Ministry of Health and Sports

Data protection and data sharing

Prof Hla Hla Win
Rector, UPH
Vice chair, IRB (UPH)

9.7.2020

What are data?

- “Recorded information, regardless of the form or medium on which it may be recorded, and includes writings, films, sound recordings, pictorial reproductions, drawings, designs, or other graphic representations, procedural manuals, forms, diagrams, work flow charts, equipment descriptions, data files, data processing or computer programs (software), statistical records, and other research data”

(NIH Grants Policy Statement)

- Data is any collection of facts, measurements, or observations used to make inferences about the world in which we live.

- Data can range from material created in a wet laboratory, such as an electrophoresis gel or a DNA sequence, to information obtained in social-science research, such as a filled-out questionnaire, video and audio recordings, or photographs.

- Data can be astronomical measurements, microscope slides, climate patterns, cell lines, field notes, soil samples, or results of statistical analyses.

What are data?

Sponsored projects – more than just research data

- Need to consider financial data and administrative data (subject to retention requirements)
- **Financial data** – budget information, record of expenditures
- **Administrative data** – project proposals, required approvals

- Once collected, data must be properly protected.
- They may be needed later:
- to confirm research findings,
- to establish priority, or
- to be reanalyzed by other researchers.

Data Protection

- Once collected, data must be protected for later use
  - to confirm research findings
  - to be reanalyzed by other researchers
- Data Storage
- Confidentiality
- Retention

Data storage

- The responsible handling of data begins with proper storage and protection from accidental damage, loss, or theft:
- Lab notebooks should be stored in a safe place.
- Computer files should be backed up and the backup data saved in a secure place that is physically removed from the original data.
- Samples should be appropriately saved so that they will not degrade over time.
- Care should be taken to reduce the risk of fire, flood, and other catastrophic events.

Confidentiality

- Some data may be subject to privacy restrictions (human subjects or confidential business information)
- a safe place that is accessible only to authorized personnel
- Using random codes
- PI has the primary responsibility for its protection.

Period of retention

- Data should be retained for a reasonable period of time to allow other researchers to check results or to use the data for other purposes.
- NIH generally requires that data be retained for 3 years following the submission of the final financial report.
- Some government programs, few Universities – 3-7 years
- Aside from these specific guidelines, however, there is no comprehensive rule for data retention

- ***Retention***
  - retained for a reasonable period of time to allow other researchers
    - to check results or to use the data for other purposes
  - no common definition
  - In general, three years should be considered a minimum in
    - academia (NIH)
  - Different universities – different data retention policies (3 to7 years)

Retention

- difficult to predict when data collected sometime in the past
  could be useful
- AIDS – researchers use stored samples to pinpoint first occurrences of HIV infections

Give careful consideration to potential future uses of your data

Data Sharing

- widely agreed that research data should be shared
- Others to replicate findings, to advance science, new research, new datasets
- when and with whom

Data Sharing

- Preliminary data (data that have not been carefully checked
  and validated) – usually should not be released
- Exception – potentially benefit the
  public eg. - unexpected side effects
  from a drug
  - unrecognized environmental health problem

Data Sharing

- Confirmed or validated data – keeping data confidential
  prior to publication is a commonly accepted practice
- NIH data sharing policy – “timely release and sharing”
- “no later than the acceptance for publication”
- Data sharing plan

Data Sharing

- Published data – generally expected that all the information about that experiment (including the final data) should be freely available
- Some journals – formally require that the data published in articles be available to other researchers upon request or stored in public databases

- Data sharing statements must indicate the following:
- whether individual deidentified participant data (including data dictionaries) will be shared;
- what data in particular will be shared;
- whether additional, related documents will be available (e.g., study protocol, statistical analysis plan, etc.);
- when the data will become available and for how long;
- by what access criteria data will be shared (including with whom, for what types of analyses, and by what mechanism).
- Illustrative examples of data sharing statements that would meet these requirements are in the Table.

