



**National Institutes of Health**  
*Turning Discovery Into Health*



## **Scientific Management Review Board**

### **DRAFT REPORT ON APPROACHES TO ASSESS THE VALUE OF BIOMEDICAL RESEARCH SUPPORTED BY NIH**

**December 13, 2013**

## EXECUTIVE SUMMARY

The Scientific Management Review Board (SMRB) was established under the National Institutes of Health (NIH) Reform Act of 2006 to advise the NIH Director and other appropriate officials on the use of certain organizational authorities reaffirmed under the same act. In July 2012, the SMRB was charged by NIH Director Francis Collins with helping to identify appropriate parameters and approaches for assessing and communicating the value of biomedical research (VOBR) supported by NIH. In response to the charge, the SMRB assembled the Working Group on Approaches to Assess the Value of Biomedical Research Supported by NIH. The Working Group conducted extensive consultations, were briefed on existing and planned assessment tools and databases, and reported their findings and recommendations to the full SMRB.

The Working Group concluded that, at the present time, the essential tools, techniques, and data systems required for developing comprehensive measurements of value at an aggregate NIH level are in early stages of development. The Working Group identified several ways to improve NIH's assessment efforts, recommending the following:

- 1) NIH should intensify its efforts to systematically, comprehensively, and strategically assess the value of biomedical research for the purposes of accountability, effective management, and public awareness. This will require a sustained investment in strengthening NIH's data infrastructure and a dedicated funding stream or mechanism to support assessment projects.
- 2) Assessments of NIH's value should draw clear connections between the generation of basic and clinical knowledge and the impact of this knowledge along differing translational pathways.
- 3) Credible, interpretable, and useful assessments of the value of NIH should be clear in attributing outcomes to all contributors and adopt a timeframe that is broad enough to include sufficient time for discovery to be applied.
- 4) NIH's assessments should be done in partnership with its many stakeholders.
- 5) NIH should establish a trans-NIH Committee on Assessments that will:
  - a. Develop a strategy to support or conduct assessments of value, including through grants or contracts with external experts;
  - b. Determine a process for strategically selecting study topics that map to a conceptual framework, including different translational pathways;
  - c. Oversee (in conjunction with NIH's recently established "Big Data" committees) NIH efforts to strengthen data needed for assessing value, including:
    - i. Identifying and gaining consensus on a core set of indicators to be included in its data infrastructure;
    - ii. Creating better data linkages with NIH's partners and hand-off sectors;
  - d. Identify promising analytical approaches and develop an assessment approach guide that outlines the factors to consider and the mix of methodologies (e.g., retrospective,

prospective, qualitative, quantitative) that should be employed in attempting to capture value; and

- e. Seek input from external experts in the development of methods and tools to improve assessments of the value of biomedical research.
- 6) Every assessment activity that NIH undertakes should begin with identifying the purpose of the study and its audiences. Assessment study designs should include diverse communication strategies to disseminate results in ways that will enhance awareness and understanding of the scientific research process among a variety of audiences.

NIH should engage experts from all appropriate sectors of our society to develop such needed tools and techniques for better assessing VOBR and utilize external partnerships in conducting assessments to assure objectivity and effectiveness. The findings and recommendations of the Working Group are described in detail in the following report.

## **I. INTRODUCTION**

The American public has entrusted the National Institutes of Health (NIH) with the Nation's largest investment in biomedical research. NIH's mission is to support scientific research aimed at enhancing health, lengthening life, and reducing illness and disability, and there is evidence that throughout its 120 years of existence, NIH has contributed to many of the scientific breakthroughs that have led to tangible improvements in the health of the public. Yet systematically and comprehensively capturing these improvements in ways that clearly link them to the public's investment in NIH remains a significant challenge.

As a steward of public funds, NIH is responsible for using its resources effectively to address the many health challenges that face our nation and the world. NIH uses a well-established, rigorous decision-making process that relies on scientific expertise and stakeholder input when reviewing proposed projects and setting research priorities. As a Federal agency, NIH is obligated to demonstrate the effectiveness of its stewardship of public funds and to communicate the results of its efforts. A better understanding of which management or research approaches are most effective could enhance NIH's priority-setting and decision-making processes and help NIH further increase the value it provides.

Today, because of increasing data sophistication and greater access to information, there are opportunities for NIH and similar agencies around the globe to better account for their efforts. In addition, the current era of constrained budgets and economic challenges requires NIH and other funding institutions to make difficult decisions about how to advance biomedical research with fewer dollars, providing further incentive to improve our understanding of which approaches tend to generate the greatest value.

Before assessment results can inform decisions or communicate value, it is first necessary to determine the diverse types of impact from biomedical research and to establish guidelines and parameters for sound assessments to ensure that the results are credible. For this reason, the NIH Director charged the NIH Scientific Management Review Board (SMRB) with helping to identify the best methods and strategies for assessing the value of NIH research. Specifically, the SMRB is charged with:

- Undertaking a comprehensive analysis of the methods and strategies used in assessing the value of biomedical research;
- Defining the fundamental principles that should underpin any strategy used in assessing the value of investing in NIH; and
- Identifying strategies used in assessing the value of NIH that are scientifically sound and reflect the diverse outcomes related to this investment.

In undertaking this charge, the SMRB created the Working Group on Approaches to Assess the Value of Biomedical Research Supported by NIH to:

- Analyze studies assessing the value of biomedical research across nations and sectors to consider the strengths and weaknesses of different approaches;

- Define a range of diverse outcomes that are or may be attributable to NIH activities, including improvements in human health, advances in knowledge, technological innovations, economic benefits, etc.;
- Identify strategies for assessing the value of biomedical research and articulate the rationale for their selection;
- Seek input from the general public and stakeholders in the biomedical research community, as well as from individuals with expertise in assessing health, scientific, technological, economic, and broader societal impacts of publicly funded research and development in the U.S. and abroad;
- Hold deliberations in conjunction with the full SMRB; and
- Report findings and conclusions to the full SMRB.

Over the course of its deliberations, members of the Working Group reviewed many studies that assess the value of the biomedical enterprise; considered the merits of models, tools, and databases currently used in value assessments; reviewed reports that survey commonly used methods in measuring research; and heard from a broad range of stakeholders and experts in scientific assessment, health, economics, international research, management, and information technology. The Working Group and the entire Board greatly appreciated the time and effort of those who made presentations to the SMRB (Appendix A). A comprehensive list of reports and studies reviewed by the Working Group can be found in Appendix B.

The findings and recommendations presented in this report address ways to strengthen NIH's ability to identify and assess the outcomes of its work so that NIH can more effectively determine the value of its activities, communicate the results of studies assessing value, ensure continued accountability, and further strengthen processes for setting priorities and allocating funds.

## **II. OVERVIEW: ASSESSING THE VALUE OF NIH**

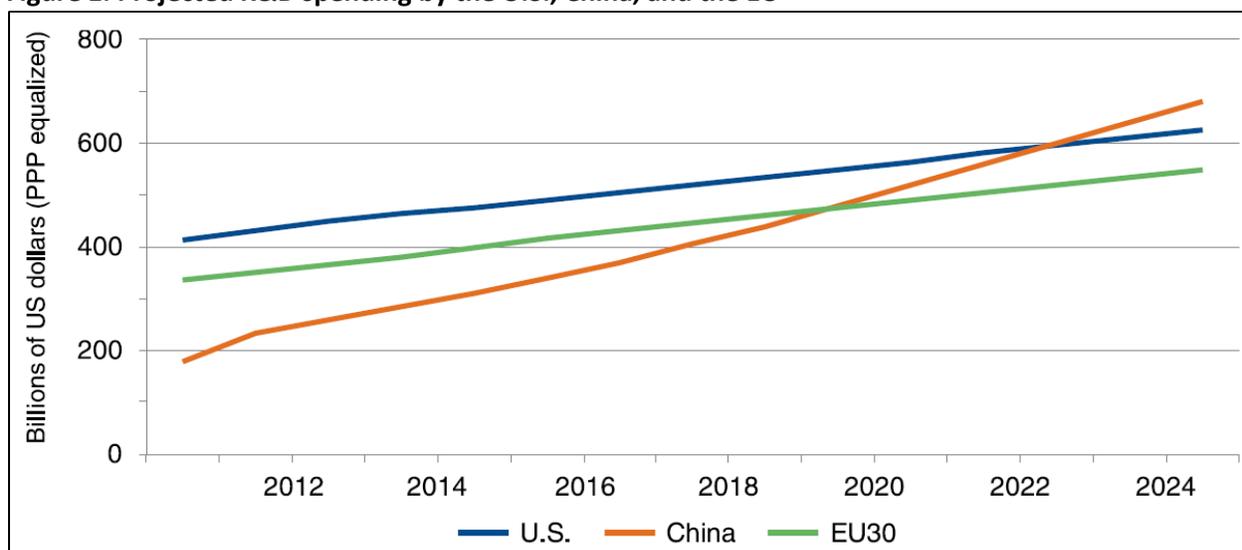
### **A. The biomedical research enterprise**

Biomedical research includes basic research on biological processes and applied research to understand and alter the course of disease and disability. Many scientific fields contribute to biomedical research, working together to develop a greater understanding of human biology and the diseases and conditions that affect humans. The ultimate goal of biomedical research is to improve human health and longevity by providing a scientific knowledge base that is applied within the health ecosystem. Successful application of this knowledge base to human health is a long and often unpredictable process. While this process is far from linear, the concept of a research continuum is helpful to understanding the significant time and effort it may take to move from discovery to health application. The continuum starts with basic research to identify fundamental biological and behavioral mechanisms and how they go awry. This understanding may then be translated into new approaches to identify and prevent, or interrupt disease processes, leading to testing of diagnostic, therapeutic, and preventive approaches

and exploring the means by which these new approaches may benefit all. The length of time from an initial scientific discovery to health impact is substantial, varying over many years or even decades.<sup>1,2</sup>

The biomedical research enterprise spans the globe and is supported by multiple sectors, including federal and state governments, academic institutions, start-up companies, industry, and philanthropic foundations. Each sector has a unique role and advances research, often through the hand-off of knowledge and intellectual property, synergy between experts in different fields, and leveraging of resources. Figure 1 demonstrates the increasing globalization of research and development (R&D),<sup>3</sup> of which biomedical R&D is an important part. This dynamic and interconnected system can advance research by encouraging collaboration and pooling resources and risk. As a result, any one scientist may receive support from a growing number of sectors and funders involved in the biomedical research enterprise, and a particular research finding may advance through multiple sectors on what is often a necessarily complex path to health application.

