*Updated 11/12/2007*

# Implementation Guidance for NIH Extramural Staff

*Policy for Sharing of Data Obtained in NIH Supported or*

*Conducted Genome-Wide Association Studies (GWAS)*

**Version 1.0**

**November 2007**

# TABLE OF CONTENTS

## I.    INTRODUCTION

As part of its mission to discover new knowledge that will lead to better health for the public, the National Institutes of Health (NIH) is interested in advancing the use of data obtained through genome-wide association studies (GWAS) to identify common genetic factors that influence health and disease.  After public consultation (http://grants.nih.gov/grants/gwas/index.htm), the NIH announced the final policy for the sharing of data obtained through NIH supported or conducted GWAS (http://grants.nih.gov/grants/guide/notice-files/NOT-OD-07-088.html).

The purpose of the NIH GWAS Policy is to facilitate broad and consistent access to NIH supported GWAS data in order to speed the translation of basic genetic research into therapies, products and procedures that benefit the public health.  The NIH believes that the full value of GWAS to the public can be realized only if the resulting genotype and phenotype datasets are made available as rapidly as possible to a wide range of scientific investigators.  Rapid and broad data access is particularly important for GWAS; these studies generally require significant resources, present challenges in analyzing the large datasets, and provide extraordinary opportunities for making comparisons across multiple studies.

Many investigators may have already read about GWAS in the September 2007 *NIH Extramural Nexus* (http://grants.nih.gov/grants/partners/0907Nexus.htm#GWAS) and may have additional questions about the policy.  NIH staff can advise other investigators and provide further information by directing them to the *NIH Guide* Notice (http://grants.nih.gov/grants/guide/notice-files/NOT-OD-07-088.html) or the other informational web sites contained in the resource section of this document.  Brochures and a PowerPoint slide set also will also be available for NIH staff to bring to scientific meetings and conferences as appropriate. This document provides implementation guidance for NIH extramural staff who are involved in GWAS activities. Additional guidance will be release specifically addressing the requirements for contracts supporting GWAS and the GWAS governance process.
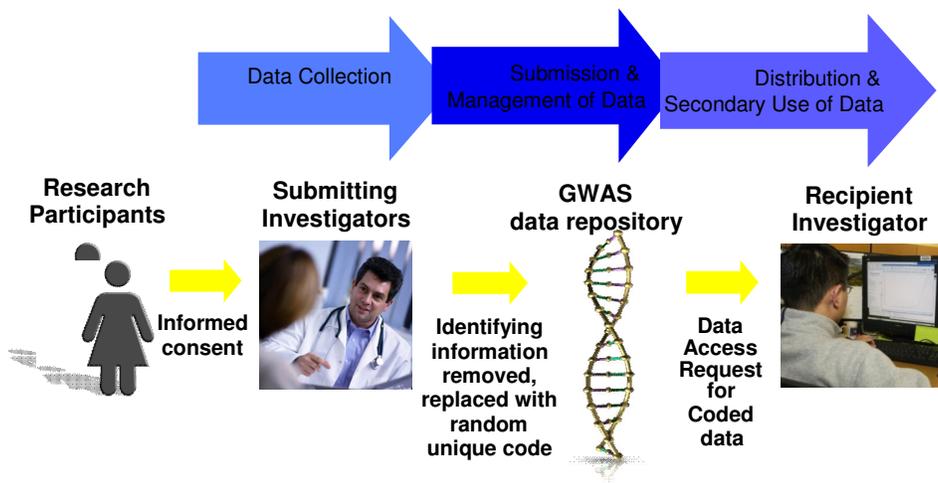
## II. APPLICABILITY

### Funding Mechanisms and Research Purpose

This policy has no direct cost threshold associated with it so all eligible grants and contracts proposing GWAS under any of the following funding mechanisms are expected to adhere to the policy. The NIH GWAS Policy applies only to mechanisms requesting support for research.

The GWAS Policy **does** apply to:

- Research project grants (Rs);
- Program projects (Ps) and SCORs (Ss);
- Cooperative research mechanisms (Us);
- Individual career development awards (Ks) that include a research component;
- S mechanisms that include a research component; and
- All other mechanisms that include a research component.

