PARTNERS HEALTHCARE | FOUNDED BY BRIGHAM AND WOMEN'S HOSPITAL AND MASSACHUSETTS GENERAL HOSPITAL

MEMORANDUM TO: Principal Investigators and Research Staff **DATE:** 2/22/15

FROM: Anne Klibanski, MD, Partners Chief Academic Officer (CAO)
Paul Anderson, MD, PhD, BWH CAO
Harry Orf, PhD, MGH Sr. Vice President-Research
Joseph Coyle, MD, McLean Chief Scientific Officer
Ross Zafonte, DO, Spaulding Vice President, Medical Affairs/Research

SUBJECT: NIH Genomic Data Sharing (GDS) Policy Guidance Memo #2:

As we indicated in Guidance Memo # 1, we are writing to provide additional information on the NIH's implementation of the revised GDS policy. Earlier this month the NIH published limited guidance on what should be included in data sharing plans submitted with grant applications or contract proposals, in addition to addressing the policy's applicability to existing projects. New information is highlighted in yellow throughout this document.

This memo supersedes Guidance Memo #1 and reflects the latest information from the NIH.

New Projects: If you submitted an application or proposal for the 1/25/15 or a subsequent deadline with a genomic data sharing plan, you do not have to take any action at this time. If the NIH has questions about your plan, these will be raised at Just-in-Time (JIT) when the hospital will be required to provide "Institutional Certification" to assure 1) the proposal for data submission and sharing is consistent with the NIH policy as well as appropriate laws and regulations, and 2) to address relevant human subjects protection issues in the Human Subjects section of the application. At JIT the hospital will also be required to certify that the IRB (or IACUC) has reviewed the data sharing plan. Remember that this is a change from the previous NIH GDS policy which required certification of institutional compliance when data were uploaded to dpGAP.

Existing Projects: In our previous memo we noted that the new policy was applicable to awards made in response to applications or proposals submitted 1/25/15 or subsequent deadlines, but that the NIH was encouraging PIs of existing awards to develop a transition plan to bring their projects into compliance with the new policy. While technically this remains unchanged, the NIH has taken encouragement one step further in its recently published guidance by instructing PIs to include an updated genomic data sharing plan in the next progress report and transition to the informed consent for broad sharing and future research purposes. We recommend at progress report time you have a discussion with your Program Officer about what would be appropriate for your project.

NIH Genomic Data Sharing (GDS) Policy Guidance Memo #2[1]

Scope & Application/Proposal Requirements

- The policy does not apply to
  - Institutional training grants (T32, T34, T35, and TL2)
  - K12 career development awards
  - Individual fellowships (F)
  - Resource grants and contracts (S)

[1] This document supersedes Guidance Memo #1.

  - Facilities or coordinating centers funded through related initiatives to provide genotyping, sequencing or other core services in support of GDS

- The policy applies to
  - Research Project grants (R)
  - Program projects (P) and SCORs (S)
  - Cooperative agreements for research (U)
  - Individual career development awards (K) that include a research component
  - S activities that include a research component
  - All other activities that include a research component

- The new GDS policy does not set a funding threshold for data sharing; you are required to submit a data sharing plan regardless of the requested funding amount.

- To determine whether your project constitutes a large-scale genomic project, please see the chart below. If you are submitting a contract proposal, you should also consult the Request for Proposal (RFP) for additional guidance in making this determination.

Type of Data	***From***	***From***
Human		
>300,000 variant sites	Genotyping, methylation, RNA	>1000 individuals
DNA Sequence	> 1 gene or similar region	>1000 individuals
DNA Sequence	>100 genes or regions of similar size	>100 individuals
Sequence	> 100 metagenomes or metatranscriptomes	Human microbiome
Animal		
>100,000 SNPs	Genotyping	>1 model organism species or strain
DNA Sequence	Whole exome or whole genome	>1 model organism species or strain
Gene expression	Transcriptome	1 or more model organism species or strain
Sequence	> 100 metagenomes or metatranscriptomes	Model organism microbiome
Microbial		
Sequence	DNA or RNA	>100 isolates of infectious organisms
Cells		
DNA methylation	Comparison of genomewide methylated sites	>10 cell types
Other		
DNA methylation	Comparison of genomewide differential methylation at single-base resolution	Within an individual or across cell types within the same subject

- If your project is large scale, you are required to submit a data sharing plan. Costs associated with data sharing may be included in the application budget.