**Figure 1. Projected R&D spending by the U.S., China, and the EU**



Source: Battelle and R&D Magazine

The U.S. leads the world in biomedical research and development spending,<sup>4</sup> with over \$130 billion total investments in 2011 (Figure 2). In 2011, the Federal government provided approximately \$40 billion in biomedical research funding, allocating approximately \$30 billion of these funds to NIH. The remainder of federal support for biomedical research is spread across more than a dozen other agencies, including

<sup>1</sup> Balas, E. Andrew and Susan A. Boren. *Managing clinical knowledge for health care improvement*. In: Bommel J, McCray AT, editors. *Yearbook of Medical Informatics 2000: Patient-Centered Systems*; 2000:65-70.

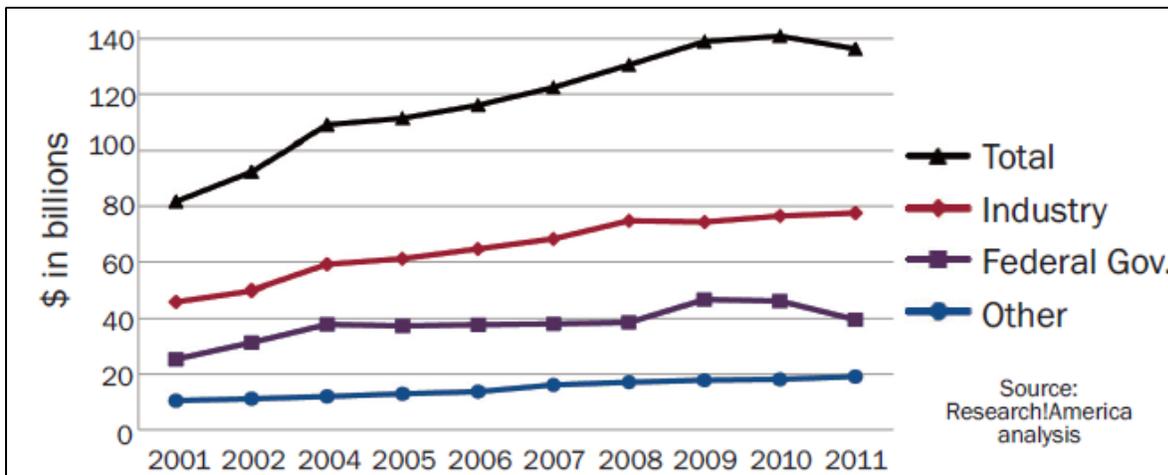
<sup>2</sup> Morris ZS, et al. *The answer is 17 years, what is the question: understanding time lags in translational research*. *Journal of the Royal Society of Medicine*; 2011, vol. 104 no. 12; pgs. 510-520.

<sup>3</sup> Battelle and R&D Magazine. (December 2013). *2014 Global R&D funding forecast*. [http://www.rdmag.com/sites/rdmag.com/files/gff-2014-5\\_7%20875x10\\_0.pdf](http://www.rdmag.com/sites/rdmag.com/files/gff-2014-5_7%20875x10_0.pdf).

<sup>4</sup> U.S. leadership in biomedical R&D is challenged by declining investment by the U.S. government and increased spending by some other countries. For more information, see <http://www.unitedformedicalresearch.com/wp-content/uploads/2012/07/Leadership-in-Dilemma-Assessing-US-International-Competitiveness-in-Biomedical-Research.pdf>.

the National Science Foundation, the Department of Defense, and the U.S. Department of Agriculture. Public investment plays a critical role in the biomedical research enterprise. Because the private sector has little incentive to conduct high-risk or early stage research that is decades away from becoming a profitable product, funding from NIH and other government agencies is essential for generating a steady stream of fundamental research discoveries. The government is uniquely positioned to invest in long-range research goals such as basic research, research that is not likely to lead to large profits (e.g., vaccine or antibiotics development), or research that addresses rare conditions. Indeed, such research is likely to slow significantly or come to a halt without government support.

**Figure 2. U.S. Health research investment by sector (\$ in billions)<sup>5</sup>**



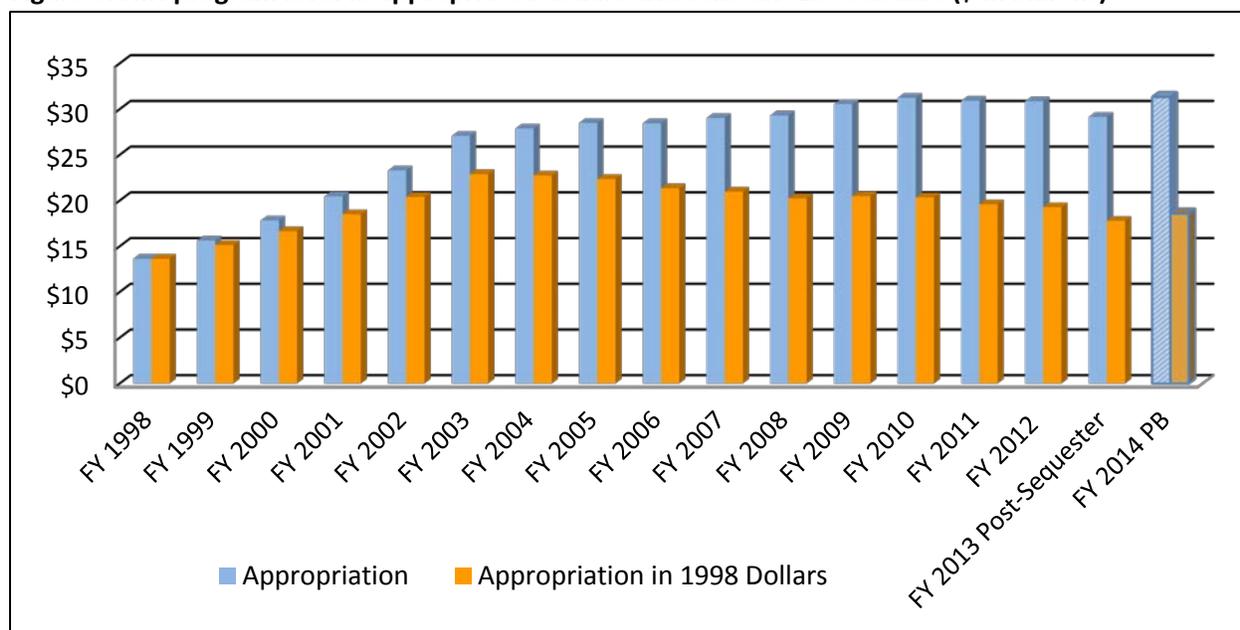
**B. NIH’s role in the biomedical research enterprise**

With its origins as a single-room “laboratory of hygiene” created in 1887 within the Marine Hospital Service,<sup>6</sup> NIH today is the largest public funder of biomedical research in the world. A Federal agency with 27 institutes and centers, NIH’s \$30.86 billion budget in fiscal year 2012 supported approximately 50,000 competitive grants to more than 300,000 researchers at more than 2,500 universities, medical schools, and other research institutions in every state and around the world.. NIH also supports intramural research projects, including approximately 1,200 principal investigators and 6,000 trainees ranging from high school students to postdoctoral and clinical fellows. Figure 3 portrays U.S. investment in NIH over 15 years.

<sup>5</sup> For more information, see [http://www.researchamerica.org/research\\_investment](http://www.researchamerica.org/research_investment).

<sup>6</sup> For more information, see <http://www.nih.gov/about/history.htm>.

**Figure 3. NIH program level in appropriated dollars and constant 1998 dollars (\$ in billions)**



NIH’s primary role is to support biomedical research through funding research projects, developing research infrastructure (e.g., facilities, databases, research equipment), and training the scientific workforce. NIH also conducts biomedical research within its own intramural research program. NIH’s Clinical Center is a unique resource for advancing clinical research and in particular for patients with diseases of unknown etiologies. Furthermore, NIH is responsible for communicating information about the research process and its results to both the research community and the broader public.

NIH’s role in nurturing the biomedical and health research enterprise goes much further than direct support of the research itself; NIH’s investment in training, infrastructure, and resources provide an essential platform for research and enduring value to the biomedical research system. Training the research workforce and future leaders in healthcare occurs at every career stage through NIH’s funding of numerous training and fellowship programs. NIH also conducts workshops, bringing together top scientists to address pressing issues in many areas of biomedical research, and establishes cooperative agreements and cross-institute initiatives to accomplish research goals. NIH’s role in advancing global health and training of future leaders of health policy, particularly in low and middle-income countries is often underappreciated.

Furthermore, NIH support for developing and maintaining databases, information systems, analytic tools, and other resources ensures that all researchers can access and analyze data to replicate or expand scientific studies. In addition, through its dissemination of resources such as the National Library of Medicine’s PubMed, NIH provides inexpensive and often free access to the scientific literature in formats accessible to scientists and the public.

Beyond the role of supporting and advancing science through various mechanisms, NIH provides resources for patients who need reliable information on a particular disease or disorder. NIH also

communicates breakthrough research findings that may affect health and helps increase public knowledge of health-related research.

NIH plays an important role in the biomedical research enterprise due to its size, its infrastructure, and its ability to convene key players from academia, industry, regulatory bodies, and around the world to address challenges and advance research. NIH can quickly leverage its resources and position to address public health emergencies, such as the HIV/AIDS crisis, SARS, or the anthrax scare in 2011.

### **C. The NIH mission**

NIH funds research with the goal of improving human health. Its mission is to:

“...seek fundamental knowledge about the nature and behavior of living systems and the application of that knowledge to enhance health, lengthen life, and reduce illness and disability.”

Achieving this mission first depends upon NIH’s ability to catalyze the generation of fundamental scientific knowledge. To do this, NIH-supported scientists collect data on the function and interactions of biological and behavioral systems and when these systems go awry.

