the community.

#### Steps for improving equity in access and accessibility of publications

Access and accessibility of publications

Journal publishers have long been collaborating with various stakeholders to develop and implement collaborative projects that enhance the public access, utility, preservation, and discoverability of materials that report on and analyze and interpret results of federally funded research. Publishers participate in a multitude of services that enhance discoverability, including ORCID, Crossref, the Committee on Publication Ethics, and provide guidelines that are not influenced by pharmaceutical companies as well as making sure conflicts of interest are accurately noted. Federal agencies should collaborate with publishers and other stakeholders to ensure minimum standards, share best practices, and minimize duplication of work.

Providing immediate access to all scientific research comes with significant issues and significant financial/labor costs of compliance. ASH wants to make sure that authors' intellectual property remains accurately presented on the worldwide stage; we are concerned that the research could be pirated by outside bodies that may misinterpret the results to suit their needs. While publisher's efforts to support free, immediate access to COVID-19 research were a boon to scientists, we also saw a rise of misuse and misunderstanding of research among the public. Strong intellectual property protections are a necessary safeguard against the acceleration of this trend. We recommend that NIH support publisher's ability to enforce copyright protection by maintaining publishers' rights in and to the content published.

#### Early input on considerations to increase findability and transparency of research

#### Consistent Guidance

There are many examples of advancements already accepted by the industry such as DOIs, ORCID, funder registries, discovery tools for content mining, and use of JATS for structured metadata to increase findability and transparency of research. If NIH wants to aggregate these data, ASH recommends collaboration with various stakeholders to create and engage in guidance for authors and publishers regarding standards to ensure best practices and minimize duplication of work.

Thank you for the opportunity to provide our thoughts regarding NIH's Public Access Plan. Please contact Suzanne Leous, MPA, Chief Policy Officer (<u>sleous@hematology.org</u>) or Nina Hoffman, Chief

Publications Officer (<a href="mailto:nhoffman@hematology.org">nhoffman@hematology.org</a>), should you have any questions regarding ASH's comments.

Sincerely,

Robert A. Brodsky, MD

A. Brods

President

I am responding to this RFI: On behalf of an organization

Name: David Mellor

Name of Organization: Center for Open Science

Type of Organization: Nonprofit research organization

Role: Institutional official

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

We believe that an interface that does not focus on journal name but rather clear results reporting can satisfy both the spirit of these open access plans (which is to increase access to research findings) and also the need to improve scholarly communication. It will do so by reducing the implication that the value of a research finding is associated with the name of the journal that publishes that finding. Such associations bias the research process by incentivizing novelty over rigor.

- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.

The increasing cost of publications- both through traditional subscription models and through the rising costs of APCs- are a cause for concern as it shunts money away from Other public benefits in higher education and scientific research. While there are reasonable concerns about placing a cap on the price per publication fee, namely that such a cap would become the new standard price for publishing, we encourage NIH to define "reasonable publication fees" in a manner that is not too ambiguous or that encourages further, unchecked growth in these fees. Specifying a maximum percentage that a proposed grant application budget would be one key signal to indicate how excessive could be defined.

AnOther key strategy to reigning in the cost of publication fees is transparency. Currently, there are very high barriers to even knowing how much money is spent on such fees. A relatively simple method to increase awareness about these costs is to disclose the amount of money earmarked for publication fees in funded grant applications. This process can be accomplished in an aggregated way, which would still provide the community with information about the total costs.

Together, these two steps (clearly defining reasonable costs and reporting how much money is spent annually on them) will help monitor the growing costs associated with publishing.

#### 4. Early input on considerations to increase findability and transparency of research.

We believe that there are two important steps that NIH can take in order to increase the findability and transparency of research outputs. The first is to include the Data Management and Sharing Plans (DMSP) as part of the reviewer-scored criteria during the grant review process. Currently, these plans merely have to pass administrative review and be deemed acceptable or appropriate for the grant to be submitted and reviewed for consideration of funding. While this is an understandable first step, it does not go far enough to ensure that these plans are truly as good as they could be. The current workflow reinforces the idea that data sharing is an administrative burden and not an integral part of the process of scientific research. Since grant reviewers are themselves experts within the discipline and of the

proposed methods, they will best be able to determine if the proposed plans are feasible, high quality, and meet the realistic expectations of the community. For example, when dealing with particularly sensitive datasets or those that include data about indigenous communities, the panel of expert reviewers are best able to know if the plan meets ethical norms and considerations of the field. Likewise, in areas where data sharing poses fewer ethical constraints, the reviewers will take that into consideration and place higher expectations for broad sharing and preservation of the data. Only through scoring proposed DMSPs will grant authors take as much care and consideration as they could with the details of the plans. This will elevate data sharing and improve data quality in ways administrative review cannot.

The second step is to assign DMSPs persistent, unique identifiers (ideally DOIs) and to make these DMSPs publicly available for all awarded grants. The rationale for not publicly posting many grant materials is fully justifiable, as the intellectual property of the proposals remain that of the submitters. However, DMSPs do not contain, and should not contain, original ideas or Other IP that can give away any competitive advantage from grantees. They should merely assist future researchers in finding research outputs from funded work by specifying where data and related items will be hosted. This will also help increase accountability with proposed data sharing plans, as there will be an easier way to determine how data are created and preserved. We believe that transparency in this setting will help readers, future researchers, and members of the public see how data are generated, stored, and reused in order to maximize the benefit of public research investments.

These two steps- making data management plans part of the scored reviewer criteria and making them publicly available- will increase the quality and accountability of data-generating research.

Email: david@cos.io

I am responding to this RFI: On behalf of an organization

Name: Karen Caputo

Name of Organization: Case Western Reserve University Libraries

Type of Organization: University

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Investigators would benefit from education and promotion of PMC manuscript submission (Green OA/repository deposit) since it eliminates financial barriers to complying with NIH's policy. When working with researchers to deposit manuscripts into our institutional repository, we often have to educate researchers on manuscript versions, publisher policies, and the differences between OA publishing and repository deposit. Some assume that once they publish their article they cannot share any version of it or can only do so if they publish open access and pay an APC. Many are happy to learn that they can publish in their preferred journals and still make their AAM available in the repository. NIH should be explicit in stating that there is no charge for manuscript deposit in PMC, and any charges from a publisher are for publishing with that journal not for complying with NIH's policy. In our resources for federally funded researchers, we are encouraging repository deposit first to comply with these policies, but it would help to have that reiterated by funders themselves.

In working with researchers at our institution, many are not aware that they can ask to retain rights, so NIH's proposal to offer rights retention language to investigators will greatly help investigators with PMC manuscript submission. In addition to the language though, specific instructions and resources on rights retention would help investigators navigate the process and understand why it is important. Our institution is considering passing a rights retention policy (Faculty OA Policy), but many researchers are confused by this process and need more explanation on how rights retention works.

NIH should also encourage investigators to consider publishing options that do not charge for publishing, such as open access journals that do not charge APCs (Diamond/Platinum Journals). Studies have found that APC costs disproportionately affect early career researchers, female researchers, and researchers from less well resourced institutions. We encourage our researchers to consider the free publishing options available to them, but many are still unaware these options exist. These options eliminate financial barriers for researchers and support more equitable publishing models.

Recently, our institution joined HELIOS. NIH should consider working with a group like HELIOS to encourage incentives for investigators who comply with public access policies.

Finally, NIH might consider providing academic libraries and research offices with tools to help researchers comply with this policy. NIH could provide academic research offices with targeted language on steps to consider to comply with this policy that they could include in their instructions to investigators when applying for and fulfilling NIH grants. NIH might also consider providing grant applicants with a list of NIH designated repositories.

### 2. Steps for improving equity in access and accessibility of publications.

Again, rights retention language and support would improve access and accessibility of publications. NIH should ensure investigators are retaining the right to make their final peer-reviewed manuscripts freely available and also reusable. Open licenses like Creative Commons licenses should also be considered since they provide free access and reuse rights. Open licenses are easy to understand for both researchers and users, so more users can access and reuse content, and more researchers can provide access to and reuse of their work. Open licenses also allow use of content on assistive devices, as well as enabling text and data mining.

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

NIH should consider publicly tracking APC fees for publications that are the result of NIH supported research whether NIH covered that cost or not. This tracking would increase transparency around these costs, reveal affected communities, and provide an understanding of how these publishing costs are taking funds away from research.

#### 4. Early input on considerations to increase findability and transparency of research.

Repository metadata varies considerably, so support of more standardization across repositories would be helpful and encourage reuse of this metadata. The U.S. Repository Network would be a good partner in this effort. As far as PIDs, NIH should consider commonly accepted external identifiers for researchers, publications, data, grants, etc. that are open and are useful outside of NIH's systems.

Email: karen.caputo@case.edu

I am responding to this RFI: On behalf of an organization

Name: Tayler Williams

Name of Organization: American Medical Informatics Association

Type of Organization: Professional org association

Role: Member of the public

- 1. How to best ensure equity in publication opportunities for NIH-supported investigators.
- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

## **Uploaded File:**

NIH-public-access-plan-comments-GPJ-submitted.pdf

Email: carrie@korrisgroup.com



April 24, 2023

Response to Request for Information (RFI): NIH Plan to Enhance Public Access to the Results of NIH-Supported Research (NIH Public Access Plan); NOT-OD-23-091

The American Medical Informatics Association (AMIA) appreciates the opportunity to comment on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research (NIH Public Access Plan); NOT-OD-23-091. AMIA is the professional home for more than 5,500 informatics professionals, representing frontline clinicians, researchers, and public health experts who bring meaning to data, manage information, and generate new knowledge across the health and healthcare enterprise. As the voice of the nation's biomedical and health informatics professionals, AMIA plays a leading role in advancing health and wellness by moving basic research findings from bench to bedside, and evaluating interventions, innovations and public policy across settings and patient populations.

AMIA offers the following comments for NIH's consideration.

#### **Embargo Period**

The goal of offering faster access to NIH-funded research publications is laudable, however, it is unclear what the impact of the policy proposed by NIH would be on both researchers and journals.

#### Specifically:

- That NIH-funded researchers might be forced to limit publication to journals willing to accept NIH's proposed policy of not allowing any period during which the publisher had exclusive rights to the publication. It is unclear how big a problem this may be, however:
  - a. This could preclude publication in more desired outlets in terms of subject matter and intended audience.
  - b. This could cause publishing delays if researchers need to submit to journals that are not their first choice in terms of intended audience, possibly reducing their

- chance of acceptance, which paradoxically could cause even longer delays until the results are publicly available.
- c. If researchers make new discoveries based on data generated under NIH-funding, but after such funding has ended, does this still apply to them?
- 2. That some journals may be adversely affected financially if they cannot recoup their costs based on subscription/membership fees. In the case of niche or highly specialized journals the impact of this loss of income could be particularly acute. While not all research is NIH-funded, a great deal of academic research is.

#### **Authors' Freedom to Choose**

#### NIH states:

As noted previously in this Plan, the NIH Public Access Policy does not affect authors' freedom to choose the vehicle or venue for publishing their results. NIH does not propose requiring authors to publish in any particular type of journal or journal with any specific type of business model (e.g., subscription model, open access model). NIH expects that NIH-supported investigators will continue to publish the results of their research consistent with their professional autonomy and judgment to advance science as efficiently and comprehensively as possible. As previously stated through this <u>Guide Notice</u>, NIH encourages authors to publish in reputable journals that follow accepted standards of publishing practices and ethics.

AMIA believes the above statement is only accurate if the chosen journal allows immediate public access. Authors may be forced into suboptimal choices when the ideal journal(s) enforce a strict embargo."

#### **Supplemental Materials**

The NIH proposal needs to be clearer about the relationship between supplemental materials and manuscripts.

## **Investigator Rights**

AMIA is concerned about the aggressive timeline for this proposal. NIH states it will 'develop language that NIH-supported investigators may use for submission with their peer-reviewed manuscripts to journals to retain rights to make the peer-reviewed manuscript available post-publication in PMC as soon as processing is complete, without an embargo period.'

Given that all federal agencies must implement the OSTP open access proposal no later than December 31, 2025, is there a timeline for NIH to develop this language for investigators?

#### **NIH RFI Questions**

1. How to best ensure equity in publication opportunities for NIH-supported investigators. The NIH Public Access Plan aims to maintain the existing broad discretion for researchers and authors to choose how and where to publish their results. Consistent with current practice, the NIH Public Access Plan allows the submission of final published articles to PMC (in cases where a formal agreement is in place) to minimize the compliance burden on NIH-supported researchers and also maintains the flexibility of NIH-supported researchers to submit the final peer-reviewed manuscript. These submission routes are allowed regardless of whether or not the journal uses an open access model, a subscription model of publishing, or other publication model. This flexibility aims to protect against concerns that have been raised about certain publishing models potentially disadvantaging early career researchers and researchers from limited-resourced institutions or under-represented groups. NIH policy already allows supported researchers to charge reasonable publishing costs against their awards. NIH seeks information on additional steps it might consider taking to ensure that proposed changes to implementation of the NIH Public Access Policy do not create new inequities in publishing opportunities or reinforce existing ones.

AMIA Response: Overall, it would be beneficial to have more of a cost-benefit analysis to justify changes that could have a negative impact on authors and on the speed with which new research might actually reach the optimal audience.

2. Steps for improving equity in access and accessibility of publications. Removal of the currently allowable 12-month embargo period for NIH-supported publications will improve access to these research products for all. As noted in the NIH Public Access Plan, NIH also plans to continue making articles available in human and machine-readable forms to support automated text processing. NIH will also seek ways to improve the accessibility of publications via assistive devices. NIH welcomes input on other steps that could be taken to improve equity in access to publications by diverse communities of users, including researchers, clinicians and public health officials, students and educators, and other members of the public.

AMIA Response: AMIA disagrees with NIH's assumption that "Removal of the currently allowable 12-month embargo period for NIH-supported publications will improve access to these research products for all." NIH defines public access as "free availability of federally funded scholarly materials to the public (including publications, data, and other research outputs). The removal of the embargo period is simply that; there is nothing to suggest this will uniformly and consistently 'improve access to these research products for all.'

3. **Methods for monitoring evolving costs and impacts on affected communities.** NIH proposes to actively monitor trends in publication fees and policies to ensure that they remain reasonable and equitable. NIH seeks information on effective approaches for monitoring trends in publication fees and equity in publication opportunities.

AMIA Response: Monitoring does not ensure that fees and policies "remain reasonable," only that it will be more quickly detected if they do not. Given the potential revenue loss journals may anticipate or experience, this policy could directly incentivize journals to increase publication fees. What is the proposed NIH redress if fees escalate unreasonably, and who would determine what is unreasonable?

4. Early input on considerations to increase findability and transparency of research. Section IV of the NIH Public Access Plan is a first step in developing the NIH's updated plan for PIDs and metadata, which will be submitted to OSTP by December 31, 2024. NIH seeks suggestions on any specific issues that should be considered in efforts to improve use of PIDs and metadata, including information about experiences institutions and researchers have had with adoption of different identifiers.

AMIA Response: More clarity is needed to understand what NIH seeks with regard to "transparency of research." We look forward to more detailed information from NIH and opportunities to comment on future NIH plans for PIDs and metadata.

Thank you for your time and consideration of these comments. If you have questions or require additional information, please contact Tayler Williams, AMIA Public Policy Manager, at <a href="mailto:twilliams@amia.org">twilliams@amia.org</a>

Sincerely,

Gretchen Purcell Jackson, MD, PhD, FACS, FACMI, FAMIA

President and Board Chair, AMIA

Gretche P Jackson

Vice President & Scientific Medical Officer, Intuitive Surgical

Associate Professor of Surgery, Pediatrics, and Biomedical Informatics, Vanderbilt University Medical Center

I am responding to this RFI: On behalf of an organization

Name: Laura Weidner

Name of Organization: Epilepsy Foundation

Type of Organization: Patient advocacy organization

Role: Patient advocate

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Costs to publish in open-access journals are shifted to authors, compared to subscription-based journals. These shifts in cost could result in publishing discrepancies, especially for underfunded and/or underresourced institutions and groups, as well as unestablished early career researchers. Additionally, these discrepancies have downstream effects, including limiting the accessibility and dissemination of research produced by these populations. The NIH should consider the limitations of open access journals faced by underprivileged groups and potential solutions to promote equity across publishing. Suggestions for improving equity in publication opportunities are as follows.

## Equity

- -The NIH should consider the creation of stakeholder working groups to gain feedback on potential disadvantages and limitations of the proposed policy. These working groups should include, but are not limited to, under-resourced groups, under-represented groups, early career investigators, and students/trainees.
- -Elaboration/clarification on allowable publication costs is necessary. Proposed requirements, stipulations, and exceptions for allowable costs should be presented to the public for feedback. It is important to ensure that the plan promotes equity and does not create unfair limitations on underprivileged groups.
- -The implementation of policies that promote equitable publishing opportunities should be considered. Potential examples include fee waivers, voucher programs, and/or discounts for under-resourced and under-privileged groups.
- -Some institutions have already entered into agreements with publishers that subsidize or even fully cover open access fees. This could have a big impact on institutions with a smaller institutional funding base and/or lead to a smaller investment in libraries. The potential impact should be monitored to ensure equity between institutions with varied resources.

#### **Preprint Servers**

-An additional pathway to increase publishing equity could be NIH support for preprint servers, as they encourage feedback, allow for rapid publication, and increase audience reach. Support for preprint servers could be done by generating discussions with publishers regarding the potential elimination of preprint restrictions. For example, ensuring unformatted pre-editorial papers are deposited in the NIH's PubMed Central repository is one effective way NIH increases access to NIH sponsored research.

#### **Data Collection**

- -Inequities and barriers in publishing opportunities that may arise from the updated policy should be monitored and publicly reported, perhaps via surveys. Potential variables of interest could include publication tracking (under-resourced/under-represented groups) and accessibility, usability, and compliance as they relate to the PMC platform.
- -NIH should consider supporting the ability to directly link published papers with publicly available data, and should encourage academic institutions to place a high value on published data sets when considering faculty for promotion.

## 2. Steps for improving equity in access and accessibility of publications.

#### Accessibility

- -NIH should ensure compliance with Section 508/Web Content Accessibility Guidelines (WCAG) by the Web Accessibility Initiative (WAI) for the PMC platform to ensure publication accessibility for all. The Foundation works in concert with Other disability organizations to ensure accessibility of all websites. For people with disabilities, accessibility of websites is a is a civil right necessary for equal opportunity. Accessibility of online information is not limited to those with sensory disabilities; many individuals with Other disabilities, such as those who use augmentative and alternative communication devices, those with intellectual and developmental disabilities, and many more find that they are unable to access online systems that are integral to modern daily life.
- -Potential methods for providing publications in multiple languages on the PMC platform should be considered and develop a plan for implementation. Multilingual options increase accessibility to science and research for those not fluent in English.

#### Training and Education

-The NIH should provide training and education on accessing publications. Potential education areas of interest include how to access and use research products, best practice on how to search and find articles of interest, and a research article overview (i.e., the different sections, what's included in each section, where to find what information, how to "read through" the science).

#### Stakeholder engagement

- -Non-profit organizations, patient societies, and community stakeholders are all involved in providing research to the public, yet often cannot afford the institutional subscription fees to access the latest scientific findings. Feedback from these entities would offer valuable insight regarding unforeseen or unexpected barriers to access.
- -It is also important to note that these stakeholders face a paywall in regard to accessing research. As a result, these organizations, who are usually the bridge between science and families, are unable to share relevant information with their communities. To combat these limitations, NIH incentivize journals to provide open access options that allow non-profits access to research articles and reviews.
- -Funding agency collaboration would promote discussions on best practices, increase equity and accessibility efforts, and encourage joint initiatives.

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

#### **Data Collection**

- -The NIH should consider stakeholder (i.e., publishers, researchers, institutions, non-profits) surveys that ask about factors that may affect equity in publishing opportunities (i.e., publication fees, open access policies/impact of publication models, paywall limitations/article access costs for non-academic organizations). Additionally, longitudinal surveys would offer the opportunity to examine trends and changes over time, which could be useful for future policy updates.
- -A publication cost analysis would identify the different components that make up publication fees. We recommend creating a publicly available report of the results to increase transparency. Such findings could also promote ideas or provide direction to the NIH on how to support researchers in the mitigation of those costs.
- -Pre-post data collection, via publisher collaboration and/or publicly available data, on publication fees and policies could provide insight on changes implemented as a result of the updated policy. Data collection would also promote consistency and transparency and could include annual or bi-annual public reports.
- -Transparency and consistency among cost and impact analyses will be important. One option to promote effective and equitable monitoring could be an NIH developed open access data analysis tool.

## 4. Early input on considerations to increase findability and transparency of research.

N/A - no response on this section.

#### **Uploaded File:**

EF-Public-Access-RFI-Comments-Intro.pdf

Email: lweidner@efa.org



The Epilepsy Foundation (hereinafter "the Foundation") is the leading national voluntary health organization that speaks on behalf of at least 3.4 million Americans with epilepsy and seizures. We foster the wellbeing of children and adults affected by seizures through research programs, educational activities, advocacy, and direct services. Epilepsy is a disease or disorder of the brain which causes reoccurring seizures affecting a variety of mental and physical functions. It is a spectrum disease comprised of many diagnoses and an ever-growing number of rare epilepsies. There are many different types of seizures and varying levels of seizure control.

Approximately 1 in 26 Americans will develop epilepsy at some point in their lifetime. As an organization committed to overcoming the challenges of living with epilepsy and accelerating therapies to stop seizures, find cures and save lives, the Foundation is keenly interested in understanding and educating the epilepsy community about the most current and relevant data pertaining to the epilepsies. We therefore thank the National Institutes of Health (NIH) for developing this proposed approach for improving public access to scholarly publications and data resulting from federally supported research and appreciate the opportunity to provide input.

In addition to our targeted recommendations for Section III of the Public Access Plan, the Foundation has two general suggestions that can be applied to all aspects of the proposed policy. First, we applaud the NIH on its commitment to providing timely educational materials to the public. Aligning with NIH standards on communicating with the public, we encourage the production of training and educational materials regarding the updated policy. Examples of training and educational materials include, but are not limited to, webinars, workshops, conferences, and mentorship programs. Appropriate training and widespread dissemination of the updated policy will be critical for researchers, organizations, and publishers alike. Secondly, the Foundation strongly recommends incorporating the variety of perspectives that will be provided by this RFI into the decision- making process. Feedback can not only assist in identifying potential issues and concerns, but can also provide a deeper understanding of the intersection and interactions among diverse stakeholders. In short, implementation of these two suggestions would improve and strengthen the policy development process.

Thank you again for the opportunity to comment. If you have any questions, please contact Vice President of Government Relations & Advocacy Laura Weidner at <a href="mailto:lweidner@efa.org">lweidner@efa.org</a>.

National Headquarters

3540 Crain Highway, Suite 675, Bowie, MD 20716

301.459.3700

I am responding to this RFI: On behalf of an organization

Name: Chris Bourg

Name of Organization: Massachusetts Institute of Technology - MIT Libraries

Type of Organization: University

Role: Institutional official

- 1. How to best ensure equity in publication opportunities for NIH-supported investigators.
- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

## **Uploaded File:**

RFI\_NIH\_plan-to-enhance-public-access\_20230423\_MITlibraries.pdf

Email: nurnberg@mit.edu

## Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research

Notice Number: NOT-OD-23-091

Name: Chris Bourg

Name of Organization: Massachusetts Institute of Technology - MIT Libraries

Type of Organization \*: University

Role: Institutional Official

NIH seeks information regarding the NIH Public Access Plan, from all interested individuals and communities, including, but not limited to, authors, investigators, research institutions, libraries, scholarly publishers, scientific societies, healthcare providers, patients, students, educators, research participants, and other members of the public. While comments are welcome on all elements of the NIH Public Access Plan, input would be most welcome on Section III related to scholarly publications and on the particular issues identified below. Comments may be entered below or attached in the next section.

1. How to best ensure equity in publication opportunities for NIH-supported investigators.

The NIH Public Access Plan aims to maintain the existing broad discretion for researchers and authors to choose how and where to publish their results. Consistent with current practice, the NIH Public Access Plan allows the submission of final published articles to PubMed Central (PMC) (in cases where a formal agreement is in place) to minimize the compliance burden on NIH-supported researchers and also maintains the flexibility of NIH-supported researchers to submit the final peer-reviewed manuscript. NIH seeks information on additional steps it might consider taking to ensure that proposed changes to implementation of the NIH Public Access Policy do not create new inequities in publishing opportunities or reinforce existing ones.

 The current incentive structures in science, higher education, and their influence on scholarly publishing more broadly, continue to be the major source of the inequities that manifest in publication opportunities for NIH-supported investigators. Across the globe there are groups working to address this important issue such as the Higher Education Leadership Initiative for Open Scholarship (HELIOS), the Declaration on Research Assessment (DORA), GRC research assessment working group, The Latin American Forum on Research Assessment (FOLEC-CLACSO), the UKRI Future research assessment program, the Dutch initiative "Room for everyone's talent", HuMetrics HSS, the Research Data Alliance (RDA) Evaluation of Research Interest Group, and multiple initiatives under the European Research Area policy agenda, as reforming research assessment is one of their 20 priority actions. The current inequity in the scholarly communications system, either in the subscription model, which blocks access to publications, or the APC model, which blocks access to publishing, is driven by incentive structures that reward the "publish or perish" mentality and reduce the assessment of research to a list of publications valued according to the journal they are published in. This sidelines the assessment of both the quality of the research itself and the actual

components that are created as part of the research process. Therefore, the NIH should amplify the importance of these efforts and seek to encourage the implementation of equitable research assessment practices and career advancement incentives that will increase equity in science and in publication opportunities.

The financial burden placed on investigators by several scholarly publication models is also an important source of inequity in publication opportunities. In most cases, the Article Processing Charge (APC) model has proven prohibitively expensive while providing very poor value for money and diverting funds that could be better applied towards more critical research needs. As Grossman and Brembs (2021) highlight, these costs often include the value of the investigator labor and research already included in their submission and the volunteer labor of editors and peer reviewers. To the extent that can be determined given the lack of cost transparency, the actual publisher value-added services make up a very small portion of the overall costs included in that charge. As stewards of taxpayer monies, the NIH should avoid being charged twice by hybrid journals.

To ensure equity in this area, it is important that the NIH make it clear that there are cost-free paths towards compliance, and that neither researchers nor institutions should feel compelled to purchase their way towards compliance. They should also provide their investigators with language and specific guidance to help ensure that authors retain the rights necessary to make their federally funded, peer-review manuscript freely available and reusable post-publication in PMC – without an embargo period – to be in compliance with the NIH's policy.

Given the inequities inherent in the pay-to-publish model prevalent among the majority of publishers of federally funded research, a shift towards greater reliance on the well-established network of repositories is critical to an equitable implementation of NIH's public access plan. (see <u>U.S. Repository Network's Desirable Characteristics of Digital Publication Repositories</u> and <u>COAR Community Framework for Good Practices in Repositories</u>). The repository path eliminates author-facing financial burdens and reduces the inequities in publication opportunities. In addition, repository services typically include a level of expertise in the curation, discovery, and reusability of scholarly content that is lacking across the commercial publishing landscape.

MIT's institutional repository has enabled 58% of all faculty publications since 2009 to be publicly available. This has been achieved only through MIT resourcing the libraries robustly enough to support monitoring, support, and outreach to faculty about the Institute's policy, and equally resourcing the repository infrastructure and the technical, metadata, and expertise needed. The cost to sustainably support public access to NIH-funded research would be significantly less if it were implemented through repositories compared to the costs of the current commercial scholarly publication models which funnels significant amounts of taxpayer dollars into commercial publishers' profit margins. Supporting repositories in this function would keep the resources within the

research system and would provide the opportunity for more equity in publication opportunities, as repositories do not usually charge authors to deposit their peer-reviewed manuscript.

2. Steps for improving equity in access and accessibility of publications.

Removal of the currently allowable 12-month embargo period for NIH-supported publications will improve access to these research products for all. As noted in the NIH Public Access Plan, NIH also plans to continue making articles available in human and machine-readable forms to support automated text processing. NIH will also seek ways to improve the accessibility of publications via assistive devices. NIH welcomes input on other steps that could be taken to improve equity in access to publications by diverse communities of users, including researchers, clinicians and public health officials, students and educators, and other members of the public.

- Reproducibility is an equity issue. All of the necessary components to reproduce an
  experiment or a study need to be accessible in meaningful ways in order to ensure
  equity of opportunity to contribute in a field. Equitable, immediate access to a publication
  is a good but insufficient step to achieve the NIH's and the OSTP stated goals.
  Resources should be allocated for infrastructural support at a systematic level to be able
  to communicate, discover, and maintain the individual components of research in
  appropriate ways for those components.
- CC BY or similar licenses should be used to ensure that legal access for adaptation for accessibility concerns is permitted from the beginning. This kind of licensing would also permit the content to be translated into other languages which increases the potential audience and impact of NIH-funded research.
- While we applaud NIH's commitment to improving guidance for submitters on supplying more human-accessible content, it is important to note that such guidance should be sufficient to ensure that the full research product (text, figures, tables, scientific notations, etc.) is accessible to minimize the dependence on consumer accessibility remediation that may be difficult or limiting due to, for example, missing contextualization. As Brinn et al. (2022) note, publication accessibility too often falls short, and the NIH's guidance needs to be comprehensive and reflective of current and evolving accessibility approaches for publications.
- As mentioned above, ensuring that NIH investigators keep the rights necessary to make their final, peer-reviewed manuscript freely available and fully reusable post-publication in PMC without an embargo period is a critical step to achieve the NIH's stated goal. At the same time, it is important to require licenses that permit computational access of these publications for further research purposes.

3. Methods for monitoring evolving costs and impacts on affected communities.

NIH proposes to actively monitor trends in publication fees and policies to ensure that they remain reasonable and equitable. NIH seeks information on effective approaches for monitoring trends in publication fees and equity in publication opportunities.

Listed publication fees, publish & subscribe agreements, and fee waivers have a substantial influence on publication opportunities in science. Evidence suggests that opportunities to publish systematically vary by race, gender of author, and characteristics of home institution. Measurement and evaluation, however, is currently obstructed by a lack of systematic open information about publication fees and author characteristics<sup>1</sup>

NIH can play a central role in addressing these gaps through specific, practical, and systematic actions. These include:

- requiring that NIH-supported publications include standard metadata documenting the standard/list APC used for that paper, the actual APC charged, and, if different, the amount of difference due to individual waiver and/or institutional agreement (e.g., publish and subscribe); and
- integrating demographic information NIH collects on awardees with data collected on publications to produce systematic public statistics on the distribution of publication fees for NIH outputs.

NIH efforts should be inclusive of research output beyond publication to include data management, sharing, and curation costs. To this end, we recommend that NIH:

- support research into data sharing costs potentially utilizing data acquired from NIHDMSP – to increase understanding and estimates of costs and to determine factors that create equity hardships;
- require archives to report charges for publicly archived data (parallel to reporting APCS);
- analyze budgets from funded grants to inform understanding of data-sharing cost trends;
   and
- refer to those disciplines that have been sharing data effectively and efficiently for many years, and adopt and adapt the practices that are working in those disciplines.
- 4. Early input on considerations to increase findability and transparency of research.

Section IV of the NIH Public Access Plan is a first step in developing the NIH's updated plan for persistent identifiers (PIDs) and metadata, which will be submitted to OSTP by December 31, 2024. NIH seeks suggestions on any specific issues that should be considered in efforts to improve use of PIDs and metadata, including information about experiences institutions and researchers have had with adoption of different identifiers.

<sup>&</sup>lt;sup>1</sup> See for a discussion Altman, Micah. "Designing Community Tracking Indicators for Open and Inclusive Scholarship." *Proceedings of the Association for Information Science and Technology* 59, no. 1 (2022): 393-397.

There is broad consensus in the scientific community that implementations of PIDs and metadata describing research inputs and outputs should be sufficiently robust to large scale (machine analysis, accountability and, where applicable, reuse of the referenced content).<sup>2</sup> Best practices in this area requires that selected PIDs' protocols and metadata formats are community-based and openly documented; each PID can be persistently, publicly, globally resolved to machine-actionable metadata; and the accompanying metadata provides sufficient information to enable direct machine and human access to the content of the described outputs (for authorized users).

To be consistent with these principles and practices, all NIH awards, awardees, and outputs should be associated with PIDs and metadata. Thus, NIH awards and any outputs from them (including preprints, publications, data, and software) should be associated with PIDs and metadata sufficient to: (a) locate its content and determine its accessibility; (b) link each output to any supporting NIH awards (and vice-versa); (c) link each individual NIH-supported publication to preprint versions, supporting data, and supporting software.

For this purpose we recommend that NIH adopt practices already used in the community for identifying and citing scientific publications, datasets and software. NIH should consider specific application of DOI's or the use of RAiDs (Research Activity Identifiers) for its awards and the research activities therein. These identifiers should resolve to standardized machine-actionable metadata, as per Starr et al. (2015). Research outputs referenced in awards' reporting should be associated with Funder Registry metadata and ORCID identifiers documenting all contributors, and include citation metadata. These outputs should include relevant data, research software, unique script or code, and other materials necessary to understand, validate, and support the research findings associated with the award.

\_

<sup>&</sup>lt;sup>2</sup> See for applications to data and software; Altman, Micah, Christine Borgman, Mercè Crosas, and Maryann Matone. "An introduction to the joint principles for data citation." *Bulletin of the Association for Information Science and Technology* 41, no. 3 (2015): 43-45.; Smith AM, Katz DS, Niemeyer KE, FORCE11 Software Citation Working Group. (2016) Software Citation Principles. PeerJ Computer Science 2:e86.DOI: 10.7717/peerj-cs.86; Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg et al. "The FAIR Guiding Principles for scientific data management and stewardship." *Scientific data* 3, no. 1 (2016): 1-9.

I am responding to this RFI: On behalf of an organization

Name: Claire Redhead

Name of Organization: OASPA

Type of Organization: Other

Type of Organization-Other: Open Access Scholarly Publishing Association

Role: Institutional official

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Opportunities to publish open access are not equitable at the moment, and OASPA believes that policy can help with this.

OASPA has studied the OA market, and is concerned about consolidation and lack of diversity in the way in which OA is being achieved. In a separate study across well over 4 million OA articles published over the last 12 years, we see that in 2022 just 10 publishers accounted for 83% of OASPA members' OA output (as reported here). Market concentration is increasing rapidly - in 2020 these figures were six publishers accounting for 75% of OA output.

Although the NIH policy allows for a number of ways in which to achieve public access, we expect that the NIH would care about this market consolidation as there are both equity concerns as well as a cost element involved. A reasonable proportion of NIH funded work is published via the Gold-OA route, which our research on the OA market tells us would primarily, therefore, be via APC payments or transformative agreements. As we argue in the following paragraph, there are legitimate concerns that these prevalent (APC) and developing (transformative) models of open access publishing tend to exclude authors of particular career stages, particular genders, and particular institutions in addition to also excluding those from certain world regions.

OASPA notes a raft of evidence and views supporting the problematic nature of the APC, from this 2020 commentary to this 2022 review and this 2022 study stating that open access is leading to closed research. OASPA also notes this 2019 blog post that asserts "unfairness lies at the core of the APC problem". This 2020 study examining content published by US-based researchers between 2014 and 2018 in over 25,000 academic journals reveals that, in general, the likelihood for a scholar to author an APC-OA article "increases with male gender, employment at a prestigious institution, association with a STEM discipline, greater federal research funding, and more advanced career stage (i.e., higher professorial rank)."

The APC is most often the 'basic unit' used to compute and derive terms around newer 'transformative' deals which increase access to OA publishing for researchers at select (mostly the best-resourced) institutions.

A predominance of these APC and 'transformative' routes to OA would have negative impacts for equity. However, it should be recognised here that for many publishers these routes are the only reliable means to achieve open access. Funding for Other (more equitable) models that could be adopted is as yet not

well established. This needs attention and structural support to enable the move to more equitable routes of immediate open access that also allow for the widest possible reuse.

More on this topic is available in OASPA's blog reporting from our first 'Equity in OA' workshop held on 7 March 2023 which brought together publisher, librarian and funder participants from a wide range of countries including the USA.

Given that 47% of articles received into PMC are via publisher-deposits from some 3000 journals [as per ref 16 here] and also given that the NIH wishes to keep a handle on costs, although the NIH policy is focused on public access, OASPA is convinced that making OA better, and most importantly, increasing equity and diversity in the routes to OA, will help the NIH's aims around equitable public access and increase publication opportunities for NIH-supported investigators.

#### OASPA's suggestions are:

Push for more responsible practice and support reasonable publishing costs

How specifically publication venues/journals could be judged as delivering equitable open access has not yet crystallized, but work that OASPA and Others are doing in this area will reveal more answers in the coming year or so. Specifically, OASPA is aiming to develop ways in which publishing organizations can demonstrate their portfolios' adherence to principles around equity and organizational efforts to increase inclusion and equity so that there is better and greater access to participation in OA publishing.

The NIH already has a plan to develop more details and supplementary information around this space. The NIH could potentially consider adding a condition around journals' adherence to certain principles of equity in achieving OA of the final published version of articles. We would love to work with federal agencies on this and/or provide inputs from the work we are doing in this area if seen to be relevant. OASPA is working on Equity in OA in parallel with the library and funder communities that are also taking steps to define such principles. See also the response to point (3) below involving a future 'Equity in OA' OASPA workshop.

Change the language around "reputable" journals mentioned in III.D.1. This word is tied up with a current, perverse, research-assessment and incentives culture. It also is a barrier for the establishment of new models and the experimentation and innovation that is needed for open access to be more broadly adopted. This language can reinforce an unnecessary drive towards higher cost OA-publication venues and greater market consolidation across publishing venues that have greater brand presence where a diversity of publication venues exist and more cost-effective routes would suffice. The very nature of 'reputable journals' is, in itself, a major contributing factor of the exclusionary research culture that is prevalent today.

Stay in touch and/or work with OASPA and Other stakeholders building equity in scholarly communications. OASPA's recent work has revealed that differentiated pricing on the basis of the ability to pay and automation in discounting and waiver practices for Gold-OA publications (that rely on APCs) would be helpful as short-term fixes in addressing equity issues. There are as yet no bases for such pricing mechanisms in the scholarly publishing industry, but potential solutions were aired and discussed last month as part of OASPA's Equity in OA workshop series. OASPA is a proven convener of stakeholders for constructive conversation and is keen to work with the NIH and Other funders and agencies to continue to develop models and solutions that foster equity in open access publishing.

(links to support the points made above can be found in the uploaded version of our response)

## 2. Steps for improving equity in access and accessibility of publications.

OASPA believes that when we focus more (or solely) on access than reuse then we all stand to lose out on the full benefits of both public access and open access. Using the paywalled/subscription route with zero-embargo deposits to PMC removes a cost barrier and broadens participation, but it's important to make sure that discoverability and re-use are maximized. Accessibility and equity should also be about making content as useful to the public as possible, and to achieve its full potential that content needs to be reusable. By also including strong requirements for PIDs and metadata, visibility of published outputs can be widened.