PARTNERS HEALTHCARE | FOUNDED BY BRIGHAM AND WOMEN'S HOSPITAL AND MASSACHUSETTS GENERAL HOSPITAL

Sample Data Sharing Plans for Large-Scale Human Genomic Studies Applications/Proposals

- *Data collection will occur after 1/25/2015: Additions highlighted in yellow.*
  We are committed to making resources and data from the proposed research available to other investigators in the research community. All data and samples prospectively collected for this research will be obtained with IRB approval and informed consent of study participants to sharing of de-identified large-scale genomic data. We will submit large-scale human *[insert type of genomic data to be shared – see table for guidance]* genomic data as well as relevant associated data (e.g., phenotype and exposure data) to an NIH-designated data repository in a timely manner, as indicated by the NIH GDS policy; typically, up to six months after data submission is initiated or at the time of acceptance of initial publication, whichever occurs first. We will also submit any information necessary to interpret the submitted genomic data, such as study protocols, data instruments and survey tools. The identities of research participants will not be disclosed to the NIH-designated registry. We will take appropriate steps to de-identify datasets in accordance with the NIH GDS policy and as approved by the IRB. Institutional Certification will be submitted and assurance by Partners HealthCare Human Research Committee (IRB) will be obtained that the data can be shared through NIH-designated data repositories, consistent with data sharing under the NIH GDS Policy.

  Insert a paragraph describing any data sharing restrictions: e.g., limited to cardiovascular disease or breast cancer; any restrictions described in the informed consent.

  Aggregate Data *will/will not* be available for submission/general research use.

- *Some data collected prior to 1/25/2015: Additions highlighted in yellow.*
  We are committed to making resources and data from the proposed research available to other investigators in the research community. All data and samples prospectively collected for this research will be obtained with IRB approval and informed consent of study participants to sharing of de-identified large-scale genomic data. We will submit large-scale human genomic data as well as relevant associated data (e.g., phenotype and exposure data) to an NIH-designated data repository in a timely manner, as indicated by the NIH GDS policy; typically, up to six months after data submission is initiated or at the time of acceptance of initial publication, whichever occurs first. We will also submit any information necessary to interpret the submitted genomic data, such as study protocols, data instruments and survey tools. The identities of research participants will not be disclosed to the NIH-designated registry. We will take appropriate steps to de-identify datasets in accordance with the NIH GDS policy and as approved by the IRB.

  Some of the specimens analyzed for data were previously collected (prior to 1/25/2015) using a consent form not consistent with the GDS Policy. We would like to submit these data, as part of this protocol, as it would be impracticable to reconsent these subjects at this time. *[Provide justification for an exception to the data sharing expectation that is pertinent to your project.]* The study will be registered in the database of Genotypes and Phenotypes (dbGAP) and the exception will be publicly explained there.

  Describe any data sharing restrictions: e.g., limited to cardiovascular disease or breast cancer; any restrictions described in the informed consent.

  Aggregate Data *will/will not* be available for submission/general research use.

IRB Review/Informed Consent/Institutional Certification

- *eIRB Application*
  The IRB is adding questions to the eIRB application to help PIs identify, and allow the IRB to capture in a systematic way, those studies that meet the NIH large-scale criteria. We anticipate the additions will appear in early 2015. In the meantime, if you have questions, please contact Megan Morash (mmorash@partners.org) of Human Research Affairs.

- *Informed Consent*
  Human Research Affairs has updated the Tissue Repository Informed Consent templates to meet the NIH GDS requirements. These are posted on the IRB web site in the Research Consents section, Tissue Bank/Repository. The Short Tissue Repository Consent Form is the preferred form. It must be given to subjects with the accompanying Fact Sheet. Please open the links below to find the forms on the Research Navigator.