Achieving NIH’s mission also requires that fundamental knowledge is applied to actionable interventions and then implemented to enhance health, lengthen life, and reduce illness and disability. Despite the dependence of its mission on implementing the evidence base, it is not in NIH’s purview to provide, monitor, or regulate health services and products. Instead, NIH must ensure that its evidence base is disseminated to the public and to those whose mission is to develop, provide, and regulate health services, products, and policy. To accomplish this, in addition to its own dissemination activities, NIH must partner with many other actors in the health ecosystem—especially that of NIH’s sister agencies within the Department of Health and Human Services.

### **D. Capturing the value of NIH’s achievements**

From the Latin *valēre*, which means “to be of worth” or “to be strong”, the value of NIH is rooted in its mission and role. When performed successfully, NIH’s activities lead to new knowledge and improvements in health, which generate additional societal gains. Therefore, for the purpose of this report, NIH’s value to society can be categorized as 1) the generation of scientific knowledge, 2) the impact of scientific knowledge on the health of the public, and 3) the broader effects of NIH-funded activities on other aspects of society.

***The value of fundamental knowledge.*** Scientific knowledge is accumulated and evaluated through a cycle of discovery, validation by peers, and eventual consensus. The impact and value of any one finding is unpredictable and will change over time as the scientific enterprise evolves. Rather than simply valuing knowledge for knowledge’s sake, much of the value of knowledge is realized when reproducible research findings are made accessible to others and can support or inform action within the biomedical research enterprise and the larger health ecosystem.

**The value of health effects.** NIH's value to the public's health is derived from its ability to generate knowledge that has relevance and applicability to health challenges. This can also be viewed as a staged process: initial application to human health, implementation and uptake into medical and community health settings, and eventual downstream effects on individual and population health status. The process of translating research to practice is far from regular or linear, and the health impact of foundational scientific discoveries often takes years and even decades to be realized.

**The value of broader societal impacts.** These broader mission-related achievements can include economic benefits, international competitiveness, technological advancements, increased scientific literacy, international collaboration, and many other effects. Furthermore, when knowledge is applied to the improvement of the health of the public, additional spillover effects can be generated: a more health-literate population chooses healthier behaviors and informed care-seeking;<sup>7</sup> healthier populations lead to averted medical costs and increased productivity (ultimately increasing GDP);<sup>8, 9</sup> advances in the pharmaceutical and biotechnology industries create more effective drugs, research tools, or medical devices; and international health efforts can serve as cornerstones of diplomacy.<sup>10, 11</sup>

## E. The purpose of assessing NIH's value

The agency has long engaged in activities to assess the value of NIH-supported research. Three general categories summarize the compelling reasons for these assessment activities: accountability, management, and communication.

**Accountability to the public.** As a Federal agency, NIH is accountable to the public and its elected representatives. It is incumbent on the agency to demonstrate effective stewardship of the resources with which the public has entrusted it. Therefore, to address accountability, NIH must engage in rigorous and systematic efforts to provide an accounting of its activities and the outcomes of those activities. Assessments done for the purpose of demonstrating accountability must be designed with an eye to the audiences most in need of the information, including the Department of Health and Human Services, the Administration, and Congress. Often these assessments are done in response to Government-wide efforts to standardize accountability measures. For example, since the 1990s NIH has responded to the requirements of the Government Performance and Results Act (GPRA) by identifying specific scientific goals and providing yearly reporting on progress towards those goals.

**Management of NIH's portfolio and activities.** Constant improvement in management of NIH's research portfolio and its activities and processes is achieved through the continual assessment of program performance, including the measurement of outputs (i.e., products of research). A better understanding

---

<sup>7</sup> Berkman ND, et al. (2004). *Literacy and health outcomes* (AHRQ Publication No. 04-E007-2). Rockville, MD: Agency for Healthcare Research and Quality.

<sup>8</sup> Murphy, Kevin M. and Robert Topel. (1999). *The Economic Value of Medical Research*. University of Chicago.

<sup>9</sup> Murphy, Kevin M. and Robert Topel. (2005). *The Value of Health and Longevity*. University of Chicago, National Bureau of Economic Research, and the George J. Stigler Center for the Study of the Economy and the State.

<sup>10</sup> Daulaire N. (2012). *The Importance of the Global Health Strategy*. American Journal of Tropical Medicine & Hygiene; Vol. 87.

<sup>11</sup> Glass, Roger. (2013) *What the United States has to gain from global health research*. JAMA 310(9); 903-4.

of all aspects of NIH's work leads to increased efficiency and effectiveness of that work. In addition, these assessments may provide insights regarding barriers to achieving value, such as research practices that result in non-reproducible findings and the slow adoption of some validated research advances by clinicians and patients.

NIH has created and continually refines a vast data infrastructure to manage its research funding. Analyses of outcomes based upon this data are used to inform NIH senior leadership and aid in their decision-making. Efforts to improve upon these assessments are in keeping with the Obama Administration's mandate to carry out an aggressive management agenda, an important component of which is to improve the generation of evidence about what works and what does not work in managing the Federal enterprise. In addition, in a climate of flat or declining budgets, there is a recognized need to identify priority funding areas. While NIH faces difficult choices in resource allocation, challenges facing human health are seemingly more complex than ever. By better understanding the results of its activities, NIH leadership can make ever-more informed decisions with an eye toward enhancing the value of its activities to the American public.

**Communicating results of assessments.** NIH must be transparent and responsive to its many partners and stakeholders, including the public, patient and provider communities, and policy makers at all levels of government. NIH should be clear in its role, in its approach, and in its impact so that its partners and stakeholders can fully understand its functions and accomplishments. Better capturing the value of NIH's endeavors and communicating that value to the public and healthcare professionals is essential.

## **F. Challenges to Accurately Assessing Value**

Both the generation of knowledge and the application of that knowledge to health, as well as the impacts of these pursuits on broader society, are vital parts of NIH's value. Given this vast scope of effects, it is a daunting task to assess NIH's progress in each of the three value streams described above: Scientific Knowledge, Health, and Broader Societal Impact. Several of the challenges are described below.

**Differing definitions of value.** Value varies depending on differing viewpoints and perspectives over time. For example, is a finding that profoundly alters the life trajectory of a small number of individuals more valuable than a finding that has a lesser benefit for multitudes of people? Greater societal goals also influence value judgments, such as the value placed on scientific freedom, reducing health disparities, or balancing public health needs with individual rights (e.g., privacy).

**Time considerations.** To assess the value of NIH's investments adequately, one must take the long view. However, a protracted timeframe introduces many variables that are difficult to ascertain and are often beyond the control of NIH. Adding to the importance of the time dimension is that realized benefits accumulate; today's discoveries can affect multiple generations. For example, research on surfactants to treat premature infant lung development not only saves the lives of babies today, but will allow those babies to grow to adulthood, contribute to society, and raise future generations. A discovery that seems to have little significance today may prove transformational in the future, and a single finding may have implications for numerous subsequent lines of inquiry. For example, the Nobel prize-winning discovery

of the polymerase chain reaction (PCR) has had implications, many of them difficult to predict, for innumerable fields of study over the course of decades of research. The full and lasting value of NIH research findings will not be evident in the short term and can multiply over time.

**Complexity of the biomedical research and health enterprises.** Numerous factors contribute to the complex processes of science and health, making accurate attribution an especially difficult obstacle to assessing the value of NIH-supported research. The diversity of research funders complicates efforts to assign credit for particular advances. As biomedical research becomes increasingly global, retracing the steps that led to a breakthrough and assigning attribution becomes even more difficult. Given that one scientist could be supported by multiple funders across the research enterprise, how can one apportion NIH's contribution to that particular scientist's accomplishments? The issues of attribution to NIH are even more complex when considering its impact on health. These include the (i) multifactorial influences that affect human health (including social determinants); (ii) challenges associated with documenting health-related behavior, their relationship to health outcomes, and the influence of research findings on behavior; (iii) lag time between initial discoveries and their impact on human health; and (iv) data limitations in linking research results to changes in practice. For example, NIH has contributed to the noted decline in U.S. deaths due to heart disease. But how much credit may be given to NIH-supported research versus other developments in health? How much of the decline in mortality rates is due to better medical intervention in the form of statins, aspirin, and beta-blockers versus changes in behavior like smoking less and eating a better diet? And for both advances in medical intervention and behavior change, what role did NIH play in bringing about these improvements? Even more challenging is the issue of attribution of broader societal impacts to NIH. How can one begin to determine the contribution of NIH to overall economic gains or to changes in the U.S. population's scientific literacy?

**Lack of agreement on metrics related to biomedical research outputs, outcomes, and impacts.** Given the aforementioned challenges, there is no universally agreed upon list of metrics that one should use in assessing the overarching value of biomedical research activities. Efforts to capture the outputs of biomedical research have looked at the quantity and quality of research patents, journal articles, and Nobel prizes, for example, but the value of such outputs only constitutes a small piece of the overall value of research. A number of studies attempt to measure longer-term research outcomes, such as changes in the cost of medical care, life expectancy, and quality of life. However, there are many assumptions inherent to these measures and little consensus on their appropriate use. For example, what value would one assign to an additional year of life?

### III. FINDINGS OF THE WORKING GROUP

Demonstrating the value of biomedical research in a systematic way is extremely challenging. While markers of progress measure some of what NIH produces, capturing the full impact of NIH's investments is a more complicated venture. Fortunately, there is growing worldwide interest in assessing the outcomes of scientific endeavors and to capitalize on competencies in communication as well as data collection and analysis. Furthermore, the need to capture the outcomes of scientific endeavors more definitively to inform the public and decision-makers increases in times of economic uncertainty.

Recommendation: NIH should capitalize on the ongoing innovations in data collection and analysis by intensifying its efforts to systematically, comprehensively, and strategically assess its value. Results of studies over the next several years can be used to demonstrate accountability, enhance management, and increase public awareness.

Such an effort requires a commitment by NIH leadership to carry out necessary assessments of the value of NIH-supported research and also significant resources to organize, develop, and support thorough, well-designed studies. To ensure the highest quality and validity of these assessments, NIH should couple internal efforts with engaging external experts in the planning and conduct of assessments of the value of NIH research.