To be truly equitable and inclusive, and to support the broadest possible human engagement (to sit alongside machine-readability and mining as well), the sharing and re-usability of outputs needs to be more specifically supported.

The NIH policy already says: "NIH will continue to promote the broadest possible reuse of its supported articles, while limiting inappropriate uses, such as redistribution of PMC content for sale." This could go further to specify that re-use licensing on deposited accepted manuscripts (AMs) and published articles should specifically articulate and facilitate appropriate reuse.

(links to support the points made above can be found in the uploaded version of our response)

## 3. Methods for monitoring evolving costs and impacts on affected communities.

In the publishing sector today, the APC is able to 'pull in' research-funder investment (albeit in the US these are often via convoluted routes, with APC monies nested in research grants or only available through trade offs - more on this within the survey findings from an October 2022 report from the American Academy for the Advancement of Science (AAAS). Nonetheless, awareness of funding requirements for APCs is established. However, there are scant (no?) routes that OASPA is aware of for equivalent support for models involving collective action, S2O or diamond routes all of which deliver OA with no researcher-facing fees for reading or publishing.

OASPA recommends greater normalization of investment for these additional routes to support a more equitable form of OA enabling greater participation. We believe that this will drastically alleviate the impacts on affected communities because: (1) with additional funder support available, more US institutions and librarians may find it easier to repurpose existing spends from paywalled to OA titles that rely on collective action and or diamond routes; (2) fewer NIH-supported researchers will see or need to deal with invoices at the individual article level.

While the NIH (and the OSTP policy) is clear that it is model-agnostic, failure to provide support (through policy and funding) for more equitable OA models such as collective action and Diamond OA will only serve to entrench the currently dominant modes of Gold OA publication (via APCs and transformative agreements) that are inherently inequitable as argued above.

OASPA will be holding future 'Equity in OA' workshops in June 2023 where we hope that multistakeholder conversations around shared principles for equity in OA agreements can be developed. We aspire to next-generation agreements and publishing practices with equity and inclusion central to their conception devised to help secure and establish equity in OA regardless of business model. It's important though to note here - as we have covered in points above - that costs are not the only barrier preventing researchers from contributing and so Other factors should be addressed alongside, such as format, language, incentives, assessment, and the notions of quality and prestige.

(links to support the points made above can be found in the uploaded version of our response)

#### 4. Early input on considerations to increase findability and transparency of research.

OASPA's suggestion is to ask that this behavior in the community of scholars is specifically and actively rewarded. The NIH has the opportunity to help build credit and benefits for those researchers who deposit data and follow open-publishing practices. An additional option would be to consider making open access and open data prerequisites for grant funding. OASPA would welcome participation from the NIH in work with institutions to build rewards and incentives for open practices into career evaluations.

This also links to better uptake of PIDs and usage of metadata, both of which contribute to the findability and transparency of research. OASPA is actively involved in initiatives which are focussed on implementing more widespread adoption of PIDs and supports the uptake of new identifiers such as ROR. The OA Switchboard, a community-led initiative founded by OASPA, is also helping to increase PIDs and participation provides a practical mechanism for improving publisher metadata. There is a timely opportunity for all of us to collaborate.

Unsurprisingly, OASPA advocates for as much openness as possible throughout the whole publication process and for all components, including citations and abstracts. OASPA has been a supporter of I4OC and I4OA since they were founded. We encourage data sharing, under FAIR principles, and actively support our members regarding data citation.

Findability and transparency of research is also directly linked to research integrity and is a key area of OASPA's work which we have always placed great importance on. It enables the ability to combat all bad actors, not just researchers. Other aspects can also support this, such as having more information available regarding peer review - we encourage NIH to think beyond current practices and to explore open access to Other research outputs connected to publishing, for example peer review reports. Encouraging this through research assessment reform will also help with proliferation of such behaviors throughout the researcher community.

(links to support the points made above can be found in the uploaded version of our response)

#### **Uploaded File:**

OASPA-response-to-NIH-RFI-2023-1.pdf

**Description:** OASPA response to NIH RFI 2023

Email: <a href="mailto:claire.redhead@oaspa.org">claire.redhead@oaspa.org</a>



## **OASPA** response to NIH RFI 2023

This is OASPA's response to the Request for Information based on this <u>policy (with revisions)</u> from the NIH as released on 21 Feb 2023.

OASPA (the <u>Open Access Scholarly Publishing Association</u>) represents a diverse community of organizations engaged in open scholarship and encourages and enables open access as the predominant model for scholarly outputs.

OASPA wishes to ensure that open access is equitable and inclusive and is keen to explore with its publisher members and library stakeholders ways to increase equity in open access publishing. Why? Because the inclusion of all researchers, including authors from developing and transition countries, and indeed from all backgrounds and life stages, is essential for advancing human knowledge and also for a successful transition to open access. Without the development of new and more equitable approaches to open access, we will not benefit from its full potential. This includes:

- Equity in pricing models for all forms of business models, otherwise authors will continue to face financial barriers to participation.
- Being a stakeholder with influence. Without the development of new and more equitable engagement models for all forms of publishing, authors will also continue to face barriers to participation.

Increasing equity is a shared challenge and no single stakeholder, group, country, or region can deliver this alone. Open access is one of the main ways in which public access mandates will be achieved and we believe that the types and modes of open access that OASPA is interested in supporting and promoting will strengthen the NIH's public-access goals, and that our ongoing work may provide insights into the questions raised by the NIH in its 2023 Request for Information. Equity as we see it is not just about having access to research outputs or the ability to reuse them, but also the ability to be able to participate fully and contribute to the global endeavor of research and scholarship.

OASPA supports the NIH policy and the direction it is taking for scholarly communication. We applaud NIH's commitments to advance the use of persistent identifiers (PIDs) and metadata and the wider benefits they will bring, including increasing equity in discoverability of research.

We strongly endorse the NIH's outlined data policy. Publications and the data underlying them are linked and so it is important that these areas are developed in parallel. OASPA's core

interest sits on the publications side, but data is key to the validity, integrity and replicability of published works.

OASPA has always called for immediate open access to scholarly outputs and so we welcome the move to remove embargos from publications - this makes a huge impact on access to scholarly research outputs. The widespread adoption of depositing the accepted manuscript into PMC will provide a catalyst to fully take advantage of the range of business models that are not based on APCs (Article Processing/Publishing Charges) or transformative agreements. Furthermore, the proposal of developing language to support authors in retaining their rights and bring clarity to the submission process for both authors and publishers - as well as clear conditions for reuse of published works - is welcomed. In the detail of this document we emphasize the importance of enabling reuse, and we call for conversation and support from the NIH around developing routes to open access that are more inclusive and equitable.

OASPA is delighted to see that the NIH policy is focussed on increasing immediate access to research. This lies at the heart of open access. OASPA also feels that scholars should not be faced either with barriers to participation or unfair costs.

The Request for Information from the NIH is focussed on these four questions:

- 1. How to best ensure equity in publication opportunities for NIH-supported investigators.
- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

OASPA and the NIH are aligned in not promoting particular business models as there is more than one way to successfully achieve open access. With our <u>recent focus on equity in open access</u>, OASPA is very encouraged to see equity as the focal point of the questions that the NIH is seeking specific responses about. On these we offer a few thoughts:

# 1. How to ensure equity in publication opportunities for NIH-supported investigators?

Opportunities to publish open access are not equitable at the moment, and OASPA believes that policy can help with this.

OASPA has <u>studied the OA market</u>, and is concerned about consolidation and lack of diversity in the way in which OA is being achieved. In a separate study across well over 4 million OA articles published over the last 12 years, we see that in 2022 just 10 publishers accounted for 83% of OASPA members' OA output (as <u>reported here</u>). Market concentration is increasing rapidly - in 2020 these figures were six publishers accounting for 75% of OA output.

Although the NIH policy allows for a number of ways in which to achieve public access, we expect that the NIH would care about this market consolidation as there are both equity concerns as well as a cost element involved. A reasonable proportion of NIH funded work is published via the Gold-OA route, which our research on the OA market tells us would primarily, therefore, be via APC payments or transformative agreements. As we argue in the following paragraph, there are legitimate concerns that these prevalent (APC) and developing (transformative) models of open access publishing tend to exclude authors of particular career stages, particular genders, and particular institutions in addition to also excluding those from certain world regions.

OASPA notes a raft of evidence and views supporting the problematic nature of the APC, from this 2020 commentary to this 2022 review and this 2022 study stating that open access is leading to closed research. OASPA also notes this 2019 blog post that asserts "unfairness lies at the core of the APC problem". This 2020 study examining content published by US-based researchers between 2014 and 2018 in over 25,000 academic journals reveals that, in general, the likelihood for a scholar to author an APC-OA article "increases with male gender, employment at a prestigious institution, association with a STEM discipline, greater federal research funding, and more advanced career stage (i.e., higher professorial rank)."

The APC is most often the 'basic unit' used to compute and derive terms around newer 'transformative' deals which increase access to OA publishing for researchers at select (mostly the best-resourced) institutions.

A predominance of these APC and 'transformative' routes to OA would have negative impacts for equity. However, it should be recognised here that for many publishers these routes are the only reliable means to achieve open access. Funding for other (more equitable) models that could be adopted is as yet not well established. This needs attention and structural support to enable the move to more equitable routes of immediate open access that also allow for the widest possible reuse.

More on this topic is available in OASPA's blog reporting from our <u>first 'Equity in OA'</u> <u>workshop</u> held on 7 March 2023 which brought together publisher, librarian and funder participants from a wide range of countries including the USA.

Given that 47% of articles received into PMC are via publisher-deposits from some 3000 journals [as per ref 16 <a href="here">here</a>] and also given that the NIH wishes to keep a handle on costs, although the NIH policy is focused on public access, OASPA is convinced that making OA better, and most importantly, increasing equity and diversity in the routes to OA, will help the NIH's aims around equitable public access and increase publication opportunities for NIH-supported investigators.

OASPA's suggestions are:

#### Push for more responsible practice and support reasonable publishing costs

How specifically publication venues/journals could be judged as delivering equitable open access has not yet crystallized, but work that OASPA and others are doing in this area will reveal more answers in the coming year or so. Specifically, OASPA is aiming to develop ways in which publishing organizations can demonstrate their portfolios' adherence to principles around equity and organizational efforts to increase inclusion and equity so that there is better and greater access to participation in OA publishing.

The NIH already has a plan to develop more details and supplementary information around this space. The NIH could potentially consider adding a condition around journals' adherence to certain principles of equity in achieving OA of the final published version of articles. We would love to work with federal agencies on this and/or provide inputs from the work we are doing in this area if seen to be relevant. OASPA is working on Equity in OA in parallel with the library and funder communities that are also taking steps to define such principles. See also the response to point (3) below involving a future 'Equity in OA' OASPA workshop.

- Change the language around "reputable" journals mentioned in III.D.1. This word is tied up with a current, perverse, research-assessment and incentives culture. It also is a barrier for the establishment of new models and the experimentation and innovation that is needed for open access to be more broadly adopted. This language can reinforce an unnecessary drive towards higher cost OA-publication venues and greater market consolidation across publishing venues that have greater brand presence where a diversity of publication venues exist and more cost-effective routes would suffice. The very nature of 'reputable journals' is, in itself, a major contributing factor of the exclusionary research culture that is prevalent today.
- Stay in touch and/or work with OASPA and other stakeholders building equity in scholarly communications. OASPA's recent work has revealed that differentiated pricing on the basis of the ability to pay and automation in discounting and waiver practices for Gold-OA publications (that rely on APCs) would be helpful as short-term fixes in addressing equity issues. There are as yet no bases for such pricing mechanisms in the scholarly publishing industry, but potential solutions were aired and discussed last month as part of OASPA's Equity in OA workshop series. OASPA is a proven convener of stakeholders for constructive conversation and is keen to work with the NIH and other funders and agencies to continue to develop models and solutions that foster equity in open access publishing.
- 2. Steps for improving equity in access and accessibility of publications: OASPA believes that when we focus more (or solely) on access than reuse then we all stand to lose out on the full benefits of both public access and open access. Using the paywalled/subscription route with zero-embargo deposits to PMC removes a cost barrier and broadens participation, but it's important to make sure that discoverability and re-use

are maximized. Accessibility and equity should also be about making content as useful to the public as possible, and to achieve its full potential that content needs to be reusable. By also including strong requirements for PIDs and metadata, visibility of published outputs can be widened.

To be truly equitable and inclusive, and to support the broadest possible human engagement (to sit alongside machine-readability and mining as well), the sharing and re-usability of outputs needs to be more specifically supported.

The NIH policy already says: "NIH will continue to promote the broadest possible reuse of its supported articles, while limiting inappropriate uses, such as redistribution of PMC content for sale." This could go further to specify that re-use licensing on deposited accepted manuscripts (AMs) and published articles should specifically articulate and facilitate appropriate reuse.

### 3. Methods for monitoring evolving costs and impacts on affected communities.

In the publishing sector today, the APC is able to 'pull in' research-funder investment (albeit in the US these are often via convoluted routes, with APC monies nested in research grants or only available through trade offs - more on this within the <u>survey findings from an October 2022 report from the American Academy for the Advancement of Science (AAAS)</u>. Nonetheless, awareness of funding requirements for APCs is established. However, there are scant (no?) routes that OASPA is aware of for equivalent support for models involving collective action, S2O or diamond routes all of which deliver OA with no researcher-facing fees for reading or publishing.

OASPA recommends greater normalization of investment for these additional routes to support a more equitable form of OA enabling greater participation. We believe that this will drastically alleviate the impacts on affected communities because: (1) with additional funder support available, more US institutions and librarians may find it easier to repurpose existing spends from paywalled to OA titles that rely on collective action and or diamond routes; (2) fewer NIH-supported researchers will see or need to deal with invoices at the individual article level.

While the NIH (and the OSTP policy) is clear that it is model-agnostic, failure to provide support (through policy and funding) for more equitable OA models such as collective action and Diamond OA will only serve to entrench the currently dominant modes of Gold OA publication (via APCs and transformative agreements) that are inherently inequitable as argued above.

OASPA will be holding future 'Equity in OA' workshops in June 2023 where we hope that multi-stakeholder conversations around shared principles for equity in OA agreements can be developed. We aspire to next-generation agreements and publishing practices

with equity and inclusion central to their conception devised to help secure and establish equity in OA regardless of business model.

It's important though to note here - as we have covered in points above - that costs are not the only barrier preventing researchers from contributing and so other factors should be addressed alongside, such as format, language, incentives, assessment, and the notions of quality and prestige.

### 4. Input on considerations to increase findability and transparency of research

OASPA's suggestion is to ask that this behavior in the community of scholars is specifically and actively rewarded. The NIH has the opportunity to help build credit and benefits for those researchers who deposit data and follow open-publishing practices. An additional option would be to consider making open access *and* open data prerequisites for grant funding. OASPA would welcome participation from the NIH in work with institutions to build rewards and incentives for open practices into career evaluations.

This also links to better uptake of PIDs and usage of metadata, both of which contribute to the findability and transparency of research. OASPA is actively involved in initiatives which are focussed on implementing more widespread adoption of PIDs and supports the uptake of new identifiers such as ROR. The <a href="OA Switchboard">OA Switchboard</a>, a community-led initiative founded by OASPA, is also helping to increase PIDs and participation provides a practical mechanism for improving publisher metadata. There is a timely opportunity for all of us to collaborate.

Unsurprisingly, OASPA advocates for as much openness as possible throughout the whole publication process and for all components, including citations and abstracts. OASPA has been a supporter of I4OC and I4OA since they were founded. We encourage data sharing, under FAIR principles, and actively support our members regarding data citation.

Findability and transparency of research is also directly linked to research integrity and is a key area of OASPA's work which we have always placed great importance on. It enables the ability to combat all bad actors, not just researchers. Other aspects can also support this, such as having more information available regarding peer review - we encourage NIH to think beyond current practices and to explore open access to other research outputs connected to publishing, for example peer review reports. Encouraging this through research assessment reform will also help with proliferation of such behaviors throughout the researcher community.

In closing, OASPA is enthusiastic about the NIH's goals and welcomes the future trajectory as outlined in this revised policy. OASPA would also like to reiterate the alignment between public access goals of the NIH (and OSTP) and OASPA's views on open access.

Equity in OA is an ongoing area of work at OASPA with future conversations and workshops actively being convened and a post reporting from our second workshop expected in the coming days. Outputs from our work so far are listed below as further reading. We would welcome working with the NIH and other federal agencies to help sculpt pathways to more equitable ways of achieving both open and public access.

Open access is one of the main ways in which public access mandates will be achieved and this is the time to lay down strong foundations for the coming years. If done right, as <u>LaToya E</u> <u>Eaves eloquently argues</u>: "Widespread open access publishing would bring about a more just distribution of knowledge within the United States and globally".

\_\_\_\_\_

### Equity in open access - further reading and resources from OASPA:

Feb 2023 <u>Briefing Document for 'Equity in OA' workshop attendees</u> (including reading list)

March 2023 <u>OASPA's Equity in OA workshop 1 report</u>

March 2023 blog - <u>Reflections from workshop #1 - the APC debate, reflections and rainbows</u>

April 2023 <u>OASPA's Equity in OA workshop 2 report</u>

This open response to NIH is also publicly available on OASPA's website: <a href="https://oaspa.org/oaspa-response-to-nih-rfi-2023/">https://oaspa.org/oaspa-response-to-nih-rfi-2023/</a>

**Submit date:** 4/24/2023

I am responding to this RFI: On behalf of an organization

Name: Jennifer Brogan

Name of Organization: Wolters Kluwer

Type of Organization: Other

**Type of Organization-Other:** Professional Publisher

1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Please see attached.

2. Steps for improving equity in access and accessibility of publications.

Please see attached.

3. Methods for monitoring evolving costs and impacts on affected communities.

Please see attached.

4. Early input on considerations to increase findability and transparency of research.

Please see attached.

#### **Uploaded File:**

Wolters-Kluwer-NIH-RFI-response-04.24.23.pdf

**Description:** Wolters Kluwer Response to National Institutes of Health (NIH) Request for Information on

the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research 04.24.23

Email: jennifer.brogan@wolterskluwer.com

### Wolters Kluwer Response to National Institutes of Health (NIH) Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research 04.24.23

#### Introduction

Thank you for the opportunity to provide feedback on the NIH's draft plan for implementing the White House Office of Science and Technology Policy Nelson Memo for funding agencies to accelerate Public Access to Federally Funded Research output.

Through this response, we are also expressing the concerns of our many professional association partners, including:

American Academy of Neurology

American Association for the Surgery of Trauma

American College of Sports Medicine

American Society for Dermatologic Surgery

American Urological Association

International Anesthesia Research Society

**Wolters Kluwer's Health Learning, Research and Practice division,** headquartered in Philadelphia, is a leading global provider of trusted clinical technology and evidence-based solutions that engage clinicians, patients, researchers, and students with advanced clinical decision support, learning, research, and clinical intelligence. We publish more than 300 society and proprietary journals in both print and electronic formats for healthcare professionals in virtually every specialty, including 70 titles within the 1<sup>st</sup> quartile (Q1) of their JCR (Journal Citation Reports) categories.

We share the objectives of the memo, and as stakeholders, we are perfectly positioned to collaborate with the NIH on the stated goals. A substantial portion of our publishing portfolio is in partnership with some of the most prestigious Professional Associations representing more than **1.2 million** medical, nursing and allied health professionals. A key driver for this response is to represent and advocate on behalf of our association partners. These partnerships are mutually beneficial; WK enables our partners to disseminate their important research at a global level and supports researchers through our precision search available through Ovid. In return, WK provides funds back to these partners, primarily through the subscription-based model so they can invest in more research and further their collective mission of improving patient outcomes.

Beyond our association partners, this response is also guided by our wide breadth of proprietary journals and reflective of the significant role scholarly and clinical resources fulfill in translating research and influencing practice behavior to improve patient outcome. While we share the goal of accelerating discovery and innovation, we must also ensure the continued sustainability of our overall journals portfolio.

We have responded to the specific NIH plan questions and have also outlined concerns to be addressed in order to ensure the success of the new policy.

**1.** How to best ensure equity in publication opportunities for NIH-supported investigators. The NIH seeks information on additional steps it might consider taking in order to ensure that the proposed changes to the implementation of the NIH Public Access Policy do not create new inequities in publishing opportunities or reinforce existing ones.

This is a challenging question as most publishing business models will need to rapidly adapt in order to adhere to the NIH's plan.

Current subscription-based publishing models support author equity by providing equal opportunity for **all** authors to submit and benefit by peer review, editorial oversight, production, and dissemination without charge. The unintended consequence of this policy change, however, will be disadvantaging unfunded authors or authors with limited financial resources, as many established publications will need to rapidly transition to a publishing model substantially supported by Article Processing Charges and other fees that will have to be borne by the authors. The deposit of the final peer-reviewed manuscript in PubMed Central (PMC) remains viable and is the best way to ensure equity in publication opportunities for NIH supported investigators.

Subscription revenue financially underwrites content types beyond original research such as review articles, clinical case studies, clinical trials, guidelines, and clinical content and commentary, as well as article extenders such as supplemental digital content, infographics, and videos. Unfortunately, Librarians representing some of the largest markets for scholarly journals have already viewed the implementation of the NIH's plan as an opportunity to drastically reduce their subscription spend without accounting for the valuable content outside funded research articles. Representatives at the NIH's Listening Session on April 12 specifically discussed this.

Flexibility is key to equity, academic freedom, and ensuring that researchers have the opportunity to best advance their discoveries to support innovation and public health. This includes flexibility in licensing. We believe that the NIH's current policy of recommending that researchers ensure their publishing agreements include the right to provide a copy of the final peer-reviewed manuscript to the NIH upon acceptance for public archiving in PMC has served the public and NLM well and should be retained. Under this new policy, that copy would now be made available without any embargo.

Authors should continue to be able to transfer copyright to the journal owner if required by the journal. Any change that requires researchers to obtain additional rights risks the unnecessary creation of inequity in publication opportunities for NIH-supported investigators. This is because some journals will need exclusive rights to support sustainable business models and continue investments needed for quality, preservation, discoverability, impact, and innovation. Should the NIH adopt a policy whereby authors are required to apply a Creative Commons public copyright license to any Author Accepted Manuscript (AAM) version arising from their submission to a journal, unintended consequences will result. For example, journal owners would likely require NIH-supported investigators to publish the version of record (VoR) open access and pay an APC. Authors not in a position to pay that APC would have limited publication venue options and the NIH policy will again have created the inequity it seeks to avoid.

Again, if the NIH wishes to pursue this type of rights retention without causing inequity, then they could consider entering into financial agreements with publishers to directly cover the fees that will be required to support such a policy. Such agreements could then allow all authors to continue to publish in the publication venue of their choice without direct fees.

In addition, the NIH should recognize the risk of creating these new inequities, especially for scientists from traditionally marginalized communities as well as early career researchers and ensure that these researchers and institutions have the funding support necessary to pay APCs should the researcher choose to publish in a journal that requires such fees.

Implementation of the NIH plan will also attract opportunistic commercial enterprises such as predatory publishers looking to take advantage of researchers and authors. Predatory publishers undermine the editorial process and threaten the validity and credibility of medical content and cause direct harm. We all have a responsibility to educate researchers and authors to ensure that they select the appropriate publication for their work so that they can benefit from the rigorous editorial process resulting in the most impactful paper.

2. Steps for improving equity in access and accessibility of publications. Removal of the currently allowable 12-month embargo period for NIH-supported publications will improve access to these research products for all. NIH welcomes input on other steps that could be taken to improve equity in access to publications by diverse communities of users, including researchers, clinicians and public health officials, students and educators, and other members of the public.

Although the NIH is not promoting one specific publishing business model, the new policy will likely result in Gold open access publishing becoming the de facto model for federally funded research. The existing model that requires deposit of the final peer-reviewed manuscript in PMC following a 12-month embargo is currently underwritten by global subscription, licensing, and advertising revenue. This model, as supported by publishers, has ensured compliance and adherence to the current requirements.

Removing the 12-month embargo as required by the new plan, undermines the ability of the society or publisher to recoup investment in content-related and infrastructure costs—such as stipends for editors; validation of publication research integrity; content recruitment, development, and enrichment through production of ancillary material such as podcasts, infographics, and videos; submission and peer review systems; editorial tools such as plagiarism detection; digital platforms; and dissemination. Furthermore, it erodes the longer-term value of global subscriptions for journals with significant amounts of federally funded content thus impacting the revenue that medical societies use for key endeavors that support the scholarly enterprise, such as funding of educational programs and research, curricula development, professional training, advocacy, and development and dissemination of guidelines that advance medical practice and improve patient care.

There is risk associated with transforming a global subscription model to a transactional model supported by research grants as publishers will need to limit their innovation on other aspects of publishing. Publishers will have to investigate new means of supporting content and infrastructure costs such as increased APCs and service-based charges to authors and a reduction in discount and waiver programs that support authors from underfunded disciplines and regions. This may result in additional fees for previously provided services such as routine manuscript deposits to PMC.

Furthermore, accessibility is not limited to making content publicly accessible. Publishers already invest in ensuring that content is available in accessible formats and disseminated via numerous channels that adapt to the user access needs. Publishers have ensured that articles are accessible in various human and machine-readable formats and are available to those with diverse needs. Publishers have created a diverse ecosystem of accessible resources available to diverse audiences with or without assistive technologies. These additional infrastructure and formatting investments are enabled by sustainable business models.

Therefore, per our response to Question 1, avoiding the unintended consequences of a policy change that requires authors to obtain or retain additional rights, will be important.

**3.** Methods for monitoring evolving costs and impacts on affected communities. The NIH proposes to actively monitor trends in publication fees and policies to ensure that they remain reasonable and equitable. NIH seeks information on effective approaches for monitoring trends in publication fees and equity in publication opportunities.

Provided the NIH policy continues to support a plurality of business models then pricing impacts on affected communities could remain reasonable. However, if the policy results in both a decline in global journal subscriptions and automatic deposit of the final peer-reviewed manuscript to PMC on behalf of authors, the Gold open access model will become the de facto route to publication and will have a cost impact on affected communities. In this case, the NIH may wish to consider entering into agreements directly with journal owners to cover publication charges as suggested in our response to Q1.

Regarding methods for monitoring publication fees, we would recommend the NIH monitor each discipline separately and avoid using average calculations. "Reasonable and equitable" pricing is subjective due to a wide number of factors. Investments in publishing services vary and are often dependent on specialty as well as the nature of the publication. For example, the investments committed to a clinical nursing journal are different from those committed to a neurology research journal, and the pricing required to maintain a highly selective journal that receives a high volume of submissions can be significantly different to pricing required to maintain less selective journals or those with lower volumes of submissions. Prices and revenue streams also vary drastically depending on the audience, circulation/reach, ranking, number of articles published, field/specialty, and distribution method (print vs. online). Furthermore, we recommend the NIH consider legal advice regarding competition and antitrust laws in relation to the influencing of industry market pricing.

Its an unfortunate reality that journals are also dealing with threats to credible scholarly content due to academic misconduct such as peer review manipulation, papermills, data and image fabrication as well as the exponential acceleration of AI (such as ChatGPT) and its ascendance into the manuscript writing process. Providing services that address these issues will likely result in higher prices.

**4.** Early input on considerations to increase findability and transparency of research. The NIH seeks suggestions on any specific issues that should be considered in efforts to improve use of PIDs (Persistent Identifier) and metadata, including information about experiences institutions and researchers have had with adoption of different identifiers.

Current PID and metadata structure is supported by publishers through sponsorship and membership in organizations such as CrossRef and ORCID. Publishers also work with National Information Standards Organization (NISO) to ensure metadata remains current and accessible and are included in the cost to prepare for content dissemination.

#### Additional Feedback on the NIH Plan

#### **Rights retention**

Wolters Kluwer is seeking further detail and clarity on the issue of rights retention. Most journals require authors to transfer or license article rights to the Publisher or Journal owner. Retention of any such rights requires careful consideration to avoid unintended consequences. Journals require exclusive rights to the content published in order to protect and receive a return on the investment in the peer-review and publication process.

Can the NIH provide context for considering RRS (Rights Retention Strategy) and engage with their publisher stakeholders for alternatives? Can the NIH agree that certain reuses are either already allowed or could be allowed without the need of a problematic RRS?

See also our responses in Questions 1, 2, and 3.

### Clarification of Depositing Author Accepted Manuscript and Logistics

How does the NIH expect that manuscripts will be deposited? While this is the author's responsibility, most publishers currently manage the process on behalf of the authors at no cost. Given the impact on the subscription business model, the responsibility for depositing will likely revert to the authors. Alternatively, publishers can continue to deposit, passing the cost onto the authors who select this path.

#### Impact of NIH on Industry

WK supports the objectives of Open Medicine and the NIH's goal of providing the output of federally funded research rapidly so that it can impact health and fuel additional research. We suggest a more measured approach that supports the evolution of the publication models that will continue to underwrite the publishing infrastructure that allows for wide interpretation and dissemination. Publishers are instrumental in funding many initiatives and entities that address research integrity and editorial ethical issues (such as Committee on Publication Ethics (COPE), make content available to underserved global communities (Hinari, Research4Life), and Data Preservation (CLOCKSS, Portico).

As research and publishing have continued to evolve at a rapid pace, commercial publishers have invested millions in platforms and tools that can support a variety of content formats and media, address research needs, enhance editorial and scientific excellence, and satisfy the content consumption expectations of a wide audience of readers, researchers, clinicians, faculty, and students, novice to expert.

**Impact on commercial publishers in partnership with society partners.** Professional organizations partner with commercial publishers for many reasons and financial return is a primary one. For many non-profit associations, publishing revenue is one of the most robust revenue channels for the non-

profit professional association. Not only does that revenue support the editorial operation and staff of the journals, but many associations also rely on this revenue for infrastructure support and funding of member benefits, educational programs and research, curricula development, professional training, advocacy efforts, and development and dissemination of guidelines that advance medical practice and improve patient care.

Risk to the financial support that associations rely on will negatively impact the editorial management of health journals as well as the numerous other educational projects sponsored by those associations. The NIH's new plan would stress this model resulting in fewer association-sponsored research grants, which in turn will ultimately decrease the volume of scholarship and scientific advancement in any given research area. The new plan would also tax the editorial administrative process resulting in a less robust peer review as outlined below.

#### Impact on Editorial Services Including Quality of Peer Review

Currently, the subscription model funds editorial offices, either directly contracted via publishers or through publishers' partnerships with associations, to manage the peer-review process, evaluation and development of content and curation of articles. Without that funding, editorial operations that produce credible, validated, accessible and timely scientific papers will be taxed as budget cuts would force consolidation at the editorial leadership level and reduction of staff at the administrative level. This would result in slower peer review and/or a less rigorous review overall.

Editorial offices and publishers are also addressing numerous other issues. Top among these is scientific and editorial misconduct such as plagiarism, data and image manipulation, conflict of interests, author impersonation or fabrication, papermill output and ethical violations. They are also actively grappling with issues of Equity, Diversity and Inclusion.

**Influx of borderline and predatory publications.** As mentioned in our response to Q1, the new plan invites the proliferation of substandard publications. Historically, journals were evaluated by purchasers and with this market-driven approach, required to demonstrate their value. Publishers were able to provide metrics to illustrate that value, including both usage data and impact factors. Removing the need to prove quality encourages a proliferation of new journals and publications. Those publications are incentivized to accept as many manuscripts as necessary without regard for quality of science or impact.

**Impact on non-funded disciplines.** Because of the scope of research funded by the NIH, the future NIH requirements will be the de facto requirements for all medical publishing. While this will open the market to many questionable publications as outlined above, it may have a negative effect on those disciplines that are now underfunded by NIH. Areas that traditionally receive the same level of funding, such as anesthesiology or ophthalmology, will have limited options as their publications are authored by those without federal funding and therefore no budget to fund APCs.

#### Summary

We appreciate this opportunity to comment and feedback on the NIH Plan. Our response seeks to provide the NIH with guidance on the way forward while highlighting certain unintended consequences

that would be detrimental to the research community, non-profit professional associations, and the scientific process itself, as well as publishers. We remain critically concerned about the potential adverse effects the plan may have on a variety of stakeholders if these concerns are not adequately assessed and addressed. To achieve our mutually held goals for open science and societal transformation, all stakeholders need to commit to work together more closely to identify and address these key challenges thoughtfully. Listening exercises should make way for open discussions.

**Submit date: 4/24/2023** 

I am responding to this RFI: On behalf of an organization

Name: Maria Gould

Name of Organization: Research Organization Registry (ROR)

Type of Organization: Other

Type of Organization-Other: Infrastructure provider

Role: Member of the public

- 1. How to best ensure equity in publication opportunities for NIH-supported investigators.
- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

### **Uploaded File:**

 $\hbox{NIH-Public-Access-RFI-Research-Organization-Registry.} pdf$ 

Email: maria@ror.org



April 24, 2023

Office of The Director National Institutes of Health 9000 Rockville Pike Bethesda, Maryland 20892

RE: ROR Comments in Response to NOT-OD-23-091, "Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research"

To Whom It May Concern:

I write on behalf of the Research Organization Registry (ROR) responding to the Request for Information (RFI) on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research issued on February 21, 2023.

ROR is a global registry of open persistent identifiers for research organizations. ROR IDs are designed to be used in publication metadata and research infrastructure to unambiguously identify the organizations researchers are affiliated with, including their employers and their funders, so that people and systems can reliably connect research outputs to organizations. ROR is operated as a joint initiative by California Digital Library, Crossref, and DataCite, three not-for-profit organizations that have deep ties to research communities as well as extensive experience building and maintaining persistent identifier services and infrastructure.

ROR IDs are being integrated into various systems wherever there is a need to identify organizations and capture affiliation metadata. These implementations reflect the importance of affiliation metadata both upstream in the research and publishing process—i.e., identifying author affiliations upon submission of a manuscript—and in downstream services and systems for discovery and tracking of research, such as Crossref metadata, scholarly indexes and databases, and repositories. ROR is the preferred identifier for use in DOI metadata for publications, datasets, and grants registered in Crossref and DataCite, it is the primary identifier supported in ORCID records for researcher affiliations, and it has been recommended in national PID policies recently announced in Australia, Canada, the Netherlands, and the United Kingdom. ROR is unique among other organization identifiers because it is freely and openly available, specifically focused on connecting research organizations to research outputs, and designed to be used with other persistent identifiers, such as DOIs, ORCID IDs, and Funder Registry IDs.

ROR supports the NIH's interest in incorporating guidance on uses of PIDs and metadata in the Public Access Plan. Our specific comments on this aspect of the RFI notice are provided below:

### 4. Early input on considerations to increase findability and transparency of research.

Section IV of the NIH Public Access Plan is a first step in developing the NIH's updated plan for persistent identifiers (PIDs) and metadata, which will be submitted to OSTP by December 31, 2024. NIH seeks suggestions on any specific issues that should be considered in efforts to improve use of PIDs and metadata, including information about experiences institutions and researchers have had with adoption of different identifiers.

Persistent identifiers are an essential building block of research infrastructure. They facilitate disambiguation, enable discovery and tracking of research, and establish connections that can reveal key insights about how research is being conducted and consumed. While persistent identifiers on their own do provide a fundamentally useful function—unambiguous identification—they can be much more powerful and meaningful when they contain rich metadata and when they are linked to a network of multiple identifiers. This power is only unlocked when the identifiers and their underlying metadata are openly available for anyone to use and reuse, and it becomes especially relevant in computational contexts. ROR is a prime example of this power.

ROR IDs can be beneficial to NIH in many ways, including:

- Disambiguating and normalizing researcher affiliations
- Disambiguating and normalizing funder information
- Discovering and tracking research outputs associated with a specific institution
- Discovering and tracking research outputs associated with a specific funder or award
- Identifying connections between research awards, research funders, research outputs, and research organizations
- Enhancing the machine-readability and overall data quality of publication metadata, which supports accessibility needs as well as computational activities such as text and data mining of publications
- Facilitating the creation of automatic tools that track policy adherence by institutions

In order to realize these benefits, we encourage NIH to consider the following concrete actions:

Require or strongly encourage DOIs for NIH datasets and inclusion of ROR IDs in
DOI metadata registered in DataCite. ROR IDs can be included in DOI metadata for
researcher affiliations, funder information, and publisher information (forthcoming in the
next version of the DataCite metadata schema). This will enhance the discoverability of
NIH data registered in DataCite, and make it possible for downstream discovery services
to use this information to more efficiently track research outputs.

• **Register DOIs for NIH awards** (e.g., via Grant IDs provided by Crossref) and include ROR IDs in the award metadata so that downstream discovery services can use this information to more efficiently track research outputs connected to specific awards.

Require or strongly encourage investigators to obtain a DMP-ID for their Data
 Management and Sharing Plan. This will ensure that the researcher's affiliation is
 automatically captured in the plan metadata in the form of a ROR ID and that the
 metadata about the plan, subsequent award, and resulting research outputs will be made
 publicly available in DataCite and downstream discovery services.

Map PubMedCentral author affiliations to ROR IDs and make this metadata
available in PMC APIs. This will make it possible to create more reliable search and
browse features by author affiliation in PMC, to query PMC deposits to report on
publications associated with specific institutions, and to connect the data to other
indexes by normalizing on ROR IDs.

 Encourage publishers to provide ROR IDs for author affiliations in DOI metadata registered with Crossref. This will accelerate pressure on publishers to make their metadata openly available and will enable more efficient discovery and tracking of research outputs by institution.

Thank you for your consideration of these comments on behalf of ROR. We look forward to continued engagement on this issue.

Sincerely,

Maria Gould, ROR Lead

California Digital Library, University of California Office of the President

**Submit date: 4/24/2023** 

I am responding to this RFI: On behalf of an organization

Name: Mary Lee Kennedy

Name of Organization: Association of Research Libraries

Type of Organization: Professional org association

Role: Institutional official

- 1. How to best ensure equity in publication opportunities for NIH-supported investigators.
- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

### **Uploaded File:**

ARL-comments-NOT-OD-23-091.pdf

**Description:** ARL comments on the NIH draft public access policy.

Email: cvitale@arl.org



April 24, 2023

NIH Office of Science Policy 6705 Rockledge Drive, Suite 630 Bethesda, MD 20892

Re: Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research (NOT-OD-23-091)

On behalf of the members of the Association of Research Libraries (ARL), thank you for the opportunity to provide comments on the National Institutes of Health (NIH) "Plan to Enhance Public Access to the Results of NIH-Supported Research." We applaud NIH for its leadership in public access, specifically its investment in PubMed Central (PMC) and the recently implemented "NIH Policy for Data Management and Sharing." ARL and its members are committed to the advancement of open scholarship and open access to accelerate scientific and medical advances and to expand diverse, public participation in federally funded research. We appreciate NIH's commitment to making the results of federally funded research widely available without embargo, leveraging persistent identifiers to support scientific integrity, and ensuring equitable access.

