

Office of the Vice President for Research and Economic Development

Guidelines for data collection, documentation, and storage

May 2014

Introduction

Proper management of research data is essential to the responsible conduct of research. These guidelines are set forth to build awareness of good data management practices and to establish in broad terms the expectations for Principal Investigators and project personnel in collecting, documenting, storing, securing, and retaining research data at UAB. Ultimately the Principal Investigator is responsible for ensuring compliance for the collection, documentation, storage, and security of the data, records, works, and information resulting from his/her research and scholarly activities. The information contained in this document is intended as a useful resource and to establish a principal component of stewardship. As such, this is not an exhaustive list of all various data and works produced at UAB and is not intended to supersede or otherwise modify the requirements of applicable laws, regulations, and UAB policies governing research data management.

Thorough data collection should include a system for the continuous evaluation and documentation of what was done, why it was done, and how it was done. This is important for supporting claims and conclusions presented in research publications. It is also critically important in case data or conclusions presented in publications or funding applications are questioned. Furthermore, systematic and thorough data collection and proper storage can be of critical importance for the documentation of discoveries that may lead to intellectual property filings, patent applications, or other opportunities.

In summary, sound data collection and storage practices ensure that when research is evaluated through publication or by reviewers of funding applications it will be deemed good science meaning, that the research is both precise and honest.

It is the responsibility of the PI to ensure that project personnel know and comply with rules and policies that govern data collection, documentation, storage, and security.

Guidelines for data collection, record keeping, storage, and retention

Data collection and record keeping

When data are collected, records should accurately represent the progress of a project and answer such questions as why, what, and how data were collected. Records should be durable and accessible but at the same time safe from tampering or falsification. For small projects, bound notebooks provide a convenient way for all research team members to keep track of data and daily activities of a project. When keeping written records, errors or changes should be marked, initialed by the investigator, and dated but never erased. This way, they can provide a quick visual account of any changes or errors that have occurred.

However, written records make it difficult to search for a specific fact or to compare observations from several sources. Also, maintaining handwritten records is not possible for larger projects such as clinical trials or epidemiological surveys.

Electronic records allow efficient access and comparison of information from different sources and across similar projects. There are numerous electronic data capture programs that allow researchers to enter, store, and audit research data. Although there are methods for protecting electronic records, the security of such records is a significant concern, (discussed below). In addition, the migration of data records to electronic files may be time consuming and may not be cost effective for large ongoing projects. Therefore, many projects employ a combination of written and electronic records, which represents a balance the risks and benefits.

In addition to record keeping, the validity of the data collected can also be affected by whether or not proper policies and procedures for research are followed on a project and an individual level. Personnel must be constantly aware of all the guidelines that might apply to the project's implementation and dissemination, including special regulations that involve human and animal subjects, hazardous materials, or other controlled biological agents. Every research team member should be aware of project guidelines and standards for collecting valid data, to ensure consistency throughout the project.

Systematic and responsible record keeping is essential to ensuring the integrity of research data. To help maintain data validity and reliability, consider these guidelines when planning or completing data collection:

- **Include notes:** Records should allow you not only to account for what occurred during the course of research but also to reconstruct and justify your findings. It is important that records include notes about what methods did or did not work, observations, and commentary on the project's progress. Keep notes according to the research team's predetermined communications and storage plan.
- **Notebooks:** For smaller projects using handwritten data, each team member should have his or her own personal notebook for recording project data, observations, etc. Entries should be made in a chronological and consistent manner --for instance, each new workday should begin on a new page. Try not to leave blank lines between entries.

Recording information: Record anything that seems relevant to the project, its data, and the standards of the project. At a bare minimum, records should include the following information:

date and time

names and roles of all team members who collected or worked with the data

materials (including lot numbers for reagents), instruments, and software (and version) used

identification number(s) to indicate the subject and/or session

data from the experiment and any pertinent observations concerning the data collection

It may also be helpful to include a summary of the day's data collection activities and a task list for the next day.

Notation of changes: Use a consistent system for noting errors or adjustments. In written records, make entries in indelible pen so that records cannot be altered or damaged. The alterations in the notebook should be initialed by the PI and dated.

Transferring information: When transferring records from written to electronic format, use a double entry system to reduce the possibility of incorrectly entered electronic data. To implement such a system, have two different individuals enter all of raw data into the software program, then cross-check the data to identify and remedy inconsistencies at the time of data entry. Use a worksheet to help track your data collection and entry activities.

