

Genome-Wide Association Studies (GWAS) Policy - Frequently Asked Questions

A. Overall Policy

1. [When is the GWAS Policy effective?](#)
2. [How does the NIH GWAS policy relate to the existing NIH data sharing policy?](#)
3. [What mechanisms are in place to monitor the implementation of the GWAS policy and to ensure that it keeps pace with evolving technological and ethical issues?](#)

B. Application Submission and Review

1. [When is the GWAS Policy effective?](#)
2. [To what types of awards does the GWAS policy apply?](#)
3. [Does the GWAS data sharing policy apply to already-funded NIH studies?](#)
4. [Are clinical trials that include a genetic association component subject to the GWAS policy?](#)
5. [How should an application be identified as a "GWAS" application?](#)
6. [How will GWAS sharing plans be reviewed?](#)
7. [Are there cases in which data sharing will not be possible? How should NIH GWAS grant applicants deal with such instances?](#)
8. [What is the process for obtaining certificates of confidentiality?](#)
9. [Can investigators seek additional resources to contact previous participants of studies to obtain additional informed consent?](#)
10. [When should investigators submit their data use certification from their institutions?](#)
11. [What should be included in a progress report for NIH-funded GWAS?](#)

C. Data Access

1. [What is dbGaP?](#)
2. [Will there be one GWAS repository or multiple GWAS repositories?](#)
3. [How will the NIH GWAS data repository include data from other existing data sources, such as the Cancer Biomedical Informatics Grid \(caBIG\)?](#)
4. [Will biological samples \(e.g. tissues, DNA, cell lines\) associated with the data included in the NIH GWAS data repository be available through the repository?](#)
5. [Can non-research entities \(e.g., law enforcement agencies, insurance companies, employers\) request access to identifiable information corresponding to phenotype and genotype data held in the NIH GWAS data repository?](#)
6. [What security measures will be in place to prevent unauthorized access to the NIH GWAS repository?](#)
7. [Will NIH intramural staff be expected to follow the same processes for submitting data and requesting access as extramural investigators?](#)
8. [Must the members of the Data Access Committees \(DACs\) all be Federal or NIH employees? What kinds of expertise will be represented on the DACs?](#)
9. [What should investigators do if they are contacted by other researchers who want to use their data but don't want to go through the GWAS repository?](#)
10. [What should be included in the annual reports of approved users of data from the NIH GWAS data repository and to whom should they be submitted?](#)
11. [What should be included in a progress report for NIH-funded GWAS?](#)

D. Data Monitoring and Oversight

1. [How will the NIH ensure that data obtained from the NIH GWAS data repository are used in a manner consistent with the informed consent provided by research participants?](#)
2. [What should be included in the annual reports of approved users of data from the NIH GWAS data repository and to whom should they be submitted?](#)

E. Data Repository

1. [What is dbGaP?](#)
2. [Will there be one GWAS repository or multiple GWAS repositories?](#)
3. [How will the NIH GWAS data repository include data from other existing data sources, such as the Cancer Biomedical Informatics Grid \(caBIG\)?](#)
4. [Will biological samples \(e.g. tissues, DNA, cell lines\) associated with the data included in the NIH GWAS data repository be available through the repository?](#)
5. [Will the NIH standardize phenotype information?](#)
6. [How will the NIH ensure that data obtained from the NIH GWAS data repository are used in a manner consistent with the informed consent provided by research participants?](#)

F. Data Submission

1. [What is dbGaP?](#)
2. [Will there be one GWAS repository or multiple GWAS repositories?](#)
3. [Will the NIH standardize phenotype information?](#)
4. [What is the IRB's role in submission of data to the NIH GWAS data repository?](#)
5. [How will updates of original data from the submitting investigators be made available to recipient investigators?](#)
6. [Can investigators conducting GWAS as part of a clinical trial limit their data submission to baseline data only \(i.e. data obtained prior to the start of the intervention\)?](#)
7. [Will biological samples \(e.g. tissues, DNA, cell lines\) associated with the data included in the NIH GWAS data repository be available through the repository?](#)
8. [Are there cases in which data sharing will not be possible? How should NIH GWAS grant applicants deal with such instances?](#)
9. [For multi-center studies, is the submitting institution expected to certify data that are contributed by data collection centers at other institutions?](#)
10. [Will the NIH GWAS repository accept GWAS data if the individuals from whom the data/samples are derived are deceased and either a\) information about the consent process or the consent forms is not available or b\) consent was silent or not inconsistent with regard to submission to the NIH GWAS Data Repository?](#)
11. [As stated in the NIH GWAS policy, the NIH expects that all submissions to the NIH GWAS data repository will include a certification by the responsible Institutional Official\(s\) of the submitting institution that the expectations of the policy have been met for submission to the NIH GWAS data repository. For multi-center studies, is the submitting institution expected to certify data that are contributed by data collection centers at other institutions?](#)
12. [Are Submitting Institutions expected to certify that data submission is consistent with applicable laws and regulations in effect at any and all locations at which data were collected?](#)
13. [Will NIH accept the submission of GWAS data to dbGaP if the data were derived from de-identified specimens that were not consented for research purposes \(e.g., residual clinical samples or cell lines\), which is permitted under the Common Rule?](#)

G. Governance

1. [What mechanisms are in place to monitor the implementation of the GWAS policy and to ensure that it keeps pace with evolving technological and ethical issues?](#)
2. [How will the NIH ensure that data obtained from the NIH GWAS data repository are used in a manner consistent with the informed consent provided by research participants?](#)

H. Publication

1. [How will the work of submitters of GWAS data be acknowledged by secondary users?](#)