Recommendation: NIH should establish dedicated and sustained resources to support the planning and conducting of assessments of the value of NIH research. These resources should be modest in comparison with the expenditure of funds for core NIH research.

#### **A. Considerations for assessing and communicating NIH's value**

##### ***NIH's value is derived from producing knowledge that can be applied to improve the public's health.***

NIH's primary value lies in the knowledge it generates, not just the immediate value of knowledge but also its "option value." Knowledge generated by NIH research creates opportunities for more research advances and for eventual translation into health and societal benefits. Much like stock options, scientific knowledge is often exercised at a later date; both have inherent value before they are applied. In this analogy, NIH-funded research generates options that others in the health ecosystem might choose to exercise. For example, a private company may work to develop new medications based on NIH-funded research, or the Centers for Disease Control and Prevention may act upon NIH research findings in their response to a public health emergency. Scientific knowledge can affect society through myriad pathways, ranging from medical practice to international relations. Tracing these pathways back to the initial point of knowledge generation presents an opportunity for NIH to define the value of that knowledge.

Given the well-appreciated time lag from discovery to application and the multitude of ways knowledge is applied, it is extremely difficult, if not impossible, to predict which scientific discoveries today will prove to be the most valuable in the future. It is also challenging to place an empirical value, especially a dollar amount, on knowledge. An essential role for NIH, therefore, is to create a diverse portfolio of viable knowledge-options both for further scientific discovery and for its implementation. NIH's value is derived from the generation of these options and the many ways in which they are exercised.

Recommendation: Assessments of NIH's value should attempt to trace NIH's contribution to the numerous translational pathways that exist and draw clear connections between the generation of knowledge and its application to health and broader societal impacts.

***Many factors need to be considered in order to accurately determine NIH's contribution to a particular outcome.*** To be unbiased and effective, an assessment must be appropriately rigorous in determining

attribution; however, assigning attribution adds complexity to the assessment process. For example, NIH works with many partners to support research, training, and infrastructure development; these partnerships must be appropriately acknowledged in any assessment. In addition, other actors in the health ecosystem are responsible for taking the knowledge generated by NIH and advancing it to affect health and other societal impacts. These other actors, and the public itself, will influence the range and depth of the value of NIH's contributions.

Unanticipated outcomes also confound assigning attribution in assessments of value. For example, a single discovery may have multiple applications, and a single application may be based upon multiple discoveries. In addition, the significant lag time that often occurs between a discovery and its application further complicates a timely and accurate assessment of the value of NIH research.

Recommendation: Credible, interpretable, and useful assessments of the value of NIH should, when possible:

- a. Acknowledge and determine the role of other players in the advancement and adoption of research findings or other outcomes of interest.
- b. Attribute outcomes entirely to NIH only when this is proven to be the case.
- c. Establish a timeframe that is broad enough to include sufficient time for discovery to be applied.

***NIH affects and is affected by many participants in the scientific and health ecosystems; therefore, many stakeholders must be considered in any attempt to assess and communicate its value.*** Thorough and accurate assessments of NIH's value require concerted and continual effort, with input from numerous participants in the scientific and health ecosystems. Many of the diverse audiences for assessments of NIH's value overlap with the participants in the process, including NIH leadership and program staff who design and conduct the value assessments as well as researchers, research institutions, patients, practitioners, the general public, policymakers, other Federal agencies, and the private and nonprofit sectors.

Recommendation: NIH's assessment strategies should reflect all stakeholder voices, and NIH should regularly seek input from stakeholders to inform its assessment activities and communicate their results.

Often, other entities in the health ecosystem, such as practitioners, patients, or other Federal agencies, develop the knowledge generated by NIH research into its end use or product. Because NIH is often not directly involved in this aspect of the research enterprise, it will be necessary to engage these hand-off sectors in the collection, validation, and linkage of data about the scientific, health, and broader societal outcomes that are produced by biomedical research.

Recommendation: NIH should seek ways to partner with other entities in government, non-governmental organizations, and the private sector in its attempt to develop tools, techniques, and databases to strengthen assessments of value.

It is especially important for NIH to take into account the effect assessments may have on researchers and research institutions when designing future assessment studies. For example, many metrics focus on the number of publications produced because data on this activity is readily available, but this focus can disincentivize the production of other valuable research outputs. Data collection and reporting for assessment purposes beyond the research question can be time consuming and require specific knowledge or understanding. The time and effort devoted to proper assessment data collection and reporting is often at the deficit of research efforts.

Recommendation: NIH should give careful consideration to the consequences of its assessment efforts, ensuring that assessment measures do not negatively influence the conduct of research.

## **B. Components of scientifically rigorous assessment approaches**

***Numerous attempts to assess aspects of NIH's value have been undertaken by NIH and by many of its stakeholders, but these efforts have not been comprehensive, systematic, or coordinated.*** NIH and its stakeholders engage in numerous activities to assess and communicate the outcomes of NIH activities, ranging from disseminating stories of discoveries to formal program evaluations. Each of the NIH Institutes, as well as the NIH Office of the Director, conducts studies to measure the outcomes of specific NIH programs and activities. Several external studies, samples of which are found in Appendix B, have also tried to assess outcomes of NIH activities. However, these studies have varying degrees of rigor and scope and, by necessity, many formed conclusions based on a multitude of assumptions.

The complexities of the scientific process, the vast array of research areas addressed, and the array of strategies that NIH employs in pursuit of its mission make it nearly impossible to assess the overall value of NIH activities. Yet, with planning and cross-agency coordination, assessments could capture a better approximation of the value of the enterprise as a whole. A comprehensive strategy to address each aspect of study design, including deciding upon the specific study question, determining the data needed, designing an appropriate methodology for the collection and analysis of the data, and effectively communicating the results of the study, must be created through a concerted and sustained effort across the agency in partnership with stakeholders.

### **1. Study topics**

A comprehensive, systematic, and coordinated approach towards assessing the value of NIH begins with identifying a representative sample of NIH activities to study that will reflect the whole of NIH and the many avenues by which NIH influences any number of societal factors. Assessments of value should attempt to measure the impact of all NIH-supported activities, including supporting and conducting research, workshops, conferences, training, public-private partnerships, infrastructure, database development and maintenance, and communication of health research and information in numerous

In addition to the range of activities NIH undertakes to carry out its mission, the scientific knowledge it generates translates into varied outcomes via many pathways. The studies included in Appendix B attempt to measure outcomes along several of these translational pathways, including NIH's contributions to changes in health measures, its impact on pharmaceutical innovation, or its influence on the U.S. economy. A more systematic approach to capturing some of these outcomes is warranted.

Recommendation: NIH should support internal and external assessment studies that are strategically and systematically selected to represent the full spectrum of NIH activities and processes, as well as the variety of translational pathways by which they affect individual health, the health of the public, and broader societal impacts.

### Study Topic Selection

During the course of their deliberations, the Working Group discussed a wide range of study topic suggestions. (Note that the suggestions are not listed in priority order.)

- Examining 25 diseases over several generations to examine the effects on and the outlook for patients
- Case studies on research areas without economic incentives
- Study of an alleged "failure" that still resulted in positive economic impact (e.g., unintended spin-off, industry avoided investing)
- Mapping NIH's investment and value in major biomedical breakthroughs (e.g., top ten scientific breakthroughs as chosen by Science magazine, Wall Street Journal list)
- Contributions of NIH to research and/or researchers of note (e.g., Nobel Prize winners)
- Specific health and disease topic areas such as antibiotics, vaccines, polio, HIV/AIDS, cardiovascular disease, Alzheimer's disease, diabetes, premature births, behavioral therapies, etc.
- Prevention strategies and campaigns (e.g., smoking, obesity)
- The impact of ARRA funding
- The effect of doubling the NIH budget
- Studies that address critiques of NIH and the larger biomedical research system

NIH must also consider several other factors when selecting the most appropriate study topics for value assessment. It will be important to consider the attributes, elements, and components necessary for a rigorous assessment study. For example, the study must be feasible, with defined indicators and quality data sources in place. The purpose of the assessment and the audience should also be considered as the study questions are formulated, and the most appropriate study design may vary according to the different goals or purposes for gathering the information. How value is assessed may also vary for different modes of research (e.g., basic, translational, clinical).

Study topics should include both research successes and research "failures," since important lessons may be learned from each. "Failure" can have multiple meanings in the context of biomedical research.

For example, research that is not well-designed or executed could be considered a true failure and a waste of resources. However, while the failure to show safety or efficacy in a clinical trial is undoubtedly disappointing, the study may well have generated critical knowledge and provided re-direction for research and development. In this case, the research yielded a valuable, if unexpected, result. Additionally, well-planned, well-performed research studies with null or negative results are not “failures,” and it is important to communicate to the public that the essence of science is the discovery of what will and will not work.

The selection of representative study topics in a systematic way will require creation of a framework that captures a concept of the whole of NIH in terms of its activities and the outcomes it aims to achieve. The development of such a framework will require a concerted, agency-wide effort and should be informed by outside experts.

Recommendation: An internal trans-NIH Committee on Assessments should be established to coordinate agency efforts in assessing its value, including devising a systematic approach for the selection of study topics. External experts should be consulted in the selection of study topics.

## 2. Data needs

***There is insufficient data collection, storage, and linkage between data sets to conduct thorough assessments of value.*** Sound data, along with a rigorous and interconnected data infrastructure, is a prerequisite for valid assessments of value. This requires NIH to identify relevant and reliable sources of data, collect and store data in a routine and highly organized fashion, link multiple data sources together, and enable access to such data in formats amenable to analysis.

Recommendation: The trans-NIH Committee on Assessments should coordinate efforts and engage nationally recognized experts to improve NIH’s data infrastructure for monitoring and assessing itself.

Recent NIH efforts in this arena are encouraging and have led to the development of data systems such as RePORTER, the Scientific Publication Information Retrieval and Evaluation System (SPIRES), the Electronic Scientific Portfolio Assistant (eSPA), and STAR METRICS.<sup>12</sup> These systems have greatly improved NIH’s ability to identify, analyze, and report shorter-term scientific outputs of its research spending, in particular publications, citations, patents, job creation, and economic activity in the private sector. However, NIH’s data infrastructure was built primarily to manage grants and contracts during their life-cycle, not to track outcomes. As such, NIH’s ability to examine medium- and long-term outputs, especially those related to health and broader societal impacts, is limited.

