

Data Management

Michael Kalichman, Ph.D.
Director, Research Ethics Program

Ethics in Scientific Research
May 19, 2008

Data Management

1. Definition

Definition of Data

If you were to publish a paper on your current area of research and, in 3 years, you were accused of having falsified or fabricated the data, what would you have to prove that you had done the work as described?

Examples of Data

- Measurements
- Entries in computer files
- Gels
- DNA sequences
- DNA libraries
- Antibodies
- Unique cell lines or reagents
- Transgenic animals
- Micrographs
- Video taped interviews
- Custom software or hardware
- Specialized methods

Data Management

1. Definition
2. Acquisition

Data Acquisition

- Are those collecting data adequately trained and motivated to do so?
- Are procedures in place to deal with problems that are likely to arise?
- Are experimental methods designed to eliminate, or at least minimize, bias?
- Are records being kept in a manner that permits future verification of what was done?

Data Acquisition: Recordkeeping

How do you keep your research records?

Best model for recordkeeping will not be the same for all areas of research.

Nearly all types of research include records that should be kept in a bound lab notebook.

Data Acquisition: Lab Notebooks

At a minimum, notebooks can provide a listing of:

- date
- investigator(s)
- what was done
- where relevant research products can be found

Supplement as needed by specialized methods of recordkeeping suitable for specific types of data, e.g.:

- computer files
- Videotapes
- gels

Data Management

1. Definition
2. Acquisition
3. Ownership

Data Ownership

Responsibilities of data ownership include:

- Understanding of implications of institutional ownership
- Sharing of data
- Retention of data

Data Ownership: Institutional ownership

In general, equipment, materials, reagents, and data **all belong to the institution** in which they are purchased or produced.

However, unless and until a marketable product is produced, the institutional ownership is **not visibly exercised**.

If profits are a possibility, issues of **technology transfer, copyright, patents, or licensing** must be negotiated with the institution.

Data Ownership: Institutional ownership

Although institution has legal standing to override, researchers have following roles and responsibilities:

- Principal investigator has right and responsibility to make all decisions about collection, use, and sharing of the data
- Original research records should be kept and maintained in the laboratory in which they were created
- Researchers leaving one laboratory for another should be entitled to take copies of their research records

**Data Ownership:
Institutional ownership**

If regulatory, proprietary, confidentiality, or other issues preclude researchers taking copies of their research records, then:

- Principal investigator is responsible to make this clear to anyone in his or her research group.

**Data Ownership:
Sharing Guidelines**

Before publication

- maintain open data policy,
- but with appropriate caution

After publication

- be prepared to grant *reasonable access* to the raw data

**Data Ownership:
Sharing Guidelines**

Reasonable access =

- in the interest of scientific inquiry
- can be accomplished without inordinate expense, and
- can be completed in a time frame that is not burdensome to either the researcher or the requestor

**Data Ownership:
Retention of Data**

When can you throw it away?

Under current Health and Human Services requirements, research records must be maintained for **at least three years after the last expenditure report** (e.g., three years after the last expense under a multi-year grant program).

for some purposes, federal regulations or institutional guidelines may require that data be retained for periods **longer than three years**

**Data Ownership:
Retention of Data**

Other considerations:

- **minimum necessary information** to reconstruct what was done
- extent to which a line of research is **still being pursued**
- **likelihood of ongoing interest** in the field of research
- continued assurance of **confidentiality for human subjects**
- **ease of locating** individual records
- **space and expense** necessary for storage of records