I. Research Participant Protections

1. [How will the NIH ensure that data obtained from the NIH GWAS data repository are used in a manner consistent with the informed consent provided by research participants?](#)
2. [What is the IRB's role in submission of data to the NIH GWAS data repository?](#)
3. [Are the data included in the NIH GWAS data repository subject to the Freedom of Information Act \(FOIA\)?](#)
4. [Can non-research entities \(e.g., law enforcement agencies, insurance companies, employers\) request access to identifiable information corresponding to phenotype and genotype data held in the NIH GWAS data repository?](#)
5. [What is the process for obtaining certificates of confidentiality?](#)
6. [Can investigators seek additional resources to contact previous participants of studies to obtain additional informed consent?](#)
7. [What is the role of the NIH DACs in considering risks to individuals, their families, and groups or populations associated with data submitted to the NIH GWAS data repository?](#)
8. [What is the role of the IRB/Privacy Board of the submitting institution in considering risks to groups or populations associated with data submitted to the NIH GWAS data repository?](#)

J. Institutional Review Board (IRB) Involvement

1. [As stated in the NIH GWAS policy, the NIH expects an IRB and/or Privacy Board to review and verify that the submission of data to the NIH GWAS Data Repository and subsequent sharing for research purposes are consistent with the informed consent of study participants from whom the data were obtained. Does this mean that the NIH expects specific elements to be included in the informed consent documents? For example, are the elements discussed in the Points to Consider for IRBs and Institutions document \(p.11-13\) intended to be an explicit consent standard that IRBs should use in assessing consent forms?](#)
2. [The NIH GWAS policy indicates that monies may be available for re-consent of participants. Does the NIH have criteria or guidelines regarding when an IRB should pursue re-consent of participants?](#)
3. [What process should IRBs or Privacy Boards use when reviewing GWAS data sharing plans for submission of GWAS data to the NIH GWAS Data Repository?](#)
4. [What is the role of the IRB/Privacy Board of the submitting institution in considering risks to groups or populations associated with data submitted to the NIH GWAS data repository?](#)
5. [Can IRBs obtain a waiver of consent for submission of data to the NIH GWAS repository?](#)

Genome-Wide Association Studies (GWAS) Policy - Frequently Asked Questions

A. Overall Policy

1. **When is the GWAS Policy effective?**

The GWAS Policy is effective for competing applications and proposals submitted for January 25, 2008 and all subsequent standard NIH receipt dates, including submission for on-going Funding Opportunity Announcements (FOAs) issued prior to January 25, 2008. All such applicants will be expected to submit a data sharing plan consistent with the GWAS Policy, and to meet the expectations of the Policy in good faith. Any submitted data sharing plans that are accepted by the funding IC may be referenced in the Notice of Award as a term and condition of the award.

The policy also applies to NIH intramural research projects that which contain GWAS and are approved for support on or after January 25, 2008.

2. **How does the NIH GWAS policy relate to the existing NIH data sharing policy?**

The data sharing expectations in the GWAS policy are similar to the expectations of the current NIH data sharing policy applying to applications requesting more than \$500,000 or more in direct costs in any single budget year (or as required by a specific FOA). Both policies reaffirm NIH support for timely sharing and distribution of biomedical research resources. However, the GWAS policy does not include a dollar threshold, and applies specifically to GWAS, with the expectation that data will be submitted to the NIH GWAS data repository.

3. **What mechanisms are in place to monitor the implementation of the GWAS policy and to ensure that it keeps pace with evolving technological and ethical issues?**

The governance structure for GWAS is described in the preamble section of the policy (<http://grants.nih.gov/grants/guide/notice-files/NOT-OD-07-088.html>). The NIH Director is ultimately responsible for overseeing the GWAS policy and its implementation. The Director will receive guidance and appropriate leadership from within the Office of the Director. A fundamental aspect of the GWAS governance plan is a working group of the Advisory Committee to the NIH Director (ACD), which is composed of external experts representing the interests of the public and the scientific community. The ACD Working Group will be responsible for informing the ACD regarding any emerging ethical or scientific issues that may be relevant to NIH efforts to maintain the highest standards in participant protection and the promotion of quality research. The ACD will then advise the NIH Director on these issues.

The NIH Data Access Committees will monitor data use practices by reviewing annual reports that it will receive from all approved users. Summary reports on data distribution, DAC processes, and GWAS data use practices will also be provided to the ACD Working Group on a regular basis.

[Back to Top](#)

B. Application Submission and Review

1. **When is the GWAS Policy effective?**

The GWAS Policy is effective for competing applications and proposals submitted for January 25, 2008 and all subsequent standard NIH receipt dates, including submission for on-going Funding Opportunity Announcements (FOAs) issued prior to January 25, 2008. All such

applicants will be expected to submit a data sharing plan consistent with the GWAS Policy, and to meet the expectations of the Policy in good faith. Any submitted data sharing plans that are accepted by the funding IC may be referenced in the Notice of Award as a term and condition of the award.

The policy also applies to NIH intramural research projects that which contain GWAS and are approved for support on or after January 25, 2008.

2. To what types of awards does the GWAS policy apply?

The GWAS policy applies to all NIH mechanisms of support for extramural research, with the exception of individual fellowships (F awards) or institutional training grants (T awards). The policy does apply to individual Career Development (K) awards, provided that the scope and aims of the project include the development of GWAS data. Comparable expectations are in place for all GWAS conducted by NIH intramural investigators.

3. Does the GWAS data sharing policy apply to already-funded NIH studies?

Non-competing (Type 5) awards are not expected to include data sharing plans that address the GWAS Policy. Applications received for competing awards (Type 1 and 2) for the January 25, 2008 receipt date (and all subsequent receipt dates) will be expected to include a data sharing plan consistent with the GWAS Policy or provide an explanation of why data sharing is not possible.