A number of promising data types and sources can be used to track research outputs. While not exhaustive, Table 1 lists dozens of research outputs mapped to the three value streams of scientific knowledge, health, and broader societal impacts that could be measured in short-, medium-, and long-

<sup>12</sup> Callahan, Kevin, et al. (2012). *Assessing the Outcomes of NIH Research Spending: Toward a Better Approach*. Bethesda, Maryland: National Institutes of Health.

term timeframes. Ultimately, these outputs are a means to an end, leading to the outcomes or end goals of biomedical research. Appendix C presents a chart that matches these outputs to their potential sources. In this chart, specific data gaps, limitations, and weaknesses that could hinder assessment are noted in red text. Some of these outputs are much easier to measure than others, and the tendency to measure what is easy has undoubtedly influenced the current use of certain outputs as indicators of progress. Although an active area of study, the identification and community acceptance of reliable and appropriate indicators of NIH's outcomes remains a challenge. Views vary as to which indicators are both feasible to collect and best reflect the products and impacts of research spending. Moreover, even if perfect data existed for each of the indicators listed in Appendix C, comprehensively measuring and analyzing all of these indicators would be impractical.

NIH should prioritize a set of indicators that can better demonstrate the value of NIH and the diverse outcomes of NIH spending. Given NIH's stature in biomedical research, the agency could be a leader in establishing and refining indicators that could provide a richer account of the science-based outputs that affect society.

Recommendation: NIH should identify and gain consensus on a core set of indicators to be included in its data infrastructure.

Table 1. Biomedical research outputs and outcomes

<b>SCIENCE</b>	
<b>Outputs</b>	
<b>Discovery</b>	Primary research articles • Conference abstracts/ presentations • Research databases and repositories (i.e., NGVB, NDAR) • Development of research resources and infrastructure • Non-peer reviewed research findings and reports (e.g., self-publishing, blogging, data sharing)
<b>Validation</b>	Patents • Licenses • IP transfer agreements (MTAs) • Development and initial clinical testing of dx, tx, and px interventions (e.g., FDA applications and Phase I, II, and III clinical trials) • Standardized research protocols • Validated data repositories (i.e., NCBI's GenBank, PDB)
<b>Consensus</b>	FDA approvals of New Drug Applications (NDAs) • Consensus Development Conferences and Bodies (i.e., National Toxicology Program) • Book chapters • Post-grad curriculum guidelines • Systematic reviews/ meta-analyses • Knowledge transfer metrics (i.e., citations, web hits)
<b>Outcomes/Goals</b>	
Fundamental scientific knowledge	
<b>HEALTH</b>	
<b>Outputs</b>	
<b>Application</b>	Advanced testing of evidence-based diagnostics • Advanced testing of evidence-based therapeutic interventions • Advanced testing of evidence-based preventive interventions and strategies • Identification of major health issues and disease risk factors
<b>Implementation</b>	Drugs to market • Dissemination and Implementation efforts, Practice guidelines, etc. • Adoption of evidence-based dx, tx, and px strategies • Public health campaigns
<b>Health measures</b>	Burden of Disease • Quality of Life • Disability • Longevity
<b>Outcomes/Goals</b>	
Living longer, healthier lives	
<b>BROADER SOCIETAL IMPACTS</b>	
<b>Outputs</b>	
<b>Proximal</b>	Support government, science, and technology jobs • Create demand for R&D supplies • International collaboration • Support academia • De-risk pre-competitive space • Cross-sector collaboration
<b>Intermediate</b>	Spur private sector activity (i.e., Biotech, Pharma) • Enhance STEM education • Communication and interpretation of findings across sectors and to public • International S&T capacity building • Spur local economy
<b>Distal</b>	Uptake and spread of technological innovations • Workforce output (longevity, health) • Workforce development • Internationally competitive science and technology sectors • GDP • Emergence of new sectors and industries • Health care costs
<b>Outcomes/Goals</b>	
Scientifically literate public • Healthcare-related cost savings • ↑ Productivity • ↑ Capacity for Innovation • ↑ Global R&D competitiveness • Diplomacy & stability through science	

Establishing core indicators presents both a hazard and an opportunity – the indicators that NIH measures are likely to influence the behavior of scientists and their research outputs. Progress indicators need to have strong credibility and acceptance across the biomedical research and practice communities, which will require a clear articulation of their limitations, predictive value, and how the agency will use such measures. The agency must take care that core indicators do not present a high risk of unproductive behavior change in the research community. On the other hand, NIH could introduce novel positive incentives related to a broad set of activities (e.g., training, mentorship, data sharing, collaboration, software development, community outreach, science communication).

Funding bodies and research organizations are enhancing and opening access to data as a means to better analyze scientific outcomes. NIH can partner with the many different science agencies that are also seeking better ways to measure their impact. Internationally, the Medical Research Council (MRC) in the United Kingdom and the Canadian Institutes of Health Research (CIHR), to cite two examples, have recently implemented advanced performance data collection and analysis systems to capture a wide range of possible research project outcomes. Universities are also increasingly engaged in finding ways to demonstrate the educational, scientific, commercial, and economic value of their research activities, especially at the regional and state level. In addition, many of NIH's partners in the health enterprise, both public and private, are also looking to better capture health outcomes. An excellent example is CDC's Data Linkage Program which links

***Opportunities arising from NIH's new administrative data collection efforts***

Three recent efforts led by the NIH's Office of Extramural Research are aimed at enriching NIH's administrative data on funded research projects and researchers. The **Research Performance Progress Report (RPPR)**, which was mandated by the Office of Management and Budget, will help address many existing data weaknesses. First, the RPPR will collect some performance data in a structured format. Second, all U.S. government agencies that fund research will use the RPPR, so the RPPR essentially establishes a set of common, standardized data elements. Both of these features of the RPPR represent marked advances in project performance reporting that may greatly improve the ability of each research agency to link and integrate performance data. Moreover, there are emerging opportunities to integrate performance-measurement systems like the RPPR with other reporting and monitoring systems, allowing NIH to leverage data infrastructure improvements to reduce reporting burden from the research community while also enhancing the utility of outcomes data for the researchers themselves. For example, RPPR will link to another newly established NIH administrative data tool, the **Science Experts Network or SciENCv**, which enables researchers to easily maintain and generate biosketches for federal grant applications and progress reports. SciENCv also provides a mechanism (via the third-party organization, ORCID) for users to generate and associate a unique international ID to their profiles. Third, NIH is engaging in a large-scale data-sharing effort with several other major biomedical research funders internationally to build the **World RePORT** database. The system integrates detailed data on funded projects into an illustrative mapping system designed to enhance communication and coordination among research funders. Now in a beta version, World RePORT displays information on projects being funded in Sub-Saharan Africa by nearly a dozen public and private funders, including the Canadian Institutes of Health Research, the UK's Medical Research Council and Wellcome Trust, and the European Commission.

together administrative data, such as the National Death Index and claims from the Centers for Medicare & Medicaid Services, with data from numerous population health surveys to uncover factors that influence health and health outcomes.

Recommendation: NIH should link its own data infrastructure with that of its many partners in the science and health ecosystems who are already tracking many outcomes of interest to NIH (e.g., CDC, U.S. Patent and Trade Organization, Food and Drug Administration).

Once a core set of outcome and impact data has been established, the agency can focus on acquiring the data, analytic tools, and skill for measuring this set of indicators rigorously. The advent of “Big Data” is providing rapidly emerging capabilities only imagined a decade or two ago. There are unprecedented opportunities to improve the collection, quality, and linkages of vast amounts of data that will enable better assessments of NIH’s value and diverse outcomes. New and emerging capabilities will make the data infrastructure more useful to those assessing biomedical research.

Recommendation: The trans-NIH Committee on Assessments should examine ways to capitalize on NIH’s many efforts in “big data” by conferring with newly established NIH Big Data committees.

There are many opportunities to improve the quality of data and the way data are collected, linked, and made available for analysis. These include:

- *Filling gaps.* Existing NIH databases could be enhanced to provide more data on relevant inputs, outputs, and outcomes. For example, many of these systems were not designed to track individuals (e.g., investigators, trainees, research project staff), making it difficult for the agency to track job placement of NIH trainees and NIH-funded researchers.
- *Linking data systems.* Linking data between existing NIH systems is currently difficult because no single data element or data field (such as project ID or investigator ID) exists within all relevant databases. For example, clinicaltrials.gov does not currently include project or investigator IDs. This is even more problematic when attempting to link between NIH and other government databases, whether they also support science (e.g., National Science Foundation, Department of Defense) or are implementing the NIH evidence-base (e.g., FDA, USPTO, and Agency for Healthcare Research and Quality). Even more complicated is linking to non-government data sources, such as those curated by research institutions and the private sector. This greatly hinders the ability to draw connections between short-term NIH outputs (e.g., publications, clinical trials, patents) and longer-term outputs (e.g., changes in health practice, industry start-ups, and FDA approvals).
- *Improving quality:* Many factors may drive inaccurate, inconsistent, and incomplete data, including the use of multiple systems for data entry, lack of data entry and reporting enforcement or incentives, and lack of validation.
- *Improving utility through data standardization.* The utility of data is often hampered by a lack of consensus on terminology and nomenclatures. Data elements may look similar across databases but be defined somewhat differently. Addressing standardization issues will require building

consensus on definitions or producing data thesauri so that data elements are comparable across databases.

- *Collecting and storing data in structured formats.* Usability of data can be greatly enhanced by collecting and storing data in more structured formats. For example, while recent progress has been made in identifying patent applications and licenses that are a direct off-shoot of research performed by NIH grantees, patent applications also contain citations to prior articles. However, these citations are not collected or stored in a structured format. Addressing this limitation could allow NIH a new avenue in understanding the role of biomedical research in establishing the evidence base for intellectual property.
- *Opening access.* Some databases are subject to limited access. By opening access to data, multiple user communities can also test the data, allowing for continual improvements in data quality.

