

# Research Data: Recording, Retention and Access Policy

## Policy Sections

- **Policy Statement**
- **Reason for Policy**
- **Who Should Know This Policy**
- **Contacts**
- **Applicable Policies & Guidelines**
- **Applicable Federal Regulations & Criteria**
- **Overview of Research Data**
- **Appendices**

**Last Revised:** January 2007

**Responsible WMC Officer**  
Executive Vice Dean for Research

**Responsible Office**  
Office of the Executive Vice Dean for Research

---

## Policy Statement

---

Weill Cornell Medical College recognizes and desires to protect the interests of researchers and the institution in matters related to the recording, retention and accessibility of research data.

---

## Reason for Policy

---

The scientific and technical data generated in the course of research and scholarly activity lie at the very heart of the research enterprise. Such data is of immediate use to researchers as they conduct their investigations, and may prove to be a resource of continuing value to the researchers and to others over time. Both researchers and Weill Cornell Medical College have responsibilities and rights with respect to research data, among them being the protection of intellectual property rights, enabling appropriate responses to questions about accuracy, authenticity, primacy and compliance with laws and regulations governing the conduct of research, and responding to the legitimate interests of research sponsors and governmental agencies. Researchers and the College share an obligation to appropriately record research data, archive data for a reasonable length of time, and make data available for review and possible reuse under appropriate circumstances. In addition, the Cornell University Library (CUL) encourages researchers to consult with CUL to consider options to ensure long-term preservation and sharing of research data.

## Who Should Know This Policy

---

- Dean
  - Senior Administration – Vice Provosts, Associate Deans
  - Director, Department Chair, Division Head
  - Faculty
  - Departmental/Divisional Administrators
  - Finance Personnel – Research Accounting, Accounting Services
  - Grants & Contracts Personnel
  - Research Compliance Personnel
  - Research Integrity Personnel
  - All Employees
  - Information Technology Personnel
  - Human Resources Personnel
  - Graduate and undergraduate students involved in scholarly or academic research
- 

### Contacts

#### WMC Policy & Administration

Roles and Responsibility Matrix Website: [http://med.cornell.edu/research/for\\_pol/forms/Comp\\_Matrix.xls](http://med.cornell.edu/research/for_pol/forms/Comp_Matrix.xls)

---

Subject	Contact	Phone	E-mail
Research Compliance Questions & Complaints. Policy Clarification and Interpretation	Assistant Dean, Research Compliance	(212) 821-0722	<a href="mailto:blp2001@med.cornell.edu">blp2001@med.cornell.edu</a>
Scientific Misconduct Questions and Investigations. Formal Dispute Resolution Regarding Request for Research Data	Dean of Graduate School	(212) 746-6900	<a href="mailto:dphajjar@med.cornell.edu">dphajjar@med.cornell.edu</a>
Research Administration Questions	Associate Dean, Research Administration	(212) 746-5979	<a href="mailto:hmlander@med.cornell.edu">hmlander@med.cornell.edu</a>
Long-term retention and sharing of research data	University Archives	(212) 255-3530	

Grants & Contracts Website: <http://www.med.cornell.edu/research/>

Roles and Responsibility Matrix Website: [http://med.cornell.edu/research/for\\_pol/forms/Comp\\_Matrix.xls](http://med.cornell.edu/research/for_pol/forms/Comp_Matrix.xls)

Direct Link: TBA

Research Compliance Website: TBA

---

### Applicable Policies & Guidelines

---

- [Misconduct Policy](#)
- [Responsible Conduct of Research](#)
- [University Policy 1.2](#), Academic Misconduct
- [University Policy 4.6](#), Standards of Ethical Conduct
- [University Policy 4.7](#), Record Retention
- [University Policy 5.4.1, Security of Information Technology Resources](#)