4. Are clinical trials that include a genetic association component subject to the GWAS policy?

Yes. The GWAS policy applies to all basic and clinical research supported by the NIH.

5. How should an application be identified as a "GWAS" application?

A cover letter indicating the presence of a GWAS component should accompany all such applications. Staff in the CSR Division of Receipt and Referral will confirm the identification and enter as a dual assignment whenever appropriate. Program staff in NIH ICs may also recommend that a dual designation be added or removed.

6. How will GWAS sharing plans be reviewed?

Peer reviewers will evaluate data sharing plans for consistency with the NIH GWAS policy and will provide comments on them in the Remarks section of summary statements. Program staff are responsible for assessing the appropriateness and adequacy of proposed data sharing plans. Program concerns regarding data sharing plans must be resolved prior to making awards.

7. Are there cases in which data sharing will not be possible? How should NIH GWAS grant applicants deal with such instances?

Submission of some GWAS data to the repository may be precluded by various factors, such as international laws, limitations in the original informed consents, concerns about harms to individuals or groups, or other cases where expectations for data submission cannot be met or restrictions on the original funding mechanism. Through the governance structure, ICs will be provided guidance on appropriate exceptions. Applicants who believe that they will be unable to share data as described in the policy will be asked to explain in detail why data sharing is not possible in the data sharing plan portion of the application. ICs make funding decisions based on both scientific excellence and program priorities and may take data sharing plans into consideration. In doing so, applications that represent the highest quality

science but do not provide for standard data deposition and distribution through the NIH GWAS data repository will be considered for funding by ICs on a case-by-case basis.

8. **What is the process for obtaining certificates of confidentiality?**

The necessary information regarding the process and the appropriate application of this protection may be found at the NIH Certificates of Confidentiality Kiosk located at <http://grants.nih.gov/grants/policy/coc/>.

9. **Can investigators seek additional resources to contact previous participants of studies to obtain additional informed consent?**

The NIH may give programmatic consideration to requests for supplemental funding to obtain additional participant consent when appropriate. Requests for such support should be directed to the appropriate IC Program Officer.

10. **When should investigators submit their data use certification from their institutions?**

The institutional certification is submitted as part of the data submission to the NIH GWAS data repository.

11. **What should be included in a progress report for NIH-funded GWAS?**

Grantees receiving support for GWAS are expected to complete a standard NIH progress report. If the competing award included a plan to submit data to the NIH GWAS data repository, the grantee should describe progress in implementing the plan. A final statement on submitting data to the repository should also be included in the final progress report prior to grant closeout. In addition to progress reports related to the receipt of any funding award for GWAS, all approved data users of GWAS datasets will be expected to submit annual reports to the appropriate NIH DAC describing the use and analysis of any GWAS dataset accessed.

[Back to Top](#)

C. Data Access

1. **What is dbGaP?**

dbGaP (Database of Genotypes and Phenotypes) was developed by the National Center for Biotechnology Information (a division of the National Library of Medicine of the NIH) to archive and distribute the results of studies that have investigated the interaction of genotype and phenotype. dbGaP will serve as the NIH GWAS data repository. Additional information on dbGaP can be found at <http://www.ncbi.nlm.nih.gov/sites/entrez?db=gap>.

2. **Will there be one GWAS repository or multiple GWAS repositories?**

The NIH will maintain one central GWAS repository. However, it is not intended to be the exclusive source for data from the NIH supported or conducted GWAS. For instance, the same data may also be held by an institutional repository or other sites, depending on the study consents and the needs of the submitting investigators.

3. **How will the NIH GWAS data repository include data from other existing data sources, such as the Cancer Biomedical Informatics Grid (caBIG)?**

The NIH GWAS data repository was designed to accommodate the data structures of many

of the GWAS that are likely to submit data to it. Nonetheless, the National Center for Biotechnology Information (NCBI) will discuss the best transfer format with all submitting investigators prior to data transmission. The same plan will be followed for importing relevant data from caBIG. In the future, the NIH GWAS data repository may contain information that will allow users to readily access more specialized resources linked to the data in the repository.

4. **Will biological samples (e.g. tissues, DNA, cell lines) associated with the data included in the NIH GWAS data repository be available through the repository?**

No. Only the genotype and phenotype data will be available through the NIH GWAS data repository.

5. **Can non-research entities (e.g., law enforcement agencies, insurance companies, employers) request access to identifiable information corresponding to phenotype and genotype data held in the NIH GWAS data repository?**

The NIH will not possess any direct identifiers within the NIH GWAS data repository, nor will the NIH have access to the link between the data keycode and the identifiable information that may reside with the primary investigators and institutions for particular studies. The NIH explicitly encourages investigators submitting GWAS to consider the potential appropriateness of obtaining a Certificate of Confidentiality (please see <http://grants.nih.gov/grants/policy/coc/>) as an added protection against any future compelled disclosure of identities for studies planning to collect genome-wide association data.

6. **What security measures will be in place to prevent unauthorized access to the NIH GWAS repository?**

Datasets are stored in the NIH GWAS data repository under strict security provisions, including multiple firewalls, separate servers, and data encryption protocols. Investigators and their sponsoring institutions seeking access to data in the NIH GWAS data repository must submit a data access request that specifies both the data to which access is sought and the planned research use, and agree to the terms of access set forth in the Data Use Certification (DUC). Investigators are approved by a Data Access Committee (DAC) for access to specific datasets for a specific use(s). Data can only be accessed through the NIH login process. In addition, the DUCs include a provision that approved users and their institutions agree to store the requested data securely and to not share the requested data with third parties.