Decisions made by NIH, one of the world's largest funders of scientific research, will influence the entire scholarly publishing ecosystem, with implications for researchers globally. ARL recommends that NIH consider the far-reaching, global impact of its policy implementation with regard to non-NIH-funded researchers in addition to those funded by NIH.

While the "NIH Plan to Enhance Public Access to the Results of NIH-Supported Research" covers publications, data, and other research outputs, our recommendations below focus primarily on publications. ARL has submitted prior comments¹ and work on behalf of its members with regard to the "NIH Policy for Data Management and Sharing."

<sup>&</sup>lt;sup>1</sup> "ARL Comments on Draft NIH Policy for Data Management and Sharing," Association of Research Libraries, January 9, 2020; "ARL Comments on Draft Genomic Data Management and Sharing Policy," Association of Research Libraries, March 9, 2022; Institutional Strategies for the NIH Data Management and Sharing Policy: Infrastructure, Policies, and Services, Association of Academic Health Science Libraries, Association of American Medical Colleges, Association of Research Libraries, September 2022.

We submit the following comments on the "NIH Plan to Enhance Public Access to the Results of NIH-Supported Research."

### 1. How to best ensure equity in publication opportunities for NIHsupported investigators

The Association of Research Libraries appreciates the framing of "publication opportunities," recognizing both publishing and access to publishing as equity issues, and recommends that NIH:

- Clarify for investigators that there is no charge for manuscript deposit into PMC, and that publishing charges by journals are not public-access compliance fees
- Work with research institutions, their libraries, and their professional associations on coordinated education to investigators on their options for costfree manuscript deposit
- Continue partnerships and experimentations with preprint services (such as the NIH Preprint Pilot) to accelerate sharing of research findings, including assigning PMC IDs to peer-reviewed preprints, and recognizing peer-reviewed preprints that are substantially similar to author-accepted manuscripts for the purposes of compliance with the policy

If a researcher chooses to accept funding from NIH or other federal R&D agencies, they must agree to grant the funding agency a nonexclusive license to their scholarly outputs funded by the grant. In this scenario, the researcher retains their copyright, unless and until they assign it to another party, such as a publisher. According to the August 2022 Nelson memo,² agency policies must describe the prerequisites needed to make publications freely and publicly available by default, including reuse rights and attribution, which has implications for the type of license that the researcher may use. Retaining copyright enables researchers to make those license choices.

#### ARL recommends that NIH:

 Provide rights-retention language (for investigators to use upon submission of manuscripts to journals) that encourages authors to retain their copyrights and assign a Creative Commons Attribution (CC BY) or similar license to their work in order to enable full reuse rights. Open licenses are easy to understand for both researchers and users, so more users can access and reuse content, and more researchers can provide access to and reuse of their work.

<sup>&</sup>lt;sup>2</sup> Alondra Nelson, "<u>Ensuring Free, Immediate, and Equitable Access to Federally Funded Research</u>," US Office of Science and Technology Policy, August 25, 2022.

• Consider using the following language, modeled after the Wellcome Trust language:

This research was funded in whole or in part by the National Institutes of Health [grant number]. For the purpose of public access, the author has applied a CC BY public copyright license to any author-accepted manuscript version arising from this submission.

According to cOAlition S funders, "In the two years or so since this [rights retention] approach was introduced by many cOAlition S funders, [the funders] are only aware of one example where a publisher rejected a manuscript due to the existence of a prior licence."<sup>3</sup>

- Develop a mechanism to ensure that funds are available post-closeout for publication expenses. Post-award publication funding may be particularly important for early-career, postdoctoral, and graduate student researchers whose publication costs may not have been factored into the original grant budget.
- Consider additional supplemental funding or new grant models to support innovative institutional services for investigators in meeting public-access requirements. ARL member institutions and their libraries help investigators navigate the various publishing options, manuscript versions, publisher policies, and the differences between public-access publishing and repository deposit.

# 2. Steps for improving equity in access and accessibility of publications ARL recommends:

- Encouraging open licenses (see above), which allow use of content on assistive devices as well as enabling text and data mining
- Requiring that all deposited manuscripts or final publications meet Web Content Accessibility Guidelines (WCAG) and Section 508 compliance standards, so publications can be properly rendered to assistive technologies

### 3. Methods for monitoring evolving costs and impacts on affected communities

Fully monitoring publication expenses will require looking beyond the grant budget line item for publication costs. Given the different mechanisms for funding publication costs (grant-based, departmental, library funds, and bundled read-and-publish agreements), the single budget line item does not entirely reflect the full range of expenses.

<sup>&</sup>lt;sup>3</sup> "Making Full and Immediate Open Access a Reality," cOAlition S, April 11, 2023.

Association of Research Libraries

#### ARL recommends:

- Surveying researchers and/or institutions at closeout for additional information on publication costs, or commissioning a study that would incorporate both researcher costs and additional data from global registries of article-processing charges (APCs) and other publication fees
- Monitoring publication trends across publication formats, including journal articles, book chapters, and other peer-reviewed publications
- Reviewing the publication costs of journal titles that NIH-supported researchers most commonly publish in

## 4. Early input on considerations to increase findability and transparency of research

The Association of Research Libraries recommends that NIH:

- Adopt the <u>Implementing Effective Data Practices</u> report recommendations from higher education associations, including the adoption of the following persistent identifiers (PIDs) at a minimum:
  - Digital object identifiers (DOIs) for each publication and research output (data, code, software, etc.)
  - Open researcher and contributor identifiers (ORCID IDs) to uniquely identify authors
  - Research Organization Registry (ROR) IDs to link authors with known organizations
  - Crossref Funder Registry IDs to associate a research output with a granting agency
  - Crossref Grant IDs to uniquely identify a research award with an author, an organization, and a funding agency

This report also provided considerations that would help support this necessary PID infrastructure. NIH could lead the following to advance the sharing of research and research data.

 NIH, in coordination and harmonization with other federal agencies, could fund the design and development of tools and services to support the use of PIDs. NIH could fund investigators developing research-related workflows and systems that enable the collection of PIDs, storage of PID

<sup>&</sup>lt;sup>4</sup> Implementing Effective Data Practices: Stakeholder Recommendations for Collaborative Research Support, Association of Research Libraries (ARL), California Digital Library, Association of American Universities (AAU), and Association of Public and Land-grant Universities (APLU), September 23, 2020.

metadata, and connections to PIDs in other systems.

- NIH, in coordination and harmonization with other federal agencies, could invest in infrastructure and initiatives that support the use of PIDs by supporting member organizations that promote open scholarly infrastructure, such as Crossref, DataCite, and ORCID; funding organizations and data repositories that follow best practices for FAIR (findable, accessible, interoperable, and reusable) data; and supporting community-led initiatives such as the Research Organization Registry and DMPTool.
- NIH, in coordination and harmonization with other federal agencies, could minimize the burden on researchers by making it easy and seamless for researchers to use PIDs by designing workflows and systems to assign and collect them automatically and by supporting PID services or data repositories within the PubMed Central platform. Finally, NIH could work with vendors of tools to require them to adopt workflows and software that automatically collect PIDs. This will be especially necessary for less-resourced institutions that may not have research librarians to provide these services.

We look forward to continued engagement with the NIH during the development of the agency's public access plan. We are happy to work with the NIH to identify ARL member institutions to participate in conversations regarding any of these specific topics. Please feel free to contact me or my colleague Cynthia Hudson Vitale, Director of Science Policy and Scholarship, (cvitale@arl.org) with any questions about these comments.

Sincerely,

Mary Lee Kennedy Executive Director **Submit date:** 4/24/2023

I am responding to this RFI: On behalf of an organization

Name: Jessica Sebeok

Name of Organization: Wiley

Type of Organization: Other

Type of Organization-Other: Publishing company

Role: Member of the public

1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Please see attached comments.

2. Steps for improving equity in access and accessibility of publications.

Please see attached comments.

3. Methods for monitoring evolving costs and impacts on affected communities.

Please see attached comments.

4. Early input on considerations to increase findability and transparency of research.

Please see attached comments.

### **Uploaded File:**

Wiley-NIH-RFI-submission\_NOT-OD-23-091\_04\_24\_2023.pdf

Description: Wiley comments on the NIH Plan to Enhance Public Access to the Results of NIH-Supported

Research

Email: jsebeok@wiley.com



April 24, 2023

Lyric Jorgenson, PhD Acting Associate Director for Science Policy NIH Office of Science Policy 6705 Rockledge Drive Bethesda, MD 20892

### RE: Response to NIH Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research

Dear Dr. Jorgenson:

Thank you for the opportunity to share our views on NIH's plan to enhance public access to the results of NIH-supported research. We appreciate this important feedback mechanism and look forward to working with NIH and other stakeholders to deliver meaningful outcomes that advance open science and research. We support the objectives set forth in the memo released by the White House Office of Science and Technology Policy (OSTP) of Ensuring Free, Immediate, and Equitable Access to Federally Funded Research and we hope to work with NIH to ensure scientists have the tools necessary to communicate their research for the advancement of science.

Founded in 1807, Wiley is one of the world's largest publishers and a global leader in research and education. For more than 215 years, we have been enabling discovery, powering education, and shaping workforces. As the nation's largest scientific and scholarly research publisher and the world's leading disciplinary society publishing partner, we are proud to publish nearly 2,000 academic journals which, together, brought more than 280,000 unique pieces of scholarship to the world in 2022. As the publishing partner for numerous scholarly societies in the United States, such as the Federation of American Societies for Experimental Biology, the American Heart Association, the American Cancer Society, the Alzheimer's Association, the American Geriatrics Society, and the Obesity Society, we publish over 13,000 NIH funded articles each year.

America's scientific leadership and competitiveness are supported by a thriving scholarly communication ecosystem of researchers and institutions, public and private. Together we are creating the tools and infrastructure to advance research in the 21st century, and ensuring this system is imbued with the values that underpin the U.S. research community – rigor and integrity; academic freedom; openness; partnership; diversity, equity, and inclusion (DE&I); and respect for innovation, commercialization and intellectual property rights.

We appreciate that the efforts described by NIH are focused on public access. It is our firm belief that in order to be truly effective, any public access policy should promote open access and open science and in doing so should:

- Endorse the final published Version of Record (VoR)<sup>1</sup> as the article format which will deliver the full benefit of open access (OA) to the scientific community;
- Include a federal funding mechanism that recognizes the cost of peer-review, editing, publication, distribution, and long-term stewardship of articles; alleviating the administrative and financial burden of publishing costs from universities, libraries, and individual researchers; and
- Leverage the many services currently provided by publishers to advance discovery and innovation, thereby avoiding a duplication of efforts and investments already made in support of open access and open science.

### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Wiley supports equitable publication opportunities for NIH-supported investigators. As a service to the researchers we work with, we currently facilitate the automatic deposit of the accepted article<sup>2</sup> into PubMed Central (PMC) after the one-year embargo for articles published under the subscription journal model and deposit the VoR into PMC immediately when articles are published under open access.

Ensure all NIH Supported Investigators have the opportunity to publish the most trusted version of their research article open access.

Open access publication of the VoR as an option for all is integral to achieving equity in publication opportunities.

The VoR provides transparent access to all of the publication ethics practices and standards that are applied to the author's manuscript both leading up to and following publication including:

- the names and affiliations of the editors;
- peer review model;
- required protocols and reporting guidelines, e.g. CONSORT, STROBE, ARRIVE; etc.
- conflict of interest policies;
- corrections;
- expressions of concern;
- retraction notices; and
- other research integrity safeguards journals have in place to uphold trust in peer review; trust in research; and ultimately trust in scientific practice.

<sup>&</sup>lt;sup>1</sup> NIH nomenclature refers to the Version of Record (VoR) as the "final published article"

<sup>&</sup>lt;sup>2</sup> NIH nomenclature refers to the accepted article as the "peer-reviewed manuscript"

None of this critical information would be made available on external repositories holding only the accepted articles.

A large percentage of federally funded authors who publish in Wiley journals are based at R1<sup>3</sup> institutions, many of which have Transformative Agreements (TAs) or open access accounts with Wiley. This means their Article Publication Charges (APCs) will be covered under those agreements. Any future policy should ensure that federally funded authors who are not covered under an institutional agreement should also have access to funding that will allow them to publish the VoR open access in order that the final, published, maintained and linked version of their work is available to the widest possible audience.

### Ensure funding is available to support a diverse publishing ecosystem that maximizes author choice.

As the Publisher of a wide range of journals and journal types, from those with highly selective publishing polices to those with more inclusive approaches, we are acutely aware that different journals have different costs and resource requirements. A mechanism to ensure that this variety of journals can continue to deliver the services they provide is vital to the ongoing diversity of the scholarly record. We urge NIH to recognise that all public access business models have costs and require some form of funding to ensure they are sustainable, be that through the subscription model or an open access model. There are no cost-free routes to public access.

Supporting a Green OA route to public access and removing embargoes without providing adequate funding for Gold OA, and alternative OA models, will severely threaten the sustainability of journals and ultimately limit publication choice for NIH funded researchers. The provision of publishing services come at a cost, irrespective of OA model, and without funding to cover these costs, many journals will struggle to remain financially viable meaning that the number of trusted publication venues will decrease, ultimately limiting publication choice for researchers.

Furthermore, it is important to note that some of our partner scholarly societies have expressed concern over the new public access policies because such policies could have negative impacts on the DE&I efforts currently underway in their communities. Revenues from many society journals are essential for the operations, services, and products key to the mission of smaller scholarly societies, including their activities relating to improving DE&I. Any reduction in journal publishing revenues could negatively impact their ability to fulfill their missions and serve the needs of their members and jeopardize the societies' ability to sustain high-quality publishing activities.

<sup>&</sup>lt;sup>3</sup> R1: Doctoral Universities – very high research activity

### Allow NIH Funded Investigators freedom to license their rights as they wish to protect the integrity of their work and publish in their venue of choice.

Agency requirements restricting NIH Funded Investigators' ability to license their rights, for example through a rights retention mandate, could significantly limit publication options. A 'one size fits all' approach to licensing could be problematic as publication requirements, and in particular reuse requirements, can vary considerably for individual Investigators depending on the subject areas and types of content that they are publishing. Allowing NIH Funded Investigators to choose how they want to license their rights and share their work would be the most equitable option.

### 2. Steps for improving equity in access and accessibility of publications.

At Wiley, we believe everyone should be able to access the research they need. We welcome opportunities to expand access to the results of federally funded research in a way that maintains research integrity and provides maximum benefit to the public and the American scientific enterprise, in coordination with federal funding agencies. We are committed to fulfilling that mission through our significant investment in open science.

### Ensure all readers have immediate access to the most trusted version of the research article that connects all other outputs including supporting data.

Open access to the VoR, supported by other research outputs, is the best way to improve equity in access and accessibility of publications. Only the final published VoR article delivers the full benefit of open access. In addition to the metrics, metadata, and context associated with the VOR outlined in our response to Question 1 above, many research artifacts from the research lifecycle such as preprints, open data, code, protocols (to name a few), are linked to the final VoR. Directing readers to the final publication, the VoR, provides transparent, linked access to all associated research artifacts which will ultimately validate the quality and integrity of the research process. An NIH Supported Investigator uploading an accepted article into a repository would not benefit from this verified interlinking of connected outputs and they would not receive the same visibility, level of engagement, and community recognition that they would otherwise achieve through the sharing of the final VoR publication. By supporting the publication of the VoR on journal platforms, readers can verify the mechanisms through which publishers support and uphold research integrity thereby ensuring trust in the authors work; continued trust in peer review, trust in research, and ultimately trust in scientific practice.

Open access is just one part of the open science ecosystem. At Wiley we are supporting open science practices and opening up more research outputs beyond the research article. Wiley was a founding member of CHORUS and the Initiative for Open Citations (I4OC) and we are enabling research data to be shared (and particularly Findable, Accessible, Interoperable, and Reusable

"FAIR" data) by defining standard criteria for repository selection, qualification and certification. We offer Transparent Peer Review making the associated peer review history openly available in a growing number of our journals. Nearly 75% of Wiley Journals currently support Preprints and via our Under Review service on Authorea we are working to streamline the early sharing of research, making the peer review process even more transparent. Making more research outputs open and reusable not only ensures integrity at every stage of the research process, but also reduces the unnecessary duplication of research, saving billions in research funding.

### 3. Methods for monitoring evolving costs and impacts on affected communities.

New costs that will arise from updated NIH public access requirements could impact equity in the scientific research community. To protect the scientific ecosystem and make it more robust and equitable, NIH should consider efforts to improve equity in research funding, understanding what sustainable high-quality publishing and repository costs entail, and educating the NIH research community to adopt cultural changes so that researchers become accustomed to including OA publishing costs in their research proposals and spending research funds on publishing in their venue of choice.

### Ensure the infrastructure and support is in place to help NIH funded Investigators and their Research Institutions budget for publishing costs.

Funding agencies such as NIH should help researchers to budget for anticipated publishing costs, and should consider creating a dedicated fund to support open access publication costs. If such a fund is fairly distributed, it could help to tackle inequality in publishing opportunities, create transparency for the monitoring of costs and impact of the new mandates, and avoid the problem of placing additional financial burdens on individual researchers and libraries.

We are aware that OA fund management can be extremely challenging and investment in tools and services to support Institutions and Funders is needed to build a sustainable and effective open science infrastructure. OAble, an open access management software solution, was developed by Knowledge Unlatched (a Wiley company) with significant stakeholder input to effectively manage the ever-growing complexity of OA activities and changing business models. Continued investment in these kinds of tools is needed to ensure Institutions and Funders can effectively manage OA funding.

Most current OA funding systems are built to accommodate corresponding author funder mandates and policies, with the assumption that the corresponding author is often also the grant recipient and responsible for funding acknowledgement. All stakeholders must be able to measure and address the administrative and open access funding burdens that would arise if coauthors are required to comply with federal agency public access policies.

### Work with publishers to develop requirements for better price and service transparency.

Wiley is actively working towards greater transparency by helping our customers to understand the value of the services we provide. We provide public access to data related to the peer review services we provide (turnaround times, acceptance rates), engagement (usage), impact (citations, media references (Altmetric), author contributions (CRediT), and re-use (scite). In 2022, Wiley provided data to the Plan S Price and Service Transparency Framework and Journal Comparison Service (JCS) with the aim of providing more transparency around the services that we provide. To date, we are the only major academic publisher participating in this initiative.

### 4. Early input on considerations to increase findability and transparency of research.

Access to high quality research for diverse, global audiences is vital for achieving open science. Information has to be discoverable. Wiley doesn't just make content available; we enable content to be found, providing the best opportunities for it to be discovered, so that the right audiences are aware of the latest research. Throughout our multi-step publishing process, we adhere to industry-accepted standards, from discoverability and archiving to presentation of published content.

### Make use of existing tools, initiatives and skills to avoid duplication of effort.

We support the FAIR principles and recognize the benefits of large-scale bibliometric analysis of research outputs that may lead to greater scientific and medical discovery. In 2022, we signed on to participate in the Initiative for Open Abstracts (I4OA) which allows for Crossref deposit and interrogation of abstract metadata. In addition, we support unrestricted access to article metadata on our publishing platform (Wiley Online Library) including abstracts, references, funder acknowledgements, data availability statements, and in many cases important contextual information like lay summaries and patient summaries.

We are continually investing in improvements and innovations in response to the evolving needs of the communities we serve while ensuring responsible, ethical publishing and preservation practices. By investing in new technologies and initiatives, we enable knowledge to be created, accessed, shared, and discovered more quickly on a global scale. Publishers have the skill and capacity to invest in maintaining the integrity of the VoR and in increasing the findability and transparency of research outputs, we recommend that NIH make best use of existing tools and initiatives to avoid unnecessary costs and duplication of effort.

Work with stakeholders to agree on a set of scholarly PIDs and open and non-proprietary metadata, to enable attribution to original publication sources.

The importance of having an appropriate and interoperable infrastructure to support these services cannot be underestimated. Crucial to this shared infrastructure is a set of commonly agreed persistent identifiers (PIDs) for researchers and organizations. These include the Open Researcher Contributor identifier (ORCID), the Research Organization Registry (RoR), and the Digital Object Identifier (DOI) for different scholarly outputs. An agreed set of scholarly PIDs and open and non-proprietary metadata, to enable attribution to original publication sources, has many benefits and is key to reproducibility and research integrity. Machine reading can then link and mine different research outputs and connect them to researchers or organizations, as well as to grants and different projects. This will enable verification, replication, discovery, and the reporting and tracking of research outputs, people, projects, and organizations.

### **Concluding Thoughts**

Wiley is committed to a future in which research is open. Last year 47% of the articles we published were done so under an open access license. We are committed to providing open access to the final, trusted version of record as the best way to accelerate open science, and to ensure that the published version remains the linchpin that connects numerous essential research outputs that add value and insight to the article itself. A sustainable transition to an open science future relies upon all stakeholders making a commitment to ensure that any future policies are financially and technically supported, and we are ready and willing to work collaboratively to make this a reality using the systems we have built to support scientific collaboration, integrity, and the research enterprise.

Where publishers are not provided the opportunity to fully participate in implementation and delivery discussions, or to collaborate on innovation and new business model development, solutions will remain inherently inequitable and our ability to achieve open science will be hindered. Ultimately, our mission is to serve researchers. We must ensure that under no circumstances will the quality of the works we publish and the valuable services that journals and societies provide to communities be compromised. In addition, we urge you to carefully review the responses you receive from smaller, less well-funded professional societies who make an important contribution to scientific and medical research in the United States. Without the diversity of specialized knowledge these societies, and their publications, contribute to the global scholarly ecosystem, the United States will begin to fall behind our global competitors.

We look forward to working with NIH and the wider scientific community on these issues and are committed to working collaboratively to develop forward-looking partnerships that strengthen research and innovation and deliver on the promise of open science. The stakes have never been higher, and we must leverage the entrepreneurial spirit of the research community and private sector to enable our country's continued leadership in the scientific enterprise.

Sincerely,

Jay Flynn

Jay Flynn Executive Vice President and General Manager - Research Wiley **Submit date: 4/24/2023** 

I am responding to this RFI: On behalf of an organization

Name: Seventeen Science Societies

Type of Organization: Nonprofit research organization

- 1. How to best ensure equity in publication opportunities for NIH-supported investigators.
- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

### **Uploaded File:**

 $Scientific\text{-}Society\text{-}Letter\_NIH\text{-}RFI\text{-}PublicAccess\_v2.pdf}$ 

**Description:** A letter signed by seventeen professional scientific societies and associations.

Email: jcarney@aaas.org

### April 24, 2023

# Joint Response to RFI on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research

As representatives of the American scientific community, the undersigned scientific societies and associations welcome the NIH's efforts to enhance public access, in line with the recent OSTP guidance aimed at making federally funded research publications and supporting data publicly available. We further applaud the NIH for emphasizing equity in its approach to public access policy development.

As a critical component of its public access plan, we urge the NIH to focus on creating an environment that balances reader access to published work with researchers' ability to publish. This will require transparency and recognition of the costs borne by researchers and research funders. We must strive to create a system wherein scientists are not required to pay additional fees to publish and where grants are not required to bear the brunt of publishing costs. Otherwise, we risk creating heavy cost burdens not only for researchers and their institutions, but also for funders of research, including taxpayers. Our organizations and others are experimenting with various sustainability models for public access, including diamond, green, and Subscribe to Open. These represent potential pathways towards the goal of optimizing equity for researchers and readers. We are open to other models that achieve NIH's equity goals.

Scientists' ability to communicate their scientific results through publication is critical to the incorporation of their expertise into the scientific enterprise and the progression of their careers. Monitoring implementation of changes to the public access policy, and how researchers and institutions pay publishing costs, will be critical to ensuring that public access plans do not create new systemic inequities or reinforce existing ones. Careful and continued study will be essential for understanding the near- and long-term effects of related changes. A study of cost effects at the researcher, institution, and enterprise levels is needed. It may also be valuable for NIH to survey researchers and institutions about publishing costs and about tradeoffs made to pay such costs.

Adaptation of federal grant agreements to require reporting on the payment of publication fees and reliance on transformative agreements (in instances where authors avoid payment of a fee because their institution has a transformative agreement with their journal of choice) represents one logical approach to monitoring fees. All analyses of and reporting on costs paid by institutions or researchers for publication should examine potential variability in costs across disciplines, career stages, and institution types, as well as variability based on researcher backgrounds.

As representatives of the scientific community, we believe we are at a crucial moment in the timeline of public access policy development, and we continue to share our view that public access should optimize equity for researchers and readers. We appreciate your consideration of these comments, and are committed to collaborating with NIH, other federal research agencies, and OSTP to develop public access policies that balance access to published work with the ability to publish.

Sincerely,

American Anthropological Association

American Association for the Advancement of Science

American Association for Anatomy

American Association for Dental, Oral, and Craniofacial Research

American Institute of Biological Sciences

American Society of Tropical Medicine and Hygiene

Association for Women in Science

Association of Independent Research Institutes

Biophysical Society

Federation of Associations in Behavioral and Brain Sciences

Human Factors and Ergonomics Society

Research!America

Sigma Xi, The Scientific Research Honor Society

Society for Research in Child Development

The American Society for Biochemistry and Molecular Biology

The American Society for Cell Biology

The Gerontological Society of America

**Submit date: 4/24/2023** 

I am responding to this RFI: On behalf of an organization

Name: Kathryn Richmond

Name of Organization: The Allen Institute

Type of Organization: Nonprofit research organization

Role: Institutional official

### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Exciting options exist for NIH to better ensure equity in publication opportunities, and these span both policy updates and broadening programmatic support and compliance.

To allow greater equity in publishing, opportunities exist upstream during publication creation and include providing additional resources for NIH-funded manuscripts in the form of support for technical writers and writing workshops. This could occur through partnering with external organizations. NIH leadership would also be pivotal in providing researchers clear guidance on rights retention given the complex landscape of copyright law and the need for authors to retain sufficient rights so that they may make their publications available in PubMed Central.

Additionally, while the proposed NIH guidance supports compliance through the archiving of articles in specific repositories (PubMed Central for NIH), we also strongly support depositing manuscripts onto preprint server(s).. In addition to the final published article, such public access must also consider all the materials required to ensure results can be reproduced. For reproducibility of results in the life sciences, we should aspire to include the following:

- Availability of the detailed research methods and procedures to generate the primary data
- Availability of all the metadata that materially affect the interpretation of results
- Availability of the full analysis details including intermediate results

The NIH and Other funders should pay attention to incentives to encourage adoption with these requirements. Examples of incentives may include data supplements on existing grants and data acquisition and reproducibility grants. And for those with a track record of equitable sharing of data, that this is taken into consideration when researchers submit for new NIH -funded grants.

### 2. Steps for improving equity in access and accessibility of publications.

Assuming the publication is freely available, an additional step to improve equity in access and accessibility of publications may be to require all NIH-supported work to include high-level plain language summaries that can be more accessible to the public, as well as to support language translation options and the ability to publish in native languages. Likewise, there could be an incentive to encourage publications to follow current data standards and best practices for their work, as well as funding to create such standards and organize data repositories. At this time, there is great potential in leveraging artificial intelligence approaches to ease the implementation path for these processes.

Lastly, NIH leadership would again be pivotal in providing researchers clear guidance on open access terms and/or utilization of licenses (ex. Creative commons options).

### 3. Methods for monitoring evolving costs and impacts on affected communities.

An important element of monitoring the publication cost landscape is creating more transparency across this dynamic area, and this could be accomplished by providing grants to study and report on such costs and their impacts.

### 4. Early input on considerations to increase findability and transparency of research.

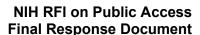
While persistent identifiers (PIDs) are helpful and allow users find and understand data and research products (particularly regarding institutions, authors, funders, and publications), ideally PIDs would be expanded to include their use for reagents, metadata, and protocols, etc. so as to enhance the findability of all research outputs.

AnOther suggestion to increase transparency in research would be to have publications include comprehensive results summaries alongside the published manuscripts for indexing on a summary website. The inclusion of such summaries would support good faith training.

#### **Uploaded File:**

FINAL-Allen-Institute-NIH-RFI-Response 4.24.23.pdf

Email: kathrynr@alleninstitute.org





### **General Information and Links:**

The National Institutes of Health (NIH) has released a Request for Information (RFI) soliciting feedback on the agency's proposed plan to enhance public access to the results of NIH-funded research. This RFI and associated draft plan was released in response to the 2022 White House Office of Science and Technology Policy (OSTP) <a href="mailto:memorandum">memorandum</a> on <a href="mailto:Ensuring Free">Ensuring Free</a>, <a href="mailto:Immediate">Immediate</a>, and <a href="mailto:Equitable Access to Federally Funded Research</a>. This 2022 guidance from OSTP directed federal agencies to update their public access policies to make publications and research stemming from federally funded research publicly accessible, without an embargo or cost. Agencies must fully implement these updated policies by December 31, 2025.

The draft NIH Plan to Enhance Public Access to the Results of NIH-Supported Research (NIH's Public Access Plan) outlines the proposed approach NIH will take to implement the 2022 OSTP guidance to enhance access to scholarly publications and scientific data resulting from NIH-funded research.

Comments must be submitted at <a href="https://osp.od.nih.gov/nih-plan-to-enhance-public-access-to-the-results-of-nih-supported-research/">https://osp.od.nih.gov/nih-plan-to-enhance-public-access-to-the-results-of-nih-supported-research/</a>. Responses will be accepted through April 24, 2023.

### Requested responses are to the following:

### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

The NIH Public Access Plan aims to maintain the existing broad discretion for researchers and authors to choose how and where to publish their results. Consistent with current practice, the NIH Public Access Plan allows the submission of final published articles to PubMed Central (PMC) (in cases where a formal agreement is in place) to minimize the compliance burden on NIH-supported researchers and maintains the flexibility of NIH-supported researchers to submit the final peer-reviewed manuscript. NIH seeks information on additional steps it might consider taking to ensure that proposed changes to implementation of the NIH Public Access Policy do not create new inequities in publishing opportunities or reinforce existing ones.

**Response**: Exciting options exist for NIH to better ensure equity in publication opportunities, and these span both policy updates and broadening programmatic support and compliance.

To allow greater equity in publishing, opportunities exist upstream during publication creation and include providing additional resources for NIH-funded manuscripts in the form of support for technical writers and writing workshops. This could occur through partnering with external organizations. NIH leadership would also be pivotal in providing researchers clear guidance on rights retention given the complex landscape of copyright law and the need for authors to retain sufficient rights so that they may make their publications available in PubMed Central.

Additionally, while the proposed NIH guidance supports compliance through the archiving of articles in specific repositories (PubMed Central for NIH), we also strongly support depositing manuscripts onto preprint server(s).. In addition to the final published article, such public access



## NIH RFI on Public Access Final Response Document

must also consider all the materials required to ensure results can be reproduced. For reproducibility of results in the life sciences, we should aspire to include the following:

- Availability of the detailed research methods and procedures to generate the primary data
- Availability of all the metadata that materially affect the interpretation of results
- Availability of the full analysis details including intermediate results

The NIH and other funders should pay attention to incentives to encourage adoption with these requirements. Examples of incentives may include data supplements on existing grants and data acquisition and reproducibility grants. And for those with a track record of equitable sharing of data, that this is taken into consideration when researchers submit for new NIH – funded grants.

#### 2. Steps for improving equity in access and accessibility of publications.

Removal of the currently allowable 12-month embargo period for NIH-supported publications will improve access to these research products for all. As noted in the NIH Public Access Plan, NIH also plans to continue making articles available in human and machine-readable forms to support automated text processing. NIH will also seek ways to improve the accessibility of publications via assistive devices. NIH welcomes input on other steps that could be taken to improve equity in access to publications by diverse communities of users, including researchers, clinicians and public health officials, students and educators, and other members of the public.

**Response:** Assuming the publication is freely available, an additional step to improve equity in access and accessibility of publications may be to require all NIH-supported work to include high-level plain language summaries that can be more accessible to the public, as well as to support language translation options and the ability to publish in native languages. Likewise, there could be an incentive to encourage publications to follow current data standards and best practices for their work, as well as funding to create such standards and organize data repositories. At this time, there is great potential in leveraging artificial intelligence approaches to ease the implementation path for these processes.

Lastly, NIH leadership would again be pivotal in providing researchers clear guidance on open access terms and/or utilization of licenses (ex. Creative commons options).

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

NIH proposes to actively monitor trends in publication fees and policies to ensure that they remain reasonable and equitable. NIH seeks information on effective approaches for monitoring trends in publication fees and equity in publication opportunities.

**Response**: An important element of monitoring the publication cost landscape is creating more transparency across this dynamic area, and this could be accomplished by providing grants to study and report on such costs and their impacts.

**4.** Early input on considerations to increase findability and transparency of research. Section IV of the NIH Public Access Plan is a first step in developing the NIH's updated plan for persistent identifiers (PIDs) and metadata, which will be submitted to OSTP by December 31,



## NIH RFI on Public Access Final Response Document

2024. NIH seeks suggestions on any specific issues that should be considered in efforts to improve use of PIDs and metadata, including information about experiences institutions and researchers have had with adoption of different identifiers.

**Response:** While persistent identifiers (PIDs) are helpful and allow users find and understand data and research products (particularly regarding institutions, authors, funders, and publications), ideally PIDs would be expanded to include their use for reagents, metadata, and protocols, etc. so as to enhance the findability of all research outputs. Another suggestion to increase transparency in research would be to have publications include comprehensive results summaries alongside the published manuscripts for indexing on a summary website. The inclusion of such summaries would support good faith training.

I am responding to this RFI: On behalf of an organization

Name: Andrew Herrin

Name of Organization: Society for Industrial and Applied Mathematics (SIAM)

Type of Organization: Professional org association

Role: Institutional official

- 1. How to best ensure equity in publication opportunities for NIH-supported investigators.
- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

#### **Uploaded File:**

SIAM-NIH-Open-Access-RFI-Response-April-2023-Final.pdf

**Description:** RFI Response

Email: andrew@lewis-burke.com



April 24, 2023

Lyric Jorgenson, PhD
Acting Director, Office of Science Policy and
Acting NIH Associate Director for Science Policy
The National Institutes of Health
6705 Rockledge Drive, Suite 630
Bethesda, MD 20892

RE: SIAM Comments Response to NOT-OD-23-091, Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research

Comments transmitted electronically via RFI Web form on April 24, 2023

Dear Dr. Jorgenson,

The Society for Industrial and Applied Mathematics (SIAM) appreciates the opportunity to provide comments in response to NOT-OD-23-091, Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research. SIAM is an international community of over 14,500 members from academia, industry, and government. Members come from many different disciplines, and all have a common interest in applying mathematics, computational science, and data science towards solving real-world problems. SIAM is committed to ensuring that the highest quality research is readily available to enhance the nation's research enterprise.

SIAM appreciates NIH's thoughtfulness in developing public access plans. Independent, non-profit, discipline-specific, professional societies and publishers such as SIAM foster focused and creative communities that enable professional development, build new research connections, and enhance workforce development, adding immense value to the scientific field and publishing ecosystem. Publishing is a critical part of SIAM's impact and also the largest source of income for the organization, which has also faced rising expenses in recent years due to the COVID-19 pandemic and other economic factors. As NIH and other federal agencies consider updated open access plans, it is critical to be cognizant of the future of small, independent, society publishers and their larger impact and importance to the research community. We appreciate NIH's commitment to broad engagement of the scientific community and nonprofit scholarly publishers of all sizes.

Supporting a balanced ecosystem of journals is crucial for the vitality of the research community. SIAM applauds the NIH Public Access Plan in allowing authors to submit the author accepted manuscript as the deposited version, rather than the version of record. This helps safeguard SIAM's income stream, allowing for continued reinvestment in the current 18 peer-

ADVANCING SCIENCE AND INDUSTRY WITH MATHEMATICS SINCE 1952

reviewed research journals, which are the leading source of knowledge for the world's applied mathematics, computational science, and data science communities. In addition, SIAM employs very liberal green open-access policies in line with the NIH's public access vision while most articles are available as preprints on relevant repositories. SIAM is appreciative that the Creative Commons Attribution License (CC-BY) is not mandated for the author accepted manuscript, remaining aligned with NIH's current Data Management and Sharing policy and not creating unnecessary barriers for authors and society publishers.

SIAM encourages NIH to continue to monitor the impact of open access models on the financial sustainability of federal grants as well as nonprofit society publishers in order to ensure continued independence and high standards. Openness must always be combined with quality and trust.

Equity must be placed at the heart of any open access funding model – it is critical, particularly in the disciplines that SIAM serves, that authors are able to publish their research in their journal of choice regardless of the author's ability to pay or the institutional open access agreements that may – or may not – be in place. Current gold Open Access models such as APCs and read & publish transformative agreements risk creating a two-tier system where researchers from well-funded institutions and/or disciplines and/or countries get open access, while the rest do not. Such barriers go against the core scholarly publishing principle that the best ideas should be published and promoted wherever they come from. SIAM is actively working on developing an open access model that is equitable, sustainable, as simple as possible to administer, and does not erect any barriers for authors globally.

SIAM commends NIH for its leadership in engaging the community and adapting the plan for public access to allow researchers and societies to produce research free from unnecessary burdens and easily accessible by the public. As the leader in industrial and applied mathematics research, SIAM stands ready to continue to be a resource as NIH crafts public access policies that impact independent society publishers and the wider publishing landscape.

Sincerely,

Dr. Suzanne L. Weekes Executive Director, Society for Industrial and Applied Mathematics

I am responding to this RFI: On behalf of an organization

Name: Karen McDonnell & Liz Borkowski

Name of Organization: Women's Health Issues

Type of Organization: University

Role: Scientific researcher

1. How to best ensure equity in publication opportunities for NIH-supported investigators.

We recommend that NIH revise its draft policy in order to avoid creating financial pressures that lead to peer-reviewed journals adopting policies and practices that reduce equitable opportunities for researchers. Please see our attached comments for details.

- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

#### **Uploaded File:**

WHI-comment-on-NOT-OD-23-091.pdf

**Description:** Comments from the editor-in-chief and managing editor of the peer-reviewed journal

Women's Health Issues

Email: borkowsk@gwu.edu



April 24, 2023

Office of the Director
National Institutes of Health
U.S. Department of Health and Human Services
9000 Rockville Pike
Bethesda, Maryland 20892

Re: Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research (NOT-OD-23-091)

As editors of the peer-reviewed journal *Women's Health Issues*, we appreciate the opportunity to comment in response to "Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research" (NOT-OD-23-091). We applaud NIH for taking steps to make agency-funded research findings immediately available to the public and recommend that NIH revise its draft policy in recognition of 1) the value that journals provide and the financial support they need to do so and 2) the common practice of generating numerous publications from a single dataset.