Table 1. Examples of Data Sharing Statements That Fulfill These ICMJE Requirements*

	Example 1	Example 2	Example 3	Example 4
Will individual participant data be available (including data dictionaries)?	Yes	Yes	Yes	No
What data in particular will be shared?	All of the individual participant data collected during the trial, after deidentification.	Individual participant data that underlie the results reported in this article, after deidentification (text, tables, figures, and appendices).	Individual participant data that underlie the results reported in this article, after deidentification (text, tables, figures and appendices).	Not available
What other documents will be available?	Study Protocol, Statistical Analysis Plan, Informed Consent Form, Clinical Study Report, Analytic Code	Study Protocol, Statistical Analysis Plan, Analytic Code	Study Protocol	Not available
When will data be available (start and end dates)?	Immediately following publication. No end date.	Beginning 3 months and ending 5 years following article publication.	Beginning 9 months and ending 36 months following article publication.	Not applicable
With whom?	Anyone who wishes to access the data.	Researchers who provide a methodologically sound proposal.	Investigators whose proposed use of the data has been approved by an independent review committee ("learned intermediary") identified for this purpose.	Not applicable
For what types of analyses?	Any purpose.	To achieve aims in the approved proposal.	For individual participant data meta-analysis.	Not applicable
By what mechanism will data be made available?	Data are available indefinitely at (*link to be included*).	Proposals should be directed to xxx@yyy. To gain access, data requestors will need to sign a data access agreement. Data are available for 5 years at a third party website (*link to be included*).	Proposals may be submitted up to 36 months following article publication. After 36 months the data will be available in our University's data warehouse but without investigator support other than deposited metadata. Information regarding submitting proposals and accessing data may be found at (*link to be included*).	Not applicable

* These examples are meant to illustrate a range of, but not all, data sharing options.

Conclusion

- Data – integral to the research process
- Data management practices – complex (including collection, sharing, and ownership) and should be discussed before data collected
- The integrity of data – paramount importance and requires careful planning

Sources

- Responsible Conduct of Research (RCR). David R. Clark, Research Compliance Officer (www. research-compliance.wsu.edu)
- Data Acquisition, Ownership, Management, and Sharing. L. Gabriel Navar, Department of Physiology, Renal and Hypertension Center of Excellence Tulane University School of Medicine, New Orleans, Louisiana
- Responsible Conduct of Research Training: Acquisition, Management, Sharing and Ownership of Data. Frank Van Breukelen, University of Nevada, Las Vegas
- Responsible Conduct of Research . Afya Bora Consortium Global Health Leadership Fellowship Program; A Distance Learning Module
- ORI Introduction to the Responsible Conduct of Research. Nicholas H. Steneck. Revised Edition August 2007

Thank You

Case example

- Drs. Kessenbaum and Wilcox are conducting a long-term, observational study of the health of pesticide applicators. The protocol calls for an initial health assessment, including a health history, physical exam, and blood and urine tests. The researchers will collect a DNA sample from cheek scrapings and collect dust samples from the applicators' clothing and hair and underneath their fingernails. After the initial health assessment, the applicators will complete yearly health surveys and undergo a full health assessment every four years. The researchers will follow the subjects for at least 25 years. Their work is funded by the NIH.

- Drs. Kessenbaum and Wilcox have been conducting their study for 15 years, and they have compiled an impressive database. They have already published more than a dozen papers from the database. Whenever they share data, they require researchers who request it to sign elaborate data-sharing agreements, which spell out clearly how the data will be used. The agreements also specify the kinds of studies that can be published using the data, which allows Drs. Kessenbaum and Wilcox to protect their interests in publishing on certain topics.

- In the past month, they have received some requests to access their database. One request has come from a pesticide company, another has come from a competing research team also studying the health of pesticide applicators, and another has come from a radical environmental group with an antipesticide agenda.

- How should Drs. Kessenbaum and Wilcox handle these requests to access their database?
- Should they refuse to share data with the pesticide company or the environmental group?
- Is it ethical to require people who request data to sign elaborate data-sharing agreements?

- Responsible conduct of research 3rd Ed (Adil E Shamoo, David B Resnik)

Questions, Comments & Suggestions