7. **Will NIH intramural staff be expected to follow the same processes for submitting data and requesting access as extramural investigators?**

Yes. NIH intramural staff will be expected to submit data for all GWAS protocols approved after January 25, 2008, including those protocols generating data from collaborations with outside institutions. The data submission is subject to the same certifications expected of extramural investigators. Access to the data submitted by intramural investigators to the NIH GWAS data repository is also subject to the same provisions as data submitted by extramural investigators.

8. **Must the members of the Data Access Committees (DACs) all be Federal or NIH employees? What kinds of expertise will be represented on the DACs?**

The members of the DACs must all be federal employees but need not all be NIH employees. The members will be selected for their expertise in areas such as the relevant scientific and

clinical disciplines (based on programmatic areas of interest), research participant protection, and privacy.

9. **What should investigators do if they are contacted by other researchers who want to use their data but don't want to go through the GWAS repository?**

Principal Investigators who submitted GWAS data to the NIH GWAS data repository may share their projects data directly with other investigators, following applicable regulatory requirements and institutional policies. However, investigators that receive the data through the NIH GWAS data repository should not provide data to any individual not included within the scope of the DAC-approved Data Access Request per the terms of access in the Data Use Certification.

10. **What should be included in the annual reports of approved users of data from the NIH GWAS data repository and to whom should they be submitted?**

Approved data users are expected to submit annual reports that include details of any significant research findings, breaches in data security or other issues related to the terms of access as specified in the Data Use Certification, and any publications or intellectual property developed through the use of data from the NIH GWAS data repository. The reports should be submitted to the Data Access Committee that approved their Data Access Request.

11. **What should be included in a progress report for NIH-funded GWAS?**

Grantees receiving support for GWAS are expected to complete a standard NIH progress report. If the competing award included a plan to submit data to the NIH GWAS data repository, the grantee should describe progress in implementing the plan. A final statement on submitting data to the repository should also be included in the final progress report prior to grant closeout. In addition to progress reports related to the receipt of any funding award for GWAS, all approved data users of GWAS datasets will be expected to submit annual reports to the appropriate NIH DAC describing the use and analysis of any GWAS dataset accessed.

[Back to Top](#)

D. Data Monitoring and Oversight

1. **How will the NIH ensure that data obtained from the NIH GWAS data repository are used in a manner consistent with the informed consent provided by research participants?**

Through the Controlled Access process for providing data access to secondary users, mechanisms are in place to minimize the likelihood of usage of GWAS data in ways that are inconsistent with the original informed consent. For instance, the NIH expects that an IRB or Privacy Board will have reviewed all proposed submissions of data to the NIH GWAS data repository and determined that the submission and subsequent sharing for research purposes are consistent with the informed consent of the study participants from whom the data were obtained. In addition, submitting institutions are expected to certify that the appropriate research uses of the data and the uses that are specifically excluded by the relevant informed consent documents are delineated. For example, the consent forms for some GWAS studies may have restrictions regarding the disease(s) to be studied. Information from the submitting institution about data use limitations will then be used by

the NIH Data Access Committees to inform their review of requests for access to determine whether a proposed use of the dataset conflicts with any informed consent limitations. Researchers approved for access to a GWAS dataset must agree not to use the data for any research other than that described in their Data Access Request.

2. **What should be included in the annual reports of approved users of data from the NIH GWAS data repository and to whom should they be submitted?**

Approved data users are expected to submit annual reports that include details of any significant research findings, breaches in data security or other issues related to the terms of access as specified in the Data Use Certification, and any publications or intellectual property developed through the use of data from the NIH GWAS data repository. The reports should be submitted to the Data Access Committee that approved their Data Access Request.

[Back to Top](#)

E. Data Repository

1. **What is dbGaP?**

dbGaP (Database of Genotypes and Phenotypes) was developed by the National Center for Biotechnology Information (a division of the National Library of Medicine of the NIH) to archive and distribute the results of studies that have investigated the interaction of genotype and phenotype. dbGaP will serve as the NIH GWAS data repository. Additional information on dbGaP can be found at <http://www.ncbi.nlm.nih.gov/sites/entrez?db=gap>.

2. **Will there be one GWAS repository or multiple GWAS repositories?**

The NIH will maintain one central GWAS repository. However, it is not intended to be the exclusive source for data from the NIH supported or conducted GWAS. For instance, the same data may also be held by an institutional repository or other sites, depending on the study consents and the needs of the submitting investigators.

3. **How will the NIH GWAS data repository include data from other existing data sources, such as the Cancer Biomedical Informatics Grid (caBIG)?**

The NIH GWAS data repository was designed to accommodate the data structures of many of the GWAS that are likely to submit data to it. Nonetheless, the National Center for Biotechnology Information (NCBI) will discuss the best transfer format with all submitting investigators prior to data transmission. The same plan will be followed for importing relevant data from caBIG. In the future, the NIH GWAS data repository may contain information that will allow users to readily access more specialized resources linked to the data in the repository.

4. **Will biological samples (e.g. tissues, DNA, cell lines) associated with the data included in the NIH GWAS data repository be available through the repository?**

No. Only the genotype and phenotype data will be available through the NIH GWAS data repository.

5. **Will the NIH standardize phenotype information?**

The NIH has no current plans to standardize phenotype information submitted to the NIH GWAS data repository, although the NIH database of Genotypes and Phenotypes (dbGaP) is

modeling and archiving existing studies with their existing designs. Since every element in dbGaP is accessioned and versioned in a public database, standardized phenotypes can be explicitly applied to dbGaP data by any of a variety of common approaches.