The GWAS Policy **does not** apply to:

- Institutional training grants (T32s, T34s, T35s and TL2s);
- K12 career-development awards (KL2s),
- Individual fellowships (Fs);
- Resource grants and contracts (Ss); or
- Linked awards derived from previously reviewed applications (KL1, KL2, RL1, RL2, RL5, RL9, TL1, UL1).
- Facilities or coordinating centers funded through related initiatives to provide genotyping or other core services in support of GWAS; or
- Linkage studies.

### Receipt Dates

This NIH policy applies to:

- Competing grant applications (type 1s, 2s and 3s, including resubmission [amended] applications) that include GWAS and are submitted to the NIH for the January 25, 2008 and subsequent receipt dates;

- Proposals for contracts that include GWAS and are submitted to the NIH for a receipt date on or after January 25, 2008; and

- NIH intramural research projects that include GWAS and are approved on or after January 25, 2008.

## III. DEVELOPMENT OF GRANT APPLICATION OR CONTRACT PROPOSAL

### Cover Letter

Applicants should state in the cover letter if they are proposing to conduct GWAS research, or plan to access GWAS data in the NIH GWAS data repository as part of the proposed research. Note that applications to access GWAS data are not considered GWAS research and do not receive the "GW" code (see below).

### Data Sharing Plan

Investigators proposing GWAS or a component that includes GWAS are expected to include a data sharing plan or an appropriate explanation for why submission to the NIH GWAS data repository will not be possible. This information should appear in the following location of the applications:

- In the PHS 398 application, **Section K. Resources of the Research Plan**; or
- In the SF424 (R&R) PDF attachment of the **PHS 398 Research Plan Component Item 17**.

The NIH expects that data derived from NIH-sponsored GWAS will be deposited in the NIH GWAS data repository (i.e., dbGaP) to be made available for appropriate research uses by the research community. In situations where sharing is not possible (e.g., due to informed consent issues, local laws and limitations, concerns about harms to individuals or groups, or other cases where the expectations for data submission cannot be met), applicants should provide a justification for why they cannot share their particular dataset (see application basics below). Not sharing is considered an exception and will be considered on a case-by-case basis.

### Research Participant Protections

As with any application involving human subjects, human subject issues must be addressed in the Human Subjects section of the grant application or proposal. Investigators should be advised that at the time of data submission to the NIH GWAS data repository an Institutional Certification that meets the specific components delineated in the GWAS Policy will be required (see Post-Award Administration below). Investigators should be advised that a Certificate of Confidentiality is often prudent to help protect participants' privacy.

### Budget

Investigators may delineate and justify any costs associated with preparing the dataset for submission in the budget section of the application. For example, costs associated with obtaining consent for data sharing, including obtaining additional consent for research involving existing data/specimens, may be included in the Budget.

### Secondary Access to GWAS Data

Investigators who propose to access individual-level GWAS data through the NIH GWAS data repository (i.e., the investigator is performing secondary data analyses), will be expected to meet data security measures and to submit an appropriate data access request, including a Data Use Certification, that is co-signed by the investigator and the designated Institutional Official. Investigators will be expected to submit annual progress reports detailing significant research findings to the Data Access Committee (DAC). Secondary users will not be expected to deposit their findings into the NIH GWAS data repository.

Investigators who plan to seek data from the NIH GWAS data repository for use in research will be expected to meet data security measures and to submit a data access request, including a Data Use Certification, that is co-signed by the investigator and the designated Institutional Official.  In the Research Design and Methods section of the Research Plan, the applicants should briefly address plans for requesting access to data and state their intention to abide by the Data Use Certification.  Applicants who intend to access data in the NIH GWAS data repository for the research proposed may wish to secure that access prior to submitting their application for NIH support.  The Design and Methods section of the Research Plan is Section D in the PHS 398 application, and Item 5 in the PDF attachment of the PHS 398 Research Plan Component.