Recommendation: NIH should improve the quality and integrity of core data and linkages through standardized data collection and established governance.

There is no single solution to strengthening data. Successfully addressing data considerations will require continual and highly coordinated effort and resources, but the payoff is worthwhile. For example, a major element of the attribution challenge includes identifying reliable links between shorter-term NIH-supported outputs and the myriad downstream effects that may also be part of the value equation. Well-designed, linked data systems can substantially ease the attribution challenge. In short, better data will promote better assessments and improve NIH's ability to optimize value.

### **3. Methodologies**

***A suite of rigorous and feasible methodologies are needed to better capture the value of NIH. Of the many assessments that have been undertaken, no single approach has proven entirely satisfactory.***

Methods and strategies for measuring the value and impact of biomedical research are quickly evolving, but no single assessment strategy can capture the full value of biomedical research. Thus, a multi-pronged approach to value assessment will be most effective.

A sample of recent assessment studies (Appendix B) include a range of methodologies, both quantitative (e.g., bibliometrics, economic analyses) and qualitative (e.g., case studies). In addition, several reports in the last few years have attempted to survey assessment methodologies, including the National Research Council's 2011 workshop on measuring the impacts of Federal investments in research<sup>13</sup> and a 2012 National Academy of Science report entitled *Best Practices in Assessment of Research and Development Organizations*.<sup>14</sup> Notable reports from international government funders and advisory councils also addressed methodologies in use or development by other countries to measure research impact,

<sup>13</sup> National Academy of Sciences. (2011). *Measuring the Impacts of Federal Investments in Research - A Workshop Summary*. Washington, D.C.: National Academies Press.

<sup>14</sup> National Academy of Sciences. (2012). *Best Practices in Assessment of Research and Development Organizations – Summary of a Workshop*. Washington, D.C.: National Academies Press.

including the UK, Australia, and Canada.<sup>15,16,17</sup> The study commissioned by the Association of American Medical Colleges, conducted by RAND Europe, presented a useful summary of methodologies and tools used by large international science funders to assess research outcomes.<sup>18</sup> According to these reports, some common categories of assessment approaches include bibliometrics, web-based metrics (e.g., altmetrics), case studies, economic analyses, peer and expert review/opinion, algorithmic clustering and text classification, and network analysis.

Many of these reports also stressed the need to consider the purpose, audience, and context when choosing the most appropriate assessment approaches. The method selected for assessment, whether prospective or retrospective, qualitative or quantitative, must produce results that can be effectively communicated to all audiences.

Narratives constructed from well-designed case studies can be especially effective illustrations of the broad impacts of biomedical research. Carefully constructed retrospective case studies can demonstrate stewardship and capture value in a range of outcomes, including the impact of a basic finding on public health, or the economic return on investment for a specific research endeavor. However, case studies must by design be narrow in focus and therefore have limitations in scope and generalizability. To use case studies most effectively to capture value, the topics for these studies should be strategically chosen to best represent the full range of NIH activities and scientific areas. Information and stories for case studies can come from a variety of sources, including the publications and communications documents of research institutions in all sectors and from around the globe.

**Recommendation:** NIH should adopt a systematic approach to designing case studies that can tell compelling and accurate stories of NIH's role in turning discovery into health.

Data and analysis can strengthen both case studies and more personal patient narratives by supporting the stories with details and context, thereby presenting a more comprehensive picture not only of the value of the research but of the obstacles to determining and attributing that value.

**Recommendation:** Analytic approaches should balance numbers with narratives, illustrating the complexities of progress, such as the time-dependence of R&D and the pivotal roles of other actors in the biomedical sciences.

Important methodological considerations include:

---

<sup>15</sup> Panel on Return on Investment in Health Research. (2009). *Making an Impact: A Preferred Framework and Indicators to Measure Returns on Investment in Health Research*. Ottawa, ON, Canada: Canadian Academy of Health Sciences.

<sup>16</sup> Rymer, Les, et al. (2011). *Measuring the Impact of Research: The Context for Metric Development, Go8 Backgrounder 23*. Turner, Australia: The Group of Eight.

<sup>17</sup> Research Excellence Framework. (2014). *Decisions on Assessing Research Impact*. Bristol, UK: Higher Education Funding Council for England.

<sup>18</sup> Guthrie, Susan, et al. (2013). *Measuring research: A guide to research evaluation frameworks and tools*. Santa Monica, CA: RAND Corporation. <http://www.rand.org/pubs/monographs/MG1217>.

- *Use complementary methods.* The combination of multiple assessment strategies and indicators can create a composite perspective on value and impact. Include quantitative and qualitative dimensions as well as both prospective and retrospective methods to assess value.
- *Choose appropriate timeframes.* Timeframes for assessment studies should allow for an appropriate period of time for the scientific process to unfold and for implementation to occur. Insufficient timeframes could seriously bias results.
- *Choose the appropriate unit of analysis.* Value and impact can be assessed at many levels: individuals, research groups, institutions, fields, and regions or countries. Many indicators are appropriate or reliable only at specific levels of analysis. For example, bibliometric analysis is generally more applicable at higher levels of aggregation, such as at the university level rather than at the level of evaluating individual scientists.
- *Identify a baseline or comparison group.* At best, assessments should establish a counterfactual (theoretical exercise to show what would happen without a particular scientific finding). However, if a counterfactual is not possible, a similar comparison group and/or baseline should be established.
- *Make use of expert interpretation and judgment.* Analyses should inform rather than replace expert judgment. Meaningful and appropriate interpretations require synthesis and judgment by experts with detailed contextual knowledge.
- *Avoid over-attribution.* Clear warning labels describing data sources, methodologies, and thoughtful discussions of study limitations should always be presented.

Making the most of these tools and methods—and ensuring that they are used appropriately—will require not only the expertise of NIH program managers and analysts but also the expertise of economists, computer programmers, demographers, epidemiologists, mathematicians, statisticians, and trained evaluators. NIH should support and encourage the development of new and improved tools, both internally and from the extramural community, by expanding collaborations with academic and industry experts in research assessment and funding researchers to develop new analytic tools.

Recommendation: The recommended trans-NIH Committee on Assessments should develop an assessment approach guide to help NIH ICs select appropriate and feasible methodological approaches based on a number of nested factors beginning with the kind and purpose of the evaluation. This guide should be updated at yearly intervals to reflect new tools or approaches. The Committee should engage external experts in the development of assessment approaches.

#### **4. Communication**

***Effectively communicating the results of assessments of the value of NIH gives the agency the opportunity to increase public understanding of the scientific research process, NIH's mission, and the ways in which biomedical research affects daily life.*** The many audiences for the assessment of NIH's value include the general public, patient communities, domestic and international policy makers, the research community, and practitioners. Each of these groups may have different needs, value different

aspects of NIH's activities, and respond to different types of assessments. Conveying information effectively to each of these audiences requires a variety of communication strategies, and all audiences could benefit from communication efforts that present the results of assessment studies against a backdrop of information about science research and the role that NIH plays in the science and health systems.

Recommendation: NIH assessment efforts should begin with identifying the purpose of the study and its audiences. Any assessment study design needs to include a plan for communication of results that includes information about the scientific research process to increase awareness and understanding of NIH's role.

Both the general public and patient communities are the ultimate recipients of NIH efforts to improve human health. These groups tend to relate to assessments communicated with narratives describing advances, treatments, and improvements that enhance quality of life. Policy makers are answerable to the public and patient communities, making them primarily concerned with NIH's contribution to improved health and good stewardship of the public's investment. They also respond to narratives, but may seek more quantitative benchmarks to aid in decision-making. Research and practitioner communities tend to focus more on rigorous and evidence-based assessments. Results of NIH assessments may have important implications for the research community by influencing their own research directions. For practitioners, results of NIH assessments may impact their adoption of evidence-based practices and changes to clinical practice guidelines.

Recommendation: NIH's strategies for communicating the results of assessments of its value should be diverse to reflect its many audiences.

- a. Stories are a particularly compelling communication tool. NIH should conduct studies that capture its impact in terms of human experiences (e.g., the impact of research for the patient), which are relatable to many audiences.
- b. Attempts to quantify NIH's impact are particularly useful for certain audiences, and these studies should be communicated with a mix of detailed analysis and contextual stories or information as well as transparency regarding the inherent limitations of quantification efforts.
- c. When using specific examples, note that they do not represent a general proof of the value of NIH-supported biomedical research.

## V. SUMMARY OF WORKING GROUP RECOMMENDATIONS

Working Group members are encouraged by ongoing assessment activities and those currently under development within and outside NIH. However, NIH leadership and the many individuals throughout NIH who are responsible for accountability, management, and communication of research will benefit from the development of a more credible, systematic, and comprehensive assessment strategy. Table 2 summarizes the Working Group's findings regarding approaches

to assess the value of biomedical research supported by NIH and recommendations for strengthening those assessments.