Short Tissue Repository Consent Form
Tissue Repository Fact Sheet

- *IRB Review*
  Given the short turnaround at JIT and the need to secure IRB approval of the protocol and the data sharing plan, we recommend that you begin the eIRB application process as soon as you receive a score that suggests you will be funded.

- *Institutional Certification at JIT*
  Your Research Management Pre-Award Grants Administrator (GA) will work with you and the IRB to provide Institutional Certification at JIT. The NIH will not issue an award without this certification.

Transition Plans for Human Genomic Data Existing Awards/Annual Research Reporting Requirement

As we indicated earlier in this memo, based on the NIH's latest guidance, you should provide an updated genomic data sharing plan in your next progress report and transition the informed consent for broad sharing and future research purposes. Before submitting your plan, we recommend that you contact Megan Morash (mmorash@partners.org) to review IRB requirements. We are working with Partners Research Management to develop a procedure for securing Institutional Certification for existing awards. The NIH has not provided any indication when or how it expects Institutional Certification. We will of course notify you as soon as we have additional information.

For ongoing studies using an older version of the Tissue Repository consent form template, we recommend you transition to the Short Tissue Repository Form and Fact Sheet. If you are unable to do so, the "long" Tissue Repository Consent Form template is available on the IRB website and meets the NIH GDS requirements. Please provide justification in your IRB amendment when updating why it is not possible to transition to the Short Form.

Long Tissue Repository Consent Form

Guidelines for Non-human Genomic Data Sharing Plans

- Your data sharing plan must include sharing of genomic data, as well as relevant associated data (e.g., phenotype and exposure data.)

- Data are expected to be made publically available no later than the date of initial publication. Earlier availability may be requested by the NIH for certain data.

- Non-human data may be made publically available through any widely used data repository, whether NIH-funded or not, such as the Gene Expression Omnibus (GEO), Sequence Read Archive (SRA), Trace Archive, Array Express, Mouse Genome Informatics (MGI), WormBase, the Zebrafish Model Organism Database (ZFIN), GenBank, European Nucleotide Archive (ENA), or DNA Data Bank of Japan (DDBJ). For more information: http://gds.nih.gov/02dr2.html

- NIH expects investigators to continue to submit data types to the same repositories they used before the effective date of the GDS Policy (e.g., DNA sequence data to GenBank/ENA/DDBJ, expression data to GEO or Array Express). Data types not previously submitted to any repositories may be submitted to these or other widely used repositories as agreed to by the funding IC.

- Data products to be shared
    - Indicate the data products that will be shared (or the amount of data shared, especially for particularly large datasets);
    - Include any analytic tools being provided, such as algorithms, code, or software;
    - What is the format of the final dataset? (e.g., Excel spreadsheets, text records, jpg images, an SQL database. Specify if there are particular tools or software required to read the data.)
    - Optional: What additional documentation will be included to allow others to use the data.

- Data access and policies
    - How will the data be disseminated? Modes of sharing include:
        - Researcher shares data upon request (shared by disk, email?) or makes data accessible through a personal website.
        - Data will be deposited at a data archive. (Name the archive or data center, mention if it is NIH-funded or has data access policies and procedures consistent with NIH data sharing policies.)
        - Data will be accessed through a data enclave (a restricted data center with controlled access, e.g., Hopkins Population Center)
    - Identify conditions for accessing data (e.g., requests for access from identified investigators working at institutions with Federal Wide Assurance) and specify policies for data re-use (e.g., signing a data sharing agreement requesting citation and secure use of data with human subject identifiers.)
    - Identify when the data will be shared (Explain any reasons for delay of sharing beyond the expectations of your community, such as patent restrictions, collaborator requirements, proprietary data from private companies);
    - Provide justification for not sharing data, such as the inability at reasonable cost to remove personal identifiers.