RECORDS RETENTION AND MANAGEMENT OF SCIENTIFIC DATA

Policy Number: 610GS

Effective Date: 6/1/16

Revised Date: 6/26/17; 4/22/19; 11/09/2020

Scope

This Policy on Records Retention and Management of Scientific Data applies to faculty, staff, predoctoral researchers, summer scholars and applicants ("Covered Individuals") of The Graduate School of the Stowers Institute for Medical Research ("The School").

Purpose

The purpose of this policy is to assist members in determining which scientific data should be kept for prescribed time periods and which data should be destroyed. Archiving original scientific data enables a retrospective audit of the validity of data if challenged. It also protects reputational risk and intellectual property. Publicly sharing published data archives also promotes scientific progress when others build upon published data by performing independent analysis and interpretation. The SGC has therefore established this policy and Policy Numbers 207GS (Records Retention and Management) and 605GS (Recording of Scientific Data) to promote economy, efficiency and discipline in the creation, organization, retention and disposal of records. This policy sets forth the informational requirements of the Institute's archive, the processes involved, and Covered Individuals' responsibilities as it relates to scientific data.

The School is included in the Stowers Group of Companies ("SGC") Organizations and participates in the archiving system. As such, the School has adopted the following policy as its own.

Policy

This policy applies to all scientific data generated by SGC Organizations, regardless of the medium or format in which they appear.

Definition

- "Scientific data," as used in this policy, refers to data resulting from the conduct of research activities, including but not limited to quantitative data (e.g., graphs, recorded numbers, instrument output of any type, including photographic materials from which measurements can be made), qualitative data (e.g., notes of any type and some types of instrument output, including photographic media), and research tools (e.g., protocols in any form, computer software). In some cases there may be original samples in (e.g., biological specimens, slides) that should be documented in order to connect them to the data described above.

Records Retention Policy Regarding Scientific Data

- Original Data Repository data as defined in this policy will be retained as long as feasible.

- Physical laboratory notebooks that have been scanned will be considered "scientific data" as covered in this policy.
- Scientific data, samples and research tools have a retention period of 7 years unless governed by the following Scope of Exceptions.

Scope of Exceptions

- Data generated using external funding must follow the records retention policy or guidance of the funding agency.
- Data that is deemed to be irrelevant or not scientifically valuable may be subject to early disposal by the members who created it. Data possibly subject to early disposal may include but is not limited to instrument failure, batch run fault, lack of controls, member separation or otherwise cannot be salvaged for use.

Review of Scientific Data

- The Information Technology department will conduct an periodic crawl of Network Data Shares and populate a list of large data files that have reached the expiry date of the retention period. Laboratories and departments will have an opportunity to determine if these large files should be retained or removed.

Original Data Repository

The Institute will maintain an Original Data Repository (ODR) with entries for Covered Research according to the Informational Requirements below. Covered Individuals must comply with the Workflow specified below in a timely fashion. ODR entries will include privately and publicly accessible information to support both internal retrospective auditing and public sharing. The Institute's Research Integrity Officer has overall responsibility for the ODR and will assign and supervise the roles of ODR Science Curator and ODR Coordinator.

Definitions

- "Original Data" – any data generated by an instrument or a person without any additional processing constitutes original data. In ambiguous cases, the Stowers Research Integrity Officer is responsible for defining Original Data based on the specific situation.
- "Reference to Primary Source" – a description allowing an interested person to trace Original Data to the location where the Original Data was first recorded. For example, the Reference to Primary Source may be a lab notebook page number, a directory path, a LIMS record number or an accession number. A directory path to a secondary recording, e.g. a second copy of a file used for processing, is not considered a Reference to Primary Source.
- "Annotation" – supplemental information that aids in the interpretation of Original Data. Examples include a description of how the Original Data has been processed or analyzed to create an image, table or movie; a description of how figure subpanels map to the Original Data; or instructions on how to open a file.

Informational Requirements
Each scientific publication of Covered Research requires registration in the Stowers' Publication & ODR Review LIMS module (as per Policy 600R, Scientific Publication). An ODR submission must correspond to a single manuscript identified by the accession number assigned to the manuscript in the Stowers' Publication & ODR Review LIMS module.