Women's Health Issues (WHI) is the peer-reviewed journal of the Jacobs Institute of Women's Health, which is based at the Milken Institute School of Public Health at the George Washington University. WHI is dedicated to improving the health and health care of all women throughout the lifespan and in diverse communities. The journal seeks to inform health services researchers, health care and public health professionals, social scientists, policymakers, and others concerned with women's health. WHI is published by Elsevier and has an impact factor of 3.053 and CiteScore of 4.2.

#### **Our Journal's Operations**

We consider our journal's mission to include supporting the development of early-career and emerging researchers. Such researchers often conduct secondary analyses of datasets their mentors have collected, and their manuscripts often require substantial revision before they are ready to publish. While our volunteer peer reviewers provide extensive constructive comments on the substance of manuscripts, our editorial team members also devote considerable effort to helping authors edit their work for clarity, precision, and readability. It is not unusual for us to spend three to four hours writing up recommendations on a single manuscript (after it has already undergone a round of revisions in

response to peer-review comments), and we do so in a manner intended to help the authors learn from the experience and improve their future writing. Of course, it is not the case that all manuscripts from early-career researchers require such intensive editing, and some manuscripts from established researchers also require extensive revision; regardless of the author's stature, we aim to help them publish polished work. Our team also proofreads typeset articles prior to final publication to catch the inevitable typographical errors that could mar the quality of published work.

Elsevier's copyeditors and typesetters also contribute to the quality of published manuscripts. In particular, they identify discrepancies between in-text citations and reference lists and thereby help authors reference others' work appropriately. By handling numerous logistical elements, from the manuscript submission system to the online posting of articles, Elsevier leaves our editorial team free to focus on working with authors to solicit, receive expert feedback on, revise, and publish high-quality research.

#### **Financial Realities**

Like many smaller journals, we operate with a small editorial staff (one part-time editor-in-chief, one part-time managing editor, and one part-time editorial assistant) and often struggle to afford our personnel costs. We are dependent on annual editorial stipend and royalty payments from Elsevier, which are calculated based on subscriptions, article downloads, and sale of Open Access licenses (for which Elsevier currently charges \$3,360 per article). At the moment, our journal does not charge publication fees, and this allows us to accept submissions from authors who do not have grant funding for the work they publish with us; we receive many submissions from doctoral students and postdocs publishing their dissertation research, and from junior faculty members seeking new grant funding.

If institutions drop Elsevier journal subscriptions because much of the material they seek to access is available for free in public repositories, our revenue from Elsevier is likely to decline and we will find it difficult, if not impossible, to continue publishing manuscripts that require substantial editorial involvement prior to publication. If many journals make this kind of calculation, early-career researchers and others who have not had the benefit of past writing instruction will likely find it harder to publish and advance their careers.

A possible response to the new policy, and one our journal will have to consider if NIH adopts the draft policy without modification, is to begin charging all authors to publish in *WHI*. Such a move would prevent us from providing equitable publication opportunities to those without grant funding, but it might become necessary for our financial survival. We anticipate that many other journals would make similar calculations, which would lead to a sharp reduction in publishing opportunities for researchers

who lack grant funding — a group disproportionately composed of early-career authors and those from marginalized racial and ethnic groups.<sup>1</sup>

#### Recommendations

Our reading of section III.A.3 of NIH's draft policy suggests that all NIH-funded authors will have to deposit their accepted manuscripts in PubMed Central (PMC) and that PMC will make them available as soon as they are published. To ensure journals' ability to survive while still welcoming submissions from early-career researchers, we recommend that NIH add to III.A.3 another avenue for policy compliance besides depositing the manuscript with PMC: publication under an Open Access license. Sales of optional Open Access licenses could replace revenue lost as institutions drop subscriptions and allow journals to continue accepting submissions regardless of authors' ability to pay publication fees.

The draft policy's statement that NIH will continue to allow reasonable publication costs for all NIH-supported or authored scholarly publications is welcome, but it is not clear that it would allow for the use of NIH grant funds to purchase optional Open Access licenses from journals such as ours that do not charge publication fees as a standard practice. This merits clarification. As noted above, we would like for NIH-funded authors to be able to publish with *WHI* and use their grant funds to purchase optional Open Access licenses as an alternative to having their work made available at PMC upon publication.

In addition, we ask that NIH recognize the number and timing of publications that use NIH data and consider additional or enhanced mechanisms to allow those who conduct later secondary analyses to use grant funds to purchase Open Access licenses. NIH-supported investigators often use grant funding to support the publications answering the grant's primary research questions, but their grants can end or be exhausted before doctoral students and other junior colleagues publish secondary findings from the same dataset. We recommend both that 1) NIH encourage investigators to include the purchase of numerous Open Access licenses in their budgets without reducing funding in other areas to allow for it and 2) NIH establish a mechanism by which authors can request such funding from NIH after a grant has ended.

Journals and their publishers perform important services to help authors publish high-quality work that advances knowledge in their topic areas. We agree that the public should have immediate access to government-funded research, but we fear that implementation of NIH's draft policy will lead to a sharp reduction in funding for the services journals provide. *Women's Health Issues* would like to continue welcoming publications from early-career and unfunded authors, but we will be unable to continue our

<sup>&</sup>lt;sup>1</sup> Taffe MA & Gilpin NW. (2021). Racial inequity in grant funding from the US National Institutes of Health. eLife, doi: 10.7554/eLife.65697.

current practices if our funding drops. A revised policy incorporating the above recommendations would better balance the public's right to access with journals' need for financial sustainability and the goal of ensuring equitable publication opportunities for a diverse group of researchers.

Thank you for this opportunity to comment in response to the RFI. If you have any questions, please contact WHI managing editor Liz Borkowski at 202-994-0034 or borkowsk@gwu.edu.

Sincerely,

Karen A. McDonnell, PhD
Editor-in-Chief, Women's Health Issues
Milken Institute School of Public Health
George Washington University

Liz Borkowski, MPH
Managing Editor, Women's Health Issues
Milken Institute School of Public Health
George Washington University

I am responding to this RFI: On behalf of an organization

Name: Tina Baich

Name of Organization: U.S. Repository Network

Type of Organization: Not applicable

1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Please see attached PDF document.

- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

#### **Uploaded File:**

USRN-NIH-RFI-Response.pdf

**Description:** RFI Response from U.S. Repository Network

Email: tina@sparcopen.org

# **U.S. Repository Network**

April 24, 2023

Office of Science Policy National Institutes of Health 6705 Rockledge Drive, Suite 630 Bethesda, MD 20892

Submitted via electronic form

Re: Notice Number NOT-OD-23-091

The U.S. Repository Network (USRN) is grateful for the opportunity to comment on the NIH's request for information regarding the agency's plan to enhance public access to the research it funds. The USRN is an inclusive community committed to advancing repositories in the U.S. through advocacy, good practices, and community building. The organization, supported by SPARC (the Scholarly Publishing and Academic Resources Coalition), is propelled by the community-driven strategic vision that an interoperable network of repositories is an essential component of our national research infrastructure, offering rapid and open access to research, and plays a crucial role in collective efforts to transform global research communications, leading to a more open, inclusive, and equitable system.

Repositories are key institutional tools that ensure access to and reuse of valuable research outputs. They support preservation; facilitate reproducibility of research, research assessment, and compliance workflows; afford new opportunities for publishing; and increase individual and institutional visibility. By enabling rapid and open access to research outputs, repositories accelerate the pace of scholarship and the social impact of research for the public good.

Acting collectively, repository hosts can leverage their power to strengthen repositories and interact with other types of services, adding value and leading to significant innovation in the landscape.

#### Question 1. How can NIH best ensure equity in publication opportunities for its investigators?

The USRN is supportive of repository deposit as a primary compliance mechanism for the NIH Public Access Policy. It is critical that NIH-funded investigators understand that they can fully comply with the agency's public access policy by depositing their author's accepted manuscripts into PubMed Central (PMC) - or any other agency approved repository - and that there is no fee required to do so.

The USRN recently released its <u>Desirable Characteristics for Digital Publication Repositories</u>, which is intended to align with the <u>Desirable Characteristics of Data Repositories for Federally Funded Research</u> already issued by the National Science and Technology Council. The USRN Desirable Characteristics have been developed with input from the US repository community, and provide community guidance,

## **U.S. Repository Network**

advance understanding of the utility and value of repositories, and, ultimately, increase the interoperability and coherence of U.S. repositories.

We are concerned that compliance mechanisms that rely on article processing charges (APCs) reinforce and/or introduce inequities within the research communication system. For instance, APCs have proven to be prohibitively expensive for individuals and their institutions, and studies have documented that APC costs disproportionately affect younger researchers, female researchers, and those at less well-funded institutions. APCs also require a diversion of funds away from the research process; investigators often must use money originally intended for materials and equipment, supporting postdocs, and for professional development opportunities including presenting research results at conferences. Ensuring ease of compliance through fast, free repository deposit mechanisms will provide an important channel to help eliminate the need for such expensive, unsustainable fees.

We appreciate NIH's commitment to avoiding further inequity as an unintended consequence of its policy, and are eager to support NIH in this commitment, providing a compliance route at no cost to investigators.

#### Question 2. What steps can NIH take to improve equity in access and accessibility of publications?

Repositories ensure that research outputs are discoverable, visible, and accessible for future use, extending beyond publications and creating an open ecosystem that will support linking, verification, and reuse of the entire corpus of research results. Moreover, repositories enhance access to research outputs through their lack of paywalls and their use of open licensing. Encouraging the use of open licenses that enable full reuse rights (such as those offered by Creative Commons) would ensure the robust utility of NIH-funded research.

NIH, with its long history of hosting PubMed Central, can share best practices for implementing robust machine-readability and text mining functionality as well as assistive technology to further improve equity in access and accessibility, with the rest of the repository community. The USRN stands ready to work with the NIH to advance equity in access and accessibility across the U.S. repository landscape.

We thank the NIH for centering equity in both its draft policy and planning process, and for providing the opportunity to submit these comments.

Sincerely,

Tina Baich Visiting Program Officer U.S. Repository Network

I am responding to this RFI: On behalf of an organization

Name of Organization: American Society of Mechanical Engineers

Type of Organization: Nonprofit research organization

Role: Institutional official

- 1. How to best ensure equity in publication opportunities for NIH-supported investigators.
- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

#### **Uploaded File:**

ASME-NIH-RFI-Response.pdf

Email: fakesp@asme.org



1828 L Street, N.W.

Suite 510

Tel 202.785.3756 Fax 202.429.9417

Washington, D.C. 20036

www.asme.org

# ASME's Response to the Request for Information (RFI) on the National institutes of Health (NIH) Plan to Enhance Public Access to the Results of NIH-Supported Research Notice Number: NOT-OD-23-091

#### **April 24, 2023**

Founded in 1880, The American Society of Mechanical Engineers® (ASME) mission is to help the global engineering community develop solutions to real world challenges facing all people and our planet. We actively enable inspired collaboration, knowledge sharing and skills development across all engineering disciplines throughout the world, while promoting the vital role of engineers in society.

With over 85,000 Members, our organization is one of the largest technical publishing operations in the world, offering thousands of titles and some of the most prestigious engineering content in 33 technical journals including the topics of biomedical and bioengineering, power generation and storage, and manufacturing to highlight a few.

ASME serves a wide-ranging engineering community through quality learning, the development of codes and standards, certifications, research, conferences and publications and other forms of outreach. We collaborate with 36 Technical Divisions, including an ASME Bioengineering Division which includes over 5,000 members who employ mechanical engineering principles in the development of many lifesaving and life-improving technologies such as robotic surgery, the artificial heart, prosthetic joints, diagnostics and numerous rehabilitation technologies and directly contribute to U.S. advances in bioengineering.

ASME is driven by global engineering communities to ensure high-quality, rigorously peer-reviewed content is accessible and freely available online for everyone. ASME journals provide extensive, diverse indexes of research articles that span the broad spectrum of engineering topics. ASME supports compliance with government and funder mandates for Open Access publication, including Plan S for European-funded research, and offers authors the option to publish their papers hybrid Open Access across all our journals or in the fully Open Access ASME Open Journal of Engineering with payment of an Article Publishing Charge (APC).

ASME continues to offer publication at no cost to an author through traditional subscription access. However, the White House Office of Science and Technology Policy's 2022 guidance, "Ensuring Free, Immediate, and Equitable Access to Federally Funded Research", would eliminate the subscription-based model, resulting in the need for new funding streams to support zero cost-to-author publication. The current 12-month embargo period allows publishers to recoup at least part of their costs by incentivizing subscriptions for readers who desire immediate access. The new OSTP policy will eliminate the ability to recoup any part of the costs incurred in publishing, leaving smaller institutions that are

dependent on this model, including many non-profit organizations with public service missions, resource constrained and marginalized.

Pre-requisites for ensuring success of the OSTP's new policy include:

- Development of economic and sociological impact study and analysis of new public costs resulting from the 2022 OSTP policy guidance
- Development of guidance to authors/researchers on how to budget for new publication and data management costs
- Development of policies to ensure researcher freedom to choose venue of publication, repository, and an appropriate re-use license
- Development of agency metrics and guidelines to support maximation of equitable access to funding

#### Responses to NIH-identified questions:

# How to best ensure equity in publication opportunities for NIH-supported investigators?

ASME supports open science by ensuring our peer reviewed scholarly publications are of the highest quality and integrity. By fostering their dissemination, we advance engineering and scientific research to ensure the United States remains globally competitive.

The peer-reviewed scholarly publications which are included in our journals are not the direct result of the expenditure of taxpayer funds; conversely, they result from a significant publisher investment. Over the years, ASME has dedicated significant resources in innovative platforms that enable exceptional digital peer-review, production, distribution, interoperability, and discovery of the latest scientific and scholarly works to ensure our publications are of the highest quality. Our Digital Collection provides unparalleled depth, breadth and quality of peer-reviewed content and includes: 33 technical journals; 26 conference proceedings (annually); 3,500 journal articles reviewed by over 8,000 subject matter expert editors (annually); and comprised of over 308,000 technical papers and 2.400,000 technical pages.

ASME's peer-reviewed journal articles are the direct result of our investments and our extensive collaborations with authors, which is why they are considered the "gold standard" of scientific communication. The ability to recoup our investment enables innovation, allows infrastructure to be developed (including archives and metadata), and provides incentives to try new approaches. Long-term stewardship of content also carries significant costs that are already being borne by publishers.

Any policy change requiring us to make our peer-reviewed publications immediately available for free without charging a fee is not economically sustainable for our organization, as well as other scholarly publishers. A new, sustainable funding model must include clear guidance on how private publication costs will be transferred to a new publicly funded model. The scholarly research and publishing enterprise is a very complex and intricate ecosystem. We must be able to recoup our investments in order to publish high quality peer reviewed journals and research articles, as well as to sustain collaborations of this nature.

#### Steps for improving equity in access and accessibility of publications.

While immediate open access is often couched in terms of expanding access in equity terms, for researchers it threatens to create a pay-to-play system benefiting well-resourced institutions and researchers. While large corporations and well-funded universities may be able to absorb new R&D publishing and administrative costs, smaller colleges and companies will struggle to function. For HBCUs, rural institutions, community colleges, and undergraduate-only programs, this policy will further strain already-tight research budgets and marginalize their contributions.

We share the goal of open access for taxpayer-funded research. However, current proposals fail to sufficiently address guidance and budget forecasting for the crucial funding mechanisms which will allow for the peer-reviewed publication of vital research. We encourage Congress and the Administration to closely coordinate with the research and scholarly publishing communities on clear guidance supporting equitable solutions to providing the necessary funding streams to meet the expanded public policy objectives of the revised OSTP Public Access policy.

#### Methods for monitoring evolving costs and impacts on affected communities.

ASME is concerned that the currently proposed OSTP guidance does not sufficiently account for transition to a model where subscriptions are largely eliminated. There is already substantial evidence of subscription cancellation and market consolidation in the face of open access mandates, both in Europe and in the United States. Assertions that expanded Open Access policy objectives can be achieved without any new costs are not supported by any exploration of the state of the scholarly publishing industry.

The National Institutes of Health recognizes the importance of seeking post-publisher peer-reviewed article versions and other additive content to satisfy OSTP's proposed open access requirements, as opposed to the author's original manuscript. Agencies should also develop planning to account for new peer-review costs, data management costs, including re-investment into expanded public-private databases, costs for maintenance of versions of record and related open access data repositories.

Further, Open Access APCs are likely to be subject to annual discretionary appropriations from Congress and individual institutional budgetary decisions. Federal agency leaders should develop transparent economic modeling to support elimination of the subscription revenue stream from scholarly publications supporting federally funded researchers, including guidance to researchers on how to account for new open access policy implementation costs. We believe helping researchers understand and budget for costs, as well as NIH and other federal agencies seeking robust and sustainable funding from agency leaders and Congress is the best way to ensure authors at all institutions have a wide array of options to communicate their research.

#### Early input on considerations to increase findability and transparency of research.

Researchers in the academic, government, and corporate sectors are generating massive quantities of data across all scientific, technical, and medical disciplines at an accelerating

rate. Increasingly, government and other funding bodies are beginning to require expanded data management plans, including in the NIH Public Access Plan. ASME currently participates in the Open Researcher and Contributor ID (ORCID) and research Organization Registry (ROR) in order to provide persistent digital identifiers that authors and research organizations own and control.

The United States world-leading professional and scholarly publishing sector provides a strong foundation for scientific integrity around the world, but this sector requires a strong enabling framework of copyrights and intellectual property protections to sustain it, especially in the face of growing technological means of undermining existing copyright protections. It is important that federal agencies do not force researchers into untenable rights or licensing agreements that could suppress researcher choice in how they communicate their research. Researchers need flexibility, including non-commercial, non-derivative versions that allow them to protect the integrity of their work. Agency requirements restricting authors' ability to license their rights, for example through a rights retention mandate, would significantly limit authors' options to bring their work to the scientific community, thereby increasing costs and limiting equity options.

An industry-university-government partnership is essential to the progress of science, engineering and education, and we look forward to working with the NIH to ensure that scientific information itself remain free from political interference to the maximum extent possible. As agencies consider societal communication of scientific and technical information, it is critical that science and engineering communicators have a healthy degree of freedom of choice in how and where they can publish, as well as separation from the appearance of undue government influence in the preparation and publication of scientific information. This issue is especially salient as society struggles with scientific disinformation and mistrust in government institutions.

The erosion of copyrights for independent technical and scholarly publishers risks driving further consolidation of the publishing industry into fewer distribution mediums, a dynamic fundamentally at odds with maintenance of a healthy, competitive, innovative, and independent scholarly publishing ecosystem.

ASME continues to accelerate public access while advancing engineering and technological research to ensure the United States remains a global leader in engineering innovation. While ASME endorses the dissemination of the results of all peer-reviewed research, including research supported by federal funding, it must be done in a manner that is sustainable for the scholarly publishing community.

I am responding to this RFI: On behalf of an organization

Name: Kacy Redd

Name of Organization: Association of Public and Land-grant Universities

Type of Organization: Professional org association

Role: Member of the public

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Public access to data used in federally funded research in peer-reviewed journals is essential for rigorous science, discovery, and the reproducibility of research. Public universities are committed to sharing the results of their research whenever possible. For this reason, the Association of Public and Land-grant Universities (APU), in collaboration with the Association of American Universities (AAU), and with funding from the National Science Foundation (NSF#1837847 and #1939279) and National Institutes of Health, held a series of workshops and conferences with researchers, senior research officers, librarians, chief information officers, and organizations supporting increasing public access to research. Many of the concerns outlined in this RFI were discussed by the research community during these convenings, and we draw upon that insight in our responses below.

NIH is a recognized world-leader in facilitating public access to research publications in the biomedical sciences with the creation, support, and management of PubMed Central. APLU appreciates that the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research allows for flexibility in where researchers publish and that the plan allows researchers to charge reasonable publishing costs to their awards. We have provided some suggestions for determining what are reasonable publication costs in a later section. The flexibility in where to publish and covering reasonable costs are critical elements in protecting our current peer review system in that the peer review system relies on the coordination of journal editors and publishers. However, these elements of the NIH plan are not necessarily sufficient to ensure equity in where and how researchers can publish their peer-reviewed scholarship. Reasonable costs might not cover all costs, which might preclude being able to publish in journals with the broadest reach and impact.

NIH could help address equity concerns for investigators and/or institutions to deposit research data by creating an agency-wide repository for data, especially for data without a current NIH-supported disciplinary repository. Such a repository or repositories would ensure that research data adheres to the FAIR principles of findability, accessibility, interoperability, and reusability of data. NIH could support both the technical infrastructure and the human infrastructure required to ensure quality data curation. This would increase standardization across the NIH directorates for research data produced in NIH-funded projects. This would also reduce the burden on any single researcher or institution on selecting an appropriate repository. An NIH-supported research data repository with expert staff (e.g. research librarians who could provide curatorial support) would reduce the cost to the investigator and/or institution, which would address many of the equity concerns related to publishing research data. The investigator community would be further helped if NIH led a coordinated effort with Other research agencies to support a common platform with common requirements. If there was such infrastructure,

investigators would only need to navigate one platform. Through this portal, NIH could also provide the aforementioned expert support services coupled with guidance and training for investigators on how to share their data so that it is FAIR and maximizes impact.

Creating critical infrastructure is only one concern of researchers and research institutions. AnOther concern that investigators have is that the publication of research may happen after the end of an award period due to the often-delayed peer review and revision process. How will NIH address cases where research outputs exceed the funding allocated in the grant or contract, or in which publications come out after the grant period has ended? A lack of funds may significantly impede the researcher's ability to publish their results in their discipline's preferred journal or deposit their research data in the discipline's preferred repository. This could, subsequently, affect the visibility and impact of the research, resulting in the marginalization of the career of researchers at emerging research institutions or less-resourced institutions. To help ensure that all researchers have the funds to appropriately and with greater impact share their findings, we encourage NIH to consider 1) allowing institutions to prepay publication costs; 2) allowing institutions to hold designated publication funds after the end of the award to pay for these publication costs; or 3) make supplemental funding available to cover these publication costs.

Additionally, the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research could help enable equity in sharing scientific data by harmonizing its requirements with the guidance given in the OSTP 2022 memo on "scientific data". The memo directs the agencies to ensure that "Scientific data underlying peer-reviewed scholarly publications [emphasis added] resulting from federally funded research should be made freely available and publicly accessible by default at the time of publication..." Currently, 'research findings' as defined by federal regulations (CFR 200.315 (e) and 45 CFR 75.322 (e)) are required to be published in a peer-reviewed scientific or technical journal. However, NIH's proposed expansion of the definition of 'scientific data' in its plan goes beyond the requirements of the OSTP memo and current regulations, potentially encompassing research data that has not necessarily been peer-reviewed. This could create uncertainty for researchers in determining when and what data needs to be shared "to validate and replicate research findings". Such a new standard for 'scientific data' may pose additional burdens on researchers, and the implications of this expansion need further clarification to ensure effective implementation of the NIH Plan while considering existing regulations and scholarly practices.

NIH could provide additional guidance to help researchers and institutions understand the impact of this plan on legal protections, retention of rights and intellectual property, and impacts on human subject protections and national security. As one concern, less resourced institutions may not have the technology transfer expertise to adequately determine whether a non-peer-reviewed data set falls under export control concerns. Who then ultimately decides which data that do not support a peer-reviewed publication is appropriate to share with the international community? Once research findings and research data enter the public domain it may be impossible to protect economically valuable information or protect against Other unintended consequences. Who then becomes liable for any adverse outcomes that could not be reasonably anticipated?

Most importantly, there is not a clear mechanism for peer review to ensure the quality of the shared data - data that the general public might access and on which draw erroneous or even harmful conclusions. That it has been funded by NIH and shared with the public will signal that it is of a certain

quality and reliability, which may not be accurate. Who becomes liable for adverse events based on sharing this non-peer-reviewed data?

Further, we are concerned that there is no reasonable estimate for the cost to share this expanded scope of data that does not underlie peer-reviewed publications. The cost to the compliance system and the burden on researchers will be great since there is no clear way to track this data and currently little benefit to the researcher to share this data.

Given these concerns, we recommend that NIH's Plan to Enhance Public Access to the Results of NIH-Supported Research and related guidance follow the OSTP guidance to ensure scientific data underlying peer-reviewed publications be made accessible.

We recognize that in our current system some valuable data is not regularly peer-reviewed and shared (i.e. negative results data). To address this challenge, APLU would be happy to work with NIH and the broader research community to address how we might increase the incentives for publishing negative results by supporting venues where the data can be peer-reviewed and ensuring such publications are valued in grant reviews and performance/promotion.

#### 2. Steps for improving equity in access and accessibility of publications.

To increase the discoverability of NIH-supported research data, NIH should support infrastructure that would enable searching all NIH-supported research data repositories via a common portal as NIH has done for peer-reviewed publications in NIH's PubMed Central.

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

We are pleased to see that NIH plans to actively monitor trends in publication fees and policies. We encourage NIH to also monitor publication outcomes that assess whether less resourced institutions, disciplines, and/or labs are increasingly locked out of publishing in the most accessed journals.

We are concerned that costs to publish will increase as publishers shift their business practices from collecting revenue from readers to collecting revenue from research awards. We encourage NIH to engage in an analysis of current article processing charges (APCs) within different disciplines and base "reasonable publishing costs" on current market rates for publications and depositing research data. This could be an annual analysis to ensure guidance on "reasonable publishing costs" is current and that publishing costs are not increasing at an excessive rate due to publishers shifting costs to investigators due to these new policies. The NIH could also engage in periodic dialogs with researchers, institutions, repositories, and publishers, especially from professional societies, to discuss what are "reasonable publishing costs".

#### 4. Early input on considerations to increase findability and transparency of research.

APLU joined the Association of Research Libraries, the California Digital Library, and the Association of American Universities in convening an NSF-supported conference in 2019 (NSF #1945938) and released a report with recommendations for data practices supporting an open research ecosystem. Through those discussions, we came to a consensus on five persistent identifiers (PIDs) that would help ensure that research data is FAIR. These were:

1. Digital object identifiers (DOIs) to identify research data, as well as publications and Other outputs

- 2. Open Researcher and Contributor (ORCID) IDs to identify researchers
- 3. Research Organization Registry (ROR) IDs to identify research organization affiliations
- 4. Crossref Funder Registry IDs to identifier research funders
- 5. Crossref Grant IDs to identify grants and Other types of research awards

We also identified recommendations that would help support this necessary PID infrastructure. NIH could lead the following to advance the sharing of research and research data.

- NIH, in coordination and harmonization with Other federal agencies, could fund the design and development of tools and services to support the use of PIDs. NIH could fund investigators developing research-related workflows and systems that enable the collection of PIDs, storage of PID metadata, and connections to PIDs in Other systems.
- NIH, in coordination and harmonization with Other federal agencies, could invest in infrastructure and initiatives that support the use of PIDs by supporting member organizations that promote open scholarly infrastructure, such as Crossref, DataCite, and ORCID; funding organizations and data repositories that follow best practices for FAIR data; supporting community-led initiatives such as the Research Organization Registry and EZDMP.
- NIH, in coordination and harmonization with Other federal agencies, could minimize the burden on researchers by making it easy and seamless for researchers to use PIDs by designing workflows and systems to assign and collect them automatically and by supporting PID services or data repositories within the PubMed Central platform. This will be especially necessary for less-resourced institutions that may not have a research librarian to provide these services.

Email: kredd@aplu.org

I am responding to this RFI: On behalf of an organization

Name: Krystal Toups

Name of Organization: COGR

Type of Organization: Professional org association

Role: Member of the public

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

To best ensure equity in publication opportunities for NIH-supported investigators, we offer the following comments and recommendations.

#### Costs:

As stated previously\*, ensuring public access to publications and research data resulting from federally funded research requires financial investments across the research enterprise. The 2022 OSTP Memorandum notably removes the 12-month embargo period, and while we understand and support the benefits of this policy change, we share in the community's expressed concerns about the potential for shifts in publishing models and increased costs with varying impacts depending on institutional characteristics. It is important that agencies plan accordingly to prevent any inequities.

Publication Cost - While NIH policy allows supported researchers to charge reasonable publishing costs against their awards, it is important to recognize that "reasonable costs" may not account for all costs or account for increased costs due to a shift in the publishing models. We share the community's concerns about the shift in the publishing model towards Article Processing Charges (APC), which is a significant fiscal and cultural change from subscription-based cost models. This shifting model forces universities to bear an ever-increasing proportion of the costs associated with publishing, including APC, subscription costs, and provision of uncompensated scholarly reviewers. Budgetary constraints may force institutions to make difficult choices about which faculty members to fund, and early career researchers, researchers from institutions with limited resources, and/or under-represented groups may be disproportionately disadvantaged. Although NIH states in the RFI that APC may be charged to NIH grants, unless supplemental funds are provided, these charges will have a significant impact on the overall project budget. We hope that agencies and OSTP will directly address these concerns, and NIH should clearly state all APC, and Other publishing costs should be budgeted accordingly in NIH grants and contracts.

Modular Budget Caps - We would like to direct NIH to COGR's recent letter\*\* that addresses the limitations of modular budget caps. COGR's December 8, 2022 letter provided support and analysis for raising the current modular cap (\$250,000) or eliminating the direct costs cap altogether (thereby allowing for all NIH-funded research to utilize the modular budget format). There has been a significant decline in the number of applications covered by modular budgets since implementation (90% in 1998 compared to 29% in 2021), and the modular cap has limited the ability to support fully all research activities in today's research environment. This is of particular concern within the context of Other recommendations being considered by NIH. Modular budgets are steadily squeezed in absorbing

increased activities, including activities for Data Management and Sharing\*\*\* and publishing costs. Increasing the modular budget cap or eliminating it together would allow researchers and institutions to account for the true costs of the project without hesitation or a need for tradeoffs to cover public access costs\*\*\*\*.

Costs Beyond the Award Period/Post-Grant Funding - One area of NIH's Plan to Enhance Public Access that requires additional clarification is recovery of scholarly publication costs that will occur after the close of a project. These costs include fees associated with storing data and costs for manuscripts published after the grant has ended. We recommend that NIH address how these costs will be covered to meet policy expectations, such as providing supplements to cover costs, including those that occur during a no-cost extension.

#### Repositories:

Reduce Burdens Associated with Scholarly Publication Deposits - The OSTP memo requires that scholarly publications are made available in agency-designated repositories. The NIH Public Access policy requires that scholarly publications be made available in PubMed Central. Some universities additionally require that publications be deposited into University repositories (i.e., eScholarship), and the best practices of some fields recommend discipline-specific repositories. Depending on the situation, a researcher may be required to deposit the same publication in four different places to comply with various policy requirements. Considering the associated administrative burden with meeting various requirements, efforts to centralize and automate deposits into a single point for researchers will reduce the burden. Further, there is a concern that publishers may shift their approach away from automatic deposits to charging fees to deposit. This will increase the associated costs and researcher burden and potentially cause noncompliance with NIH's public access policy. To help reduce this burden, NIH should consider the following: 1) assume a larger role in creating a single central federal repository for public access, and 2) clarify whether PubMed Central meets the OSTP requirement.

- \* May 6, 2020 Joint Association Letter to OSTP on Public Access RFI <a href="https://www.cogr.edu/sites/default/files/OSTP-RFI-Public-Access-AAU-APLU-COGR-formatted.pdf">https://www.cogr.edu/sites/default/files/OSTP-RFI-Public-Access-AAU-APLU-COGR-formatted.pdf</a>
- \*\*December 2022 NIH Modular Grant Application and Award Process Letter

  <a href="https://www.cogr.edu/sites/default/files/FINAL%20COGR\_Modular%20Tabak%20Letter%20November">https://www.cogr.edu/sites/default/files/FINAL%20COGR\_Modular%20Tabak%20Letter%20November</a>
  %202022%20%288002%29.pdf
- \*\*\*December 19, 2022

https://www.cogr.edu/sites/default/files/DMS\_COGR\_Policy\_Review\_Dec19\_final.pdf

\*\*\*\* https://www.aaas.org/sites/default/files/2022-10/OpenAccessSurveyReport Oct2022\_FINAL.pdf?utm\_label=&utm\_medium=twitter&utm\_source=social&utm\_campaign=AAAS

#### 2. Steps for improving equity in access and accessibility of publications.

We are encouraged by NIH's plan to continue making articles available in human and machine-readable forms to support automated text processing to improve the accessibility of publications. NIH should work with the community to develop procedural improvements to ensure that articles are broadly available through assistive devices.

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

We appreciate NIH acknowledging the importance of monitoring trends in publication fees and associated policies to ensure that they remain reasonable and equitable. As described above, we are concerned about an adverse shift in publication models that may increase costs and impact early-career researchers, researchers from institutions with limited resources, under-represented groups, and researchers without federally funded research. As such, it is important for NIH to monitor trends and act, if publishing fees increase to ensure that researchers do not face undue burdens to publish. In this regard, we recommend coordination across NIH units, including OSP and OPERA, to ensure efficient practices are developed that reduce burden. To monitor costs, we recommend that NIH perform an assessment to identify equitable funding models. We are particularly concerned that increased costs and burden may disincentivize researchers to publish, leading to a decreased number of publication outputs.

#### 4. Early input on considerations to increase findability and transparency of research.

COGR supports NIH's efforts to increase the findability and transparency of research by engaging through community dialogue for proposed policies related to PIDs and metadata. A particular area of importance is promoting cross-agency coordination to ensure consistency of agency plans and minimize compliance burden. We look forward to engaging with NIH further on this topic.

NIH should create template language, leveraging existing author addenda created by stakeholders and best practice organizations that may be utilized by researchers and institutions during the publication process to retain not only the right to publicly share an accepted manuscript but to create derivative works and to distribute the peer-reviewed manuscript under an open license even when publishing in a subscription journal (III.C.1). One example of this is the SPARC\*\*\*\*\* addendum. NIH should also consider encouraging licenses to permit sharing and reuse (i.e., Creative Commons and Other similar protocols) that enable broad circulation of scholarly publications. To maximize the impact of the above, NIH should consider what mechanisms and processes could be put in place to encourage researchers to use the provided template language and select less restrictive licenses. Finding ways to give researchers cause to use such resources would go a long way toward equitable compliance and ensuring the impact of funded research results is maximized.

\*\*\*\*\*https://sparcopen.org/our-work/author-rights/brochure-html/

**Uploaded File:** COGR-Response-to-NOT-OD-23-091.pdf

**Description:** Please see the attached letter for additional comments. We greatly appreciate the opportunity to comment and NIH's efforts.

Email: ktoups@cogr.edu



April 24, 2023

NIH Office of Science Policy 6705 Rockledge Drive, Suite 630 Bethesda, MD 20892

**Re:** Request for Information (RFI) on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research [NOT-OD-23-091]

Submitted Electronically to:

 $\underline{https://osp.od.nih.gov/nih-plan-to-enhance-public-access-to-the-results-of-nih-supported-research/}$ 

#### To Whom It May Concern:

COGR is an association of over 200 public and private U.S. research universities and affiliated academic medical centers and research institutes. COGR concerns itself with the impact of federal regulations, policies, and practices on the performance of research conducted at our member institutions. As recipients of a significant portion of NIH extramural research programs, COGR's members value the opportunity to respond to this request. The White House Office of Science and Technology Policy (OSTP) memo<sup>1</sup> sets forth requirements to increase access to publications and data resulting from federally funded research, and the NIH RFI NOT-OD-23-091 outlines NIH's plans to address this directive. As recipients of federally funded research, ensuring public access to publications and research data resulting from supported research is core to our mission as research universities and a responsibility we take seriously. We look forward to continuing to engage with the community and the agencies on this important topic and offer the following comments.

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators

To best ensure equity in publication opportunities for NIH-supported investigators, we offer the following comments and recommendations.

<sup>&</sup>lt;sup>1</sup> "Ensuring Free, Immediate, Equitable Access to Federally Funded Research" - <a href="https://www.whitehouse.gov/wp-content/uploads/2022/08/08-2022-OSTP-Public-Access-Memo.pdf">https://www.whitehouse.gov/wp-content/uploads/2022/08/08-2022-OSTP-Public-Access-Memo.pdf</a>

#### Costs

As stated previously<sup>2</sup>, ensuring public access to publications and research data resulting from federally funded research requires financial investments across the research enterprise. The 2022 OSTP Memorandum notably removes the 12-month embargo period, and while we understand and support the benefits of this policy change, we share in the community's expressed concerns about the potential for shifts in publishing models and increased costs with varying impacts depending on institutional characteristics. It is important that agencies plan accordingly to prevent any inequities.

Publication Cost – While NIH policy allows supported researchers to charge reasonable publishing costs against their awards, it is important to recognize that "reasonable costs" may not account for all costs or account for increased costs due to a shift in the publishing models. We share the community's concerns about the shift in the publishing model towards Article Processing Charges (APC), which is a significant fiscal and cultural change from subscription-based cost models. This shifting model forces universities to bear an ever-increasing proportion of the costs associated with publishing, including APC, subscription costs, and provision of uncompensated scholarly reviewers. Budgetary constraints may force institutions to make difficult choices about which faculty members to fund, and early career researchers, researchers from institutions with limited resources, and/or underrepresented groups may be disproportionately disadvantaged. Although NIH states in the RFI that APC may be charged to NIH grants, unless supplemental funds are provided, these charges will have a significant impact on the overall project budget. We hope that agencies and OSTP will directly address these concerns, and NIH should clearly state all APC, and other publishing costs should be budgeted accordingly in NIH grants and contracts.

**Modular Budget Caps** – We would like to direct NIH to COGR's recent letter<sup>3</sup> that addresses the limitations of modular budget caps. COGR's December 8, 2022 letter provided support and analysis for raising the current modular cap (\$250,000) or eliminating the direct costs cap altogether (thereby allowing for all NIH-funded research to utilize the modular budget format). There has been a significant decline in the number of applications covered by modular budgets since implementation (90% in 1998 compared to 29% in 2021), and the modular cap has limited the ability to support fully all research activities in today's research environment. This is of particular concern within the context of other recommendations being considered by NIH. Modular budgets are steadily squeezed in absorbing increased activities, including activities for Data Management and Sharing<sup>4</sup> and publishing

<sup>&</sup>lt;sup>2</sup> May 6, 2020 Joint Association Letter to OSTP on Public Access RFI - https://www.cogr.edu/sites/default/files/OSTP-RFI-Public-Access-AAU-APLU-COGR-formatted.pdf

<sup>&</sup>lt;sup>3</sup> December 2022 NIH Modular Grant Application and Award Process Letter <a href="https://www.cogr.edu/sites/default/files/FINAL%20COGR">https://www.cogr.edu/sites/default/files/FINAL%20COGR</a> Modular%20Tabak%20Letter%20November%202022%20%28002%29.pdf

<sup>&</sup>lt;sup>4</sup> December 19, 2022 https://www.cogr.edu/sites/default/files/DMS COGR Policy Review Dec19 final.pdf

costs. Increasing the modular budget cap or eliminating it together would allow researchers and institutions to account for the true costs of the project without hesitation or a need for tradeoffs to cover public access costs<sup>5</sup>.