Investigators should be referred to the NIH GWAS data Repository (dbGaP) website (http://www.ncbi.nlm.nih.gov/sites/entrez?db=gap).  Open access to basic summary information is available through the public site.  Investigators seeking access to controlled datasets should be referred to the appropriate Data Access Committee (DAC) for specific procedures to access such data.

## IV. PRE-REVIEW

### Identifying Application as GWAS

#### *Investigators*

Applicants should state in the cover letter if they are proposing to conduct GWAS research, or plan to access GWAS data in the NIH GWAS data repository as part of the proposed research.

#### *NIH Staff*

The NIH has established the "GW" code in IMPAC to identify and track those funding agreements (grants, cooperative agreements, and contracts as applicable) that contain GWAS or a GWAS component.  NIH Staff of the CSR Division of Receipt and Referral will add this code as a dual assignment when they detect GWAS.  However, Program Officials (POs) are ultimately responsible for ensuring proper assignment of the "GW" code.  For a multicomponent application, if any component included GWAS, the entire application should have the "GW" dual code.  Investigators may have indicated in their submission cover letter, as outlined above, that their application contains GWAS.

Investigators proposing to access individual-level GWAS data from the NIH GWAS data repository (i.e., the investigator is a secondary user), should not be coded "GW."

##### REMOVING THE "GW" CODE

**The "GW" code should be removed if:**

- A study is limited to a certain portion of the genome;
- A study is following up on a specific association identified from a previously-conducted genome-wide screen;

- The GWAS does not involve the human genome;
- The application proposes secondary data use only; or
- The policy is not otherwise applicable based on the criteria defined above.

POs should contact investigators who inappropriately indicate that their application is for support of GWAS for clarification.

### ADDING THE "GW" CODE

**The "GW" code should be added if**:

- An application proposes any study of genetic variation across the entire human genome that is designed to identify genetic associations with observable traits (such as blood pressure or weight), or the presence or absence of a disease or condition.[1]

If this is done before review, both the investigator and the Scientific Review Officer (SRO) of the relevant review committee should be informed.

POs should contact investigators who do not designate their application as GWAS to ensure their understanding of the GWAS Policy. If a "GW" code is removed or added, the PO should advise the SRO responsible for the review of the application. The 901 form should be used to delete or add "GW" codes. POs should inform the Division of Receipt and Referral of CSR if their Funding Opportunity Announcements (FOAs) (especially RFAs or PARs) will require all applications to include GWAS, so that the Division of Receipt and Referral can ensure that all responding applications will be coded "GW."

## V. PEER REVIEW

### Data Sharing Plans

GWAS applications submitted for the January 25, 2008 due date, and all subsequent receipt dates, will be expected to contain a data sharing plan consistent with the GWAS Policy and the implementation guidance released in Fall 2007 (or provide reasons why data sharing is not possible).

GWAS applications submitted for the January 25, 2008 due date and thereafter that lack a GWAS data sharing plan (or reasons why sharing is not possible) will be assigned for review. SROs should mention the GWAS Policy, as appropriate, along with other policies in a general reminder that is sent to all applicants for a given review meeting. If an applicant realizes after submission that a GWAS data sharing plan should be added to the GWAS application, and submits one, the SRO should accept the plan as updated information, make it available to the reviewers, and make sure that it becomes a part of the official grant file. If no GWAS data sharing plan is submitted before review, the absence of the Plan should be recorded in an Administrative Note.

Peer reviewers will be instructed to evaluate the reasonableness of GWAS data sharing plans, or explanations of why submission to the NIH GWAS data repository is not possible. SROs will provide the reviewers' comments in

---

[1]To meet the definition of a GWAS, the density of genetic markers and the extent of linkage disequilibrium should be sufficient to capture (by the $r^2$ parameter) a large proportion of the common variation in the genome of the population under study, and the number of samples (in a case-control or trio design) should provide sufficient power to detect variants of modest effect.

the Administrative Notes section of the summary statement.  Unless specified in a FOA, the adequacy of data sharing plans or of the investigator's explanation of why submission is not possible will *not* be addressed in the review criteria and will *not* affect merit ratings.