Data from gels or blots: It is especially important to have a system in place for documenting and storing data from gels or blots. What type of gel, what and how much sample was loaded into each lane, the buffer, the setting on the power supply, and other detailed information should be clearly documented. The original gels or blots should be given unique identifiers and archived. Images of the original gel or blot should reference this unique identifier. When figures are produced from gels or blots for publication, there should be an entry in the lab notebook or some other central laboratory listing that references the original gel, and links the image to the page or pages in a lab notebook that documents the specific experiment.

Storage

Once data have been collected and recorded, the next concern is data storage. Data storage is crucial to a research project for the following reasons:

- Properly storing data is a way to safeguard your research.
- Data may need to be accessed in the future to explain or augment subsequent research.
- Other researchers might wish to evaluate or use the results of your research.
- Stored data can establish precedence in the event that similar research is published.
- Storing data can protect researchers in the event of misconduct allegations.
- Storing data properly can facilitate IP disclosures or support patent filings.

Type and Amount of Data to Retain

Lab notebooks, original data, and any other documentation of experiments and scholarly activities are the property of the institution and must be retained. Data that requires documentation could include protocols followed, survey instruments, informed consent documents and other information important to the proper and safe conduct of the research. With the consent of the PI, graduate students or postdoctoral trainees may take copies of their data when they leave the laboratory and/or

University, but the original data must be retained. Generally speaking, enough data should be retained so that the findings of a project can be reconstructed with ease. For many projects, this means that all the raw data, relevant statistics and analyses from this data should be saved, along with all notes or observations. Furthermore, if research involves the use of biological specimens, care should be taken to retain them unless they degrade to the point that they are no longer useable.

Electronic Data

The key issues for electronic data storage are thorough documentation to allow data to be appropriately used in the future and storage format that is easily adaptable to evolving computer hardware and software. There are some advantages that are unique to electronic data storage, including:

- Rapid access to the data
- Fast read/write rates
- Low cost
- Ability to archive the data
- Removability
- A backup system, such as storing data

Data and retention and storage

Retention periods for research data/records vary based on whether they are subject to federal regulations, state regulations, contractual agreements, and/or publisher requirements. Where multiple different retention periods govern a single research record/project/program, the PI/PD, co-researcher(s), the author(s), and/or department personnel will maintain research records for the longest period applicable. In some cases, this may be permanent. In all cases, research records should be maintained at least 3 years in accordance with the State of Alabama's RDA. Regardless of retention period, the research data/records must be appropriately maintained/archived to ensure its integrity and must be accessible in the event of a records request or inquiry. Under University of Alabama Board of Trustees Rule 105, Preservation and Ownership of Records and Files, all records created, held, or received by UAB employees or others in the scope of their work for or on behalf of UAB are the property of UAB. In the event any such research records custodian leaves UAB employment, that individual shall reach an appropriate agreement for ongoing retention of the research records with his/her department and, as applicable, the PI/PD. The Council on Governmental Relations publication entitled "Access to, Sharing and Retention of Research Data: Rights & Responsibilities" provides more specific direction. See <http://www.cogr.edu/viewDoc.cfm?DocID=151888>. Of particular note, under DHHS regulation 42 CFR § 93.105(b)(1) (http://ori.dhhs.gov/sites/default/files/42_cfr_parts_50_and_93_2005.pdf), each investigator and/or department personnel shall retain research data/records related to his/her original works, including lab notebooks, as long as he/she cites, uses, or references those original works. At the conclusion of such citation, use, or reference of those original works, research data/records shall then

be retained for the applicable retention period as dictated by federal regulations, state regulations, contractual agreement, and/or publisher requirements. For example, this could mean the State of Alabama's RDA requirement of 3 years to the federal regulation for patents of 20 years or even longer. For reference see UAB's Records Retention Policy (<http://www.uab.edu/policies/content/Pages/UAB-AD-POL-0000708.aspx>) and the attached Records Retention Schedule.

Continued Storage

Once the minimum storage period has passed, the PI must decide whether to continue storing the data. Although data can be kept indefinitely, a PI must evaluate the benefits and risks of extended storage. On the one hand, one never knows when data might be needed. On the other hand, continued storage of confidential data may increase the risk of possible violation. The monetary cost of retention and security are additional concerns.

Destroying Data

When the decision has been made to end data storage, data should be thoroughly and completely destroyed. Effective data destruction ensures that information cannot be extracted or reconstructed. Many document storage companies now offer onsite shredding and secure destruction of written and electronic records. For electronic data specifically, software products are available.

Approved



Vice President for Research and Economic Development

Date