Costs Beyond the Award Period/Post-Grant Funding – One area of NIH's Plan to Enhance Public Access that requires additional clarification is recovery of scholarly publication costs that will occur after the close of a project. These costs include fees associated with storing data and costs for manuscripts published after the grant has ended. We recommend that NIH address how these costs will be covered to meet policy expectations, such as providing supplements to cover costs, including those that occur during a no-cost extension.

#### **Repositories**

Reduce Burdens Associated with Scholarly Publication Deposits – The OSTP memo requires that scholarly publications are made available in agency-designated repositories. The NIH Public Access policy requires that scholarly publications be made available in PubMed Central. Some universities additionally require that publications be deposited into university repositories (i.e., eScholarship), and the best practices of some fields recommend discipline-specific repositories. Depending on the situation, a researcher may be required to deposit the same publication in four different places to comply with various policy requirements. Considering the associated administrative burden with meeting various requirements, efforts to centralize and automate deposits into a single point for researchers will reduce the burden. Further, there is a concern that publishers may shift their approach away from automatic deposits to charging fees to deposit. This will increase the associated costs and researcher burden and potentially cause noncompliance with NIH's public access policy. To help reduce this burden, NIH should consider the following: 1) assume a larger role in creating a single central federal repository for public access, and 2) clarify whether PubMed Central meets the OSTP requirement.

#### 2. Steps for improving equity in access and accessibility of publications.

We are encouraged by NIH's plan to continue making articles available in human and machinereadable forms to support automated text processing to improve the accessibility of publications. NIH should work with the community to develop procedural improvements to ensure that articles are broadly available through assistive devices.

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

<sup>5</sup> https://www.aaas.org/sites/default/files/2022-10/OpenAccessSurveyReport Oct2022 FINAL.pdf?utm label=&utm medium=twitter&utm source=social&utm c ampaign=AAAS

# Request for Information (RFI) on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research [NOT-OD-23-091]

We appreciate NIH acknowledging the importance of monitoring trends in publication fees and associated policies to ensure that they remain reasonable and equitable. As described above, we are concerned about an adverse shift in publication models that may increase costs and impact early-career researchers, researchers from institutions with limited resources, under-represented groups, and researchers without federally funded research. As such, it is important for NIH to monitor trends and act, if publishing fees increase to ensure that researchers do not face undue burdens to publish. In this regard, we recommend coordination across NIH units, including OSP and OPERA, to ensure efficient practices are developed that reduce burden. To monitor costs, we recommend that NIH perform an assessment to identify equitable funding models. We are particularly concerned that increased costs and burden may disincentivize researchers to publish, leading to a decreased number of publication outputs.

#### 4. Early input on considerations to increase findability and transparency of research.

COGR supports NIH's efforts to increase the findability and transparency of research by engaging through community dialogue for proposed policies related to PIDs and metadata. A particular area of importance is promoting cross-agency coordination to ensure consistency of agency plans and minimize compliance burden. We look forward to engaging with NIH further on this topic.

NIH should create template language, leveraging existing author addenda created by stakeholders and best practice organizations that may be utilized by researchers and institutions during the publication process to retain not only the right to publicly share an accepted manuscript but to create derivative works and to distribute the peer-reviewed manuscript under an open license even when publishing in a subscription journal (III.C.1). One example of this is the SPARC<sup>6</sup> addendum. NIH should also consider encouraging licenses to permit sharing and reuse (i.e., Creative Commons and other similar protocols) that enable broad circulation of scholarly publications. To maximize the impact of the above, NIH should consider what mechanisms and processes could be put in place to encourage researchers to use the provided template language and select less restrictive licenses. Finding ways to give researchers cause to use such resources would go a long way toward equitable compliance and ensuring the impact of funded research results is maximized.

#### **Additional Comments**

**Harmonization** – Policy harmonization across agencies is needed to incentivize researchers to engage in the open sharing of research outputs, assist institutions in compliance, and help maintain equity across funding agencies and researchers. One possible solution is the creation of more one-stop-shop access points for researchers that integrate grantee and funder operating procedures and

<sup>&</sup>lt;sup>6</sup> https://sparcopen.org/our-work/author-rights/brochure-html/

Request for Information (RFI) on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research [NOT-OD-23-091]

requirements. One illustrative example is the PASS System developed by Johns Hopkins University, which is making great strides in simplifying the reporting, sharing, and compliance components of federally funded research.

COGR appreciates the opportunity to comment on this Request for Information. Please contact Krystal Toups at <a href="mailto:ktoups@cogr.edu">ktoups@cogr.edu</a> if you have questions.

I am responding to this RFI: On behalf of an organization

Name: Christine Marie Battle

Name of Organization: American Association for Cancer Research

Type of Organization: Nonprofit research organization

Role: Institutional official

- 1. How to best ensure equity in publication opportunities for NIH-supported investigators.
- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

**Uploaded File:** 

NIH-RFI-NOT-OD-23-091.docx

#### American Association for Cancer Research Response to NIH RFI NOT-OD-23-091

The American Association for Cancer Research (AACR) welcomes the opportunity to comment on the Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research (NOT-OD-23-091). The American Association for Cancer Research (AACR) is the first and largest cancer research organization dedicated to accelerating the conquest of cancer. The mission of the American Association for Cancer Research is to prevent and cure all cancers through:

- Research
- Education
- Communication
- Collaboration
- Science Policy and Advocacy
- Funding for Cancer Research

Through its programs and services, the AACR fosters research in cancer and related biomedical science; accelerates the dissemination of new research findings; promotes science education and training; and advances the understanding of cancer etiology, prevention and early detection, diagnosis, and treatment throughout the world. At this writing, the AACR has more than 54,000 members residing in 130 countries and territories, including laboratory, translational, and clinical researchers; other health care professionals; and cancer advocates.

The AACR publishes 10 scientific journals, nine of which are hybrid and one of which is a gold open access journal. These journals span the breadth of cancer research and are written by and for members of the communities served by the AACR. The revenue generated by these journals represents one of three major revenue sources that are instrumental in helping the Association to fulfill its mission.

While we acknowledge the public access rules that were put forward by the White House Office of Science and Technology Policy (OSTP) in August of last year and appreciate that publishers will need to address the new requirements, we remain concerned that immediate, free public availability of the Author Accepted Manuscript will erode several existing business models on which the AACR publications—and many other scholarly association publications—depend for their financial health. During his term, a past AACR President often said, "No money, no mission." It is certainly the case that AACR initiatives and ongoing programs, many of which are focused on educating and training a diverse pipeline of future cancer researchers, depend in part on the revenues produced by the in-house publishing unit. Because the AACR is a self-publishing society, any surplus revenues are immediately invested back into the research communities that the Association serves. These investments take the form of research grants, travel awards, special conferences, the formation of task forces and workshops, among other activities.

Requiring deposit of the Author Accepted Manuscript signals that there is value placed on the peer review process. That said, quality publication requires quality peer review, which comes at a cost to the AACR, both in terms of staff time and in terms of underlying systems that support the peer review process. In other words, the Author Accepted Manuscript has already benefited from investments that publishers have made. Given the rapid proliferation of misinformation and disinformation and the lack of trust in the medical sciences, it is vitally important that publishers continue to invest in the peer review process. Perhaps the original manuscript, prior to peer review, would be suitable to fulfill the requirement of making federally funded research immediately freely available. This approach would be similar to the preprint pilot that PubMed Central has undertaken, and original manuscripts could be similarly marked as "not peer reviewed."

It would be enormously useful to have a dialogue regarding the infrastructure and costs required to produce a quality, scientific publication. For example, after final publication, publishers provide support for corrections and retractions that might be needed. This sometimes involves lengthy institutional investigations and can involve litigation in extreme cases. The goal is always to ensure that the published record is corrected as needed. Newer screening tools have helped to identify numerous errors prior to publication, but these tools come at a cost. Another example is the cost of ensuring that all content is backed up in a deep archive. This, in fact, is a MEDLINE preservation requirement and, again, publishers bear the cost.

The AACR also supports its authors and their critical research efforts in ways that continue to ensure the scientific integrity and reliability of the version of record. We would welcome a deeper discussion in this regard.

Publishers have long supported sustainable models that ensure broad access to content while also acknowledging that any model must respect the fact that there are real costs involved in the publishing process. We respectfully ask that the NIH undertake a true economic impact study that would include a view of the direct and indirect expenses incurred by the many small and medium-sized publishers that will be impacted by the 2022 OSTP Memorandum.

Respectfully submitted,

Christine Battle, Publisher/Vice President, Scientific Publications--on behalf of the AACR christine.battle@aacr.org

I am responding to this RFI: On behalf of an organization

Name: Agnes Balla

Name of Organization: University of California Office of the President

Type of Organization: University

Role: Institutional official

- 1. How to best ensure equity in publication opportunities for NIH-supported investigators.
- 2. Steps for improving equity in access and accessibility of publications.
- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

#### **Uploaded File:**

 $\hbox{UC-Comment-Letter-on-NIH-Public-Access-Plan\_final.pdf}$ 

**Description:** University of California system comment letter

Email: agnes.balla@ucop.edu

#### UNIVERSITY OF CALIFORNIA

BERKELEY • DAVIS • IRVINE • LOS ANGELES • MERCED • RIVERSIDE • SAN DIEGO • SAN FRANCISCO



SANTA BARBARA • SANTA CRUZ

OFFICE OF THE VICE PRESIDENT - RESEARCH AND INNOVATION

OFFICE OF THE PRESIDENT 1111 Franklin Street, 11th Floor Oakland, California 94607-5200

Submitted through: <a href="https://osp.od.nih.gov/nih-plan-to-enhance-public-access-to-the-results-of-nih-supported-research/">https://osp.od.nih.gov/nih-plan-to-enhance-public-access-to-the-results-of-nih-supported-research/</a>

April 24, 2023

Office of The Director National Institutes of Health 9000 Rockville Pike Bethesda, Maryland 20892

RE: UC Comments in Response to NOT-OD-23-091, "Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research"

Dear Sir or Madam:

I write on behalf of the University of California (UC) system responding to the Request for Information (RFI) on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research issued on February 21, 2023. The UC system is comprised of ten campuses, six academic health centers, and three affiliated U.S. Department of Energy national laboratories.

The UC is committed to cultivating open research practices and values public and immediate access to scholarly publications, data and code. This systemwide commitment is demonstrated by the Academic Senate and Presidential open access policies, the Faculty <u>Declaration of Rights and Principles to Transform Scholarly Communication</u>, and the university's work to transition away from subscription-based scholarly communications towards sustainable, open access publishing models. UC is also actively involved in the development of community-led open infrastructure for data sharing and scholarly journal publishing to further support open access to research results.

The UC system supports NIH's plan to remove the allowable embargo period that submitters may select when submitting articles to PubMed Central (PMC) under the current <u>Public Access Policy</u>. In our comments, we respond to the specific topics presented in the RFI as well as provide feedback on NIH's plan regarding persistent identifiers and metadata.

# 1. Strategies to ensure equity in publication opportunities for NIH-supported investigators

UC has identified two areas which significantly impact equity in publications opportunities for NIH-supported researchers:

NIH Plan to Enhance Public Access to the Results of NIH-Supported Research UC Comments April 24, 2023 Page | 2

**Burdens associated with multiple scholarly publication deposits**: Depending on the situation, a researcher may be required to deposit the same publication in multiple repositories to comply with institutional, funder, and federal policy requirements. This creates significant administrative burden in addition to posing compliance risks. We ask that NIH work with universities and publishers to automate multi-repository deposits as a single workflow to decrease burden on authors. Through automation, an author publishes their work as part of the normal publication process, and the article is then routed to the appropriate repositories to ensure high rates of policy compliance.

**Publication Cost**: It is critical that NIH continue to fund charges for open access publishing. These costs have equity implications, particularly for early career researchers working toward tenure who need publications but have limited funding costs for publications. It also disproportionately impacts researchers from institutions with limited resources and/or under-represented groups. For example, R2 universities may not be as successful identifying institutional funding sources to pay for open access if the grant cannot. UC asks that NIH prominently and consistently remind grantees to consider their publishing needs when finalizing their budgets. As a further step, NIH should provide supplemental funding for these costs, which often happen after an award has closed.

#### 2. Steps to improve equity in access and accessibility of publications

Accessibility Considerations for Color-Blind Audiences: Ensuring the accessibility of research for all readers, regardless of disability, should be a shared commitment among authors, funders, and publishers. Case in point: According to the American Academy of Ophthalmology, 8% of men and 0.5% of women are color blind. By those estimates, it would seem this affects some 13 million Americans. In addition, technologies like screen readers or other devices cannot necessarily correct this issue if figures and images in publications aren't otherwise made accessible. There are glasses and contact lenses which can provide correction, but these are not affordable to all who desire to read research articles and their associated figures. As such, NIH should consider including in their guidance for authors the requirement that figures to be accessible to a color-blind audience. This goal can be achieved simply by using symbols as well as colors to distinguish data sets on a graph, something already available in most software.

#### 3. Methods for monitoring evolving costs and impacts on affected communities

UC appreciates NIH acknowledging the importance of monitoring trends in evolving costs and associated policies to ensure that they remain reasonable and equitable. As described above, we are concerned about publication costs and the impact this may have on early career researchers, researchers from institutions with limited resources, and under-represented groups. We recommend NIH perform an assessment to identify equitable funding models. Along with any monitoring efforts, we ask that NIH act in a timely manner if publishing fees increase, offering solutions for researchers to appropriately fund publication costs. We also ask NIH to work to put pressure on publishers who charge excessive fees for publication.

UC also recommends that NIH provide guidance to the research community on how to recover the costs of scholarly publications that will occur after the close of a project. This guidance should

NIH Plan to Enhance Public Access to the Results of NIH-Supported Research UC Comments April 24, 2023 Page | 3

consider costs associated with storing data and costs for manuscripts published after the grant has ended. We recommend that NIH address how these costs will be covered to meet policy expectations, such as providing supplements to cover costs, including those that occur during a no cost extension.

Lastly, the research community needs infrastructure to support a thriving ecosystem of scholarly publications and sound data management practices. This type of infrastructure would be most effective if institutions (or a collection of institutions) created and maintained the infrastructure, rather than leaving researchers to individually try and build the infrastructure. Therefore, we strongly suggest that NIH explore ways to support institutions to build and scale needed infrastructure. NIH has provided this support in other policy areas, such as the SMART IRB platform designed to ease common challenges associated with initiating multisite research and to provide a roadmap for institutions to implement the NIH Single IRB Review policy.

# 4. Input on considerations to increase findability and transparency of research

Licenses for Sharing and Reuse: NIH should consider allowing licenses for permitting sharing and reuse (i.e., Creative Commons and other similar protocols) that enable broad circulation of scholarly publications. Depending on the license selected, researchers may translate an article to engage with international colleagues; collaborate to undertake large-scale computational analysis; and re-publish a report in a thematic collection. Such licenses typically grant permission for reuse alongside contractual obligations for attribution; even in cases where they do not, citation remains an expectation of research universities and societies.

To further encourage and promote a more equitable approach, NIH should consider the following:

- 1. Provide resources and template language that may be utilized by researchers and institutions seeking to retain not only the right to publicly share an accepted manuscript, but also to distribute this peer-reviewed manuscript under an open license even when publishing in a subscription journal (III.C.1).
- 2. Allow publication fees to be incorporated into grant budgets only in cases where a Creative Commons or other specified open license will apply (III.D.1). Where a license other than Creative Commons Attribution 4.0 International (CC BY) is used, applicants should be able to justify this choice.
- 3. Clarify in plain language the use of PMC content (noting examples of inappropriate uses, such as redistribution of PMC content for sale, as this may be allowed by licenses applied to submitted manuscripts (III.C.2)).

Lastly, NIH's existing <u>Plan for Increasing Access to Scientific Publications and Digital Scientific Data from NIH Funded Scientific Research</u>, published February 2016, states that "NIH is also exploring the possibility of using the government use license specified in 45 CFR 74.36 to help make papers public. Under these terms, the government has a royalty-free, nonexclusive and irrevocable right to reproduce, publish, or otherwise use the work for federal purposes, and to authorize others to do so." We ask that NIH share the results of that exploration with the research community.

NIH Plan to Enhance Public Access to the Results of NIH-Supported Research UC Comments April 24, 2023 Page | 4

# 5. Input on NIH's Plan regarding Persistent Identifiers and Metadata

Persistent identifiers (PIDs) and open metadata are crucial tools in enhancing research discoverability and transparency while saving time and resources, improving data quality, and generating valuable insights. However, to realize these benefits, widespread adoption is needed. While UC has seen growing awareness of and interest in PIDs, the speed of implementation remains slow, and stakeholders lack incentives to accelerate implementation. Stakeholders also lack concrete guidance on which PIDs to use. NIH has a unique opportunity to promote best practices with PIDs in their own systems and workflows, encourage broad adoption, and unlock new opportunities for discovery, insights, and innovation. As a concrete step towards these best practices, UC suggests that NIH recommend the following set of openly available core PIDs:

- Digital Object Identifiers (DOIs) for research outputs (articles, datasets, preprints, other works)
- ORCID IDs for researchers
- Research Organization Registry (ROR) IDs for research institutions
- Funder IDs or ROR IDs for research funders
- Grant IDs (a form of DOI) for research grants
- DMP-IDs (also a form of DOI) for data management plans

Thank you for the opportunity to comment on this important issue and we look forward to continued engagement on this issue as further policies and other guidance are developed. If you have any questions regarding UC's comments, please contact Agnes Balla, Director, Research Policy Analysis and Coordination, at <a href="mailto:Agnes.Balla@ucop.edu">Agnes.Balla@ucop.edu</a>.

Sincerely,

Deborah Motton, Ph.D.

**Executive Director** 

Research Policy Analysis and Coordination University of California, Office of the President

I am responding to this RFI: On behalf of myself

Name: Carrie Nelson and Cameron Cook

Name of Organization: University of Wisconsin-Madison

Type of Organization: University

Role: Institutional official

## 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

We encourage NIH to recognize that its policies are part of the global ecosystem of research and publishing. While NIH researchers are likely to comfortably be able to cover APCs or Other publishing charges with their funding, Other researchers with funding from Other agencies or without funding do not have the same advantage. While this is a great step in supporting all NIH researchers in making their work publicly accessible, it will continue to add to or potentially widen the gap with researchers without such funding. These discrepancies in resources affect where those researchers can publish, who will then read and cite them, and then subsequently their job, promotion, and funding prospects. This will continue to disadvantage early career researchers, historically excluded researchers, and researchers from Other countries across the globe.

NIH should consider working with publishers to negotiate a cap for, or Other model for managing and justifying, APC charges. While the NIH policy doesn't require a researcher to publish via gold or green OA, without any oversight publishers may take advantage of an unchecked market, those with a lack of literacy in publishing and public access policy terminology/processes, or identify ways to monetize reducing burden on the researchers via offering deposit in PMC. We are already seeing movements that suggest this is a real possibility - Springer Nature has already announced that only articles published openly will be deposited into PMC automatically.

An unchecked market could also add further stress to library budgets at research institutions. Library budgets are pressured to maintain existing purchases and subscriptions but are also beginning to be pressured to become a major mechanism for supporting publishing costs with read/write agreements, and subvention funds. This is especially challenging as library budgets serve and are set by their institutions, making it hard for libraries to both forecast and anticipate major changes or absorb extra costs. While libraries generally want to be a partner in shifting the mechanisms of open publishing, these costs continue to increase and compete with collection budgets, electronic resource budgets that continue to climb yearly, and Other staffing and resource costs. Researchers and libraries should not bear the brunt of the impact from this policy change.

There also continue to be concerns on costs -

If researchers do not publish until after close of project, they won't have funding for their publications

If costs for publishing continue to come out of the same budget as their Other research requests, they will continue to have to make unnecessary choices between their work and their options for publishing.

## 2. Steps for improving equity in access and accessibility of publications.

We encourage NIH to create standard language that authors are required to use when signing with publishers that allows them to retain their copyrights. SPARC and Other stakeholder and best practice organizations have existing author addenda language that is built in collaboration with Creative Commons and Other community members. Adapting SPARCs existing language would be best and encouraging use of it would enable easier compliance with this policy as this language requires that publishers provide a PDF final copy to the researcher who can then deposit the work. It also allows them to retain rights to make derivative works which would maximize the impact of federally funded research results. We also recommend NIH to consider encouraging researchers to use the least restrictive creative commons license possible and work with publishers to discourage use of the non-derivative (ND) clause of Creative Commons licenses.

### 3. Methods for monitoring evolving costs and impacts on affected communities.

High and unrestrained growing costs of publication fees negatively impact institutional and library spending on Other research resources.

NIH should identify opportunities for supporting early-career and historically excluded researchers. They should identify ways to work with associations or societies to connect with these populations and provide listening opportunities for feedback from those communities.

### 4. Early input on considerations to increase findability and transparency of research.

NIH should require researchers to both have and then use their ORCID ID when depositing in PMC. Encouraging the use and interoperability of standard and widely-adopted persistent identifiers will be beneficial for all stakeholders.

Email: <a href="mailto:carrie.nelson@wisc.edu">carrie.nelson@wisc.edu</a>

I am responding to this RFI: On behalf of an organization

Name: Anali Maughan Perry

Name of Organization: Arizona State University -ASU Library

Type of Organization: University

Role: Institutional official

## 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Researchers and authors have come to associate paying publishers an article processing charge with making their research open or publicly accessible, despite this being only one business model. We recommend the NIH make explicit that the NIH does not require authors to pay any fee in order to comply with the NIH policy. The manuscript submission option for publications is the most affordable and equitable compliance mechanism, since there is no additional cost for the investigator to deposit in PubMed Central. Rather, institutions and libraries can build in support for workflows to assist researchers with the deposit process across federal agencies.

The NIH could provide a service to researchers by providing clear language and processes that researchers could use when working with publishers, to ensure they retain the rights they need to make their final, peer-reviewed manuscript freely and openly available in PMC without an embargo period.

Additionally, incentives and rewards for researchers need to be adjusted to better reflect desired changes in behavior and practice. For example, continuing to reward and privilege publications in high impact journals will serve to reinforce the status quo. This presents a conflicting message to researchers, when they are faced with competing demands to make their work publicly available by their funder, but evaluated on publication in certain outlets at their institutions. The Higher Education Leadership Initiative for Open Scholarship (HELIOS) working groups are developing new paths to address this at the institution level, and the NIH should collaborate with HELIOS to harmonize these efforts to reward public access compliance.

## 2. Steps for improving equity in access and accessibility of publications.

As stated above, the NIH could greatly assist researchers by providing clear language and processes that researchers could use when working with publishers, to ensure they retain the rights they need to make their final, peer-reviewed manuscript freely and openly available in PMC and specifically addressing the rights needed to support automated text processing and improving accessibility. All publications resulting from NIH-funded research should carry open licenses to fully enable future use and reuse. A CC-BY license, or functional equivalent, is the best way to ensure equity in access and accessibility, not only through less-restrictive dissemination, but also by explicitly enabling adjustments to format to allow for computational analysis, text and data mining, and adapting to assistive technologies both now and in the future.

Without explicit licenses giving permission to modify and redistribute research, libraries must rely on copyright exceptions to fully remediate content to support universal design principles. Many

institutions do not have the copyright expertise to feel confident making these evaluative, and often ambiguous, decisions, resulting in risk-avoidant behavior.

Finally, at ASU Library, we are required to make case-by-case consultations with researchers on whether computational analysis can be used for any given database or vendor platform, which is time-consuming and inefficient. Policies and practices around permissions for text and data mining vary widely and are inconsistent, resulting in significant barriers to enabling twenty-first century research practices. Our efforts to advocate for changes in license terms to support our researchers results in small steps forward. The NIH could make a significant advance in this area by requiring open licenses that enable computational research and discovery.

## 3. Methods for monitoring evolving costs and impacts on affected communities.

Publishing models that require authors to pay an APC for journal publication present significant publication barriers for many researchers. The rising cost of APCs prove prohibitive to individuals and their institutions, resulting in fewer opportunities for publications. APC costs disproportionately affect younger researchers, female researchers, and those at less well-funded institutions, who are less likely to have secured research funding. APCs also require a diversion of funds away from the research process; investigators often must use money originally intended for materials and equipment, supporting postdocs, and for professional development opportunities.

We recommend that the NIH monitor costs associated with APCs to ensure that federal research dollars are being spent as intended on research and that the costs of publishing are not creating arbitrary barriers to entry for researchers. As previously stated, the NIH should be explicit that authors are not required nor expected to pay any publication fee in order to comply with the NIH policy.

## 4. Early input on considerations to increase findability and transparency of research.

Where possible, NIH should require the use of existing external identifiers (DOIs for data sets and DMPs, ORCIDs for publications, RORs for institutions, etc.) along with continued requirements for internal identifiers (PMCIDs, GeneBank Accession numbers, etc.). Having consistent and standard identifiers promotes adoption and interoperability, which makes workflows and systems less complicated for all stakeholders.

Email: anali.perry@asu.edu

I am responding to this RFI: On behalf of an organization

Name: Makyba Charles-Ayinde

Name of Organization: American Association for Dental, Oral, and Craniofacial Research

Type of Organization: Professional orgassociation

Role: Institutional official

## 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

The American Association for Dental, Oral, and Craniofacial Research (AADOCR) is the leading professional community for multidisciplinary scientists who advance dental, oral, and craniofacial research. We appreciate the opportunity to share our thoughts on the request for information on the National Institutes of Health (NIH) plan to enhance public access to the results of NIH-supported research. AADOCR recognizes and applauds the NIH's efforts toward providing public access to scholarly publications and data resulting from the research it supports. Further, the AADOCR commends the White House's Office of Science and Technology Policy (OSTP) efforts to ensure free, immediate, and equitable access to federally funded research in a manner that maintains scientific integrity and reproducibility of research. To respond to this request for comments, AADOCR engaged its Science Information Committee and its Board of Directors.

Increasing access to publications and data resulting from federally funded research offers many benefits to the scientific community and the public. However, there are costs associated with reviewing, editing, and publishing manuscripts that will need to be financed1. There are concerns that the publication costs from longstanding print journals may rise quickly and significantly for open-access articles. Publishing open access articles involves an open access publication fee (APC) which varies per journal2. Over the past few years, the cost for this option has increased rapidly. Significant fees for publication are now becoming normalized (apart from COVID-related articles). Several journals, such as Cell, eLife, Anatomical Record, and Nature Neuroscience, have notably increased their APC with costs reaching up to ~\$12,000 USD per manuscript3. Budgeting \$12-15,000 per year within a grant would be a substantial cost for investigators and may potentially affect the output of a researcher if they quickly exceed their publication allotment.

These high publication costs are especially challenging for early-career researchers who may feel greater pressure to publish their research more frequently, researchers within smaller institutions or organizations with limited resources, Historically Black Colleges and Universities, researchers who do not qualify or are not selected for grants to assist with APCs, and those utilizing micro or seed grants. Researchers at institutions with a student body < 10,000 students were three times as likely to find it very difficult to obtain funds for APCs as their counterparts at larger institutions, adjusting for gender, race, and length of time conducting research4. These prohibitive financial barriers may result in meaningful research going unpublished. Therefore, AADOCR supports a subsidized approach to the publication of open access articles where the funding institutions absorb a larger percentage of the APCs required to increase access to the articles. This will ensure that the NIH's Public Access Plan does not

result in scientists bearing the brunt of publishing costs through substantially higher fees passed on to them by journals.

AADOCR also supports a federally managed public registry for NIH funded studies to provide access to the results/data from these studies. The format of this registry may be similar to clinicaltrials.gov, and accessible to the general public. The public registry should provide a platform for all NIH funded researchers to deposit their results including unpublished negative data. Researchers will be required to include all experimental details and will be helpful to increase findability and transparency of research. It will also be helpful to include progress reports on available grants to ensure that analyses of studies without publications are publicly available. Although there is currently a Grantome interface, there are several challenges associated with that platform including difficulty navigating the interface (large number of unrelated or unwanted results), difficulty updating the result/publication section, and inability to include the researchers' website or data sharing links.

### 2. Steps for improving equity in access and accessibility of publications.

AADOCR supports reducing the knowledge gaps that exist with researchers and publication availability and access. Sharing information about publication availability with researchers/universities, organizations, and schools to increase awareness that these resources are available freely to them is critical to improving equity. Additionally, streamlining the NIH grant process and better publicizing mechanisms to access NIH funding that can substantially support APCs or waive publishing fees will benefit smaller institutions or early-career researchers without large grants.

Researchers from underserved populations, including early career researchers, those from historically excluded backgrounds, and those at less research-intensive institutions, do not have assured access to open access publication funds. Research has also shown gender disparities in funding for APCs as females were three times as likely to use grant funds to pay for APCs when compared to their male counterparts4. This diversion of funds comes at the expense of Other career advancement options such as professional development, equipment, and materials. This continues to further perpetuate disparity gaps in the biomedical workforce. Therefore, AADOCR supports NIH dedicating publishing resources for researchers from underrepresented populations and providing guidance to program officers on addressing equity in publication opportunities.

- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

AADOCR supports providing an option to filter the search by grant funding / IC / mechanism. This provides a utility to search by researcher and identify which publications are from the funded grants in PubMed or Other biomedical literature search engines.

#### **Uploaded File:**

AADOCR-Comments NIH-Plan-to-Enhance-Public-Access.pdf

Email: mcayinde@iadr.org



April 24, 2023

Lyric Jorgenson, PhD
Acting Director, Office of Science Policy
Acting NIH Associate Director for Science Policy
6705 Rockledge Drive, Suite 630
Bethesda, MD 20892 USA

Re: National Institute of Health Office of Science Policy Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research.

via website: <a href="https://osp.od.nih.gov/nih-plan-to-enhance-public-access-to-the-results-of-nih-supported-research/">https://osp.od.nih.gov/nih-plan-to-enhance-public-access-to-the-results-of-nih-supported-research/</a>

The American Association for Dental, Oral, and Craniofacial Research (AADOCR) is the leading professional community for multidisciplinary scientists who advance dental, oral, and craniofacial research. We appreciate the opportunity to share our thoughts on the request for information on the National Institutes of Health (NIH) plan to enhance public access to the results of NIH-supported research. AADOCR recognizes and applauds the NIH's efforts toward providing public access to scholarly publications and data resulting from the research it supports. Further, the AADOCR commends the White House's Office of Science and Technology Policy (OSTP) efforts to ensure free, immediate, and equitable access to federally funded research in a manner that maintains scientific integrity and reproducibility of research. To respond to this request for comments, AADOCR engaged its Science Information Committee and its Board of Directors.

Increasing access to publications and data resulting from federally funded research offers many benefits to the scientific community and the public. However, there are costs associated with reviewing, editing, and publishing manuscripts that will need to be financed¹. There are concerns that the publication costs from longstanding print journals may rise quickly and significantly for open-access articles. Publishing open access articles involves an open access publication fee (APC) which varies per journal². Over the past few years, the cost for this option has increased rapidly. Significant fees for publication are now becoming normalized (apart from COVID-related articles). Several journals, such as *Cell*, *eLife*, *Anatomical Record*, and *Nature Neuroscience*, have notably increased their APC with costs reaching up to ~\$12,000 USD per manuscript³. Budgeting \$12-15,000 per year within a grant would be a substantial cost for investigators and may potentially affect the output of a researcher if they quickly exceed their publication allotment.

These high publication costs are especially challenging for early-career researchers who may feel greater pressure to publish their research more frequently, researchers within smaller institutions or organizations with limited resources, Historically Black Colleges and Universities, researchers who do not qualify or are not selected for grants to assist with APCs, and those utilizing micro or seed grants. Researchers at institutions with a student body < 10,000 students were three times as likely to find it very difficult to obtain funds for APCs as their counterparts at larger institutions, adjusting for gender, race, and length of time conducting research<sup>4</sup>. These prohibitive financial barriers may result in meaningful research going unpublished. Therefore, *AADOCR supports a subsidized approach to the publication of open access articles* where the funding institutions absorb a larger percentage of the APCs required to increase access to the articles. This will ensure that the NIH's Public Access Plan does not result in scientists bearing the brunt of publishing costs through substantially higher fees passed on to them by journals.

AADOCR also supports a federally managed public registry for NIH funded studies to provide access to the results/data from these studies. The format of this registry may be similar to <a href="clinicaltrials.gov">clinicaltrials.gov</a>, and accessible to the general public. The public registry should provide a platform for all NIH funded researchers to deposit their results including unpublished negative data. Researchers will be required to include all experimental details and will be helpful to increase findability and transparency of research. It will also be helpful to include progress reports on available grants to ensure that analyses of studies without publications are publicly available. Although there is currently a Grantome interface, there are several challenges associated with that platform including difficulty navigating the interface (large number of unrelated or unwanted results), difficulty updating the result/publication section, and inability to include the researchers' website or data sharing links.

## Steps for improving equity in access and accessibility of publications

AADOCR supports reducing the knowledge gaps that exist with researchers and publication availability and access. Sharing information about publication availability with researchers/universities, organizations, and schools to increase awareness that these resources are available freely to them is critical to improving equity. Additionally, streamlining the NIH grant process and better publicizing mechanisms to access NIH funding that can substantially support APCs or waive publishing fees will benefit smaller institutions or early-career researchers without large grants.

Researchers from underserved populations, including early career researchers, those from historically excluded backgrounds, and those at less research-intensive institutions, do not have assured access to open access publication funds. Research has also shown gender disparities in funding for APCs as females were three times as likely to use grant funds to pay for APCs when compared to their male counterparts<sup>4</sup>.

This diversion of funds comes at the expense of other career advancement options such as professional development, equipment, and materials. This continues to further perpetuate disparity gaps in the biomedical workforce. Therefore, **AADOCR supports NIH dedicating publishing resources for researchers from underrepresented populations** and providing guidance to program officers on addressing equity in publication opportunities.

<u>Early input on considerations to increase findability and transparency of research</u> **AADOCR supports providing an option to filter the search by grant funding / IC / mechanism**. This provides a utility to search by researcher and identify which publications are from the funded grants in PubMed or other biomedical literature search engines.

# Support international collaboration and interoperability

AADOCR encourages NIH to support the compatibility of research platforms with existing global frameworks for sharing scientific knowledge and use common standards that are consistent with existing projects. For example, non-commercial open access platforms in Latin America, such as Redalyc, SciELO and AmeliCA, have provided software applications, interoperability, and discoverability to researchers. Similar platforms exist in Europe as well (e.g. Open Research Europe). The NIH should also take steps to encourage public-private collaboration to enhance interoperability between their platforms, reduce duplication of existing mechanisms, and allow for the repurposing of data for collaborative research.

AADOCR appreciates the opportunity to provide comments on the request for information on the NIH plan to Enhance Public Access to the results of NIH-supported research. AADOCR stands ready to work with NIDCR to flesh out mechanisms through which public access to publications and dental, oral, and craniofacial data can be increased.

If you have any further questions, please contact Dr. Makyba Charles-Ayinde, Director of Science Policy, at <a href="mailto:mcayinde@iadr.org">mcayinde@iadr.org</a>.

Sincerely,

Christopher H. Fox, DMD, DMSc

Chief Executive Officer

Alexandre Viera, DDS, MS, PhD

President

<sup>1</sup>van Mil, J.W.F. (2019). Open Access, At What Costs? *Int J Clin Pharm* **41**, 385–386. <a href="https://doi.org/10.1007/s11096-019-00806-6">https://doi.org/10.1007/s11096-019-00806-6</a>

<sup>2</sup>Nature Portfolio. (2022). Fees for Publishing in an "Open Choice" Journal. Retrieved from: <a href="https://support.nature.com/en/support/solutions/articles/6000137677-fees-for-publishing-in-an-open-choice-journal.">https://support.nature.com/en/support/solutions/articles/6000137677-fees-for-publishing-in-an-open-choice-journal.</a>
<sup>3</sup>Du Jingshan S (2022). Opinion: Is Open Access Worth the Cost? Retrieved from: <a href="https://www.the-scientist.com/critic-at-large/opinion-is-open-access-worth-the-cost-70049">https://www.the-scientist.com/critic-at-large/opinion-is-open-access-worth-the-cost-70049</a>

<sup>4</sup>American Association for the Advancement of Science. (2022). AAAS Survey: Many Researchers Face Difficulty Paying Open Access Fees. Retrieved from: <a href="https://www.aaas.org/news/aaas-survey-many-researchers-face-difficulties-paying-open-access-fees">https://www.aaas.org/news/aaas-survey-many-researchers-face-difficulties-paying-open-access-fees</a>.

I am responding to this RFI: On behalf of an organization

Name: Dan Valen

Name of Organization: Figshare (Digital Science)

Type of Organization: Other

**Type of Organization-Other:** Research Software

Role: Member of the public

### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

At Figshare, we believe equity in publishing begins with openness and transparency. One way to strengthen the NIH Public Access policy and create more equity in publication opportunities could be to require or encourage NIH-supported investigators to submit their pre-peer reviewed publications to a preprint server (such as the ones listed by ASAPbio: https://asapbio.org/preprint-servers).

Posting preprints to a preprint server increases the visibility of research more quickly than publishing in traditional journals and also provides broader exposure. Preprints have the added benefit of reaching those with and without access to expensive journals or journal databases and removes the need for researchers to wait for a peer reviewed publication.

The NIH Plan for Scientific Data addresses many of the requirements needed to not only encourage investigators and NIH-supported researchers to make data underlying a publication available, but also to treat research data as a 'first-class' research object, in turn allowing researchers to build on pre-existing research. As part of this effort, it would be beneficial to index scientific data published in repositories in PubMed as well as to ensure links between publications and datasets in PubMed metadata.

One final aspect of the new Public Access Policy would be to consider developing criteria that ensure transparency and fairness in the selection and review of articles for publication in NIH-supported journals, regardless of the authors' affiliations or backgrounds.

## 2. Steps for improving equity in access and accessibility of publications.

One of the biggest barriers to access of articles across PubMed is the majority of content on the platform is in English. The language barrier can be limiting or seen as a limitation to research from non-English-speaking countries or researchers and users in the US whose first language is not English. With the removal of the 12-month embargo for all NIH-supported publications, content will be available early and as full text. It would be fantastic to explore ways to not only provide access to articles in human and machine-readable forms but also to provide multilingual support. In addition, continuing to refine the user interface so PubMed can be easily searched by non-experts, such as patients and members of the public, would facilitate greater access to this publicly funded content.

## 3. Methods for monitoring evolving costs and impacts on affected communities.

There are a few ways to monitor evolving costs and the subsequent impacts on affected communities. Namely, the NIH could conduct regular surveys or questions around publication fees from publishers to

ensure they remain reasonable and equitable. Similarly, the NIH can consult with stakeholders and community members to help identify emerging issues related to publication fees from publishers.

The NIH can also look to analyze trends and data related to publication fees of NIH-sponsored research and publications to help inform future policies and provide a benchmark for the impact of making research openly available.