Table 2. Summary of Working Group findings and conclusions

	FINDINGS	CONCLUSIONS AND RECOMMENDATIONS
<b>Assessing the Value of NIH</b>	It is extremely difficult to demonstrate the value of biomedical research as a whole and even harder to ascertain NIH’s specific contribution.	1. NIH should intensify its efforts to systematically, comprehensively, and strategically assess the value of biomedical research for the purposes of accountability, effective management, and public awareness. This will require a sustained investment in strengthening NIH’s data infrastructure and a dedicated funding stream or mechanism to support assessment projects.
	NIH’s value is derived from producing knowledge that can be applied to improve the public’s health.	2. Assessments of NIH’s value should draw clear connections between the generation of basic and clinical knowledge and the impact of this knowledge along differing translational pathways.
<b>Considerations for assessing and communicating NIH’s value</b>	Many factors need to be considered in order to determine accurately NIH’s contribution to a particular outcome.	3. Credible, interpretable, and useful assessments of the value of NIH should be clear in attributing outcomes to all contributors and adopt a timeframe that is broad enough to include sufficient time for discovery to be applied.
	NIH affects and is affected by many participants in the scientific and public health ecosystems; therefore, many stakeholders must be considered in any attempt to assess and communicate its value.	4. NIH’s assessments should be done in partnership with its many stakeholders.
<b>Components of scientifically rigorous assessment approaches</b>	Numerous attempts to assess aspects of NIH’s value have been undertaken by NIH and by many of its stakeholders, but these efforts have not been comprehensive, systematic, or coordinated.	5. NIH should establish a trans-NIH Committee on Assessments that will: <ul style="list-style-type: none"> <li>a. Develop a strategy to support or conduct assessments of value, including through grants or contracts with external experts;</li> <li>b. Determine a process for strategically selecting study topics that map to a conceptual framework, including different translational pathways;</li> <li>c. Oversee (in conjunction with NIH’s recently established “Big Data” committees) NIH efforts to strengthen data needed for assessing value, including:                             <ul style="list-style-type: none"> <li>i. Identifying and gaining consensus on a core set of indicators to be included in its data infrastructure;</li> <li>ii. Creating better data linkages with NIH’s partners and hand-off sectors;</li> </ul> </li> <li>d. Identify promising analytical approaches and</li> </ul>
	There is insufficient data collection, storage, and linkage between data sets to conduct thorough assessments of value.	
	A suite of rigorous and feasible methodologies is needed to improve assessments of the value of NIH. Of the many assessments that have been undertaken, no single approach has proven entirely	

FINDINGS	CONCLUSIONS AND RECOMMENDATIONS
satisfactory.	develop an assessment approach guide that outlines the factors to consider and the mix of methodologies (e.g., retrospective, prospective, qualitative, quantitative) that should be employed in attempting to capture value; and  e. Seek input from external experts in the development of methods and tools to improve assessments of the value of biomedical research.
Effectively communicating the results of assessments of the value of NIH gives the agency the opportunity to increase public understanding of the scientific research process, NIH’s mission, and the ways in which biomedical research affects daily life.	6. Every assessment activity that NIH undertakes should begin with identifying the purpose of the study and its audiences. Assessment study designs should include diverse communication strategies to disseminate results in ways that will enhance awareness and understanding of the scientific research process among a variety of audiences.

As noted previously, the ultimate goal in better assessing the value of biomedical research supported by NIH is to use this information to enhance that value. Having a clear picture of value across the spectrum of biomedical research could enable decision-makers to better identify which research will have the greatest impact on health, the economy, and other aspects of society, or to note impediments to achieving value. However, the Working Group finds that assessments of biomedical research value are not yet comprehensive or systematic enough to provide a foundation for making management decisions regarding how to enhance that value. The Working Group recommends that as comprehensive assessments of value become more rigorous, NIH should study how best to use the results of these assessments to enhance value.

**VI. SMRB CONCLUSIONS – TO BE DETERMINED**

## **APPENDIX A**

### **Speakers and Dates**

#### **January 14, 2013**

- William B. Rouse, Ph.D., Chair of Economics and Engineering, Stevens Institute of Technology
- Andrew A. Toole, Ph.D., Economist, Economic Research Service, U.S. Department of Agriculture
- Simon Tripp, Senior Director, Technology Partnership Practice, Battelle Memorial Institute

#### **March 13, 2013**

- Richard Ikeda, Ph.D., Director, Office of Research Information Systems and Research, Condition, and Disease Categorization, Office of Extramural Research, Office of the Director, NIH

#### **March 19, 2013**

- Mark Rohrbaugh, Ph.D., J.D., Director, Office of Technology Transfer, Office of the Director, National Institutes of Health
- George Santangelo, Ph.D., Director, Office of Portfolio Analysis, Division of Program Coordination, Planning, and Strategic Initiatives, Office of the Director, National Institutes of Health

#### **June 4, 2013**

- James M. Anderson, M.D., Ph.D., Deputy Director for Program Coordination, Planning, and Strategic Initiatives, NIH
- Irma Arispe, Ph.D., Director, Office of Analysis and Epidemiology, National Center for Health Statistics, Office of Surveillance, Epidemiology, and Laboratory Services, Centers for Disease Control and Prevention
- James W. Curran, M.D., M.P.H., James W. Curran Dean of Public Health, Rollins School of Public Health, Emory University
- David M. Cutler, Ph.D., Otto Eckstein Professor of Applied Economics, Harvard University
- Barbara Entwisle, Ph.D., Vice Chancellor for Research and Kenan Distinguished Professor, University of North Carolina at Chapel Hill
- Laurel Haak, Ph.D., Executive Director, ORCID
- Della M. Hann, Ph.D., Deputy Director, Office of Extramural Research, NIH
- Mark B. McClellan, M.D., Ph.D., Director of the Engelberg Center for Health Care Reform and Leonard D. Schaeffer Chair in Health Policy Studies, Brookings Institution
- Robert H. Topel, Ph.D., Professor in Urban and Labor Economics, University of Chicago Booth School of Business
- Elias A. Zerhouni, M.D., President of Global R&D, Sanofi

#### **June 5, 2013**

- David Baker, Founder and Executive Director, CASRAI, Canada
- Margaret Blume-Kohout, Ph.D., Senior Research Economist, New Mexico Consortium

- Jason Priem, Ph.D. Student, School of Information and Library Science, University of North Carolina–Chapel Hill

**August 8, 2013**

- Ann Bonham, Chief Scientific Officer, AAMC
- Susan Guthrie, Senior Analyst, RAND Europe
- Tony Mazzaschi, Senior Director, Scientific Affairs, AAMC
- Steven Wooding, Research Leader, RAND Europe

**August 13, 2013**

- J. Michael McGinnis, M.D., M.P.P., IOM Senior Scholar and Executive Director of the IOM Roundtable on Value & Science-Driven Health Care
- Bhaven Sampat, Ph. D., Associate Professor, Department of Health Policy and Management, Mailman School of Public Health, Columbia University

**August 22, 2013**

- Julia I. Lane, Ph.D., Senior Managing Economist, American Institutes for Research

**September 26, 2013**

- Pritty Joshi, Ph.D., Health Science Policy Analyst, Office of Extramural Research, National Institutes of Health
- Luci Roberts, Ph.D., Director of Planning and Evaluation, Office of Extramural Research, National Institutes of Health

**October 24, 2013**

- Alan Leshner, Ph.D., CEO, American Association for the Advancement of Science
- Peter Orszag, Ph.D., Vice Chairman of Corporate and Investment Banking and Chairman of the Financial Strategy and Solutions Group, Citigroup
- Harvey Fineberg, M.D., Ph.D., President, Institute of Medicine
- Elaine Gallin, Ph.D., Partner, QE Philanthropic Advisors
- Kathryn Ahlport, M.S.P.H., Executive Director, Health Research Alliance
- Carl Rhodes, Ph.D., Senior Scientific Officer, Howard Hughes Medical Institute
- Marie Nierras, Ph.D., Assistant Vice President, International Partnerships, Juvenile Diabetes Research Foundation
- Mary Woolley, President, Research!America
- Antoinette Royster, Member, NIH Clinical Center Patient Advisory Board
- Jerry Sachs, Member, NIH Clinical Center Patient Advisory Board
- Ian Viney, Ph.D., Head of Evaluation, Medical Research Council, UK
- Philip Yeo, M.S., M.B.A., Chairman, SPRING Singapore

**November 19, 2013**

- John Burklow, Associate Director for Communications and Public Liaison, National Institutes of Health
- Sally Rockey, Deputy Director for Extramural Research, National Institutes of Health

**December 5, 2013**

- MRC Greenwood, Ph.D.

**APPENDIX B**  
**Literature**

Battelle Technology Partnership Practice. (2011). *Economic Impact of the Human Genome Project*.

Blume-Kohout, Margaret E. (2012). *Does Targeted, Disease-Specific Public Research Funding Influence Pharmaceutical Innovation?* University of New Mexico.

Callahan, Kevin, et al. (2012). *Assessing the Outcomes of NIH Research Spending: Toward a Better Approach*. Bethesda, Maryland: National Institutes of Health.

Chatterjee, Anusuya and Ross C. DeVol. (2012). *Estimating Long-Term Economic Returns of NIH Funding on Output in the Biosciences*. Milken Institute.

Committee on Prospering in the Global Economy of the 21st Century: An Agenda for American Science and Technology and Committee on Science, Engineering, and Public Policy. (2007). *Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future*. National Academy of Sciences, National Academy of Engineering, and Institute of Medicine of the National Academies.

Cutler, David M., Allison B. Rosen, and Sandeep Vijan. (2006). *The Value of Medical Spending in the United States, 1960–2000*. Harvard University, University of Michigan, National Bureau of Economic Research, and the Ann Arbor Veterans Affairs Medical Center.

Families USA's Global Health Initiative. (2008). *In Your Own Backyard: How NIH Funding Helps Your State's Economy*.

Goldfarb, Brent. (2008). *The Effect Of Government Contracting on Academic Research: Does the Source of Funding Affect Scientific Output?* University of Maryland.

Health Economics Research Group (HERG), Brunel University. (2008). *Medical Research: What's it worth?* Office of Health Economics (OHE), RAND Europe.

Hendrix, Dean. (2008). *An Analysis of Bibliometric Indicators, National Institutes of Health Funding, and Faculty Size at Association of American Medical Colleges Medical Schools, 1997–2007*. State University of New York at Buffalo.

Jacob, Brian and Lars Lefgren. (2007). *The Impact of Research Grant Funding on Scientific Productivity*. National Bureau of Economic Research.

Johnston, S. Claiborne, John D. Rootenberg, Shereen Katrak, Wade S. Smith, and Jacob S. Elkins. (2006). *Effect of a US National Institutes of Health Programme of Clinical Trials on Public Health and Costs*. University of California, San Francisco.

Murphy, Kevin M. and Robert Topel. (1999). *The Economic Value of Medical Research*. University of Chicago.

Murphy, Kevin M. and Robert Topel. (2005). *The Value of Health and Longevity*. University of Chicago, National Bureau of Economic Research, and the George J. Stigler Center for the Study of the Economy and the State.

Stevens, Ashley J. et al. (2011). *The Role of Public-Sector Research in the Discovery of Drugs and Vaccines*. Boston University, Norwegian Radium Hospital Research Foundation, and NIH.

Toole, Andrew A. (2007). *Does Public Scientific Research Complement Private Investment in Research and Development in the Pharmaceutical Industry?* Rutgers University.