Certifications will be required from institutions before they submit[2] datasets to the NIH GWAS data repository or request data from it.  Although Institutional Certifications are not required at the time of submission of a grant application, they are acceptable as Appendix items.

## Human Subjects

As with all applications involving human subjects research, Scientific Review Groups will evaluate the plans for protection of human subjects from research risks (including any risks associated with submission to the NIH GWAS data repository and subsequent sharing) and plans for inclusion of minorities, women and children in the proposed research.  IRBs should be cognizant at the time they review an application of investigator plans for submission of data to the NIH GWAS data repository since it will affect informed consent documents and institutional certification will be required at the time of data submission.

# VI. POST-REVIEW ADMINISTRATION

## Review of Data Sharing Plans

NIH staff must consider the adequacy of a sharing plan prior to making an award, consistent with programmatic goals.  Assuming an application is identified as NIH GWAS with the "GW" code and is being considered for funding, Program staff will need to review the sharing plan for consistency with the policy.  Some considerations include:

- Individual reviewer's remarks on the data sharing plan in the summary statement;

- The types of data to be collected and proposed for deposit in the NIH GWAS data repository;

- Other information that will be sent with the data—e.g., codebooks, study manuals, protocols, survey instruments;

- The adequacy of submitted information to make the data useful for secondary users;

- Timeline for submission of data to the NIH GWAS data repository;

- Plans to obtain a Certificate of Confidentiality (may be included in **Section E, Human Subjects** section and only referenced in the sharing plan).  Investigators should be encouraged to obtain Certificates of Confidentiality, as appropriate, to enhance privacy protection of participants;

---

[2] The certification should assure that:
- The data submission is consistent with all applicable laws and regulations [5], as well as institutional policies;
- -The appropriate research uses of the data and the uses that are specifically excluded by the informed consent documents are delineated;
- The identities of research participants will not be disclosed to the NIH GWAS data repository; and
- An IRB and/or Privacy Board, as applicable, 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 datasets is consistent with the standards outlined above;
- 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.

- Plans to obtain an Institutional Certification;

- Security measures and access procedures if the investigator is proposing submission to other repositories in addition to the NIH GWAS data repository;

- Plans to use a repository exclusive of the NIH GWAS data repository, as well as the justification for excluding the NIH GWAS data repository;

- Participant protection issues (These issues may be addressed in **Section E, Human Subjects** and only referenced in the sharing plan) and;

- If appropriate, justification for not sharing data with the NIH GWAS data repository.

  *Circumstances beyond an investigator's control may preclude sharing (e.g., due to informed consent issues, local laws and limitations, concerns about harms to individuals or groups, or other cases where the expectations for data submission cannot be met). These issues should be balanced against the initial investment and overall future value of the study when an IC decides to make an award. Additional guidance on appropriate principles to consider in reviewing requests for exceptions to the NIH GWAS Policy will be developed through the trans-NIH GWAS governance process and disseminated to NIH staff.*

### Plans that are Inadequate

If a proposed data sharing plan is inadequate or otherwise does not address meeting programmatic goals, POs should contact the investigator and negotiate an acceptable plan, if possible. Program staff also should alert appropriate Grants Management Officer (GMO) if a revised plan is expected.

Investigators who submit a revised plan must have the plan countersigned by the business official and sent to the GMO who will incorporate it into the Notice of Award (NOA).

If an award absolutely must be made and an acceptable plan has not been received (e.g., end of the fiscal year), a special term may be used for a restricted award (see below).

If it is not possible to negotiate an acceptable GWAS sharing plan, the IC may decide to not fund the application. POs should consult with senior IC officials to discuss such situations.

## GWAS Program Considerations

POs should consider the following issues and address them, as appropriate:

- Does the grantee's plan for sharing GWAS data include data sharing through the NIH GWAS data repository or is there an acceptable explanation why such sharing is not possible?
- Were any administrative notes in the summary statement related to the GWAS data sharing plan adequately addressed?
- Were any participant protection issues related to the NIH GWAS Policy identified by program staff adequately addressed?
- Is a statement that the data will be shared will shared consistent with the policy, referencing the application or other appropriate documents for the specific plan required in the NOA?