The benefits of open science are widely seen as a positive contribution to both the research process and wider society as a whole. One topic that is often not addressed is the economics of open data -- namely, how can we ensure that sustainable data sharing practices are accessible and equitable for researchers across diverse fields, institutions, and geographic regions? Tracking the compliance and costs not only of the publication but also of the accompanying data should be under consideration. Researchers are now encouraged to plan and budget for "funders for data management and sharing activities" and institutions are also increasingly providing resources for these efforts. NIH programs could track how funds are budgeted and spent on these activities and also survey academic institutions that receive NIH funding on the resources they are providing to meet these needs.

### 4. Early input on considerations to increase findability and transparency of research.

To increase findability and transparency of research outputs, the NIH may consider adopting a standardized system of PIDs, metadata, and specific ontologies across research objects to make content more discoverable and linkable across platforms and repositories outside of the NIH. The Generalist Repository Ecosystem Initiative (GREI) is currently addressing this among generalist data repositories, but it would be helpful to reinforce this work and promote the adoption of PIDs and standard metadata by researchers, institutions, discipline-specific repositories, and publishers to increase the interoperability and accessibility of research content. PIDs and metadata should also leverage existing community standards and initiatives to increase standardization such as the DataCite metadata schema, ORCID and ROR identifiers, and Make Data Count metrics. PIDs and metadata that allow for easy tracking and linkage with specific NIH funding sources (grants, awards, contracts) would be especially valuable for the research output community to incorporate as a common standard and support linking research outputs to funding sources, which would also facilitate tracking of data sharing and public access at the NIH and institutional levels.

The NIH should continue to consult with stakeholders and community members to identify any specific issues or use cases related to PIDs and metadata that may need to be addressed to improve the use and adoption of these tools.

Email: dan@figshare.com

I am responding to this RFI: On behalf of an organization

Name: Angela Cochran

Name of Organization: American Society of Clinical Oncology

Type of Organization: Professional org association

Role: Medical provider

## 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

ASCO and the NIH-funded researchers within our membership are concerned that if a zero-embargo green open access model (whereby manuscripts accepted for publication in journals that report on NIH funded research are made available on PubMed Central immediately upon

publication) is no longer available due to journals having to convert to author-pays gold open access models, underfunded researchers will be shut out of publishing their research in journals. This will create more disparities in the research pipeline and in evidence-based care that puts all patients at risk.

Looking ahead, we are concerned that in an environment where the majority of publications require fees for publishing, manuscript output (the main driver used by researchers to show the impact of the funding) will decline. Grant awards, already often insufficient for the intended research, will also be unlikely to fully cover publication expenses, putting the funds needed to conduct research in direct competition with funds needed to publish research.

This unintended, though highly likely, scenario will have a disproportionate adverse impact on early career researchers, whose grants typically are smaller and have less room to accommodate expenses not directly applied to the research itself. These predictable impacts are the reason that ASCO urges the NIH to fully consider and account for the ramifications this proposed plan will have on all levels of grantees. (Please see full letter attached.)

## 2. Steps for improving equity in access and accessibility of publications.

ASCO's journals provide value to clinicians, researchers, institutions, and funders by facilitating

high-quality peer-review and integrity checks on all published materials. We request that the NIH refrain from applying broad re-use licenses to the PubMed Central deposited papers as it will have the unanticipated, undesired, and paradoxical effect of diminishing the quality of content made available to the profession and the public.

Instead, to preserve equity in publishing opportunities across our journals, we are committed to providing a green open access model for as long as financially sustainable. However, if a zero-embargo green policy is coupled with broad re-use rights, we will not be able to afford to maintain a green route for author compliance. (Please see full letter attached.)

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

Certain grants do not permit use of funds for publication fees. Therefore, ASCO recommends that NIH exempt certain types of infrastructure-related grants (e.g., cancer center support grants, TSAs, NCORPs) and teaching grants (K awards, T awards) from reporting funding to journals and thus requiring deposit.

The broad reach and impact of this proposed plan will be a challenge to implement and enforce if

compliance is mandated for all NIH funded authors regardless of how much funding they received or how small a role any given individual plays in a research project or manuscript. The NIH should instead apply a minimum threshold of funding and/or level of participation by authors and researchers before subjecting the papers to the proposed mandate.

We encourage the NIH to publish clear guidance, on which circumstances qualify submitted papers to claim NIH funding, and on the conditions that invoke a requirement to comply with the public access mandate. More and better communication to grantees and Other stakeholders regarding expected compliance is essential with the planned zero embargo policy. (Please see full letter attached.)

### 4. Early input on considerations to increase findability and transparency of research.

Please see full letter attached.

#### **Uploaded File:**

ASCO-Comments-NIHopenaccRFI04242023-signed4-24-2023revised.pdf

**Description:** ASCO NIH RFI Comments

Email: angela.cochran@asco.org



#### **CHAIR OF THE BOARD**

Everett E. Vokes, MD, FASCO

#### **PRESIDENT**

Eric P. Winer, MD, FASCO

#### PRESIDENT-ELECT

Lynn M. Schuchter, MD, FASCO

#### PAST PRESIDENT

Lori J. Pierce, MD, FASTRO, FASCO

#### **TREASURER**

Elizabeth A. Mittendorf, MD, PhD, MHCM, FASCO

#### TREASURER-ELECT

Elizabeth R. Plimack, MD, MS

#### **DIRECTORS**

Frederick Chite Asirwa, MD, MBChB Ethan M. Basch, MD, MSc, FASCO Lisa A. Carey, MD, FASCO

Mariana Chavez Mac Gregor, MD,

MSc, FASCO
Tara O. Henderson, MD,

MPH, FASCO

Carolyn B. Hendricks, MD, FASCO

Taofeek K. Owonikoko, MD, PhD

Ann H. Partridge, MD, MPH, FASCO

Enrique Soto Pérez De Celis, MD, MSc, PhD, FASCO

Gladys I. Rodriguez, MD

Michael A. Thompson, MD, PhD. FASCO

Karen M. Winkfield, MD, PhD. FASCO

#### **EX-OFFICIO MEMBERS**

CHAIR, CONQUER CANCER BOARD

Howard A. Burris III, MD, FACP, FASCO

### CHIEF EXECUTIVE OFFICER

Clifford A. Hudis, MD, FACP, FASCO

#### **Via Electronic Submission**

April 24, 2023

Lawrence A. Tabak, DDS, PhD Acting Director National Institutes of Health 9000 Rockville Pike Bethesda, MD 20892

Response to NIH Request for Information – "The NIH Plan to Enhance Public Access to the Results of NIH-Supported Research" (NOT-OD-23-091)

Dear Dr. Tabak:

Thank you for the opportunity to respond to the Request for Information (RFI) on the National Institutes of Health (NIH) Plan to Enhance Public Access to the Results of NIH Supported Research. The American Society of Clinical Oncology (ASCO) shares the NIH goal of making practice-changing research available quickly and in a way that is sustainable and maintains the high bar for publishing practice-changing research.

Our comments focus on three key points within the RFI questions:

- Research integrity must be protected to advance quality patient care.
- NIH policy should not continue or worsen existing disparities, including by creating barriers to scientific publishing activities for underfunded researchers.
- NIH policies should be compatible with a sustainable future for scientific journals to maintain the integrity of the peer-review process.

ASCO has a long history of serving as an authoritative resource for oncology professionals, general practitioners, patients, and caregivers and of advancing patient care. We record, contextualize, and disseminate practice-changing information through our journals, widely attended conferences, educational resources, and through multiple professional communication channels.

Our clinical journals expediently publish thoroughly peer-reviewed research that has profound impact on patients with cancer and their families. Through



our journals we also disseminate clinical practice guidelines that inform treatment decisions, identify research gaps, influence insurance design, drive hospital and clinical operations, and—ultimately—impact and improve public health.

Diligent peer review, management and public disclosures of conflicts of interest, and data and figure integrity checks are vital parts of a responsible publication process. Threats to the integrity of the content, such as plagiarism, paper mills, inappropriate AI-generated content, and fraudulent data, are always present and require steady attention. While no system is perfect, peer-review increases the opportunity to mitigate these risks and protect the general public from ensuing harms.

Indeed, peer-review and the broader publication process are critical to the production of a final product that researchers and clinicians can rely upon as they conduct vital research and deliver evidence-based care.

All of this requires resources that are likely to be endangered if publishers lose the revenue that currently sustains this work. Such losses could occur in the form of cancelled subscriptions, insufficient total article processing charge (APC) income, and lost licensing fees for approved reuse of content, among others. Each publisher will have their own budgetary tipping point when decreased revenues force a decision to discontinue vital services now protecting the integrity of research published in our journals, but all will face this challenge and all will have to make cost-saving changes to maintain a viable publishing enterprise.

There are additional challenges to the viability of quality publishing that will also risk increasing disparities. For example, ASCO and the NIH-funded researchers within our membership are concerned that if a zero-embargo green open access model (whereby manuscripts accepted for publication in journals that report on NIH funded research are made available on PubMed Central immediately upon publication) is no longer available due to journals having to convert to author-pays gold open access models, underfunded researchers will be shut out of publishing their research in journals. This will create more disparities in the research pipeline and in evidence-based care that puts all patients at risk.

Looking ahead, we are concerned that in an environment where the majority of publications require fees for publishing, manuscript output (the main driver used by researchers to show the impact of the funding) will decline. Grant awards, already often insufficient for the intended research, will also be unlikely to fully cover publication expenses, putting the funds needed to conduct research in direct competition with funds needed to publish research.

This unintended, though highly likely, scenario will have a disproportionate adverse impact on early career researchers, whose grants typically are smaller and have less room to accommodate expenses not directly applied to the research itself. These predictable impacts are the reason that ASCO urges the NIH to fully consider and account for the ramifications this proposed plan will have on all levels of grantees.



There are other concerns to consider. Certain grants do not permit use of funds for publication fees. Therefore, ASCO recommends that NIH exempt certain types of infrastructure-related grants (e.g., cancer center support grants, CTSAs, NCORPs) and teaching grants (K awards, T awards) from reporting funding to journals and thus requiring deposit.

The broad reach and impact of this proposed plan will be a challenge to implement and enforce if compliance is mandated for all NIH funded authors regardless of how much funding they received or how small a role any given individual plays in a research project or manuscript. The NIH should instead apply a minimum threshold of funding and/or level of participation by authors and researchers before subjecting the papers to the proposed mandate.

We encourage the NIH to publish clear guidance, on which circumstances qualify submitted papers to claim NIH funding, and on the conditions that invoke a requirement to comply with the public access mandate. More and better communication to grantees and other stakeholders regarding expected compliance is essential with the planned zero embargo policy.

Lastly, ASCO's journals provide value to clinicians, researchers, institutions, and funders by facilitating high-quality peer-review and integrity checks on all published materials. We request that the NIH refrain from applying broad re-use licenses to the PubMed Central deposited papers as it will have the unanticipated, undesired, and paradoxical effect of diminishing the quality of content made available to the profession and the public. Instead, to preserve equity in publishing opportunities across our journals, we are committed to providing a green open access model for as long as financially sustainable. However, if a zero-embargo green policy is coupled with broad re-use rights, we will not be able to afford to maintain a green route for author compliance.

We thank you for this opportunity to share our comments on the proposed plan and look forward to working with you to assure policies that sustain reliable, equitable, high quality scientific content. Please contact Angela Cochran at <a href="mailto:angela.cochran@asco.org">angela.cochran@asco.org</a> or Shimere Williams Sherwood at Shimere.Sherwood@asco.org with any questions and for further discussions.

Sincerely,

Eric P. Winer, MD, FASCO

President, American Society of Clinical Oncology (ASCO)

I am responding to this RFI: On behalf of myself

Name: Jonathan Saunders

Name of Organization: UCLA, Department of Neurology

Type of Organization: University

Role: Scientific researcher

## 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

The steps towards openness in the 2022 OSTP Memorandum and subsequent notices like NOT-OD-23-091 are admirable steps to use the power of the NIH as a funding body to set standards for equity in public research. The proposals as written seem to be "fighting the last war," however, focused on closed-access publication without considering the significant shift in market structure as traditional scientific publishers have transformed into data brokers.

It is impossible to ignore the role of for-profit academic publishers as a primary source of inequity when considering these policies -- without their prior model of subscription-based access, there would be no need for these policies at all. We cannot play coy and pretend to be market neutral when considering how scientific publishing should work: for-profit scientific publishing, now largely an oligopoly owned by a handful of information conglomerates, is an ethical catastrophe, and if we intend to grasp at the root of the problem we need to contend with the ways their business models distort the practice of science at every stage.

The publishing oligopoly has had ample time to prepare for the shoe of universal open access to drop, and if their shareholder-facing communications are any indication, they have already fully accounted for it and adapted their business models accordingly. They have been focused heavily on shifting their default strategy from subscription-based publication to author-pays APC-driven open access, as this proposal tacitly endorses. This model is \*intrinsically inequitable,\* as it is explicitly designed to shift the burden of payment from libraries to individual researchers, and more closely align the cost of publication with the benefits accrued through the prestige associated with a journal brand. At the point when (1) there is \*any\* gradient of APCs such that high-prestige journals like Nature and Cell have a higher cost, and (2) publications in high-prestige journals are a necessity for grant funding and promotion, the system is fundamentally inequitable. Worse, by atomizing the ability to negotiate with publishers, shifting from libraries and library consortia to individual researchers, we neutralize the power of some of the few organizations capable of pushing back against the for-profit publishers by embracing a positive feedback loop where researchers have every incentive to slide the slippery slope of rising APCs in order to retain their employment.

If this proposal leaves the for-profit publishing apparatus largely intact, it will enter the history of half-measures made in deference to the publishing oligopoly that leave the problem perpetually unsolved. One can only imagine the state of every field of research from pharmaceuticals to astrophysics if we had the courage in 1999 to implement the full version of Harold Varmus' vision for PubMed Central, displacing for-profit publishing entirely with free to publish, free to read research as the norm. What could the world be like if we had 20 years of experimenting with open research dissemination, rather

than spending the dawn of the information era hobbled by broken systems accessible to a vanishingly small and privileged few? Will we be looking back in anOther 20 years wishing we had the courage to end for-profit publishing now?

The very framing of this RFI as being focused on open access publication rather than the infrastructure of our communication demonstrates that we are missing the implications of the shift in the business models of the major for-profit publishers towards "surveillance publishing." The next era of scholarly communication battles will be about \*infrastructure.\* Profit models are consolidating around collecting user data and repackaging it into bibliometrics and informatics platforms like so-called "research intelligence" tools like RELX's SciVal. With the requirement for open data, we will face anOther period of enclosure where there is a less clear distinction between publishing, data sharing, and computation. As written, the NIH would directly create a new triple-pay system in the very policy that is intended to address the prior one: if NIH's STRIDES project is the intended model, NIH pays cloud providers for discounts so that researchers can pay to archive their data as well as pay to export it.

The infrastructure of scientific communication is a fraction of the complexity of that which will be required for universal open data: it is trivial to start a new journal-like website, it is not so trivial to create a new server farm for storing bulk data. The inequity from APCs will be orders of magnitude greater as the process of science congeals into a series of pay-to-use platforms that skim public funding at every stage from grant proposal through data collection, analysis, and publication. The NIH discusses monitoring funding inequity for publication, but is it prepared to handle the broader inequities from the capture of research information infrastructure by a handful of cloud platform giants? Who, exactly, will have the funding necessary to pay for tools that produce clean data, to hire the data scientists to manage it, and to pay the costs of cloud storage and computation? Plainly, the NIH stands to slice off an increasing fraction of its budget to orbiting information rentiers rather than directly funding research, and the dream of universal information access will always be out of reach beyond some exorbitant hosting bill.

The landscape of options that would truly make a more equitable and robust scientific process is wide open, and all of them mean taking a meaningful stand in favor of a public information commons and against for-profit private ownership of information infrastructure. Rather than a single recommendation, I urge the NIH to reorient this and future proposals towards a nonprofit, publicly-owned informational commons. Requiring that all publishers must be operated as nonprofits is one first step. A fixed and decreasing cap on APCs to sunset pay-to-publish models in favor of so-called "diamond" open access is anOther. Publishing venue-agnostic grant decisions are anOther. Addressing the next generation of infrastructure needs equitably requires that we look beyond the "Platform as a Service" model articulated in NIH's 2018 strategic plan for data science where public research bodies outsource and rent basic infrastructure from cloud providers. A full technical evaluation is of course out of scope of this RFI, but a system of peer to peer infrastructure that can leverage resources from individual computers through institutional and federal systems without dependence on cloud providers would be capable of addressing inequity as well as realizing the ambitions of information access articulated in these proposals.

I and Others have written elsewhere and are working on these systems.

2. Steps for improving equity in access and accessibility of publications.

The greatest hindrance to accessibility of scientific publications is not technical (though the ailing infrastructure of the traditional publications is some decades behind the rest of the web), but the socio-economic construct of traditional journals themselves. The form of the scientific journal article is entirely unlike how the vast majority of non-scientists interact with information, and is structured by an industry that maintains its profit by strategically suppressing semantic organization in favor of using journal brands as the primary organization principle to maintain the effect of their prestige. It is prestigious to publish in Nature because people will read it. People read Nature papers because there are no effective means of finding research based on its content, leaving scientists to organize dissemination in ad-hoc media like Twitter or be dependent on downstream patches like Google Scholar.

If the NIH is serious about making scientific research more accessible to non-scientists, it must address the ways that research incentives uniformly encourage publication of impenetrable prose in domain- or prestige-limited venues in favor of promoting alternative means of organizing scientific communication, including peer review and publication. We need to not only make it easier for everyone to make sense of the scientific record, we must also reckon with how our incentive structures cause the scientific record to be so difficult to make sense of in the first place.

Accessibility for people that need assistive technologies can \*only be helped\* by taking more direct control over our infrastructures of communication. Rather than being beholden to the structure imposed by journals, we should directly address the technologies and social systems that structure scientific communication as part of a holistic project of information accessibility.

## 3. Methods for monitoring evolving costs and impacts on affected communities.

If the NIH agrees to step in and offset exorbitant APCs in prestige journals in the name of equity, particularly without clear language about what counts as a "reasonable" cost, it sends the message that it is willing to pay any price that the publishers demand. The framing of monitoring evolving costs indicates that the NIH is aware that this policy will increase publication costs, and those increases will inequitably affect researchers outside of the highest echelons of funding and prestige. We do not need to accept this as an inevitability --- there are multiple routes towards explicitly avoiding an APC-driven publishing market, and towards creating a peer to peer data infrastructure that avoids outsized cost burdens for marginalized researchers.

## 4. Early input on considerations to increase findability and transparency of research.

It is critical to understand the history of PIDs and how they structure and reinforce the for-profit publishing system, advantaging larger players and disadvantaging independent alternatives. The DOI system itself was created in response to NIH's 1999 push for PubMed Central in order to preserve the publishing industry's dominance in assigning identifiers --- and thus what can be counted as research. The decades of research on persistent identifiers show that decentralized alternatives like the ARK or IPFS's CID work, and we should prioritize identifiers that can be created and structured by any researcher, rather than controlled by a centralized authority. Critical research on ontologies and metadata also show their intrinsically political nature, which also points towards tooling to express metadata rather than the current approach taken by NIH's Biomedical Translator project of creating quasi-universal ontologies to be mapped onto.

I am available for further comment on this and the rest of the responses to this RFI, and I appreciate any time taken to read this.

**Description:** The NIH should directly oppose a for-profit APC-driven publication system and cloud research infrastructure, and instead focus efforts on building truly public information infrastructures.

Email: j@nny.fyi

I am responding to this RFI: On behalf of an organization

Name: Cable Green

Name of Organization: Creative Commons

Type of Organization: Nonprofit research organization

Role: Institutional official

## 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

- We note that the manuscript submission option for publications is the most affordable and equitable compliance mechanism as it is free for the investigator to deposit in PubMed Central (PMC).
- We recommend that NIH state explicitly that there is no charge for complying with NIH's policy.
- Depositing a final peer reviewed manuscript in PMC is free of charge, legal, and ensures that the researcher is in full compliance with the NIH Public Access policy.
- Any fees charged by a publisher are for publication in that particular journal not for compliance with NIH's Public Access Policy.
- NIH should offer clear language and processes that investigators can use to retain rights to make their peer reviewed manuscript freely available and fully reusable post-publication in PMC without an embargo period. Specific instructions for doing this effectively and emphasizing that it is compliant with copyright and journal policies will help authors comply with the policies, make federally-funded research reusable, and further support NIH's goal to ensure equity in publishing.
- NIH should encourage the use of publication channels that do not present financial (or Other) barriers to researchers, including non-Article Processing Charge (APC) supported open access journals, preprint servers, and Other emerging community-driven options. Journal business models requiring authors to pay a fee for journal publication present significant publication barriers for many researchers.
- The NIH should work with the higher education community to align research assessment and career advancement incentives (e.g., promotion and tenure) to support scientific publication models that actively promote equity in publication opportunities.

## 2. Steps for improving equity in access and accessibility of publications.

- The OSTP Nelson memorandum asks agencies to "make federally funded publications, data, and Other such research outputs and their metadata...findable, accessible, interoperable, and reusable, to the American public and the scientific community in an equitable and secure manner." To fulfill the reusability requirement, all publications resulting from NIH-funded research should carry standard international open licenses, and NIH or authors should explicitly retain the rights needed to authorize those open licenses.
- Placing the most current version of the CC BY license or its functional equivalent on a publication is the best way to ensure that publications can be freely accessed and fully reused.

- Open licensing advances research, enables text and data mining to foster further scientific investigation, educational reuse, translations into Other languages, and computational uses, as well as use of content on assistive devices.
- NIH should offer clear language that investigators can use to specifically retain rights to make their final peer-reviewed manuscript freely available and fully reusable (under the CC BY license or its functional equivalent) post-publication in PMC without an embargo period.

### 3. Methods for monitoring evolving costs and impacts on affected communities.

- We note once again there is no cost for complying with the NIH Public Access Policy when using the manuscript submission option.
- However, NIH should be aware that models requiring authors to pay an Article Processing Charge (APC) fee for journal publication present significant publication barriers for many researchers.
- The rising cost of APCs often prove prohibitive to individuals and their institutions, resulting in fewer opportunities for publications. Studies have documented that APC costs disproportionately affect younger researchers, female researchers and those at less well-funded institutions.
- APCs also require a diversion of funds away from the research process; investigators often must use money originally intended for materials and equipment, supporting postdocs, and for professional development opportunities including presenting research results at conferences.
- We recommend that the NIH monitor costs associated with APCs with special attention to how costs are distributed along disaggregated data variables for different marginalized groups to ensure federal research dollars are being spent as intended on research and that the costs of publishing are not creating arbitrary barriers to entry for researchers, and the ultimate availability of publishing opportunities for researchers at traditionally underrepresented institutions and in less-well-resourced disciplines.
- The NIH should monitor the cost of APCs levied on its investigators. Data collection on the amount spent to publish NIH-funded research regardless of the source would increase transparency and insight into how these fees affect various communities including the potential impacts of publishing opportunities on traditionally underrepresented institutions and in less-well-resourced disciplines.

## 4. Early input on considerations to increase findability and transparency of research.

- Ensuring that the results of NIH-funded research along with metadata containing information about who conducted the research, where it was done, and with what resources is an important component of the NIH Public Access Plan. This requires NIH to articulate clear expectations about the use of Persistent Identifiers (PIDs) throughout the research process.
- Where possible, NIH should require the use of existing external identifiers (DOIs for publications, data sets, and DMPs, ORCIDs for researchers, RORs for institutions, etc.) along with continued requirements for internal identifiers (PMCIDs, GeneBank Accession numbers, etc.)
- Because similar identifiers will be required to be used by all federal agencies as a result of the OSTP Memorandum, NIH should coordinate its efforts with Other participants in interagency working groups,

including the National Science and Technology Council's (NSTC) Subcommittee on Open Science, to identify best practices and potential standards.

- NIH should also consider collaboration with a standards body, such as the National Information Standards Organization (NISO), to help develop a set of standards and framework for a national PIDs strategy to facilitate smooth implementation.

-----

Creative Commons thanks the NIH for updating its policy to eliminate embargoes, improve data sharing and enhance reuse rights to publicly funded research outputs. Openly-licensed research accelerates the pace of discovery, reduces information sharing gaps, encourages innovation, and promotes reproducibility. We appreciate the opportunity to comment on this draft plan, and we are eager to assist in its eventual rollout.

Email: <a href="mailto:cable@creativecommons.org">cable@creativecommons.org</a>

I am responding to this RFI: On behalf of an organization

Name: James C. Appleby

Name of Organization: The Gerontological Society of America

Type of Organization: Professional org association

Role: Institutional official

## 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

As a critical component of its public access plan, we urge the NIH to focus on creating an environment that balances reader access to published work with researchers' ability to publish. This will require transparency and recognition of the costs borne by researchers and research funders. We must strive to create a system wherein scientists are not required to pay additional fees to publish and where grants are not required to bear the brunt of publishing costs. Otherwise, we risk creating heavy cost burdens not only for researchers and their institutions, but also for funders of research, including taxpayers.

Rigorous peer review and expert editorial efforts ensure GSA continues to be a trusted, reliable, and credible source for scientific publications in gerontology and geriatrics. Through a publishing agreement with a scholarly publisher, GSA can reinvest the revenue from the GSA journals to ensure the quality of the publications.

## 2. Steps for improving equity in access and accessibility of publications.

Scientists' ability to communicate their scientific results through publication is critical to the incorporation of their expertise into the scientific enterprise and the progression of their careers. Monitoring implementation of changes to the public access policy, and how researchers and institutions pay publishing costs, will be critical to ensuring that public access plans do not create new systemic inequities or reinforce existing ones. Careful and continued study will be essential for understanding the near- and long-term effects of related changes. Study of cost effects at the researcher, institution, and enterprise levels is needed. It may also be valuable for NIH to survey researchers and institutions about publishing costs and about tradeoffs made to pay such costs.

## 3. Methods for monitoring evolving costs and impacts on affected communities.

Adaptation of federal grant agreements to require reporting on the payment of publication fees and reliance on transformative agreements (in instances where authors avoid payment of a fee because their institution has a transformative agreement with their journal of choice) represents one logical approach to monitoring fees. All analyses of and reporting on costs paid by institutions or researchers for publication should examine potential variability in costs across disciplines, career stages, and institution types, as well as variability based on researcher backgrounds.

## 4. Early input on considerations to increase findability and transparency of research.

We are pleased that through our current publisher, GSA provides some or all metadata for all authors. GSA appreciates efforts underway such as the requirement for individuals supported by research training, fellowship, research, education, and career development awards to have Open Research and

Contributor Identifiers and exploring the use of the digital object identifier system. GSA looks forward to the opportunity to provide continued input as systems to increase findability and transparency of research are developed.

# **Uploaded File:**

230424-RFI-NOT-OD-23-091-NIH-Public-Access-GSA.pdf

**Description:** RFI Response

Email: pdantonio@geron.org



1220 L Street NW, Suite 901, Washington, DC 20005-4018 • 202-842-1275 • www.geron.org

April 24, 2022

Lawrence A. Tabak, DDS, PhD Acting Director National Institutes of Health

Re: Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-

Supported Research NOT-OD-23-091

Submitted electronically via https://www.regulations.gov/commenton/CMS-2022-0113-1871

Dear Director Tabak:

On behalf of The Gerontological Society of America (GSA), I write to thank the NIH for seeking public input on the "NIH Plan to Enhance Public Access to the Results of NIH-Supported Research" (NIH Public Access Plan). GSA appreciates NIH's efforts to update its Public Access Plan to meet the provisions of the 2022 OSTP Memoranda related to scientific data.

GSA is the oldest and largest interdisciplinary organization devoted to research, education, and practice in the field of aging. The mission of GSA is to cultivate excellence in interdisciplinary aging research and education to advance innovations in practice and policy. GSA's 5,400 members include gerontologists, physicians, nurses, pharmacists, social workers, behavioral & social scientists, biologists, demographers, economists, and many other disciplines. These experts study all facets of aging with a life-course orientation. The multidisciplinary nature of the GSA membership is a valued strength, enabling the Society to provide a 360-degree perspective on the issues facing all of us as we age. GSA publishes some of the longest-running, leading peer-reviewed international journals in gerontology and geriatrics.

GSA is pleased to offer the following comments:

1. How to best ensure equity in publication opportunities for NIH-supported investigators.

As a critical component of its public access plan, we urge the NIH to focus on creating an environment that balances reader access to published work with researchers' ability to publish. This will require transparency and recognition of the costs borne by researchers and research funders. We must strive to create a system wherein scientists are not required to pay additional fees to publish and where grants are not required to bear the brunt of publishing costs. Otherwise, we risk creating heavy cost burdens not only for researchers and their institutions, but also for funders of research, including taxpayers.

Rigorous peer review and expert editorial efforts ensure GSA continues to be a trusted, reliable, and credible source for scientific publications in gerontology and geriatrics. Through a publishing agreement with a scholarly publisher, GSA can reinvest the revenue from the GSA journals to ensure the quality of the publications.

## 2. Steps for improving equity in access and accessibility of publications.

Scientists' ability to communicate their scientific results through publication is critical to the incorporation of their expertise into the scientific enterprise and the progression of their careers. Monitoring implementation of changes to the public access policy, and how researchers and institutions pay publishing costs, will be critical to ensuring that public access plans do not create new systemic inequities or reinforce existing ones. Careful and continued study will be essential for understanding the near- and long-term effects of related changes. Study of cost effects at the researcher, institution, and enterprise levels is needed. It may also be valuable for NIH to survey researchers and institutions about publishing costs and about tradeoffs made to pay such costs.

## 3. Methods for monitoring evolving costs and impacts on affected communities.

Adaptation of federal grant agreements to require reporting on the payment of publication fees and reliance on transformative agreements (in instances where authors avoid payment of a fee because their institution has a transformative agreement with their journal of choice) represents one logical approach to monitoring fees. All analyses of and reporting on costs paid by institutions or researchers for publication should examine potential variability in costs across disciplines, career stages, and institution types, as well as variability based on researcher backgrounds.

### 4. Early input on considerations to increase findability and transparency of research.

We are pleased that through our current publisher, GSA provides some or all metadata for all authors. GSA appreciates efforts underway such as the requirement for individuals supported by research training, fellowship, research, education, and career development awards to have Open Research and Contributor Identifiers and exploring the use of the digital object identifier system. GSA looks forward to the opportunity to provide continued input as systems to increase findability and transparency of research are developed.

On behalf of GSA, thank you for the opportunity to provide input. GSA shares the view of many other members of the scientific community that public access should optimize equity for researchers and readers. Thank you for your consideration. We look forward to continuing to work with NIH and other federal agencies to develop policies that balance access to published work with the ability to publish.

If you have additional questions regarding these matters and the comments offered herein, please contact Patricia M. D'Antonio, Vice President, Policy and Professional Affairs at <a href="mailto:pdantonio@geron.org">pdantonio@geron.org</a> or Judie Lieu, Vice President, Publications and Professional Resources at <a href="mailto:jleu@geron.org">jleu@geron.org</a>.

Sincerely,

James C. Appleby, BSPharm, MPH, ScD (Hon)

Chief Executive Officer

James C. Appleby

I am responding to this RFI: On behalf of an organization

Name: Katie Grady

Name of Organization: American College of Radiology

Type of Organization: Professional org association

Role: Member of the public

## 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

NIH and NIH-funded researchers have a duality of interest in publication of research project results, but may have a different interest in how, when, and where those results should be published. While the NIH would likely prefer earlier reporting for transparency and access to data, researchers funded through NIH may prefer later reporting to allow for greater time and ability to complete and evaluate primary and secondary endpoints, and toxicity. To address this duality of interests, the NIH would be best served to encourage (or require) early reporting of select findings, using a templated reporting process to ensure consistency. Peer-reviewed journals are unlikely to accept these preliminary reports, so consideration could be given to an internal NIH-developed outlet, similar to the www.clinicaltrials.gov product. Reporting to/through the new product could be required, but the data provided would not preclude subsequent submission to peer-reviewed journals or meetings. Time of submission could be following completion of the initial Data Safety Monitoring Board (DSMB) review. Templated required data, to provide sufficient information for the public and providers/researchers, could include a brief outline of the trial (phase, study, control arms, inclusion and exclusion criteria, statement regarding generalizability of findings, statement of DSMB findings, statement as to meaning (or lack of meaning) of the data at this juncture, analysis of the current findings, and a statement regarding next steps. Journals are unlikely to willingly give up editorial control or consider reducing their peer-review processes or quality criteria, so novel approaches are necessary to ensure ultimate access to study results, especially for negative or discontinued trials, which journals have historically been unwilling to publish. The NIH could launch a publication for manuscripts reporting only NIH-funded research or could consider supplement(s) to existing NIH journals limited to manuscripts reporting NIH-funded research. This could be done through various Institutes and Centers (ICs). The NIH could consider support through its ICs for publication of supplements limited to NIH-funded research manuscripts. Finally, the cost of submitting articles for publication creates inequity based upon the resources available to researchers. Evaluation of this cost process should be considered to improve equity for investigators.

## 2. Steps for improving equity in access and accessibility of publications.

Except for final manuscripts at study completion, all interim reports should be open access and in a process that not only allows but requires interim reporting of NIH-funded research.

## 3. Methods for monitoring evolving costs and impacts on affected communities.

Providing a source for interim reporting directly through the NIH will allow for significant cost reduction/control.

4. Early input on considerations to increase findability and transparency of research.

Reporting of interim reports could be available through hot links on www.clinicaltrials.gov, and Other sites listing NIH-funded research.

Email: kgrady@acr.org

I am responding to this RFI: On behalf of an organization

Name: Katie Steen-James

Name of Organization: SPARC

Type of Organization: Professional org association

## 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

SPARC strongly supports the OSTP Memorandum's emphasis on ensuring equity in contributing to, accessing, and benefitting from the results of federally funded research, and we appreciate NIH's specific attention on how to ensure equity in publication opportunities for its funded investigators. As the research process has shifted to the digital environment, a wide variety of channels designed to support more rapid, frequent, and iterative communication of research findings have emerged.

It is vital that researchers have compliance options that do not present them with financial barriers. To that end, NIH should make it clear that investigators can fully comply with its public access policy by depositing their author's accepted manuscripts into PubMed Central (PMC) or any Other agency-approved repository—and that there is no charge to do so.

In its guidance, it is important for NIH to make clear that any fee that investigators may be asked to pay is a publication fee, and not a fee required by NIH to comply with its policy. It is critical that investigators do not conflate compliance with article processing charges (APCs), which create significant barriers for less-well-resourced investigators and institutions to make their research available.

There are a growing number of communications options that provide free, immediate access to research outputs that do not rely on unnecessary and unsustainable author-side charges for investigators. NIH should actively encourage the use of publication channels that do not present financial barriers, including non-APC supported open access journals, publications from non-profit University presses, and scholarly society publishers that allow repository deposit and full reuse of author manuscripts, preprint servers, and Other emerging community-driven options.

We also note that institutional repositories run by libraries and Other research institutions generally do not charge authors to deposit articles or manuscripts, and could play an important role in easing compliance burdens on investigators, improving discoverability of research outputs, and providing long term preservation support. We recommend that NIH engage with the U. S. Repository Network (which recently released the "Desirable Characteristics of Digital Publication Repositories" document) to identify additional repositories that meet NIH's criteria for depositing publications.

To accelerate and sustain equitable research communication practices in the long term, it is critical that research assessment and career advancement incentives be updated to actively promote equity in publication opportunities. NIH should look for opportunities to better align its awards process with equitable research communication practices, such as avoiding journal-based metrics and recognizing preprints. It would also be useful for NIH to engage with ongoing efforts designed to address this important area, including The NASEM Roundtable on Aligning Incentives for Open Science and the Higher Education Leadership Initiative for Open Scholarship (HELIOS).

NIH's efforts to ensure equity in publication opportunities for its investigators naturally align with the critical work of the National Science & Technology Committee's (NSTC) Subcommittee on Equitable Data. SPARC strongly supports the Subcommittee's work to "Build Capacity for Robust Equity Assessment for Policymaking and Program Implementation" and recommends NIH coordinate the implementation of its public access plan with the NSTC Subcommittee and the Department of Health and Human Services' (HHS) Equity Action Plan. Additionally, the public access plan should be included in HHS' equity assessments and disparity impact strategies.

## 2. Steps for improving equity in access and accessibility of publications.

The early stages of the COVID-19 pandemic demonstrated the importance of full reuse rights when, after prompting by global leaders, publishers made COVID-related articles immediately available in PMC under article-level licenses that allowed for full reuse and secondary analysis. Within the first two weeks, these articles had been accessed or downloaded over 2 million times-greatly accelerating the rate of discovery, speeding the translation of science, and increasing the community's understanding of the virus. This temporary shift in practice highlights the need for a permanent change making federally funded research publications both immediately available and fully reusable in order to provide much broader, real-time returns on taxpayer investments in scientific research.

The OSTP Memorandum asks agencies to "make federally funded publications, data, and Other such research outputs and their metadata...findable, accessible, interoperable, and reusable, to the American public and the scientific community in an equitable and secure manner." To fulfill the reusability requirement, NIH should ensure that all publications resulting from NIH-funded research carry open licenses and that NIH authors can explicitly retain the rights needed to authorize those open licenses, regardless of whether authors deposit an author accepted manuscript or a final published article. To this end, placing a CC BY license or its functional equivalent on a publication is the best way to ensure that publications can be freely accessed and fully reused.

NIH should ensure that it obtains sufficient rights to provide the public with the full benefits of the research that it funds. In particular, as the OSTP Memorandum directs, the public should be able to access final peer-reviewed accepted manuscripts freely, without embargo or delay, and under terms that make them fully reusable. The agency should seek to achieve this result in a manner that minimizes complexity and burden in compliance by grantee institutions and individual researchers.

## 3. Methods for monitoring evolving costs and impacts on affected communities.

At SPARC, we are deeply concerned about the financial barriers that author-side fees, particularly Article Processing Charges (APCs), present to authors and the significant additional negative effects these have on the research ecosystem. APCs are rising very rapidly in price, driving an overall increase in the cost of research communication that presents a growing risk of tradeoffs in diverting funds away from the research process itself. The diversion could negatively affect the budget needed for materials and equipment, supporting postdocs, and professional development opportunities including presenting research results at conferences.

APCs create prohibitive barriers to publication that negatively impact many researchers, especially in instances where publishing in particular APC-based journals is viewed as important for career advancement. This results in fewer opportunities for individual researchers to share their results with

the scientific community and the public. This is extremely troubling from an equity perspective, as studies have documented that APC costs disproportionately affect younger researchers, female researchers, and those at less well-funded institutions.

It is important for NIH to be aware of these impacts, and to actively monitor the impacts of any publication charges across demographic groups in the research ecosystem. For example, NIH should establish a baseline understanding of the environment by collecting data on the number and makeup of its current funding recipients who are charging publication fees as direct costs to their research grants and analyzing that data across different demographics (e.g., minority-serving institutions (MSIs), EPSCoR-eligible institutions, IDeA-eligible institutions, researchers in less-well-resourced disciplines, etc.)