6. **How will the NIH ensure that data obtained from the NIH GWAS data repository are used in a manner consistent with the informed consent provided by research participants?**

Through the Controlled Access process for providing data access to secondary users, mechanisms are in place to minimize the likelihood of usage of GWAS data in ways that are inconsistent with the original informed consent. For instance, the NIH expects that an IRB or Privacy Board will have reviewed all proposed submissions of data to the NIH GWAS data repository and determined that the submission and subsequent sharing for research purposes are consistent with the informed consent of the study participants from whom the data were obtained. In addition, submitting institutions are expected to certify that the appropriate research uses of the data and the uses that are specifically excluded by the relevant informed consent documents are delineated. For example, the consent forms for some GWAS studies may have restrictions regarding the disease(s) to be studied. Information from the submitting institution about data use limitations will then be used by the NIH Data Access Committees to inform their review of requests for access to determine whether a proposed use of the dataset conflicts with any informed consent limitations. Researchers approved for access to a GWAS dataset must agree not to use the data for any research other than that described in their Data Access Request.

[Back to Top](#)

F. Data Submission

1. **What is dbGaP?**

dbGaP (Database of Genotypes and Phenotypes) was developed by the National Center for Biotechnology Information (a division of the National Library of Medicine of the NIH) to archive and distribute the results of studies that have investigated the interaction of genotype and phenotype. dbGaP will serve as the NIH GWAS data repository. Additional information on dbGaP can be found at <http://www.ncbi.nlm.nih.gov/sites/entrez?db=gap>

2. **Will there be one GWAS repository or multiple GWAS repositories?**

The NIH will maintain one central GWAS repository. However, it is not intended to be the exclusive source for data from the NIH supported or conducted GWAS. For instance, the same data may also be held by an institutional repository or other sites, depending on the study consents and the needs of the submitting investigators.

3. **Will the NIH standardize phenotype information?**

The NIH has no current plans to standardize phenotype information submitted to the NIH GWAS data repository, although the NIH database of Genotypes and Phenotypes (dbGaP) is modeling and archiving existing studies with their existing designs. Since every element in dbGaP is accessioned and versioned in a public database, standardized phenotypes can be explicitly applied to dbGaP data by any of a variety of common approaches.

4. **What is the IRB's role in submission of data to the NIH GWAS data repository?**

The responsible institutional official of the submitting institution should provide a certification that, among other things, an IRB and/or (as applicable) Privacy Board, reviewed and verified that:

- The submission of data to the NIH GWAS data repository and subsequent sharing for research purposes are consistent with the informed consent of study participants from whom the data were obtained;
- The investigator's plan for de-identifying data is consistent with the Policy standards;
- It has considered the risks to individuals, their families, and groups or populations associated with data submitted to the NIH GWAS data repository; and
- The genotype and phenotype data to be submitted were collected in a manner consistent with 45 C.F.R. Part 46.

5. **How will updates of original data from the submitting investigators be made available to recipient investigators?**

Datasets will have version numbers. Investigators with NIH DAC-approved access privileges will be able to obtain updates to datasets for which they are approved. Investigators whose approved access period has expired may re-apply for access by submitting a renewal Data Access Request.

6. **Can investigators conducting GWAS as part of a clinical trial limit their data submission to baseline data only (i.e. data obtained prior to the start of the intervention)?**

Yes, provided that the NIH-funded GWAS research is limited to the baseline data. Baseline data may be considered a separate dataset and investigators may submit new or updated datasets as they become available.

7. **Will biological samples (e.g. tissues, DNA, cell lines) associated with the data included in the NIH GWAS data repository be available through the repository?**

No. Only the genotype and phenotype data will be available through the NIH GWAS data repository.

8. **Are there cases in which data sharing will not be possible? How should NIH GWAS grant applicants deal with such instances?**

Submission of some GWAS data to the repository may be precluded by various factors, such as international laws, limitations in the original informed consents, concerns about harms to individuals or groups, or other cases where expectations for data submission cannot be met or restrictions on the original funding mechanism. Through the governance structure, ICs will be provided guidance on appropriate exceptions. Applicants who believe that they will be unable to share data as described in the policy will be asked to explain in detail why data sharing is not possible in the data sharing plan portion of the application. ICs make funding decisions based on both scientific excellence and program priorities and may take data sharing plans into consideration. In doing so, applications that represent the highest quality science but do not provide for standard data deposition and distribution through the NIH GWAS data repository will be considered for funding by ICs on a case-by-case basis.

9. **For multi-center studies, is the submitting institution expected to certify data that are contributed by data collection centers at other institutions?**

No. The submitting institution need not certify that the expectations of the Policy are met for data collected by other institutions within its multi-center arrangement. The NIH understands that the submitting institution is not necessarily the local institution or IRB of record for all data collected in a multi-site trial. However, the submitting institution should assure the NIH

through the submission of one certification document that it believes, based on either its own review or assurance from other institutions, that the expectations of the Policy are met for the entire dataset. Further, the submitting institution should explicitly identify within that document any data use limitations that apply to the submitted dataset or subsets of such data collected at all sites. In obtaining assurance from other sites in a multi-site study, the submitting institution should retain copies of any information it receives from other data collection sites.

10. **Will the NIH GWAS repository accept GWAS data if the individuals from whom the data/samples are derived are deceased and either a) information about the consent process or the consent forms is not available or b) consent was silent or not inconsistent with regard to submission to the NIH GWAS Data**

Repository? Yes. If the submitting institution, in concert with its IRB and/or privacy board, finds that submission of the data to the NIH would be appropriate (e.g., the consent form does not preclude data sharing) and meets the other expectations of the Institutional Certification specified within the policy, then such data would be accepted by the NIH GWAS Repository. In considering submission of such data, IRBs and institutions may wish to refer to the criteria defined within 45 CFR 46.116(d) for the issuance of a waiver of informed consent. Although the GWAS database does not currently involve human subjects research under 45 CFR 46, the criteria might provide a useful framework to institutions, IRBs, and privacy boards in considering submission of data from deceased individuals.