## Budgets

Costs associated with preparing a GWAS dataset for submission to the NIH GWAS data repository should be delineated and justified in the Budget section of an application. For research involving existing data/specimen

resources, costs associated with obtaining additional consent for data sharing may be included as part of the Budget section, if the institution has determined that it is necessary for data sharing through the NIH GWAS data repository.  As with other applications, the reasonableness of the proposed budget will be evaluated in relation to the proposed work.

## VII. AWARD NEGOTIATION AND ISSUANCE

NIH GMOs are responsible for assuring that all grant awards conform to statutory authority, regulations, policy directives, and administrative guidelines and include appropriate terms and conditions of award.

### *Prior to Award*

Prior to the award, GMOs should receive assurance from POs that data sharing plans in GWAS applications are documented and approved by the PO, or that the investigator's explanation of why data sharing is not possible is acceptable to the PO.

The universal grants management checklist will be modified prior to September, 2008 when NIH GWAS awards are expected to be made, to include an appropriate question regarding GWAS data sharing plans.

### *Notice of Award - Special Terms and Conditions*

When a data sharing plan is mutually agreed upon for a GWAS application and accepted by the PO, the GMO may include a special term and condition.  A suggested term is:

**SPECIAL TERM AND CONDITION:  RESOURCE SHARING**

Dissemination of study data will be in accord with the Grantee's accepted GWAS plan as stated in the [letter/e-mail dated _____, and/or page(s) _____ of the application].  Failure to adhere to the sharing plan as mutually agreed upon by the Grantee and the NIH/IC may result in Enforcement Actions as described in the NIH Grants Policy Statement under Administrative Requirements.

The following suggested term may be used for a restricted award, which should occur only at the end of the Fiscal Year and in rare and exceptional circumstances.  It is best to have an acceptable Resource Sharing Plan in place prior to award.

**SPECIAL TERM AND CONDITION:  RESTRICTED AWARD**

This award is being issued with full restriction, pending IC acceptance of the grantee's GWAS data sharing plan.  No funds may be drawn down from the payment management system and no obligations may be made against Federal funds, until such time that the recipient has received official acceptance by the NIH GMO via a revised NOA indicating acceptance of the plan and removing this restriction.  The required plan must be submitted within 60 days of the issuance of this award.  Failure to comply with this requirement may result in Enforcement Actions as described in the NIH Grants Policy Statement under Administrative Requirements.

*Updated 11/12/2007*

## VIII. POST-AWARD ADMINISTRATION

### Annual Progress Reports

Noncompeting Continuation Applications (type 5s) should include descriptions of study progress and detail significant research findings.  If the competing award included a plan to share data with NIH GWAS data repository, the grantee should describe progress in implementing that plan.   GMOs should ensure that the PO has reviewed and documented whether the progress is consistent with the approved data sharing plan.  If comments on the sharing plan are missing from a progress report, the PO should contact the investigator to elicit a progress report addendum addressing progress in implementing the plan, and any other concerns of the PO.  Addenda should be signed by the institution's business office and submitted to the GMO, who will ensure placement in the NOA.  The grantee may be reminded that failure to include a description of the mutually agreed upon NIH GWAS sharing activities in the progress report constitute failure to provide a complete progress report.  Incomplete progress reports may result in delay of continued funding.

In pre-and post-award review of the application/continuation, GMOs must document in the official grant file that the Institute or Center (IC) PO has reviewed and concurred with the investigator's plans for submitting data to the repository, or has accepted the investigator's explanation why data sharing is not possible.

For competing awards, options similar to those described above are available when there is a concern that an investigator is not adhering to their sharing plan, particularly if the IC issued the award with GWAS sharing term in the NOA.

A final statement on submitting data to the repository should be included in the final progress report, or earlier, if the plan is implemented prior to close-out.