Data collection on the amount spent to publish NIH-funded research regardless of the source would increase transparency and insight into how these fees affect various communities - including the potential impacts on publishing opportunities.

### 4. Early input on considerations to increase findability and transparency of research.

Ensuring that the results of NIH-funded research along with metadata containing information about who conducted the research, where it was done, and with what resources is an important component of the NIH Public Access Plan.

To complement continued requirements for internal identifiers (PMCIDs, GenBank accession numbers, etc), NIH should require the use of external persistent identifiers (PIDs). Specifically, NIH should adopt DOIs for publications, data sets, and DMPs, ORCIDs for researchers, and RORs for institutional affiliations, all of which are nonproprietary community standards for each identifier type. NIH should also explore the use of the DOI system to overlay NIH's current unique identifiers for awards.

Because similar identifiers will be required to be used by all federal agencies as a result of the OSTP Memorandum, NIH should coordinate its efforts with Other participants in interagency working groups, including the National Science and Technology Council's (NSTC) Subcommittee on Open Science, to identify best practices and potential standards. NIH also should consider collaboration with standards bodies, such as the National Information Standards Organization (NISO), to develop a framework and set of standards for a national PIDs strategy to facilitate smooth implementation.

Given the growing centrality of PIDs in research infrastructure, it is essential that the NIH and Other federal agencies only adopt nonproprietary identifier types that enable the broadest possible use and allow anyone to leverage this information in new and innovative ways.

#### **Uploaded File:**

NIH-RFI-SPARC-Response.pdf

**Description:** Answers to the four questions and additional comments in letter format with hyperlinks

Email: katie@sparcopen.org



1201 Connecticut Ave NW, #608 • Washington, DC 20036 (202) 630-5090 • <a href="https://www.sparcopen.org">www.sparcopen.org</a>

April 24, 2023

Office of Science Policy National Institutes of Health 6705 Rockledge Drive, Suite 630, Bethesda, MD 20892

Submitted via electronic form

Re: Notice Number NOT-OD-23-091

Thank you for the opportunity to submit comments on NIH's draft Public Access Plan. SPARC is a non-profit advocacy organization that supports systems for research and education that are open by default and equitable by design. Our membership includes over 200 academic and research libraries across the U.S., serving institutions ranging from large research intensive universities to community colleges. We believe that sharing knowledge is a human right, and that everyone should be able to access and contribute to the knowledge that shapes our world. Our members are committed to supporting equitable systems of research and education, and we appreciate the opportunity to comment on NIH's draft plan to implement the landmark 2022 OSTP Memorandum on Ensuring Free, Immediate, and Equitable Access to Federally Funded Research.

NIH's draft plan provides a strong foundation for the agency and the public to fully realize the benefits of open science. Eliminating the existing 12-month embargo enables free and fast access to the results of the more than \$40 billion in biomedical research that the agency funds each year. As noted by the OSTP Memorandum, the 12-month embargo limits "immediate access of federally funded research results to only those able to pay for it or who have privileged access through libraries or other institutions. Financial means and privileged access must never be the prerequisites to realizing the benefits of federally funded research that the American public deserves." Removing the embargo will open up endless opportunities for new collaborations, accelerate the speed of critical discoveries, and improve lives.

Our responses to the four questions outlined in the Request for Information (RFI) offer additional steps for NIH to consider to further strengthen its plan and to address inequities in the research ecosystem.

# Question 1: How can NIH best ensure equity in publication opportunities for its investigators?

SPARC strongly supports the OSTP Memorandum's emphasis on ensuring equity in contributing to, accessing, and benefitting from the results of federally funded research, and we appreciate NIH's specific attention on how to ensure equity in publication opportunities for its funded investigators. As the research process has shifted to the digital environment, a wide variety of channels designed to support more rapid, frequent, and iterative communication of research findings have emerged.

It is vital that researchers have compliance options that do not present them with financial barriers. To that end, NIH should make it clear that investigators can fully comply with its public access policy by depositing their author's accepted manuscripts into PubMed Central (PMC) or any other agency-approved repository—and that there is no charge to do so.

In its guidance, it is important for NIH to make clear that any fee that investigators may be asked to pay is a publication fee, and **not a fee required by NIH to comply with its policy.** It is critical that investigators do not conflate compliance with article processing charges (APCs), which create significant barriers for less-well-resourced investigators and institutions to make their research available.

There are a growing number of communications options that provide free, immediate access to research outputs that do not rely on unnecessary and unsustainable author-side charges for investigators. NIH should actively encourage the use of publication channels that do not present financial barriers, including non-APC supported open access journals, publications from non-profit university presses, and scholarly society publishers that allow repository deposit and full reuse of author manuscripts, preprint servers, and other emerging community-driven options.

We also note that institutional repositories run by libraries and other research institutions generally do not charge authors to deposit articles or manuscripts, and could play an important role in easing compliance burdens on investigators, improving discoverability of research outputs, and providing long term preservation support. We recommend that NIH engage with the U. S. Repository Network (which recently released the "Desirable Characteristics of Digital Publication Repositories" document) to identify additional repositories that meet NIH's criteria for depositing publications.

To accelerate and sustain equitable research communication practices in the long term, it is critical that research assessment and career advancement incentives be updated to actively promote equity in publication opportunities. NIH should look for opportunities to better align its awards process with equitable research communication practices, such as avoiding journal-based metrics and recognizing preprints. It would also be useful for NIH to engage with ongoing efforts designed to address this important area, including <a href="https://doi.org/10.1007/nih.gov/nih.go

NIH's efforts to ensure equity in publication opportunities for its investigators naturally align with the critical work of the National Science & Technology Committee's (NSTC) Subcommittee on Equitable Data. SPARC strongly supports the Subcommittee's work to "Build Capacity for Robust Equity Assessment for Policymaking and Program Implementation" and recommends NIH coordinate the implementation of its public access plan with the NSTC Subcommittee and the Department of Health and Human Services' (HHS) Equity Action Plan. Additionally, the public access plan should be included in HHS' equity assessments and disparity impact strategies.

# Question 2: What steps can NIH take to improve equity in access and accessibility of publications?

The early stages of the COVID-19 pandemic demonstrated the importance of full reuse rights when, after prompting by global leaders, publishers made COVID-related articles immediately available in PMC under article-level licenses that allowed for full reuse and secondary analysis. Within the first two weeks, these articles had been accessed or downloaded over 2 million times—greatly accelerating the rate of discovery, speeding the translation of science, and increasing the community's understanding of the virus. This temporary shift in practice highlights the need for a permanent change making federally funded research publications both immediately available and fully reusable in order to provide much broader, real-time returns on taxpayer investments in scientific research.

The OSTP Memorandum asks agencies to "make federally funded publications, data, and other such research outputs and their metadata...findable, accessible, interoperable, and reusable, to the American public and the scientific community in an equitable and secure manner." To fulfill the reusability requirement, NIH should ensure that all publications resulting from NIH-funded research carry open licenses and that NIH authors can explicitly retain the rights needed to authorize those open licenses, regardless of whether authors deposit an author accepted manuscript or a final published article. To this end, placing a <a href="CC BY license">CC BY license</a> or its functional equivalent on a publication is the best way to ensure that publications can be freely accessed and fully reused.

NIH should ensure that it obtains sufficient rights to provide the public with the full benefits of the research that it funds. In particular, as the OSTP Memorandum directs, the public should be able to access final peer-reviewed accepted manuscripts freely, without embargo or delay, and under terms that make them fully reusable. The agency should seek to achieve this result in a manner that minimizes complexity and burden in compliance by grantee institutions and individual researchers.

# Question 3: How can NIH best monitor evolving costs, specifically publication fees, and impacts on affected communities?

At SPARC, we are deeply concerned about the financial barriers that author-side fees, particularly Article Processing Charges (APCs), present to authors and the significant additional negative effects these have on the research ecosystem. APCs are <u>rising very rapidly in price</u>,

driving an overall increase in the cost of research communication that presents a growing risk of tradeoffs in diverting funds away from the research process itself. The diversion could negatively affect the budget needed for materials and equipment, supporting postdocs, and professional development opportunities including presenting research results at conferences.

APCs create prohibitive barriers to publication that negatively impact many researchers, especially in instances where publishing in particular APC-based journals is viewed as important for career advancement. This results in fewer opportunities for individual researchers to share their results with the scientific community and the public. This is extremely troubling from an equity perspective, as <a href="studies">studies</a> have documented that APC costs disproportionately affect younger researchers, female researchers, and those at less well-funded institutions.

It is important for NIH to be aware of these impacts, and to actively monitor the impacts of any publication charges across demographic groups in the research ecosystem. For example, NIH should establish a baseline understanding of the environment by collecting data on the number and makeup of its current funding recipients who are charging publication fees as direct costs to their research grants and analyzing that data across different demographics (e.g., minority-serving institutions (MSIs), <a href="EPSCoR-eligible">EPSCoR-eligible</a> institutions, <a href="IDEA-eligible">IDEA-eligible</a> institutions, researchers in less-well-resourced disciplines, etc.)

Data collection on the amount spent to publish NIH-funded research regardless of the source would increase transparency and insight into how these fees affect various communities - including the potential impacts on publishing opportunities.

### Question 4: Early input on considerations to increase findability and transparency of research.

Ensuring that the results of NIH-funded research along with metadata containing information about who conducted the research, where it was done, and with what resources is an important component of the NIH Public Access Plan.

To complement continued requirements for internal identifiers (PMCIDs, GenBank accession numbers, etc), NIH should require the use of external persistent identifiers (PIDs). Specifically, NIH should adopt DOIs for publications, data sets, and DMPs, ORCIDs for researchers, and RORs for institutional affiliations, all of which are nonproprietary community standards for each identifier type. NIH should also explore the use of the DOI system to overlay NIH's current unique identifiers for awards.

Because similar identifiers will be required to be used by all federal agencies as a result of the OSTP Memorandum, NIH should coordinate its efforts with other participants in interagency working groups, including the National Science and Technology Council's (NSTC) Subcommittee on Open Science, to identify best practices and potential standards. NIH also should consider collaboration with standards bodies, such as the National Information Standards Organization (NISO), to develop a framework and set of standards for a national PIDs strategy to facilitate smooth implementation.

Given the growing centrality of PIDs in research infrastructure, it is essential that the NIH and other federal agencies only adopt nonproprietary identifier types that enable the broadest possible use and allow anyone to leverage this information in new and innovative ways.

SPARC appreciates the opportunity to provide comments, and we applaud the agency for its continued leadership in ensuring public access to taxpayer funded research. We look forward to working with the agency to fully accomplish the goals outlined in the OSTP Memorandum and to leverage the full value and utility of NIH-funded research.

Sincerely,

Katie Steen-James

FatieMSteen James

Manager of Public Policy & Advocacy

Heather Joseph

**Executive Director** 

**Submit date: 4/24/2023** 

Name: Juliane Baron

Name of Organization: Federation of Associations in Behavioral and Brain Sciences

**Type of Organization:** Nonprofit research organization

Role: Scientific researcher

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

FABBS appreciates the NIH's commitment to equity in publication opportunities. We share these concerns. FABBS joins the broad scientific community (see the American Academy of Arts and Sciences' statement, which FABBS has signed onto) in pointing out the limitations of current publication models and encouraging NIH to continue to explore alternatives to subscription models and pay to publish fees.

The Behavioral Medicine Research Council issued a statement on Open Science (<a href="https://psycnet.apa.org/fulltext/2023-60199-001.html">https://psycnet.apa.org/fulltext/2023-60199-001.html</a>) in Health Psychology and Behavioral Medicine, establishing a commitment to open science, identifying challenges and providing guidance on open science practices.

# 2. Steps for improving equity in access and accessibility of publications.

FABBS cautions that public access will not automatically translate directly to equitable access. Beyond the ability to download a particular article, numerous Other barriers - scientific training, time and ability to translate research to policy or practice - prevent the public from fully understanding academic language and methodology. In addition to a range of scientific training.

- 3. Methods for monitoring evolving costs and impacts on affected communities.
- 4. Early input on considerations to increase findability and transparency of research.

FABBS strongly supports the goals of increasing findability and transparency of research. Maximizing the potential of data sharing will require significant planning and effort to standardize terms, methods, and measures in the behavioral and cognitive sciences.

FABBS applauds the Office of Behavioral and Social Sciences Research for issuing a Notice of Intent to Publish a Funding Opportunity Announcement for Accelerating Behavioral and Social Science through Ontology Development and Use (U01)(<a href="https://grants.nih.gov/grants/guide/notice-files/NOT-OD-23-089.html">https://grants.nih.gov/grants/guide/notice-files/NOT-OD-23-089.html</a>). This critical effort builds upon guidance from the National Academies of Science, Engineering, and Medicine's (NASEM) report on Ontologies in the Behavioral Sciences, of which FABBS is a sponsor. (<a href="https://nap.nationalacademies.org/catalog/26464/ontologies-in-the-behavioral-sciences-accelerating-research-and-the-spread">https://nap.nationalacademies.org/catalog/26464/ontologies-in-the-behavioral-sciences-accelerating-research-and-the-spread">https://nap.nationalacademies.org/catalog/26464/ontologies-in-the-behavioral-sciences-accelerating-research-and-the-spread</a>)

By way of illustration, please see 'Limitations of the Sum-and-Alpha Approach to Measurement in Behavioral Research' (McNeish,

2022)(https://journals.sagepub.com/doi/10.1177/23727322221117144).

**Description:** The Federation of Associations in Behavioral and Brain Sciences (FABBS) is a coalition of 29 scientific societies and 60 academic departments that share an interest in equitably advancing the rigor, impact, and accessibility of our disciplines. FABBS scie

Email: jbaron@fabbs.org

**Submit date: 4/24/2023** 

I am responding to this RFI: On behalf of an organization

Name: Michael Keller

Name of Organization: Booz Allen Hamilton, Inc.

Type of Organization: Other

Type of Organization-Other: Consulting Firm

Role: Institutional official

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Booz Allen has prior experience in reviewing, planning, and implementing equitable access plans and advising on the benefits and drawbacks of not considering equity at each phase of research, including publication. For example, we assessed potential or perceived impacts of disclosure of protected health information (PHI) for research purposes on access to healthcare services by HIPAA-covered entities and real and perceived barriers to use of PHI in research in underrepresented populations. Based on such experience, we reviewed NIH's proposed Public Access Policy and provide the following recommendations relevant to equitable access to publication.

NIH should encourage grantees to share research results with the broader community and demonstrate how these studies impact the community and how communities could use NIH studies to improve health. Communities need demonstrations to access NIH search databases and the developers of these NIH resources should be utilized to make the findings accessible to all levels of understanding. Community engagement involves various participants in the publication process. Understanding the role of faith-based entities and educational institutions in community-wide activities is critical to monitoring the impact of research on communities of color (CoCs). Issues of national importance, such as COVID-19, intimate partner violence, opioid misuse, school shootings, police brutality, and Other socio-political and economic topics that NIH grants support, should be prioritized for CoCs to monitor the impact.

Supplementary to the above, there is no mention of a communication strategy for the new policy and subsequent method of implementation in the documentation. To ensure equity, communicating the new policy in its entirety and relevant impacts to all researchers at NIH will be necessary. In addition, external communication that NIH is revising is its Public Access Policy to ensure equitable access to NIH-funded research would be beneficial to informing all possible users. The external communication can extend to underrepresented populations, Historically Black Colleges and Universities (HBCUs), small research organizations, and Others, increasing the accessibility to information.

### 2. Steps for improving equity in access and accessibility of publications.

To maximize equity in access to publications by diverse communities of users, Booz Allen recommends that NIH promote adaptive technologies and strategies centered around usability and accessibility, navigation, and content. Broad-based adoption of these assistive techniques by authors and publications will increase inclusivity among diverse communities of users by promoting equal access to, and engagement around, critical research and practice in the health and life sciences. The following

recommendations are not exhaustive; however, our team understands the added importance of addressing accessible format design elements including page layout, graphics, and charts.

Based on industry best practices, Booz Allen recommends that NIH should promote and fund user-centered studies to deepen the research on how to best use technology to make scholarly publications accessible to people with different learning styles and disabilities. These studies could evaluate methods of improving search systems such as PubMed with usability evaluations tools, submission accessibility guidelines, and document navigation tools to make results more findable to broader communities.

Booz Allen further recommends that NIH should engage with journal publications to establish guidelines promoting maximum accessibility so that authors may reach the largest community of users for their discipline, such as writing broadly to reach a mix of early graduate students and early careerists, and to employ descriptors and expressions to engage non-visual users who are highly reliant on descriptive text. NIH should also engage with journal publications to employ industry-accepted adaptive technology that will support users with visual, auditory, and perceptual disabilities.

### 3. Methods for monitoring evolving costs and impacts on affected communities.

NIH's research agenda requires monitoring costs and how those costs impact communities affected by NIH research and recipients of NIH awards. To effectively monitor publication costs (e.g., fees, increases, actual cost, and profit margins), there will need to be an ongoing monitoring and evaluation plan accessible to the general public. Since transparency and equity are related, the monitoring and evaluation plan for NIH should have several components/steps in place to ensure that equity planning is sustainable - these steps include surveying the current publication data that is available, identifying what is unknown, creating scales, metrics, and performance outcomes. After setting up processes to collect this data, NIH will be able to effectively monitor the evolving cost and impacts.

Biomedical and life science research scientific journals have a wide range in scope, collection size, and acceptance criteria. As a result, the publication process has a range of costs that must be understood before monitoring begins. Before costs can be monitored, expenses must first be tracked, understood, and then agreed upon. Booz Allen recommends that NIH perform an inventory or surveying of journals that publish studies with NIH funding followed by designing a metric to comprehensively evaluate publication costs and assign a score to journals based on this metric. Finally, continuous data quality audits should be implemented to ensure data are accurate and accessible.

Every level of the publication process is associated with policies related to who is allowed to peer review articles to the revised and resubmit process, procedures, and policies directly related to cost. All NIH publications need to review the question of who benefits and have a clear understanding and definition of what is "fair" and "equitable". NIH policies involving funding for publication costs should be reviewed, and every instance that demonstrates inherent disadvantages for less privileged populations should be tracked and flagged for future updates. In addition to this policy review, Booz Allen recommends that NIH should bucket the types of policies and evaluate them based on their impact.

To monitor the impact of NIH research and publication access on communities of color, NIH should develop indicators of impact/success to determine the effect of publication policies and NIH-funded research on communities of color, train the relevant practitioners on these metrics, and then revisit the

results to understand if improved access to publication opportunities and scientific results have impacted the relevant communities.

# 4. Early input on considerations to increase findability and transparency of research.

Recently, a Booz Allen team of researchers and analysts completed an NIH-funded initiative for the Office of Data Science Strategy (ODSS) in which the team developed a competency framework to guide biomedical and behavioral researchers through how to prepare their data to adhere to the metadata-driven Findable, Accessible, Interoperable, and Reusable (FAIR) principles and AI-readiness criteria. As part of this effort, Booz Allen interviewed professors and researchers from leading universities and minority serving institutions across various research areas to identify gaps in knowledge about data sharing standards. The following recommendations outline opportunities based on Booz Allen's discoveries from these interviews that would aid NIH in its efforts to strengthen research findability and transparency through knowledge sharing and expansion of new metadata standards and resources.

Booz Allen recognizes that NIH seeks to collect and make publicly available appropriate metadata associated with scholarly publications and data at the time of deposit in a public access repository. This has been challenging because researchers often do not know which metadata or metadata ontology standards they should use. To bridge these gaps, Booz Allen recommends developing a metadata ontology dictionary that would guide interested parties to terminologies that PubMed Central (PMC) officially recognizes.

Booz Allen also recognizes that NIH is interested in discovering innovative ways to instruct federally funded researchers to obtain digital persistent identifiers (PIDs) to maximize the findability of the research they share on PMC. From an instructional standpoint, Booz Allen recommends that NIH develops trainings for FAIR and TRUST principles, as many professors and researchers have a general lack of awareness of these principles, which are closely tied to the NIH Data Management and Sharing (DMS) Policy.

In accordance with the 2022 OSTP Memorandum, NIH seeks to elevate transparency about integrity of scientific research that was paid for with taxpayer dollars. Booz Allen recommends that PMC expand its taxonomy of PIDs to include metadata that would indicate the reproducibility of findings in publications. Currently, PIDs include information about authorship, funding, affiliation, and development status of federally funded research. However, there is no easily searchable indication of how many, and which researchers, labs, and institutions have reproduced the findings of publications. Booz Allen recommends that NIH develop PID requirements for reproducibility of findings and integrate them into a PMC such that publications with stronger reproducibility rankings would appear higher in search results. This would generate incentive for researchers to promote collaborative science by seeking opportunities with Other labs because it would contribute to them gaining more exposure on PMC.

Email: black rebecca@bah.com

**Submit date: 4/24/2023** 

I am responding to this RFI: On behalf of myself

Name: Robin Puett

Type of Organization: University

Role: Scientific researcher

## 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

Multiple steps toward ensuring equity are needed, such as addressing publication bias, providing equal opportunities for null or negative findings to be published. The burden of high costs for open access journal publications, as well as publication costs in general should also be considered to ensure equity, particularly as results may continue to be reported after grants have ended. In addition, novel venues for free and low cost dissemination of research results would allow for more funding to go directly to conducting the research. These venues should include a rigorous peer-review process and should be structured to provide measurable impact for tenure and promotion reviews.

### 2. Steps for improving equity in access and accessibility of publications.

Given that distinguishing between rigorous peer-reviewed journals and disinformation outlets is challenging for individuals who are outside the field of interest, more resources should be directed toward the translation of scientific journal articles into digestible messages for a lay audience and requiring the inclusion of limitations and uncertainties. More resources should also be directed toward educating the public on how to find rigorous peer-reviewed science, distinguishing it from disinformation and critiquing it based on scientific methods. One effort toward communicating science that is commendable is the Frontiers for Kids translation of rigorous scientific articles for kids with editing by kids. The education is for scientists to translate results and for kids to be empowered and education in the scientific method.

An important related topic - is to consider how Chatgpt and AI are going to impact scientific research reporting - distinguishing disinformation will become more difficult and rigorous scientific processes may be shortcut.

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

One hidden cost is the inordinate amount of time required for submitting manuscripts (entering each first and each last name of each co-author in separate fields, followed by each affiliation, etc with large teams of scientists). The time required seems mundane but often several hours may go toward entering these data which could be spent more directly on teaching/research/engaging with the community. The technology for submission seems very antiquated.

Also publication costs are not easy to find - and an increase may also be a surprise after the manuscript has been accepted if open access is required. Any monitoring system will be challenged to find the real costs for all journals- perhaps requiring journals to be more up front with costs would be a starting point.

#### 4. Early input on considerations to increase findability and transparency of research.

Most researchers that I am aware of do this work with the ultimate goal of serving the public (which includes us our families and friends), providing them info and improving their health. However researchers also have to treat the generosity of study participants with the upmost privacy, respect and confidentiality. This ensures that future health research will happen. With the everchanging tech environment, chatgpt, AI, high performance computing providing easier ways to search for datasets, combine them and reverse engineer variables that are not technically PHI or considered identifiers - the potential for identifying study participants is an increasing and evolving risk. Researchers do not have the appropriate tools/background and universities do not have the personnel/resources to ensure that all potential identifiers are scrubbed for use of the data in perpetuity. If all identifiers are scrubbed, this often makes the data useless. For example several GIS layers which contain specific timed information can be combined and reverse engineered to isolate fairly small geographic locations - when combined with Other data, study participant identification becomes much riskier. Datasets that are publicly available via digital format will never be ensured of destruction, however tech is ensured to advance exponentially. How can confidentiality and anonymity be ensured forever? Greater consideration of risks, resources, current and future tech, limitations, and requirements for informing study participants of these changes should occur before data sharing requirements. For example, Other agencies have provided a substantial funding increase in recognition of the amount of resources required to ensure data are safely made publicly available and are archived in time limited fashion and with different levels of restriction based on types of data. Even basic DUAs require monitoring and often require the names, positions and human subjects certifications of anyone accessing the data. Findability and transparency of research is an admirable goal but also is maintaining the confidentiality and anonymity of study participants who are generously sharing their lives to help improve everyone's.

**Submit date: 4/24/2023** 

I am responding to this RFI: On behalf of myself

Name: Gary McDowell

Name of Organization: Lightoller LLC

Type of Organization: Other

Type of Organization-Other: Academic Consulting

Role: Member of the public

#### 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

I would like to encourage NIH to reconsider the use of the word "maintain" when describing the "broad discretion for researchers and authors to choose how and where to publish their results". I have the perspective that the current situation is somewhat restrictive, and could be expanded by NIH.

I think, in response to this, many will be thinking of the "right" of a researcher to publish in whatever magazine they wish; that right has always existed, and will continue to exist under the proposed changes. The question is who will pay for it. I am a taxpayer who currently is not guaranteed access to federally-funded research, that I need for my work, at the point of publication. I would therefore suggest to NIH that the burden can be removed, with a simultaneous introduction of more choices of how and what to publish, through greater incentives for using preprints. I would like to suggest greater recognition of preprints as a method for NIH to ensure compliance, to allow authors a choice of where to publish that extends beyond simply the requirement to publish in magazines.

One could make the argument that we could make an ecosystem for sharing research and data using preprints, perhaps on a federal server, which would serve perfectly well as a medium for evaluating researchers and remove the currently flawed reliance on Impact Factors and prestige, which have been shown to be biased and subjective. In such a world, academics could then pay to publish their research in magazines from their own pocket, if they so wished. Given that researchers already carry out peer review as voluntary labor, it's not clear to me why the taxpayer needs to be paying so much money to publish a certain kind of research in magazines, just because that is what academics at universities and research institutions have decided is their preferred method for evaluation and promotion.

I believe that NIH is motivated to maximize the opportunities for communication of results by its grantees, and grantees should be motivated to ensure that as much data as possible can be shared with the wider community. The current system for communicating research outputs, relying on magazine articles to publish work, restricts what can be communicated. The use of these articles by the academic community, including by funders such as NIH, to evaluate a scientist and thus determine their career path, then incentivizes publishing only in a certain manner.

Much of the work that is carried out by researchers will, by its very nature, not be in the form of novel, positive data that can be formed into the narrative structure favored by magazine publishers. There is negative data that is collected; small experiments that don't fit into a larger narrative; and of course large datasets that may not easily be evaluated by a single team. Publishing this kind of work is not currently incentivized in the traditional academic environment and it means that much of the data

funded by the taxpayer may remain within individual laboratories for the simple reason that the academic community has decided not to value this, even though it may be perfectly valid research. This may have effects, such as reducing the efficiency of biomedical research. For example, numerous labs across the country may be e.g. attempting to purify exactly the same troublesome protein, leading to potentially many laboratories trying out the same technique or experiment over and over again with no success - or, perhaps, simply taking longer to get to a successful outcome. All this is due to a lack of prior knowledge and information being published. The knowledge exists, maybe across multiple labs, but for some reason is not being shared despite the obvious efficiency it could produce for scientist and taxpayer alike. In addition, this system of only rewarding positive novel results also selects for (at best) luck, and (at worst) cherry-picking (or even outright falsifying) data, because a career is dependent not on the actual result, but a positive one. It is not designed to select for merit, as scientists cannot possibly predict whether their hypotheses will be correct all the time, and only a lucky few will chance upon the right problem to work on. Many very thorough and brilliant scientists may have been lost to academic research simply because they have not produced the particular kind of research being selected for in magazine article publishing.

I would like to provide some insights from my own experience of moving from the traditional academic environment to my consulting role. The work of a consultant is extremely similar to the work of an academic: I carry out research to solve a particular problem and am paid money by stakeholders to carry out that research (including for the taxpayer, on federally-funded research grants). The transition to this kind of work has been very simple given my previous academic experience.

There is one striking difference. When I communicate my results, my priority is to communicate all of my data and findings in a clear and concise manner to the stakeholders who paid for my work in order to help them solve a problem. This is in contrast to my previous experience as an NIH-funded postdoc.

If, as a consultant, I were to behave in the manner of a University academic, I would not write-up all of my analyses, but only those that I chose based on a narrative story of positive results. I would then publish this in a magazine, in the form of an article written in a esoteric style according to the desires of the magazine. I would then tell the people who paid me that they have to pay to read that work in the magazine, and would complain loudly about my "academic freedom" if there were moves to make me do Otherwise.

Obviously, I would not last long in the consulting business if I followed this model. This is, in part, why it is such a relief to have left the University environment and be able to do what I originally intended in science - work with people to solve challenging problems using research methods and data analysis. I am confident that there are many NIH-funded academics who feel the same way about their ability to communicate research.

The nature of my work is very similar to my time at universities; it is the incentives that are different. As a contractor I am not expected to publish magazine articles, as this is not a practice that is part of a consultant's work. In this way, I should note, publishing magazine articles is not an activity of all scientists - just a cultural practice followed by academics. However, I am also an academic, and as such I do see great value in sharing work through scholarly communication formats such as preprints, to allow for evaluation and improvement of the work as part of scholarly discourse. I am a firm believer in the principle of peer review, and in improvement of work through communicating knowledge through successive versions of analysis and interpretation, with updates as and when I receive feedback. The

change of incentive structures, by operating in a slightly different system but performing very similar work, has allowed me to think in different ways about how to communicate ALL of our data and work.

It was always my goal through my academic training to make sure as much of my data that had been gathered and analyzed appropriately and methodically was released for someone else to use; but this was harder to do when I worked in an academic University environment because of the publication structure and incentives. It is easier now to release data and analyses freed from this restrictive structure, and I enjoy research and review processes more because of this freedom.

I believe it is in NIH's interest to prioritize incentivizing taxpayer-funded researchers to similarly communicate as much of the data and work that they achieved with NIH support as possible - even if it does not have a clear "big story" to go with it. I have always maintained that any small or strange result that doesn't fit into my story could be of use to someone else, my ignorance about the problems Others work on should not determine what I think is fit, or not, to publish, if it is an experiment done well. It is not for me to determine what of my publicly-funded work could, or should, be of use to someone else.

As an academic consultant, I would love to see more public sharing of work by Others in my community of contractors. But the current magazine publishing process is long and tedious and takes up valuable time that could be spent on Other more important work, and it has little reward for those of us who are not assessed on our magazine-article publishing. Impact factor does not matter in my line of work; neither do tedious conversations about who needs what authorship where for their next career step, nor trying to fit your work into the restrictive structure of a particular magazine. When working with my academic colleagues, it is always disappointing when we move from talking about our science, to talking about the magazine publishing process and the careerist motives and strategies needed.

I also want to add that in my line of research, there is a lot of focus on trying to involve more students in scholarly communication as a way of educating them about the scientific process, but also to work on increasing their sense of identity as scientists, and sense of belonging in the process - important factors in encouraging students to follow biomedical and scientific research pathways. But there is an active conversation about how it is hard to engage students in writing articles for magazines, because they (in my view, correctly) see the current form of magazine articles as esoteric and only for future career goals. They are not viewing magazine articles as a way to actually communicate science with Others; and in many cases it seems that they are forming this impression not least because academics are reinforcing the idea that these magazine articles are credential-enhancing products, not a means for sharing results and advancing knowledge.

In all, I would urge NIH to incentivize and promote more innovative ways of sharing work, not least because the system with preprints is not without its difficulties. Carrying out peer review, or being part of peer review communities and providing feedback that will be incorporated into a work in development, still need work, support and innovation. I think there is great value to the general principles of scholarly communication, and of peer review, that need adapting and revising away from the focus of curation of magazine articles, and back towards ensuring validity and constant improvement of research.

#### 2. Steps for improving equity in access and accessibility of publications.

As an American taxpayer and small business owner, the work I carry out supporting training and education of future generations of scientists, some of it federally-funded, is affected by my current inability to access newly-published federally-funded research legally. I am therefore extremely grateful for the removal of the current embargo.

I rely on federally-funded research to carry out my work. My work itself covers issues related to early career researchers, including their participation in and education about the communication of scientific research.

Tax dollars contributed by myself and Other Americans are used by NIH-supported researchers to publish their work, and to fund institutional library subscriptions to access the work of Others - at institutions that can afford to pay these subscriptions. Therefore the taxpayer currently pays for academics to publish their work, and then a privileged subset are able to gain access to the work of Others. Meanwhile the taxpayer is left out in the cold.

I would encourage continued use of the system on PMC to allow access to articles. I do want to make clear that there are academics who have insisted that members of the public can always email corresponding authors for a copy of the manuscript; this is clearly not an appropriate recommendation as response rates are very poor, and of course there should be effort to ensure that the public can access the work they fund as easily as possible. I thank NIH for their work on this as a priority.

Under the current system, anyone who is not in one of the institutions that can afford subscriptions to journals currently faces barriers to timely access to this work. Access to federally-funded research is not extended to all who support its development. Patients, patient advocates, small-business owners - we are all excluded from reading and using this important work. In addition, access isn't even granted to all academic researchers and students. Access to specific magazines, in which scientists publish their articles, is dependent on the ability of a University to be able to pay the subscriptions. Not all institutions are able to afford subscriptions. There are therefore thousands of students and researchers at American institutions of higher education who cannot access work needed to carry out their research and education. Shockingly, the Nelson memo is a great win for education and research at American universities themselves, and will allow greater access to those students, some of whom I have had occasion to work with.

I would ask NIH to require researchers to publish using a CC BY or less restrictive license. It is most useful for educational purposes if articles are not just free to read, but are truly open access. Free to read articles restrict the ability to work with the material in an authentic way, and is restrictive. For students and educators alike to make full use of research articles, it is important to ensure free and open licensing for articles and images.

### 3. Methods for monitoring evolving costs and impacts on affected communities.

I would encourage NIH to look into the DocMaps Framework (https://docmaps.knowledgefutures.org), which I once worked on but am no longer affiliated with. This is a project by the Knowledge Futures Group to develop a community-endorsed framework for capturing valuable context about the processes used to create documents in a machine-readable way. Please see "The DocMaps Framework for representing assertions on research products in an extensible, machine-readable, and discoverable

format (<a href="https://www.biorxiv.org/content/10.1101/2021.07.13.452204v1">https://www.biorxiv.org/content/10.1101/2021.07.13.452204v1</a>)". Policies and fees associated with articles could be examples of metrics mapped onto articles.

Peer review and preprint policies are unclear at most major journals (<a href="https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0239518">https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0239518</a>), and many of these policies affect equity in the magazine publishing world, which is an opaque and subjective process rife with bias. The Royal Society of Chemistry, for example, released a report demonstrating that the process of publishing magazine articles is rife with bias against women, at every step of the process (<a href="https://www.chemistryworld.com/news/rsc-report-finds-publishing-pipeline-hinders-women/4010608.article">https://www.chemistryworld.com/news/rsc-report-finds-publishing-pipeline-hinders-women/4010608.article</a>).

There are a number of reasons NIH should be keeping a close eye, or supporting efforts to do so, on magazine publishers. For example, consider the role of early career researchers in peer review. In biomedicine it is common practice for a PI, as an invited reviewer, to pass a manuscript from a magazine on to graduate students and postdocs to carry out the review, sometimes under the guise of training, but often not reported to the magazine. We gathered data and published an analysis "Co-reviewing and ghostwriting by early-career researchers in the peer review of manuscripts" (<a href="https://elifesciences.org/articles/48425">https://elifesciences.org/articles/48425</a>), showing that it was indeed common that ECRs would undertake review with no credit, and receiving no feedback, hence negating the claim that this is a "training exercise". I will note that when my colleagues and I have been presenting or communicating about this work, a very common request from NIH-funded postdocs is that we move on to looking at the same phenomenon with NIH grants.

Ghostwriting is a form of plagiarism, and we have provided recommendations to multiple stakeholders, including magazines, about how to fix this problem, in "How to bring peer review ghostwriters out of the dark" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8101444/). However, I am disappointed to report that some magazines have doubled down against taking action to deal with this. In particular, there are efforts to require graduate students and postdocs to undertake training before being "allowed" to review, whereas someone with a faculty appointment does not have to undertake training. This very clear gatekeeping is likely to be affected by the fact that the faculty population is much less diverse than the graduate student and postdoc populations, and it is sending a harmful message that reinforces that graduate students and postdocs are incompetent. The notion of "expertise" is highly subjective, and as such is likely to be affected by the typical biases we have come to know within academia.

This is just one example of a policy that I would encourage NIH to engage in tracking. As a taxpayer, I am very concerned about how effectively tax dollars are being spent at magazines to publish articles, not least because of the evidence for biased pools of peer reviewers, and subjective decisions by editors, that are gatekeeping the very resources used to help - or hinder - the career progress of early career researchers. For example, eLife found that interventions were needed to ensure that early career reviewers in the reviewer database were actually being used and selected by editors; even when we see ECRs being added to a reviewer database, it isn't enough, because editors can pick reviewers using subjective assessments (<a href="https://elifesciences.org/inside-elife/958c61d1/early-career-reviewers-reflections-on-focused-inclusion-in-reviews-at-elife">https://elifesciences.org/inside-elife/958c61d1/early-career-reviewers-reflections-on-focused-inclusion-in-reviews-at-elife</a>). This isn't restricted to career stage - faculty peer review frequency also appears to vary with perceived institutional prestige. Just last week, I attended a national conference where a journal editor stated publicly that they do not "need" to check the

reviewer database when receiving a submission - they "already have the names in their head". Clearly such a limited and subjective pool of reviewers will give a limited perspective on academic works, and it's not a system that I'm very pleased to see the taxpayer supporting. I encourage NIH to signal that magazines need to have very good justifications for their high costs to the taxpayer. After attending the Peer Review Congress in 2022, I managed to came away with a lower opinion - and a much more evidence-based one - of magazines and their peer review processes than before. Their claims that they are providing value for money do not seem to stand up to much scrutiny - not least when one views their activities through the lens of equity.

I want to highlight that the NIH also has a working group on postdocs, and a key reason that researchers are stagnating in postdoctoral roles, and therefore a possible factor in why increasingly graduate students are choosing not to undertake an academic postdoc, is time taken for magazines to publish their articles. At this precise moment I myself am working on the second request for revisions, for a paper submitted 7 months ago. The major motivation for our group publishing this article in a magazine is because we need to support our graduate student author in their academic career aspirations. I mention this not only because this is actually a very normal timeline for the review process, that is somehow acceptable to the academic community. It is plainly ridiculous that people's careers are being held up not because of any training needs, but because of the inability of magazines to fulfill their role. This is costing the taxpayer money not only in the lengthy publishing process, but also because a significant number of these researchers are themselves are supported by taxpayer funding, and are now stagnating longer than needed at the taxpayer's expense. This is a clear opportunity for NIH to recognize preprints from graduate students and postdocs for use in evaluation of productivity, as the length of time a magazine takes to publish its articles is out of the control of any individual early career researcher, and should not be a deciding factor in selection of future faculty.