11. **As stated in the NIH GWAS policy, the NIH expects that all submissions to the NIH GWAS data repository will include a certification by the responsible Institutional Official(s) of the submitting institution that the expectations of the policy have been met for submission to the NIH GWAS data repository. For multi-center studies, is the submitting institution expected to certify data that are contributed by data collection centers at other institutions?**

No. The submitting institution need not certify that the expectations of the Policy are met for data collected by other institutions within its multi-center arrangement. The NIH understands that the submitting institution is not necessarily the local institution or IRB of record for all data collected in a multi-site trial. However, the submitting institution should assure the NIH through the submission of the certification document that it believes, based on either its own review or assurance from other institutions, that the expectations of the Policy are met for the entire dataset. Further, the submitting institution should explicitly identify within that document any data use limitations that apply to the submitted dataset or subsets of such data collected at all sites. In obtaining assurance from other sites in a multi-site study, the submitting institution should retain copies of any information it receives from other data collecting sites.

12. **Are Submitting Institutions expected to certify that data submission is consistent with applicable laws and regulations in effect at any and all locations at which data were collected?**

No. Submitting Institutions are expected to certify that data submission is consistent with applicable laws and regulation relevant to their specific activities, e.g., home state law, home institutional policies, etc. Submitting Institutions may assume that all prior data transfers from data collection sites to the Submitting Institution (e.g., a data coordinating center) were conducted according to any applicable laws relevant to the those organizations at the time of the original data transfer. The NIH, however, does expect that all data were collected in accord with 45 C.F.R. Part 46. As discussed in a separate FAQ, this assurance in multi-site studies can be made on the basis of a direct review of study materials by the Submitting Institution or based on information or assurance provided to the Submitting Institution by data collecting organizations.

13. **Will NIH accept the submission of GWAS data to dbGaP if the data were derived from de-identified specimens that were not consented for research purposes (e.g., residual clinical samples or cell lines), which is permitted under the [Common Rule](#)?**

Under the NIH GWAS Policy, there is an expectation that the submission of data to dbGaP is consistent with the informed consent originally provided by the research participant. However, this expectation does not currently apply to de-identified clinical samples and cell lines. Data derived from these sources may be submitted to NIH genomic data repositories if the submitting institution determines that the other expectations defined in the [GWAS Policy](#) for the Institutional Certification document are met. Note that in the draft Genomic Data Sharing (GDS) Policy that has been issued for public comment, NIH has proposed a change in its expectations in this regard. The draft GDS Policy states that informed consent for future research use and broad sharing should be obtained for de-identified cell lines or clinical specimens that are generated or collected after the effective date of the final policy.

[Back to Top](#)

G. Governance

1. **What mechanisms are in place to monitor the implementation of the GWAS policy and to ensure that it keeps pace with evolving technological and ethical issues?**

The governance structure for GWAS is described in the preamble section of the policy (<http://grants.nih.gov/grants/guide/notice-files/NOT-OD-07-088.html>). The NIH Director is ultimately responsible for overseeing the GWAS policy and its implementation. The Director will receive guidance and appropriate leadership from within the Office of the Director. A fundamental aspect of the GWAS governance plan is a working group of the Advisory Committee to the NIH Director (ACD), which is composed of external experts representing the interests of the public and the scientific community. The ACD Working Group will be responsible for informing the ACD regarding any emerging ethical or scientific issues that may be relevant to NIH efforts to maintain the highest standards in participant protection and the promotion of quality research. The ACD will then advise the NIH Director on these issues.

The NIH Data Access Committees will monitor data use practices by reviewing annual reports that it will receive from all approved users. Summary reports on data distribution, DAC processes, and GWAS data use practices will also be provided to the ACD Working Group on a regular basis.

2. **How will the NIH ensure that data obtained from the NIH GWAS data repository are used in a manner consistent with the informed consent provided by research participants?**

Through the Controlled Access process for providing data access to secondary users, mechanisms are in place to minimize the likelihood of usage of GWAS data in ways that are inconsistent with the original informed consent. For instance, the NIH expects that an IRB or Privacy Board will have reviewed all proposed submissions of data to the NIH GWAS data repository and determined that the submission and subsequent sharing for research purposes are consistent with the informed consent of the study participants from whom the data were obtained. In addition, submitting institutions are expected to certify that the

appropriate research uses of the data and the uses that are specifically excluded by the relevant informed consent documents are delineated. For example, the consent forms for some GWAS studies may have restrictions regarding the disease(s) to be studied. Information from the submitting institution about data use limitations will then be used by the NIH Data Access Committees to inform their review of requests for access to determine whether a proposed use of the dataset conflicts with any informed consent limitations. Researchers approved for access to a GWAS dataset must agree not to use the data for any research other than that described in their Data Access Request.

[Back to Top](#)

H. Publication

1. **How will the work of submitters of GWAS data be acknowledged by secondary users?**

The NIH expects that all secondary users of GWAS datasets will acknowledge the contributing investigators who conducted the original studies, the funding organization(s) that supported the work, and the GWAS data repository in all resulting oral or written presentations, disclosures, or publications of the analyses. This policy will be agreed to by all approved data users through their agreement to abide by the terms of use described within the Data Use Certification for each NIH GWAS dataset.