- [University Policy 6.2.10](#), Establishment of College-level Academic Employee Grievance Procedures
- [Grievance Procedures](#) for Graduate Students Relating to Graduate Education and Support
- [Patent Policy](#)
- [Copyright Policy](#)
- [Lab Notebook Guidelines](#)
- [HIPPA Requirements](#)
- [IRB Human Subjects Training](#)
- [Grants & Contracts](#)
- [Material Transfer Agreements](#)
- [IACUC](#)
- [Material Transfer Agreements](#)
- [Invention Disclosure Form](#)
- [Principal Investigator Eligibility](#)
- [Code of Academic Integrity](#)
- [Campus Code of Conduct](#)
- [Faculty Handbook](#)
- [Graduate School Code of Legislation](#)
- [Program Management by Principal Investigators: Award Closeout Procedures](#)
- [Proposal Development, Processing, and Submission Procedures](#)
- [Access to and Retention of Research Data](#), COGR
- [Weill Research Archive Storage](#)

---

## **Applicable Federal Regulations & Criteria**

---

NIH October 1, 2003 Policy Statement encourages the “timely release and sharing” of final results from NIH-supported research. NIH May 2, 2004 Policy Statement on sharing and distributing of unique model organisms.

NSF Policy on the dissemination and sharing of research results (GPM Section 734).

The Health Insurance Portability and Accountability Act of 1996 at 45 CFR 164.530 requires a covered entity to retain the documentation required by the Privacy Rule for six years from the date of its creation or the date when it last was in effect, whichever is later. Also, the HHS and the U.S. Food and Drug Administration (FDA) Protection of Human Subjects Regulations (45 CFR part 46 and 21 CFR parts 50 and 56, respectively) may apply to health services research. In addition, if health-related research involves electronic PHI, covered entities must also be consider the requirements of the HIPAA Security Rule (45 CRF, part 160 and Part 164, subparts A and C).

Public Law 107-188, the Public Health Security and Bioterrorism Preparedness Response Act of 2002, the Select Agent regulations, require the retention of administrative as well as scientific data under 42 CFR 73.15 and 7 CFR 331.14 for three years

The Export Administration Regulations, in 15 CFR 762, requires all records required to be kept by the EAR must be retained for five years from the latest of the following times:

- (1) The export from the United States of the item involved in the transaction; or
- (2) Any known reexport, transshipment, or diversion of such item.

Public Health Service Policies on research Misconduct at 42 CFR Parts 50 and 93 require that an investigation report and all records used to support the report must be maintained in a secure manner by the institution for 7 years after completion of the proceeding or the completion of any PHS proceeding involving the research misconduct allegation.

The 1999 Shelby Amendment to OMB Circular A- 110 defines data as: “The recorded factual material commonly accepted in the scientific community as necessary to validate research findings, but not any of the following: preliminary analyses, drafts of scientific papers, plans for future research, peer reviews, or communications with colleagues.”

---

## **Overview of Research Data Recording, Retention and Access Policy**

---

Weill Cornell Medical College has established requirements for recording, retaining, and accessing research data in order to protect its researchers and their efforts as well as to ensure that the College is in full compliance with applicable laws, regulations and contractual obligations.

The scientific and technical data generated in the course of research and scholarly activity lie at the very heart of the research enterprise. Such data is of immediate use to researchers as they conduct their investigations, and may prove to be a resource of continuing value to the researchers and to others over time. Both research investigators and the College have responsibilities and rights with respect to research data, among them being the protection of intellectual property rights, enabling appropriate responses to questions about accuracy, authenticity, primacy and compliance with laws and regulations governing the conduct of research, and responding to the legitimate interests of research sponsors and governmental agencies. For these reasons, research investigators and the College share an obligation to appropriately record research data, archive it for a reasonable length of time, and make it available for review and possible reuse under appropriate circumstances.

---

### **Procedures**

*Research Data: Recording, Retention and Access: Procedures*

[Direct Link: TBA](#)

---

### **Appendices**

---

#### **Appendix A – Frequently Asked Questions**

##### **Why are data held in such a way that certain individuals working on a research study do not have access to personal information about the material?**

Genetic information is very personal and private, and people participate in such research because they hope that it will help them and their families, as well as others with similar conditions. But, if someone's genetic predisposition to a disease were released to a health-insurance company or an employer, discrimination against that person could occur. People will agree to participate in research on the condition that their medical information is protected against inappropriate disclosures.