Toole, Andrew A. (2012). *The Impact of Public Basic Research on Industrial Innovation: Evidence from the Pharmaceutical Industry*. Economic Research Service at USDA.

United for Medical Research. (2011). *An Economic Engine: NIH Research, Employment, and the Future of the Medical Innovation Sector*.

United for Medical Research. (2012). *NIH's Role in Sustaining the U.S. Economy: A 2011 Update Authored by Dr. Everett Ehrlich*.

## APPENDIX C

## Biomedical Research Outputs and Outcomes with Measurement and Assessment Tools

## (1) Scientific Impacts

OUTPUTS	
DISCOVERY	<ul style="list-style-type: none"> <li>• <b>Primary research articles</b> (Assessment tools: Pubmed, SCOPUS, ISI, eSPA, SPIRES, RePORTER, RPPRs, etc.)</li> <li>• <b>Conference abstracts/ presentations</b> (Assessment tools: NIH Research Performance Progress Reports (RPPRs), <i>Tools lacking?</i>)</li> <li>• <b>Research databases and repositories (i.e., NGVB, NDAR)</b> (Assessment tools: <i>Individual metrics but no comprehensive tools</i>)</li> <li>• <b>Development of research resources and infrastructure</b> (Assessment tools: Resource and infrastructure grants; RPPRs)</li> <li>• <b>Non-peer reviewed research findings and reports (e.g., self-publishing, blogging, data sharing)</b> (Assessment tools: RPPRs (self-reported websites, audio/video products, protocols, software, etc.); <i>Tools lacking? (Altmetrics approaches are emerging)</i>)</li> </ul>
VALIDATION	<ul style="list-style-type: none"> <li>• <b>Patents</b> (Assessment tools: USPTO database; eSPA and RePORTER (linking publications to patents); RPPRs; FDA Orange Book; iEdison)</li> <li>• <b>Licenses</b> (Assessment tools: USPTO database; FDA Orange Book; iEdison)</li> <li>• <b>IP transfer agreements (MTAs)</b> (Assessment tools: Individual University Admin. Databases (<i>lack aggregate databases</i>))</li> <li>• <b>Development and initial clinical testing of dx, tx, and px interventions (e.g., FDA applications and Phase I, II, and III clinical trials)</b> (Assessment tools: Clinicaltrials.gov, RePORTER (linking projects to clinicatrials.gov), ExPORTER, RPPRs, NME and IND applications in FDA Orange Book; <i>Comprehensive tools?</i>)</li> <li>• <b>Standardized research protocols</b> (Assessment tools: <i>Publications databases, but no comprehensive tools</i>)</li> <li>• <b>Validated data repositories (i.e., NCBI's GenBank, PDB)</b> (Assessment tools: Example: NCBI analytical tools; <i>Individual metrics but no comprehensive tools</i>)</li> </ul>
CONSENSUS	<ul style="list-style-type: none"> <li>• <b>FDA approvals of New Drug Applications (NDAs)</b> (Assessment tools: FDA Orange Book)</li> <li>• <b>Consensus Development Conferences and Bodies (i.e., National Toxicology Program)</b> (Assessment tools: <i>Individual metrics but no comprehensive tools</i>)</li> <li>• <b>Book chapters</b> (Assessment tools: Publication databases (<i>comprehensive?</i>))</li> <li>• <b>Post-grad curriculum guidelines</b> (Assessment tools: <i>Tools lacking?</i>)</li> <li>• <b>Systematic reviews/ meta-analyses</b> (Assessment tools: Pubmed, Cochrane Database; <i>No comprehensive tools</i>)</li> <li>• <b>Knowledge transfer metrics (i.e, citations, web hits, etc.)</b> (Assessment tools: SPIRES, eSPA (linking projects to citations); Thomson Reuters ISI and ScienceWire; Elsevier's Scopus; Google Scholar; Journal Impact Factor; H-Index; "Altmetrics" (web-driven computational approaches))</li> </ul>
OUTCOMES	
	Fundamental Scientific Knowledge

**Biomedical Research Outputs and Outcomes with Measurement and Assessment Tools  
(2) Health Impacts**

OUTPUTS	
<b>APPLICATION</b>	<ul style="list-style-type: none"> <li>• <b>Advanced testing of evidence-based diagnostics</b> (Assessment Tools: RePORTER, ExPORTER, RPPRs, Patent applications (USPTO), IC- provided data, FDA, other HHS agencies; <i>No comprehensive tools?</i>)</li> <li>• <b>Advanced testing of evidence-based therapeutic interventions</b> (Assessment Tools: Clinicaltrials.gov (Phase IV trials), RePORTER (linking projects to clinicaltrials.gov), ExPORTER, RPPRs, FDA Orange Book (approvals of NDAs); <i>Comprehensive tools?</i>)</li> <li>• <b>Advanced testing of evidence-based preventive interventions and strategies</b> (Assessment Tools: Clinicaltrials.gov (Phase IV trials), RePORTER (linking projects to clinicaltrials.gov), ExPORTER, RPPRs, FDA Orange Book (approvals of NDAs); <i>Comprehensive tools?</i>)</li> <li>• <b>Identification of major public health issues and disease risk factors</b> (Assessment Tools: CDC (NCHS), WHO)</li> </ul>
<b>IMPLEMENTATION</b>	<ul style="list-style-type: none"> <li>• <b>Drugs to market</b> (Assessment Tools: FDA approvals, Industry Reports)</li> <li>• <b>Dissemination and Implementation efforts, Practice guidelines, etc.</b> (Assessment Tools: HHS dissemination and implementation activities; AHRQ's National Guideline Clearinghouse; Public Health agency recommendations; Professional medical associations guidelines; <i>Tools to track research citations in policies/guidelines?</i>)</li> <li>• <b>Adoption of evidence-based dx, tx, and px strategies</b> (Assessment Tools: Healthcare claims data (i.e., Medicaid); AHRQ, HHS evaluations; mHealth apps to monitor adherence; <i>Tools in development?</i>)</li> <li>• <b>Public health campaigns</b> (Assessment Tools: HHS activities (e.g., CDC); NGO activities; <i>No comprehensive tools</i>)</li> </ul>
<b>HEALTH MEASURES</b>	<ul style="list-style-type: none"> <li>• <b>Burden of Disease/Quality of Life</b> (Assessment Tools: QALYS, DALYs, HRQoL, etc. - CDC (NCHS), WHO, survey data (NSHAP, NHANES, etc.))</li> <li>• <b>Disability</b> (Assessment Tools: Activities of Daily Living (ADLs), Mini Mental State Exam (MMSE), etc.; CDC (NCHS) WHO, survey data)</li> <li>• <b>Longevity</b> (Assessment Tools: Life expectancy; NCHS, CDC, WHO, survey data)</li> </ul>
OUTCOMES	
	<p><b>Living longer, healthier lives</b></p>

**Biomedical Research Outputs and Outcomes with Measurement and Assessment Tools**

**(3) Broader Societal Impacts**

OUTPUTS	
<b>PROXIMAL</b>	<ul style="list-style-type: none"> <li>• <b>Support government, science, and technology jobs</b> (Assessment Tools: NIH Budget Office, RPPRs, contract invoicing, STAR METRICS; <i>Tools lacking?</i>)</li> <li>• <b>Create demand for R&amp;D supplies</b> (Assessment Tools: Purchase requests, RPPRs, STAR METRICS; <i>Tools lacking?</i>)</li> <li>• <b>International collaboration</b> (Assessment Tools: NIH funding and cooperative agreements for int'l activities, RPPRs, Fogarty database, <i>Tools lacking?</i>)</li> <li>• <b>Support academia</b> (Assessment Tools: NIH funding; RePORTER; STAR METRICS; RPPRs)</li> <li>• <b>De-risk pre-competitive space</b> (Assessment Tools: R&amp;D investment by Pharma and Biotech industry); <i>Tools lacking?</i>)</li> <li>• <b>Cross-sector collaboration</b> (Assessment Tools: Material Transfer Agreements (MTAs); <i>Tools lacking?</i>)</li> </ul>
<b>INTERMEDIATE</b>	<ul style="list-style-type: none"> <li>• <b>Spur private sector activity (i.e., Biotech, Pharma)</b> (Assessment Tools: FDA approvals, patents, industry reports (PhRMA), Bureau of Labor Statistics)</li> <li>• <b>Enhance STEM education</b> (Assessment Tools: NSF Report on Science and Engineering Indicators; NAS, DOE)</li> <li>• <b>Communication and interpretation of findings across sectors and to public</b> (Assessment Tools: IC provided-data, ASPE and NIH evaluations; CDC (NCHS))</li> <li>• <b>International S&amp;T capacity building</b> (Assessment Tools: <i>Tools lacking?</i>)</li> <li>• <b>Spur local economy</b> (Assessment Tools: <i>Tools lacking?</i>)</li> </ul>
<b>DISTAL</b>	<ul style="list-style-type: none"> <li>• <b>Uptake and spread of technological innovations</b> (Assessment Tools: <i>Tools lacking?</i>)</li> <li>• <b>Workforce output (longevity, health)</b> (Assessment Tools: CDC (NCHS), WHO)</li> <li>• <b>Workforce development</b> (Assessment Tools: NSF Report on Science and Engineering Indicators)</li> <li>• <b>Internationally competitive science and technology sectors</b> (Assessment Tools: NSF Report on Science and Engineering Indicators; OECD)</li> <li>• <b>GDP</b> (Assessment Tools: Bureau of Labor Statistics, Department of Commerce statistics)</li> <li>• <b>Emergence of new sectors and industries</b> (Assessment Tools: Bureau of Labor Statistics, Department of Commerce statistics)</li> <li>• <b>Health care costs</b> (Assessment Tools: Federal data sources; Commercial data sources)</li> </ul>
OUTCOMES	
	<ul style="list-style-type: none"> <li>• <b>Scientifically literate public</b></li> <li>• <b>Healthcare-related cost savings</b></li> <li>• <b>↑ Productivity</b></li> <li>• <b>↑ Capacity for Innovation</b></li> <li>• <b>↑ Global R&amp;D competitiveness</b></li> <li>• <b>Diplomacy &amp; stability through science</b></li> </ul>