[Back to Top](#)

I. Research Participant Protections

1. **How will the NIH ensure that data obtained from the NIH GWAS data repository are used in a manner consistent with the informed consent provided by research participants?**

Through the Controlled Access process for providing data access to secondary users, mechanisms are in place to minimize the likelihood of usage of GWAS data in ways that are inconsistent with the original informed consent. For instance, the NIH expects that an IRB or Privacy Board will have reviewed all proposed submissions of data to the NIH GWAS data repository and determined that the submission and subsequent sharing for research purposes are consistent with the informed consent of the study participants from whom the data were obtained. In addition, submitting institutions are expected to certify that the appropriate research uses of the data and the uses that are specifically excluded by the relevant informed consent documents are delineated. For example, the consent forms for some GWAS studies may have restrictions regarding the disease(s) to be studied. Information from the submitting institution about data use limitations will then be used by the NIH Data Access Committees to inform their review of requests for access to determine whether a proposed use of the dataset conflicts with any informed consent limitations. Researchers approved for access to a GWAS dataset must agree not to use the data for any research other than that described in their Data Access Request.

2. **What is the IRB's role in submission of data to the NIH GWAS data repository?**

The responsible institutional official of the submitting institution should provide a certification that, among other things, an IRB and/or (as applicable) Privacy Board, reviewed and verified that:

 - The submission of data to the NIH GWAS data repository and subsequent sharing for research purposes are consistent with the informed consent of study participants from whom the data were obtained;
 - The investigator's plan for de-identifying data is consistent with the Policy standards;
 - It has considered the risks to individuals, their families, and groups or populations associated with data submitted to the NIH GWAS data repository; and
 - The genotype and phenotype data to be submitted were collected in a manner consistent with 45 C.F.R. Part 46.
3. **Are the data included in the NIH GWAS data repository subject to the Freedom of Information Act (FOIA)?**

As an agency of the Federal government, the NIH is required to release government records in response to requests under the federal Freedom of Information Act (FOIA), unless the records are exempt from release under one of the FOIA exemptions. The NIH believes that release of unredacted GWAS datasets in response to a FOIA request would constitute an unreasonable invasion of personal privacy under FOIA Exemption 6, 5 U.S.C. 552 (b)(6). Therefore, among the safeguards that the NIH foresees using to preserve the privacy of research participants and the confidentiality of their genomic data in the NIH GWAS data repository is the redaction of individual-level genotype and phenotype data from any disclosures made in response to FOIA requests and the denial of requests for unredacted datasets. It is important to note, however, that FOIA affords requesters an opportunity to contest an agencies determination.
4. **Can non-research entities (e.g., law enforcement agencies, insurance companies, employers) request access to identifiable information corresponding to phenotype and genotype data held in the NIH GWAS data repository?**

The NIH will not possess any direct identifiers within the NIH GWAS data repository, nor will the NIH have access to the link between the data keycode and the identifiable information that may reside with the primary investigators and institutions for particular studies. The NIH explicitly encourages investigators submitting GWAS to consider the potential appropriateness of obtaining a Certificate of Confidentiality (please see <http://grants.nih.gov/grants/policy/coc/>) as an added protection against any future compelled disclosure of identities for studies planning to collect genome-wide association data.
5. **What is the process for obtaining certificates of confidentiality?**

The necessary information regarding the process and the appropriate application of this protection may be found at the NIH Certificates of Confidentiality Kiosk located at <http://grants.nih.gov/grants/policy/coc/>.
6. **Can investigators seek additional resources to contact previous participants of studies to obtain additional informed consent?**

The NIH may give programmatic consideration to requests for supplemental funding to obtain additional participant consent when appropriate. Requests for such support should be directed to the appropriate IC Program Officer.
7. **What is the role of the NIH DACs in considering risks to individuals, their families, and groups or populations associated with data submitted to the NIH GWAS data**

repository?

The NIH GWAS policy states that all submissions to the NIH GWAS data repository should be accompanied by a certification from the responsible Institutional Official(s) of the submitting institution that clearly delineates any Data Use Limitations, including limitations to prevent uses that might increase the risks to individuals, their families, groups or populations, if any (see above FAQ). The NIH DACs use the Data Use Limitation(s) provided in the Institutional Certification by the submitting institution to assess proposed secondary uses of the data. Through this assessment, the DACs determine whether or not a Data Access Request conforms to the uses allowed by the certification. In addition to ensuring that secondary uses conform to the Data Use Limitation(s), DACs may consider potential harm (e.g., stigmatization) to groups or populations in their deliberations of Data Access Requests. The DACs may investigate requests where the proposed research use appears inconsistent with the stated limitations, the research intent is unclear, or there is concern about potential harm to groups or populations. DACs may seek additional information from data requesters or consult with relevant technical or ethical experts to inform DAC deliberations as needed.

8. **What is the role of the IRB/Privacy Board of the submitting institution in considering risks to groups or populations associated with data submitted to the NIH GWAS data repository?**

According to the NIH GWAS policy, data submissions to the NIH GWAS repository should be accompanied by a certification from the responsible Institutional Official(s) of the submitting institution that assures, among other things, that an IRB and/or Privacy Board, as applicable, reviewed the proposed submission and verified that the risks to individuals, their families, and groups or populations associated with data had been considered. When necessary, the certification can include Data Use Limitations. The NIH included this expectation in the policy to be sure that an IRB/Privacy Board, based on information known to it at the time of the review, considered the potential for subsequent studies to pose harms to groups and/or populations. The NIH acknowledges that IRBs/Privacy Boards cannot know at the time of review what secondary uses of the data may be proposed, their delineation of the limitations on data use provides guidance for the NIH DACs to help minimize the risks associated with data sharing. The IRB/Privacy Board should express any concerns about potential uses in the Data Use Limitations by indicating secondary uses of the data that would not be appropriate because they could increase the risk to individuals, their families, and groups or populations. For example, the nature or sensitivity of some data included within a study may merit special consideration because it contains potentially stigmatizing genetic, phenotypic, behavioral or social traits. The NIH recognizes that assessing the potential for harms to groups or populations is challenging, especially given the breadth of research questions that could be pursued with the data, and that the best analytical approach to assess potential harms is the subject of considerable ethical study and debate. In addition, certain populations may have unique cultural perspectives or special expectations related to their participation in research (e.g., the need for the consent of their community). By considering these issues when providing Data Use Limitations to the NIH, IRBs/Privacy Boards can inform Data Access Committees about issues relevant to minimizing group harms, as appropriate.