With respect to preprints, I would encourage NIH to consider federal funding for a community preprint infrastructure. I would also like to take the opportunity to point out that as academic researchers already review each Other's work for free, they could publish preprints on a federal government server and then review each Other's work all for free, and this would save the taxpayer a lot of money. It would also come with the benefit of being able to publish various kinds of research, experiments, figures, data and metadata. It could therefore be less restrictive, and much cheaper, than the current magazine publishing model.

#### 4. Early input on considerations to increase findability and transparency of research.

NIH should require everyone to have an ORCID. I would like to point out that many foundations and Other funders already require ORCIDs, and it is my understanding that ORCIDs provide the only feasible means of satisfying upcoming federal policies, and will likely be required of all agencies anyway. NIH should also require the use of ORCIDs by its funded institutions, to allow connection of institutional data with their researchers, funding and publications for NIH-funded research.

NIH should assign DOIs to grants to allow them to be citable products.

NIH should index all preprints, and not just those supported by NIH investigators.

I would encourage NIH to participate in, and ensure interoperability with, global initiatives and efforts in Other countries.

Again, I would encourage NIH to look into the DocMaps Framework (<a href="https://docmaps.knowledgefutures.org">https://docmaps.knowledgefutures.org</a>). This is a project by the Knowledge Futures Group to develop a community-endorsed framework for capturing valuable context about the processes used to create documents in a machine-readable way. Please see "The DocMaps Framework for representing assertions on research products in an extensible, machine-readable, and discoverable format (<a href="https://www.biorxiv.org/content/10.1101/2021.07.13.452204v1">https://www.biorxiv.org/content/10.1101/2021.07.13.452204v1</a>)".

Email: info@lightoller.org

**Submit date:** 4/27/2023

I am responding to this RFI: On behalf of an organization

Name: Douglas White

Name of Organization: American College of Rheumatology

Type of Organization: Professional org association

**Role:** Scientific researcher

1. How to best ensure equity in publication opportunities for NIH-supported investigators.

See attached

2. Steps for improving equity in access and accessibility of publications.

See attached

3. Methods for monitoring evolving costs and impacts on affected communities.

See attached

4. Early input on considerations to increase findability and transparency of research.

See attached

**Uploaded File:** 

FINAL\_ACR\_Comments\_on\_RFI\_NIH\_Roadmap\_Open\_Access\_04.24.23.pdf



#### Empowering rheumatology professionals to excel in their specialty

600 Maine Avenue, SW • 6th Floor • Washington, DC 20024 Phone: (404) 633-3777 • Fax (404) 633-1870 • www.rheumatology.org

April 24, 2023

Lyric Jorgenson, PhD
Acting NIH Associate Director for Science Policy
Office of Science Policy
National Institutes of Health
Department of Health and Human Services
200 Independence Avenue, S.W.
Washington, D.C. 20201

Submitted electronically

RE: [NOT-OD-23-091] Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research

Dear Acting Director Jorgenson,

The American College of Rheumatology (ACR), representing over 7,700 rheumatologists and rheumatology interprofessional team members, appreciates the opportunity to provide comment on the National Institute of Health (NIH) Office of Science Policy (OSP) request for information on the plan for public access to NIH-supported research. While we recognize and support the underlying intent of the open-access policies, we have significant concerns regarding unintended consequences stemming from the implementation of the policy that will adversely affect peer review and young researchers. Therefore, ACR urges a two-year delay in implementing this new policy to allow the scientific community to prepare for the significant impacts and mitigate the deleterious effects this policy will have.

There are undoubtedly advantages to the Office of Science and Technology Policy (OSTP) policy and NIH roadmap. New knowledge described in published scientific manuscripts will become immediately available to researchers, scientists, and the lay public without a subscription. In theory, this should allow efforts to replicate results and the application of new scientific and clinical knowledge faster. However, the proposed NIH roadmap requires additional scrutiny, as it might not achieve the intended goals.<sup>1</sup>

NIH states that they are purposely not requiring a specific publication model and giving funded researchers *options where they may publish* – in Hybrid or Gold OA journals. They further state that deposition of either the accepted peer-reviewed manuscript or the final version of record in a Funder's repository or PubMed is compliant with the stated goal of *immediate*, *cost-free access*. However, by not favoring a publication model, this position creates unintended complications by equating Green and Gold OA.

Immediate access to the accepted manuscript, without embargo, is expected to prompt widespread subscription cancellations. This is problematic because competition among journals for subscriptions incentivizes high-quality peer review. In a publishing landscape where revenue is fundamentally linked to the number of published manuscripts (as opposed to the subscriptions sold), the incentive for high-quality peer review is diminished. This jeopardizes the ACR's and similar organizations' ability to invest in high-quality peer-reviewed journals

that our readers in the medical community rely on. Moreover, zero embargo Green OA undermines the value of the Version of Record that Hybrid and Gold OA journals provide. Multiple versions of a manuscript deposited in PubMed and elsewhere will predictably create chaos and confusion in terms of version control and in the oversight of the scientific record, leading to multiple versions of persistent manuscript identifiers and metadata. Pressure on subscription revenue will further accelerate the transition to fully Gold OA, penalizing early-career researchers and researchers from lower-income countries without funding or robust academic records. Publishers may also opt to create a new revenue stream, in an attempt to offset short-term subscription losses, by charging authors to deposit accepted manuscripts in the funder's repository, further disadvantaging researchers with limited resources.

The economic reality is that publishers will have to recoup their expenses and make at least some financial profit to survive. If a 12-month embargo is removed and published articles are made immediately available, journal subscription revenue will dwindle, and publishers will be forced to move toward Gold Open Access. Publishing an article in an ACR journal with immediate Open Access is currently associated with publication fees ranging from (USD) \$3080 to \$4940, which falls on the authors to pay. Thus, the new OSTP policy could potentially push the financial burden of making publications immediately available to the authors. In its current form, the NIH roadmap does not suggest or create resources for these additional publication fees, and the NIH has not clarified whether they will pay these fees. In a very real and practical sense, authors will be forced to use research budgets to fund this new mandate. <sup>2</sup>

The OSTP policy has the potential to increase inequity in science. Scientists will be forced into a pay-to-publish model. This will be manageable for researchers with substantial funds; however, many researchers will find these fees prohibitive. When funds are unavailable, publishing completed work will be delayed or abandoned, hindering the dissemination of new knowledge – precisely the opposite of the desired OSTP and NIH policy goals. Moreover, junior scientists who often have limited funds will be impacted more than established senior scientists. Researchers from countries with more limited resources will not have a chance to publish in prestigious journals, forced by the new policy to switch from a subscription to a Gold Open Access model.<sup>3</sup>

From the perspective of the Publisher, an expanded pay-to-publish model will only be sustainable by increasing the volume of accepted manuscripts. This will negatively impact rigor and reproducibility in scientific publications and further burden an already shrinking reviewer pool. <sup>4</sup> We continue to see many predatory journals, and the new policy will accelerate the move to low-quality scientific publications.

Publications from the ACR and other medical and scientific societies provide an important platform to disseminate the most significant advances in specific medical and scientific fields. Historically, some of the most impactful and paradigm-shifting work has been published in society journals, where external, rigorous, scientific peer review is critical. Unfortunately, encouraging a pay-to-publish model puts society journals (and medical societies) at substantial financial risk while jeopardizing scientific excellence in biomedical research. <sup>5</sup>

The ACR believes this policy and the subsequent roadmap will create unintended negative consequences in equity, quality, scientific record oversight, and financial sustainability. We strongly support public and immediate access to medical and scientific advances. However, we strongly believe that the OSTP policy and NIH roadmap, as currently articulated, will create negative consequences for the future of scientific research. We urge a more careful examination of the updated policy, a more extended time to hear concerns from medical societies and the public and consideration of alternatives that can increase access to scientific publications while maintaining quality. <sup>6</sup> Please contact Amanda Grimm Wiegrefe, MScHSRA, ACR Director of Regulatory Affairs, at <a href="mailto:awiegrefe@rheumatology.org">awiegrefe@rheumatology.org</a> with questions.

Sincerely,

1.11s

Douglas White, MD, PhD President, American College of Rheumatology

1,2,3,4,5,6. Sawalha AH, Solomon DH, Allen KD, Katz P, Yelin E. Immediate Open Access: The Good, the Bad, and the Impact on Academic Society Publishing. Arthritis Rheumatol 2023;75. In Press.

**Submit date:** 4/27/2023

I am responding to this RFI: On behalf of an organization

Name of Organization: Fully OA Publishers

Type of Organization: Not applicable

1. How to best ensure equity in publication opportunities for NIH-supported investigators.

On public repositories, we believe the NIH Public Access Plan rightly encourages and prioritises the widest possible choices for researchers as they relate to publishing venue, as well as the principles of academic freedom. We think the Plan strikes the right balance by making PubMed Central (PMC) a convenient and compliant repository for research without privileging or mandating it.

On the fairness of the article processing charge (APC), it is worth noting this charge is not an inevitable component of Gold Open Access (OA) publishing. Indeed, we recognize that in some cases, it is not the preferred or most sustainable price structure for researchers, funders, libraries, and research institutions. And while we, like others in the publishing industry, think the APC model is a good one, we are not in principle wedded to it. We are continually in touch with institutional partners to find solutions that meet their needs. For APCs to remain affordable, there must be fair competition on a level playing field between legacy publishers and pure open access publishers or other innovative platforms, and researchers should be rewarded to use publication funds responsibly. So called "transformative agreements" or Read&Publish agreements, where legacy publishers sell journals to libraries with subscription fees that bundle access to back-articles with coverage of APCs to publish in their journals, are in our view anticompetitive as they encourage researchers to publish in legacy titles regardless of the APC-level. Full OA publishers have nothing to "transform" so they are not included in such agreements. Instead of enabling a true competition between pure OA publishers and legacy publishers, transformative agreements subsidise publication in legacy titles and contribute to a oligopolistic publishing ecosystem by ignoring the fact that researchers may disseminate their work with other publishers (including pure open access publishers) or platforms more cost-effectively. Frameworks such as Plan P (planp.science) address the APC problem with creating a

transparent market place for publication opportunities for researchers after they made their preprint available to the public, and also support a multipayer environment, where the APC is ultimately covered by both the institution and funders.

On the additional steps the NIH might take to ensure new inequities are not created, or existing ones reinforced, we believe the NIH should

- Implement policies that make sure that institutions and libraries offer equitable publication opportunities by creating, supporting, or mandating institutional open access funds that support cost-effective peer-review and publication in all accredited open access venues, outlawing transformative agreements without the presence of a generic institutional open access fund that supports open access publication in any accredited OA journal. "Accreditation" could use existing "white-lists" such as DOAJ or OASPA membership, or be the results of an institutional/federal procurement/RFI process to create an institutional list of "accredited" OA journals that receive a APC subsidy
- encourage researchers to publish in the Gold OA model on the basis that the public funding of public access is efficient, scalable, and delivers value for money.
- Encourage researchers to make their publications available as preprints first
- Find mechanisms that support a multipayer model, where the costs APCs are shared between institutions and funders, and to make billing processes as frictionless as possible for researchers.

In our view, Gold OA publishing is one of the most effective ways of securing that outcome. It offers a simple, transparent, and competitive way to unlock the benefits of fully accessible science; and it enables researchers, agencies, universities, libraries, and repositories to fulfil both the NIH Public Access Policy and the OSTP guidance. Publishing in a Gold OA journal immediately facilitates the transfer of articles to a repository, with metadata in machine-readable formats. In this model, there are no embargoes and no superfluous or costly bundled services that are common in "hybrid" or "transformative" subscription options offered by legacy commercial publishers.

On public value for money, new federal guidelines seek public access but do not specify

delivery models. We agree that openly accessible science can - and should - be delivered by more than one publishing model. We welcome competition if it spurs innovation and the amount of rigorous science accessible to all.

But in judging those delivery models, federal agencies must make a robust and transparent assessment and comparison for efficiency, scalability, and public value for money - guided by the objective of discoverability that underpins public access.

For example, public access known as "Green Open Access (OA)" clearly removes some barriers and does not create or perpetuate inequity. But the mechanisms for finding, reading, and sharing Green OA files vary widely, and the level of peer-review is not always clear. Substantial new funding will be required just to bring that variance down and lift standards for discoverability, with new investment in infrastructure for metadata enrichment. Those institutions unable to fund that investment are likely to face the continued cost pressure of paywall subscriptions that might only minimally ease search and discovery.

So, it is vital that the funding of public access is as efficient, scalable, and as good a value for money as possible, and in our view, Gold OA publishing is one of the most effective ways of securing that outcome. It offers a simple, transparent, and competitive way to unlock of the benefits of fully accessible science.

#### 2. Steps for improving equity in access and accessibility of publications.

On the 12-month embargo, we strongly welcome the NIH's decision to end it on publications. We believe that so-called Transformative Agreements (TAs) were worthwhile in their conception as a means of smoothing the transition to fully open access science, but in their execution have become a blunt instrument.

TAs lack transparency, have complex bundles of services making it all but impossible to judge value for money, and come with no contractual commitment to a move to full open access (Green, Gold, or otherwise) within a binding deadline.

Most of these agreements are large scale "read and publish" or hybrid deals. Publishers will often allow authors to appear in their hybrid journals without being charged, if their institutions pay, while at the same time they maintain the amount of science they publish behind paywalls.

We believe TAs help prop up the market dominance of legacy publishers by controlling the pace of transition to fully open access science. The worldwide scientific publishing oligopoly is a market estimated to be around US \$27 billion.1 The five largest paywall publishing houses2 have captured more than half of it.3

On the basis the NIH seeks equity in access as well as transparency in costs, backed by financial sustainability, we believe Gold OA publishers can deliver.

On automated text processing, assistive devices, and other inclusionary measures, we fully support the NIH's position. We consistently invest in measures that improve the accessibility of our publications. Many such requirements were mandated by the Coalition S initiative, which this group fully supports, and which saw wide-ranging and progressive open access policies adopted by funders in the US, in the United Kingdom and across Europe.

We firmly back public policies that promote equity of opportunity, the ability both to read and

publish research, and the scientific rigor, academic freedom, institutional values, and personal and professional recognition that underpin success.

We are committed to increasing research access, knowledge resources, and educational opportunities for all, especially for those groups, nations, and individuals who are historically marginalized, underrepresented, or disadvantaged.

On institutional success, we work hard to build communities and tackle the inadequacies and inequities often characterizing research dissemination. The shift toward open access represents an opportunity to expand access to knowledge in a significant way across academic institutions of all stripes, as well as small businesses and the public.

We would urge the NIH to draw on its influence to see that library, research, and educational institutions commit to investing in open access so that all parties can source sufficient funding for publishing. Several equitable open publishing models are readily available. It cannot be right if colleges and universities are encouraged to maintain robust publications budgets for subscriptions and then asked to make cuts to open access. Many institutions initially supported open access with the hope that it will reduce library costs for subscriptions, and signed statements like the Compact for Open Access Equity (COPE,

http://www.oacompact.org/), which vowed that there will always be institutional support to

help with APCs; unfortunately, in many cases such institutional funds are no longer available as libraries make deals with traditional publishers that fund only their APCs (<a href="https://scholarworks.duke.edu/open-access/cope/">https://scholarworks.duke.edu/open-access/cope/</a>).

We believe there is enough funding in the system to make the transition to open access complete. But that funding can only be unlocked with public sector, policymaker, and buyer leadership, on the basis we look beyond legacy publishing models that have been responsible for a decades-long cost explosion in scholarly publishing.4 With the right policies and incentives, agencies can help drive the value of taxpayer-funded investment and spur innovation.

#### 3. Methods for monitoring evolving costs and impacts on affected communities.

On financial costs, we welcome the NIH's interest in the commercial drivers of scholarly publishing, particularly in matters of access or equity.

Since our inception as a born-digital publisher, we have sought to reduce or remove financial and operational burdens facing researchers. The governing principle of all scholarly publishing should be that the researchers have the most freedom possible to focus on their research. And so, all publishers compete to lower administrative and process-based burdens.

While the dissemination of research requires a complex ecosystem, we believe a wide-scale shift to open access would allow libraries and research institutions to free substantial resources now tied up in (paywall) subscriptions, and to apply those resources to researchers' publishing costs.

A strong signal or directive from the NIH that research institutions should commit these freed-up funds - as well as grant money ringfenced for publication - to the widespread and immediate sharing of research would have a profound and positive impact on the drive to fully open access science.

On the perceived relative fairness of pricing regimes, and as we say in response to Question 1, it is worth noting the APC is not an inevitable component of Gold OA publishing. While we think the APC model is a good one - not least because it brings greater costs transparency for monitoring purposes - we are not in principle wedded to it. We are continually in touch

with institutional partners to find solutions that meet their needs. We are seeking to shift the funding paradigm to help authors cover the fair and actual cost of publishing, to make scientific knowledge accessible to the widest possible audience.

Within an APC framework, we have expanded our portfolio of institutional models to meet the tailored needs of our customers (with, for example, institutional partnerships for research-intensive "publish" organizations as well as high consumption "read" institutions and societies). Our success indicates a range of pricing regimes can meet the needs of a range of customers and institutions.

The publishing industry at large is experimenting with pricing regimes and introducing new ones in its drive to innovate. Though the nomenclature varies - advance annual payment, fixed fee, flat fee, multi-payer, Subscribe 2 Open, waivers - all of these seek to offer more cost-efficient and sustainable alternatives to libraries' subscription expenditure.

#### 4. Early input on considerations to increase findability and transparency of research.

On data sharing, we fully back the NIH's effort through its Public Access Plan to spur a better and more consistent use of PIDs and metadata. In driving this effort, the NIH is providing critical leadership in the scholarly publishing ecosystem.

Moreover, we welcome the NIH's focus on the findability and transparency of research.

Open data drives scientific and technological innovation and spurs collaboration; is critical to driving efficiency and scaling innovation; and in uniform standards can be verified, reproduced, and built upon.

If data is transparent and open to scrutiny and evaluation, it follows that trust and confidence in science are more likely to be sustainable. The infrastructure for open data is readily available and an increasingly frequent resource; and many large-scale repositories already exist to make data open. Examples include Figshare, a commercial, field-agnostic repository; field-specific, non-profit databases like the society-supported FlowRepository for cytometry data and the commercial Protein Data Bank; and federally backed databases like NIH's data repositories.

On data repositories, substantial funding will be required for operation and upgrades. And in the absence of funding committed to scaling up PMC, we would back a federated approach that focuses on shared standards and access across multiple repositories. By way of illustration, we deposit the full text or metadata of our 230-plus journals in more than 20 repositories when we publish articles.

As a group of fully OA publishers, together we have made thousands of peer-reviewed articles available online immediately, without embargo. Our starting point - and end point - is ease of discovery.

In simple terms, an article that cannot be found, cannot be shared, and cannot be cited also cannot spur vital collaboration and breakthrough. Publishing in a Gold OA journal unlocks discoverability. The articles and underlying data are transferred to a repository such as PubMed Central or stored in commercial or other non-profit databases.

Moreover, the metadata from Gold OA journals come in XML files and other machine-readable formats to meet FAIR data standards of findability, accessibility, interoperability, and reuse. The metadata includes persistent identifiers such as that of ORCID for author identification, a Digital Object Identifier (DOI) for the article itself, and tags to the relevant grant funding or research institution. And compliance with JATS DTD for XML and other PMC-recommended tagging enables an even more efficient search and discovery experience.

Open science is all about transparency and the quality of science is expected to increase if transparency increases, e.g. by publishing protocols. While it is common to publish and register clinical trials, NIH could do more to make other forms of research more transparent. In terms of identifying protocols or grant proposals, some signatories of this letter have pioneered the use of a new persistent identifiers (PID) called IRRID (International Registered Report Identifier, <a href="https://irridregistry.org/">https://irridregistry.org/</a>), which uses the DOI system to link protocols and grant proposals (RR1) to results papers (RR2). If a protocol or grant proposal is published with a DOI, the IRRID in the results paper links back to the protocol. Together, RR1 and RR2 form "registered reports", which is the idea that scientists should publish the protocol or proposal of their work first, and then the results paper, which should be published regardless of whether the findings are negative or positive. NIH as funding agency could encourage protocol and proposal publication by

- making peer-review reports from NIH review committees openly accessible under a Creative Commons licence if the principal investigator and reviewers agree
- encouraging NIH-funded researchers to formally publish their protocols and grant proposal if they are successful so they receive a DOI and a IRRID
- encoruage or mandate to cite the protocol or grant proposal using a IRRID in the abstract of any results paper

The new federal guidelines seek public access without specifying delivery models, and we agree openly accessible science can - and should - be delivered by more than one publishing model. We welcome competition if it spurs innovation and the amount of rigorous science accessible to all.

But in judging delivery models, we believe federal agencies must make a robust and transparent assessment and comparison across efficiency, scalability, and public value for money - guided by the objective of discoverability that underpins public access.

For example, public access known as "Green Open Access (OA)" clearly removes some barriers and does not create or perpetuate inequity. But the mechanisms for finding, reading, and sharing Green OA files vary widely. Substantial new funding will be required just to bring that variance down and lift standards for discoverability, with new investment in infrastructure for metadata enrichment. Those institutions unable to fund that investment are likely to face the continued cost pressure of paywall subscriptions that might only minimally ease search and discovery.

So, it is vital that the funding of public access is as efficient, scalable, and as good a value for money as possible, and in our view, Gold OA publishing is one of the most effective ways of securing that outcome. It offers a simple, transparent, and competitive way to unlock of the benefits of fully accessible science.

We think it is possible to achieve the fullest possible access to our collective knowledge - for fairer outcomes in all parts of society - in a business model that is cost-effective, commercially sustainable, and underpinned by private sector innovation.

We stand ready to support the NIH and its partners in the federal government. It is vital we back this effort for open science and meet the public appetite for accountability,

transparency, and trust.

# **Uploaded File:**

 $NIH\_call\text{-}submission\_from\_Fully\_OA\_group.pdf$ 

# Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research

April 24, 2023

# **Headline summary**

We welcome the chance to respond to this important <u>request for information</u> from the National Institutes of Health (NIH). The <u>Fully OA Publishers</u> group currently comprises nine publishers dedicated to Open Access, disseminating high-quality research and data to the broadest possible public audience. A significant portion of the science we publish is Federally Funded, and all of it is peer-reviewed, globally shared and free to read. This submission comes from the 3 Fully OA Group Publishers referenced at the end of the document, although we are also aware other members of the Fully OA Group have submitted replies directly to NIH as well.

Our shared mission is to make all science open – so that we can collaborate better and innovate faster, for fairer and more equitable outcomes in all parts of society. That is a key social purpose of our businesses.

So, we fully support the August 2022 OSTP (Office of Science and Technology Policy) guidelines. We think the NIH has posed critical questions in this request for information, not least about the findability and transparency of research.

As a group of fully OA publishers, we have made hundreds of thousands of peer-reviewed articles available online immediately, without embargo. Our starting point – and end point – is ease of discovery.

In simple terms, an article that cannot be found, cannot be shared, and cannot be cited, clearly cannot spur collaboration and breakthrough. Publishing in a Gold OA journal unlocks discoverability. The articles and underlying data are transferred to a repository such as <a href="PubMed Central">PubMed Central</a> or stored in commercial or other non-profit databases. The metadata come in XML files and other machine-readable formats to meet <a href="FAIR">FAIR</a> data standards</a> of findability, accessibility, interoperability, and reuse. And that data includes persistent identifiers such as that of <a href="ORCID">ORCID</a> for author identification, a Digital Object Identifier (DOI) for the article itself, and tags to the relevant grant funding or research institution.

The new federal guidelines seek public access but do not specify delivery models. We agree that openly accessible science can – and should – be delivered by more than one publishing model. We welcome competition if it spurs innovation and the amount of rigorous science accessible to all.

But in judging those delivery models, federal agencies must make a robust and transparent assessment to compare them for efficiency, scalability, and public value for money – guided by the objective of discoverability that underpins public access.

For example, public access known as "Green Open Access (OA)" (which includes depositing preprints of drafts, submitted or accepted manuscripts on preprint servers) clearly removes some barriers and does not create or perpetuate inequity. But the mechanisms for finding, reading, and sharing Green OA files vary widely, and provenance, e.g. the level of peer-review or endorsement of the scientific community, is not always clear. Substantial new funding will be required just to bring that variance down and lift standards for discoverability, with new investment in infrastructure for metadata enrichment. Those institutions unable to fund that investment are likely to face the continued cost pressure of paywall subscriptions that might only minimally ease search and discovery.

So, it is vital that the funding of public access is as efficient, scalable, and as good a value for money as possible, and in our view, Gold OA publishing is one of the most effective ways of securing that outcome. It offers a simple, transparent, and competitive way to unlock the benefits of fully accessible science.

We think it is possible to achieve the fullest possible access to our collective knowledge – for fairer outcomes in all parts of society – in a business model that is cost-effective, commercially sustainable, and underpinned by private sector innovation. We stand ready to support the NIH and its partners in the federal government. It is vital we back this effort for open science and meet the public appetite for accountability, transparency, and trust.

#### **Full response**

Our detailed responses to the NIH's framing (in italics) are set out here.

# 1. How to best ensure equity in publication opportunities for NIH-supported investigators.

On public repositories, we believe the NIH Public Access Plan rightly encourages and prioritises the widest possible choices for researchers as they relate to publishing venue, as well as the principles of academic freedom. We think the Plan strikes the right balance by making PubMed Central (PMC) a convenient and compliant repository for research without privileging or mandating it.

On the fairness of the article processing charge (APC), it is worth noting this charge is not an inevitable component of Gold Open Access (OA) publishing. Indeed, we recognize that in some cases, it is not the preferred or most sustainable price structure for researchers, funders, libraries, and research institutions. And while we, like others in the publishing industry, think the APC model is a good one, we are not in principle wedded to it. We are continually in touch with institutional partners to find solutions that meet their needs.

For APCs to remain affordable, there must be fair competition on a level playing field between legacy publishers and pure open access publishers or other innovative platforms, and researchers should be rewarded to use publication funds responsibly. So called "transformative agreements" or Read&Publish agreements, where legacy publishers sell journals to libraries with subscription fees that bundle access to back-articles with coverage of APCs to publish in their journals, are in our view anticompetitive as they encourage researchers to publish in legacy titles regardless of the APC-level. Full OA publishers have

nothing to "transform" so they are not included in such agreements. Instead of enabling a true competition between pure OA publishers and legacy publishers, transformative agreements subsidise publication in legacy titles and contribute to a oligopolistic publishing ecosystem by ignoring the fact that researchers may disseminate their work with other publishers (including pure open access publishers) or platforms more cost-effectively.

Frameworks such as <u>Plan P</u> (planp.science) address the APC problem with creating a transparent market place for publication opportunities for researchers after they made their preprint available to the public, and also support a multipayer environment, where the APC is ultimately covered by both the institution and funders.

On the additional steps the NIH might take to ensure new inequities are not created, or existing ones reinforced, we believe the NIH should

- Implement policies that make sure that institutions and libraries offer equitable publication opportunities by creating, supporting, or mandating institutional open access funds that support cost-effective peer-review and publication in all accredited open access venues, outlawing transformative agreements without the presence of a generic institutional open access fund that supports open access publication in any accredited OA journal. "Accreditation" could use existing "white-lists" such as DOAJ or OASPA membership, or be the results of an institutional/federal procurement/RFI process to create an institutional list of "accredited" OA journals that receive a APC subsidy
- encourage researchers to publish in the Gold OA model on the basis that the public funding of public access is efficient, scalable, and delivers value for money.
- Encourage researchers to make their publications available as preprints first
- Find mechanisms that support a multipayer model, where the costs APCs are shared between institutions and funders, and to make billing processes as frictionless as possible for researchers.

In our view, Gold OA publishing is one of the most effective ways of securing that outcome. It offers a simple, transparent, and competitive way to unlock the benefits of fully accessible science; and it enables researchers, agencies, universities, libraries, and repositories to fulfil both the NIH Public Access Policy and the OSTP guidance. Publishing in a Gold OA journal immediately facilitates the transfer of articles to a repository, with metadata in machine-readable formats. In this model, there are no embargoes and no superfluous or costly bundled services that are common in "hybrid" or "transformative" subscription options offered by legacy commercial publishers.

On public value for money, new federal guidelines seek public access but do not specify delivery models. We agree that openly accessible science can – and should – be delivered by more than one publishing model. We welcome competition if it spurs innovation and the amount of rigorous science accessible to all.

But in judging those delivery models, federal agencies must make a robust and transparent assessment and comparison for efficiency, scalability, and public value for money – guided by the objective of discoverability that underpins public access.

For example, public access known as "Green Open Access (OA)" clearly removes some barriers and does not create or perpetuate inequity. But the mechanisms for finding, reading, and sharing Green OA files vary widely, and the level of peer-review is not always clear. Substantial new funding will be required just to bring that variance down and lift standards for discoverability, with new investment in infrastructure for metadata enrichment. Those institutions unable to fund that investment are likely to face the continued cost pressure of paywall subscriptions that might only minimally ease search and discovery.

So, it is vital that the funding of public access is as efficient, scalable, and as good a value for money as possible, and in our view, Gold OA publishing is one of the most effective ways of securing that outcome. It offers a simple, transparent, and competitive way to unlock of the benefits of fully accessible science.

# 2. Steps for improving equity in access and accessibility of publications.

On the 12-month embargo, we strongly welcome the NIH's decision to end it on publications. We believe that so-called Transformative Agreements (TAs) were worthwhile in their conception as a means of smoothing the transition to fully open access science, but in their execution have become a blunt instrument.

TAs lack transparency, have complex bundles of services making it all but impossible to judge value for money, and come with no contractual commitment to a move to full open access (Green, Gold, or otherwise) within a binding deadline.

Most of these agreements are large scale "read and publish" or hybrid deals. Publishers will often allow authors to appear in their hybrid journals without being charged, if their institutions pay, while at the same time they maintain the amount of science they publish behind paywalls.

We believe TAs help prop up the market dominance of legacy publishers by controlling the pace of transition to fully open access science. The worldwide scientific publishing oligopoly is a market estimated to be around US \$27 billion. The five largest paywall publishing houses have captured more than half of it.

On the basis the NIH seeks equity in access as well as transparency in costs, backed by financial sustainability, we believe Gold OA publishers can deliver.

On automated text processing, assistive devices, and other inclusionary measures, we fully support the NIH's position. We consistently invest in measures that improve the accessibility of our publications. Many such requirements were mandated by the Coalition S initiative, which this group fully supports, and which saw wide-ranging and progressive open access policies adopted by funders in the US, in the United Kingdom and across Europe.

We firmly back public policies that promote equity of opportunity, the ability both to read and publish research, and the scientific rigor, academic freedom, institutional values, and personal and professional recognition that underpin success.

We are committed to increasing research access, knowledge resources, and educational opportunities for all, especially for those groups, nations, and individuals who are historically marginalized, underrepresented, or disadvantaged.

On institutional success, we work hard to build communities and tackle the inadequacies and inequities often characterizing research dissemination. The shift toward open access represents an opportunity to expand access to knowledge in a significant way across academic institutions of all stripes, as well as small businesses and the public.

We would urge the NIH to draw on its influence to see that library, research, and educational institutions commit to investing in open access so that all parties can source sufficient funding for publishing. Several equitable open publishing models are readily available. It cannot be right if colleges and universities are encouraged to maintain robust publications budgets for subscriptions and then asked to make cuts to open access. Many institutions initially supported open access with the hope that it will reduce library costs for subscriptions, and signed statements like the Compact for Open Access Equity (COPE, <a href="http://www.oacompact.org/">http://www.oacompact.org/</a>), which vowed that there will always be institutional support to help with APCs; unfortunately, in many cases such institutional funds are no longer available as libraries make deals with traditional publishers that fund only their APCs (https://scholarworks.duke.edu/open-access/cope/).

We believe there is enough funding in the system to make the transition to open access complete. But that funding can only be unlocked with public sector, policymaker, and buyer leadership, on the basis we look beyond legacy publishing models that have been responsible for a decades-long cost explosion in scholarly publishing.<sup>4</sup> With the right policies and incentives, agencies can help drive the value of taxpayer-funded investment and spur innovation.

# 3. Methods for monitoring evolving costs and impacts on affected communities.

On financial costs, we welcome the NIH's interest in the commercial drivers of scholarly publishing, particularly in matters of access or equity.

Since our inception as a born-digital publisher, we have sought to reduce or remove financial and operational burdens facing researchers. The governing principle of all scholarly publishing should be that the researchers have the most freedom possible to focus on their research. And so, all publishers compete to lower administrative and process-based burdens.

While the dissemination of research requires a complex ecosystem, we believe a wide-scale shift to open access would allow libraries and research institutions to free substantial resources now tied up in (paywall) subscriptions, and to apply those resources to researchers' publishing costs.

A strong signal or directive from the NIH that research institutions should commit these freed-up funds – as well as grant money ringfenced for publication – to the widespread and

immediate sharing of research would have a profound and positive impact on the drive to fully open access science.

On the perceived relative fairness of pricing regimes, and as we say in response to Question 1, it is worth noting the APC is not an inevitable component of Gold OA publishing. While we think the APC model is a good one – not least because it brings greater costs transparency for monitoring purposes – we are not in principle wedded to it. We are continually in touch with institutional partners to find solutions that meet their needs. We are seeking to shift the funding paradigm to help authors cover the fair and actual cost of publishing, to make scientific knowledge accessible to the widest possible audience.

Within an APC framework, we have expanded our portfolio of institutional models to meet the tailored needs of our customers (with, for example, institutional partnerships for research-intensive "publish" organizations as well as high consumption "read" institutions and societies). Our success indicates a range of pricing regimes can meet the needs of a range of customers and institutions.

The publishing industry at large is experimenting with pricing regimes and introducing new ones in its drive to innovate. Though the nomenclature varies – advance annual payment, fixed fee, flat fee, multi-payer, Subscribe 2 Open, waivers – all of these seek to offer more cost-efficient and sustainable alternatives to libraries' subscription expenditure.

# 4. Early input on considerations to increase findability and transparency of research.

On data sharing, we fully back the NIH's effort through its Public Access Plan to spur a better and more consistent use of PIDs and metadata. In driving this effort, the NIH is providing critical leadership in the scholarly publishing ecosystem.

Moreover, we welcome the NIH's focus on the findability and transparency of research. Open data drives scientific and technological innovation and spurs collaboration; is critical to driving efficiency and scaling innovation; and in uniform standards can be verified, reproduced, and built upon.

If data is transparent and open to scrutiny and evaluation, it follows that trust and confidence in science are more likely to be sustainable. The infrastructure for open data is readily available and an increasingly frequent resource; and many large-scale repositories already exist to make data open. Examples include <a href="Figshare">Figshare</a>, a commercial, field-agnostic repository; field-specific, non-profit databases like the society-supported <a href="FlowRepository">FlowRepository</a> for cytometry data and the commercial <a href="Protein Data Bank">Protein Data Bank</a>; and federally backed databases like NIH's <a href="data repositories">data repositories</a>.

On data repositories, substantial funding will be required for operation and upgrades. And in the absence of funding committed to scaling up PMC, wewould back a federated approach that focuses on shared standards and access across multiple repositories. By way of illustration, we deposit the full text or metadata of our 230-plus journals in more than 20 repositories when we publish articles.

As a group of fully OA publishers, together we have made thousands of peer-reviewed articles available online immediately, without embargo. Our starting point – and end point – is ease of discovery.

In simple terms, an article that cannot be found, cannot be shared, and cannot be cited also cannot spur vital collaboration and breakthrough. Publishing in a Gold OA journal unlocks discoverability. The articles and underlying data are transferred to a repository such as <a href="PubMed Central">PubMed Central</a> or stored in commercial or other non-profit databases.

Moreover, the metadata from Gold OA journals come in XML files and other machine-readable formats to meet <u>FAIR data standards</u> of findability, accessibility, interoperability, and reuse. The metadata includes persistent identifiers such as that of <u>ORCID</u> for author identification, a Digital Object Identifier (DOI) for the article itself, and tags to the relevant grant funding or research institution. And compliance with JATS DTD for XML and other PMC-recommended tagging enables an even more efficient search and discovery experience.

Open science is all about transparency and the quality of science is expected to increase if transparency increases, e.g. by publishing protocols. While it is common to publish and register clinical trials, NIH could do more to make other forms of research more transparent. In terms of identifying protocols or grant proposals, some signatories of this letter have pioneered the use of a new persistent identifiers (PID) called IRRID (International Registered Report Identifier, https://irridregistry.org/), which uses the DOI system to link protocols and grant proposals (RR1) to results papers (RR2). If a protocol or grant proposal is published with a DOI, the IRRID in the results paper links back to the protocol. Together, RR1 and RR2 form "registered reports", which is the idea that scientists should publish the protocol or proposal of their work first, and then the results paper, which should be published regardless of whether the findings are negative or positive. NIH as funding agency could encourage protocol and proposal publication by

- making peer-review reports from NIH review committees openly accessible under a Creative Commons licence if the principal investigator and reviewers agree
- encouraging NIH-funded researchers to formally publish their protocols and grant proposal if they are successful so they receive a DOI and a IRRID
- encoruage or mandate to cite the protocol or grant proposal using a IRRID in the abstract of any results paper

The new federal guidelines seek public access without specifying delivery models, and we agree openly accessible science can – and should – be delivered by more than one publishing model. We welcome competition if it spurs innovation and the amount of rigorous science accessible to all.

But in judging delivery models, we believe federal agencies must make a robust and transparent assessment and comparison across efficiency, scalability, and public value for money – guided by the objective of discoverability that underpins public access.

For example, public access known as "Green Open Access (OA)" clearly removes some barriers and does not create or perpetuate inequity. But the mechanisms for finding, reading, and sharing Green OA files vary widely. Substantial new funding will be required just to bring

that variance down and lift standards for discoverability, with new investment in infrastructure for metadata enrichment. Those institutions unable to fund that investment are likely to face the continued cost pressure of paywall subscriptions that might only minimally ease search and discovery.

So, it is vital that the funding of public access is as efficient, scalable, and as good a value for money as possible, and in our view, Gold OA publishing is one of the most effective ways of securing that outcome. It offers a simple, transparent, and competitive way to unlock of the benefits of fully accessible science.

We think it is possible to achieve the fullest possible access to our collective knowledge – for fairer outcomes in all parts of society – in a business model that is cost-effective, commercially sustainable, and underpinned by private sector innovation.

We stand ready to support the NIH and its partners in the federal government. It is vital we back this effort for open science and meet the public appetite for accountability, transparency, and trust.

Publishers in the Fully OA group submitting this response include:

- 1. Frontiers (stephan.kuster@frontiersin.org)
- 2. Ubiquity Press (brian.hole@ubiquitypress.com)
- 3. JMIR Publications, 130 Queens Quay E, Ste. 1100, Toronto M5A 0P6, Canada, support@jmir.org