### Data Submission

#### *Coordinate investigators and NCBI*

Administrators at the National Center for Biotechnology Information (NCBI) I need some time to design data bins that are specific to a project.  To do so, they will need to understand the phenotype and assessment tools that will be used.  POs should put the investigator in contact with the NCBI as soon after award to ensure a useful submission to the NIH GWAS data repository.

#### *Advise investigators who are ready to submit their data to the NIH GWAS data repository*

Data can be shared incrementally if there are several components in the discovery process.  For example, baseline clinical data and the genomic data can be shared before experimental clinical experiments are completed, as long as the integrity of the clinical trial is not compromised.

When an investigator is ready to submit data to the NIH GWAS data repository, the PO should request and review the Institutional Certification.  The certification should assure that:

- The data submission is consistent with all applicable laws and regulations, as well as institutional policies;

- The appropriate research uses of the data and the uses that are specifically excluded by the informed consent documents are delineated;

- The identities of research participants will not be disclosed to the NIH GWAS data repository; and

- An IRB and/or (as applicable) Privacy Board, reviewed and verified that:

    o 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;

    o The investigator's plan for de-identifying datasets is consistent with the standards outlined in the NIH GWAS Policy,

    o It has considered the risks to individuals, their families, and groups or populations associated with data submitted to the NIH GWAS data repository; and

    o The genotype and phenotype data to be submitted were collected in a manner consistent with 45 C.F.R. Part 46.

When the PO has established that the appropriate Institutional Certification is in place for data submission, they should contact the NCBI and facilitate submission of the data by the grantee to dbGaP.


## IX. GLOSSARY

**GWAS**  For purposes of the policy, a genetic association study in which the density of genetic markers and the extent of linkage disequilibrium is sufficient to capture a large proportion of the common variation in the human genome in the population under study, and the number of specimens genotyped provides sufficient power to detect variants of modest effect.

**DAC**  Data Access Committee.  Committees established by each funding IC that are charged with reviewing and making decisions about access requests for GWAS data deposited in the NIH GWAS data repository. The DAC will review requests and determine whether access is only for research uses that are consistent with the participants' informed consent and that the recipient agrees to comply with the terms of the Data Use Certification (see below).

**DUC**  Data Use Certification.  This document specifies the terms and conditions under which a requestor will be given access to GWAS data within the NIH GWAS data repository.

**dbGaP**  Database of Genotype and Phenotype is a database developed by the National Center for Biotechnology Information (a division of the National Library of Medicine) to archive and distribute the results of studies that have investigated the association of genotypes and phenotypes. dbGaP currently serves as the NIH GWAS data repository.

*Updated 11/12/2007*

## X. WEB RESOURCES

### *NIH Guide* Notices

- May 15, 2006 (NOT-OD-06-071) - Notice to Applicants for NIH Genome-Wide Association Studies.
- August 30, 2006 (NOT-OD-06-094) - Request for Information (RFI): Proposed Policy for Sharing of Data obtained in NIH supported or conducted Genome-Wide Association Studies (GWAS).
- October 20, 2006 (NOT-OD-07-012) - Extended Comment Period for the Proposed Policy for Sharing of Data Obtained in NIH Supported or Conducted Genome-Wide Association Studies (GWAS).
- October 20, 2006 (NOT-OD-07-013) - NIH Town Hall Meeting on the Proposed Policy for Sharing of Data Obtained in NIH Supported or Conducted Genome-Wide Association Studies (GWAS).
- August 28, 2006 (NOT-OD-07-088) - Policy for Sharing of Data Obtained in NIH Supported or Conducted Genome-Wide Association Studies (GWAS)

### Other Relevant Resources

- Federal Register Notice – GWAS Policy:  http://edocket.access.gpo.gov/2007/pdf/E7-17030.pdf
- NIH GWAS web site: http://grants.nih.gov/grants/gwas/
- dbGaP:  http://www.ncbi.nlm.nih.gov/sites/entrez?db=gap
- NIH Grants Policy Statement:  http://grants.nih.gov/grants/policy/nihgps_2003/index.htm
- NIH Intellectual Property Policy Page:  http://grants.nih.gov/grants/intell-property.htm
- NIH Resource Sharing website:  http://sharing.nih.gov