##### **Who can take research information from the institution?**

In federally sponsored research, the university owns the data but allows the principal investigator (PI) on the grant to be the steward of the data. The PI takes responsibility for the collection, recording, storage, retention, and disposal of data. Data and data books collected by undergraduates, graduates, and postdoctoral fellows on a research project belong to the grantee institution, and students should not take

their data—although retaining copies of data is allowed, with permission. With industry-funded research, data can belong to the sponsor.

### **Why should I share my final research data?**

Data sharing achieves many important goals for the scientific community, such as

- reinforcing open scientific inquiry
- encouraging diversity of analysis and opinion,
- promoting new research, testing of new or alternative hypotheses and methods of analysis
- supporting studies on data collection methods and measurement
- facilitating education of new researchers
- enabling the exploration of topics not envisioned by the initial investigators
- permitting the creation of new datasets by combining data from multiple sources.

### **Who benefits from data sharing?**

Everyone benefits, including investigators, funding agencies, the scientific community, and, most importantly, the public. Data sharing provides more effective use of NIH resources by avoiding unnecessary duplication of data collection. It also conserves research funds to support more investigators. The initial investigator benefits, because as the data are used and published more broadly, the initial investigator's reputation grows.

### **What do you mean by unique data?**

By "unique data" we mean data that cannot be readily replicated. Examples of studies producing unique data include: large surveys that are too expensive to replicate; studies of unique populations, such as centenarians; studies conducted at unique times, such as a natural disaster; studies of rare phenomena, such as rare metabolic diseases.

### **What kinds of data are candidates for sharing?**

Potentially all kinds of data are candidates for sharing, but unique data are especially important. Some biologic sciences already have data-sharing plans in place, such as genetic mapping. But other basic science data are also amenable to sharing. Data from human subjects (e.g., surveys, clinical studies) also can be shared if the identity and privacy of research participants can be protected.

### **How soon after data collection am I obliged to share the final data?**

Recognizing that the value of data often depends on their timeliness, data sharing should occur in a timely fashion. NIH expects the timely release and sharing of data to be no later than the acceptance for publication of the main findings from the final dataset. This time point will be influenced by the nature of the data collected. Data from small studies can be analyzed and submitted for publication relatively quickly. If data from large epidemiological or longitudinal studies are collected over several discrete time periods or waves, data should be released in waves as data become available or main findings from waves of the data are published. NIH recognizes that the investigators who collected the data have a legitimate interest in benefiting from their investment of time and effort. NIH continues to expect that the initial investigators may benefit from the first and continuing use, but not from prolonged exclusive use. While NIH also understands that an institution's desire to exercise its intellectual property rights may justify a need to delay disclosure of research findings, a delay of 30 to 60 days is generally viewed as a reasonable period for such activity.

### **Can I share data with colleagues under my own auspices?**

Yes. Your data-sharing plans should indicate the criteria for deciding who can receive your data and whether or not you will place any conditions on their use. Data should be made as widely and freely available as possible while safeguarding the confidentiality of the data and privacy of participants. You should not place limits on the questions or methods others might pursue nor should you require co-authorship as a condition for receiving the data.

**How can I protect the privacy of my subjects?**

It is the responsibility of the investigators, their IRB, and their institution to protect the rights of participants and the confidentiality of their data. Data should be redacted to strip all individual identifiers, and effective strategies should be adopted to minimize risk of disclosing a participant's identity. Options to protect privacy include: withholding part of the data, statistically altering the data in ways that will not compromise secondary analyses, requiring researchers who seek data to commit to protect privacy and confidentiality, and providing data access in a controlled site, sometimes referred to as a data enclave. Some investigators use hybrid methods, releasing a redacted dataset for general use but providing access to more sensitive data through a user contract or data enclave. In most instances, sharing data is possible without compromising participant confidentiality and privacy.