[Back to Top](#)

J. IRB Involvement

1. **As stated in the NIH GWAS policy, the NIH expects an IRB and/or Privacy Board to review and verify that the submission of data to the NIH GWAS Data Repository and subsequent sharing for research purposes are consistent with the informed consent of study participants from whom the data were obtained. Does this mean that the NIH expects specific elements to be included in the informed consent documents? For example, are the elements discussed in the Points to Consider for IRBs and Institutions document (p.11-13) intended to be an explicit consent standard that IRBs should use in assessing consent forms?**

No. The NIH is not a regulatory agency and does not provide explicit guidelines or criteria that IRB should follow in reviewing consent forms prior to data submission. The Points to Consider for IRBs and Institutions is provided for educational purposes only and should not be considered a statement of NIH policy. The examples included within the document do not represent an explicit consent standard or checklist for IRBs/ Privacy Boards. Rather, the list includes a range of issues that IRBs may want to take into consideration when reviewing informed consent documents for the purposes of submission of data to an NIH repository, and that investigators may wish to refer to when preparing consent documents for prospective studies. The examples included in the Points to Consider document are not intended to be proscriptive, nor do they represent an exhaustive list of all of the issues that may be appropriate for IRBs to consider in specific scenarios. Each research project and consent document is unique and local IRBs are in the best position to evaluate the potential benefits and risks of submission of genetic data to an NIH data repository. Therefore, local IRBs should determine the significance, as it relates to data submission, of the presence or absence of particular elements within the informed consent documents.

2. **The NIH GWAS policy indicates that monies may be available for re-consent of participants. Does the NIH have criteria or guidelines regarding when an IRB should pursue re-consent of participants?**

No. The NIH acknowledges that the local IRB will have the greatest familiarity with specific studies and the associated consent documents. Therefore, the NIH does not provide guidelines or criteria that IRBs must follow regarding when to seek explicit consent for data submission to the NIH GWAS repository and subsequent sharing. The local IRB should decide if re-consent of participants is warranted. During the public comment period for the NIH GWAS policy, some respondents noted that additional resources would likely be required in cases where IRBs determined that re-consent was needed for data sharing. In response to those comments, the NIH decided to provide programmatic consideration, on a case-by-case basis, to requests from investigators for funding to support efforts to seek re-consent from participants. The statements in the NIH GWAS policy are intended only to inform investigators, institutions, and IRBs that funding might be available, at the discretion of the NIH Institute funding the GWAS study, for re-consent of participants.

3. **Will the NIH GWAS repository accept GWAS data if the individuals from whom the data/samples are derived are deceased and either a) information about the consent process or the consent forms is not available or b) consent was silent or not inconsistent with regard to submission to the NIH GWAS Data Repository?**

Yes. If the submitting institution, in concert with its IRB and/or privacy board, finds that submission of the data to the NIH would be appropriate (e.g., the consent form does not preclude data sharing) and meets the other expectations of the Institutional Certification specified within the policy, then such data would be accepted by the NIH GWAS Repository. In considering submission of such data, IRBs and institutions may wish to refer to the criteria defined within 45 CFR 46.116(d) for the issuance of a waiver of informed consent. Although the GWAS database does not currently involve human subjects

research under 45 CFR 46, the criteria might provide a useful framework to institutions, IRBs, and privacy boards in considering submission of data from deceased individuals.

4. **As stated in the NIH GWAS policy, the NIH expects that all submissions to the NIH GWAS data repository will include a certification by the responsible Institutional Official(s) of the submitting institution that the expectations of the policy have been met for submission to the NIH GWAS data repository. For multi-center studies, is the submitting institution expected to certify data that are contributed by data collection centers at other institutions?**

No. The submitting institution need not certify that the expectations of the Policy are met for data collected by other institutions within its multi-center arrangement. The NIH understands that the submitting institution is not necessarily the local institution or IRB of record for all data collected in a multi-site trial. However, the submitting institution should assure the NIH through the submission of the certification document that it believes, based on either its own review or assurance from other institutions, that the expectations of the Policy are met for the entire dataset. Further, the submitting institution should explicitly identify within that document any data use limitations that apply to the submitted dataset or subsets of such data collected at all sites. In obtaining assurance from other sites in a multi-site study, the submitting institution should retain copies of any information it receives from other data collecting sites.

5. **Can IRBs obtain a waiver of consent for submission of data to the NIH GWAS repository?**

The issuance of a waiver of informed consent, as described in 45 CFR 46.116(d), would not be appropriate for considerations related to certification of data for submission to the NIH GWAS data repository because de-identified GWAS data submitted to dbGaP do not fit the regulatory definition of human subjects data and therefore do not fall under the regulations described in 45 CFR 46. However, the NIH recognizes that the criteria for issuance of a waiver of informed consent might provide a useful framework to consider for institutions and IRBs/privacy boards during their deliberations regarding certifying submission of GWAS data (e.g., in cases where existing consent forms may be silent or not inconsistent with regard to submission to the NIH GWAS Data Repository). Institutions and their IRBs/Privacy Boards may refer to the criteria during their discussions of GWAS proposals to assist them in working through questions related to the appropriateness of data submission, although the actual issuance of a waiver is not applicable.