Required submissions
- *Original Data*. The ODR submission must provide copies of all Original Data underlying the results of a published manuscript.
- *Reference to Primary Source*. Each instance of Original Data must include an associated Reference to Primary Source.
- If data submitted in a public repository complies with the requirements for Original Data as defined in this policy, then provision of an accession number for the public repository may substitute for the Original Data and Reference to Primary Source.

Optional submissions
- *Annotation*. The association between the published result, the Original Data, and the Reference to Primary Source should achieve the standard of ready comprehension by a practicing scientist who is not an expert in the specific field of the manuscript. The Institute's Research Integrity Officer or a designee is responsible for interpreting and applying this standard. Supplemental information in an Annotation may be used to achieve the standard.
- *Private note*. The submission may include a private note to the Research Integrity Officer, the ODR Science Curator or the ODR Coordinator.

Restrictions requested by authors
Authors may request that certain parts of the Original Data be restricted from public access.

Workflow

Prior to Manuscript Submission
- A manuscript's Corresponding Author is responsible for designating a Covered Individual as the manuscript's ODR-Responsible Author. If a manuscript's Corresponding Author is not a Covered Individual then the Corresponding Author's responsibilities outlined in this policy transfer to the most senior author who is a Covered Individual.
- The ODR-Responsible Author for a manuscript must register a manuscript using the Stowers' Publication & ODR Review LIMS module prior to manuscript submission. The registration process will generate an ODR accession number.
- The ODR-Responsible Author for a manuscript must include the following statement in the acknowledgements section of the corresponding manuscript prior to submission:
- *Original data underlying this manuscript can be accessed from the Stowers Original Data Repository at http://www.stowers.org/research/publications/LIBPB-XXXX* (insert the accession number).

- Example: Original data underlying this manuscript can be accessed from the Stowers Original Data Repository at http://www.stowers.org/research/publications/LIBPB-1234.

Prior to Manuscript Acceptance

- After registration and notification through the LIMS system, the ODR Coordinator will set up an ODR working data storage folder.
- The ODR-Responsible Author is encouraged to use the ODR working data storage space to collect and organize copies of Original Data during manuscript preparation and the editorial review process.

Following Manuscript Acceptance

- The ODR-Responsible Author must prepare an ODR submission within 10 working days of manuscript acceptance and verify its accuracy and completeness.
- The ODR Science Curator must review an ODR submission for compliance with the Informational Requirements and provide a written summary to the ODR-Responsible Author and Corresponding Author in a timely fashion. The primary objective of the curator's review is to confirm that each figure component is associated with Original Data.
- The ODR-Responsible Author must address any issues raised in review within 5 working days after the ODR Science Curator's review is complete. In responding to the review, the ODR-Responsible Author may seek further assistance from the ODR Science Curator.
- The ODR Coordinator must publicly release the submission in the ODR following publication of a manuscript. For most publications, the public release will be accomplished within seven (7) working days of the manuscript's publication date.
- The Research Integrity Officer will review all policy compliance exception requests and will grant or deny them individually. At any time in the process, the ODR-Responsible Author or Corresponding Author may request exceptions to the informational requirements or workflow. Such exceptions may include, for example, a request to exclude, or flag as private, data normally required by this policy or additional time if justifiable.

Additional Responsibilities

The ODR Coordinator is responsible for monitoring timely adherence to this workflow, providing guidance and reminders as necessary, and troubleshooting system performance problems (i.e., LIMS issues).

The Research Integrity Officer and designees have the right to audit ODR submissions and any cited data and request revisions at any time.

Data Accessibility

Prior to public release of an ODR submission, access to the content of an ODR submission is restricted to authors/lab members, members of Library Services, the Scientific Director, the Research Integrity Officer, and designees of the Research Integrity Officer.

Public release of an ODR submission discloses the manuscript's Authors, Affiliations, Abstract, Original Data, and Annotations. The Reference to Primary Source and Private Note fields will not be disclosed publicly.

In cases where data is difficult to partition and consists of a mixture of relevant and irrelevant data, the authors may request exceptional treatment of the data. The Research Integrity Officer will decide whether to grant any such requests.

The Institute will provide an archive of these records in its ODR as long as feasible

This policy was approved by the GSSIMR Board of Directors on September 5, 2018.
This policy was last updated by the GSSIMR Board of Directors on November 09, 2020.
This policy will be reviewed by the GSSIMR Board of Directors in 2022.