**Can institutions and investigators subject to the Federal Health Insurance Privacy and Portability Act (HIPAA) Privacy Rule share data in accord with the NIH Data Sharing policy?**

Yes. NIH recognizes that data sharing may be complicated or limited, in some cases, by institutional policies or local IRB rules, as well as by local, state and Federal laws and regulations like the Privacy Rule. To protect the rights and privacy of people who participate in NIH-sponsored research, data intended for broader use should be free of identifiers that would permit linkages to individual research participants, and exclude variables that could lead to deductive disclosure of the identity of individual subjects. When data sharing is limited, applicants should explain such limitations in their data sharing plans.

**What are the issues in data storage and retention?**

The National Institutes of Health require that grant recipients keep all data for three years beyond the time that the final expenditure report of the grant is reported. But different agencies and academic societies have different requirements, and universities usually make it the principal investigator's responsibility to abide by the rules. The National Science Foundation has a similar requirement. However, the American Psychological Association expects its members to retain data for a minimum of five years, and different universities also require data retention for varying periods of time, usually for a minimum of five years. Legally, data need to be retained for patent protection and in case there are any misconduct allegations pending based on the data.

**What are data?**

Data are a collection of facts, measurements, or observations used to make inferences about the world we live in. Data can range from material created in a wet laboratory, such as an electrophoresis gel or a DNA sequence, to that obtained in social-science research, such as a filled-out questionnaire, videotapes, and photographs. Data can be microscope slides, cell lines, climate patterns, soil samples, astronomical measurements, and spectrographic analyses. Custom software or hardware and specialized methods can be data, too.

## Appendix B – Definitions

Term	Definition
<b>Data</b>	For the purpose of this policy, the term data is defined as recorded information regardless of form or the media on which it may be recorded. Data may be observational, experimental, or computationally derived, and may be raw, preliminary, or final. Data includes, but are not limited to, technical data, digitally recorded data, laboratory worksheets, memoranda, notes or exact copies thereof that are the result of original observation and activities of a study and are necessary for reconstruction and evaluation of the report or publication of the study. In clinical investigations, this term also includes case history records and study protocol.
<b>Researcher</b>	Faculty, staff, and students conducting research, extension and public service, scholarly, or other activity carried out in furtherance of their University Responsibilities.
<b>Principal Investigator</b>	The individual(s) with primary responsibility for the conduct of a Sponsored Project. This responsibility includes the intellectual conduct of the project, fiscal accountability, administrative aspects, and adherence to policies and regulations. (See <a href="#">Principal Investigator Eligibility</a> ). A Sponsored Project may have multiple individuals as PIs who share the authority and responsibility for leading and directing the project, intellectually and logistically. Each PI is responsible and accountable for the proper conduct of the project including adherence to the provisions of this policy.
<b>Sponsored Project</b>	A research extension or instruction activity having a defined scope of work, funded in whole or in part from support provided by an external organization (sponsor), and carried out according to terms and conditions of a signed award agreement, typically a grant, contract, or cooperative agreement, for a specific budgeted financial compensation.
<b>University Appointment</b>	A position with the University, whether academic or non academic, full or part time, paid or unpaid, continuing or limited term, including academic appointments with a modified title (visiting, courtesy, adjunct, etc.), students on assistantships and fellowships, and student employees.
<b>University Responsibilities</b>	Activities conducted in the course of a University Appointment, or in the case of a student, in the course of enrollment in a college or the Graduate School, including those activities carried out under a grant, contract or award made to the University by an external organization.
<b>Senior Investigator</b>	In the context of this policy, the generally recognized leader of a research group or scholarly activity. In the case of a Sponsored Project, this is typically the Principal Investigator(s) responsibility.
<b>Intellectual Property</b>	Property (such as patents, trademarks, and copyright) which is the product of invention or creativity. Intellectual property laws confer specific legal rights, which authors, inventors and other Intellectual Property holders may hold, or exercise.