11-8-2013

Using Cases to Teach Research Data Management

Donna Kafel
University of Massachusetts Medical School, Donna.Kafel@umassmed.edu

Follow this and additional works at: http://escholarship.umassmed.edu/lib_articles

Part of the Information Literacy Commons, and the Scholarly Communication Commons

This work is licensed under a Creative Commons Attribution-Share Alike 3.0 License.

Repository Citation
http://escholarship.umassmed.edu/lib_articles/162

This material is brought to you by eScholarship@UMMS. It has been accepted for inclusion in Library Publications and Presentations by an authorized administrator of eScholarship@UMMS. For more information, please contact Lisa.Palmer@umassmed.edu.
New England Collaborative Data Management Curriculum

Using Cases to Teach Research Data Management

Donna Kafel, MLIS
Lamar Soutter Library
UMass Medical School

This work is licensed under a Creative Commons Attribution-ShareAlike 3.0 Unported License.
Selecting a Case

✧ Discipline

✧ Research setting (e.g. clinic, lab)

✧ Specific data management concept or issue
Cases

Clinical Health Study:
- Outcomes from Orthopedic Implant Surgery
- Studying Vitamin D as an Augmentation of Treatment for Bipolar Depression

Health Study in lab:
- Combining Data from 10 Years of Research for Retrospective Studies on the Effects of Exercise and Diet on the Risk of Diabetes
Cases (cont’d)

Biomedical lab:
• Regeneration of Functional Heart Tissue in Rats

Qualitative Behavioral Health Study:
• Improving End of Life Care in African Americans

Engineering Test Lab:
• Characterizing a Component of a Rocket Engine Used to Control Satellites in Orbit
New England Collaborative Data Management Curriculum

- How to teach RDM using cases
- Summary of teaching points preceding each case
- Discussion questions following each case
- Research Data Management Concepts and Issues Illustrated in the Teaching Cases
- Integrated as activities in lesson plans
Case Analysis using Simplified Data Management Plan

1. Types of data
2. Contextual details (metadata) needed to make data meaningful to others
3. Storage, Backup, and Security
4. Provisions for Protection/Privacy
5. Policies for re-use
6. Policies for access and sharing
7. Plan for archiving and preservation of access
UK Data Archive Research Data Lifecycle

http://www.data-archive.ac.uk/create-manage/life-cycle
A data management case study in biomedical engineering research
“Advances knowledge in engineering, biology, and medicine, and improves human health through cross-disciplinary activities that integrate principles of engineering sciences with medical sciences and clinical practice.”

3Imperial College London. N.D.
http://www3.imperial.ac.uk/pls/portallive/docs/1/51182.PDF
Experiment: delivery of stem cells on a biological fibrin microthread to areas of damaged heart tissue in one rat’s heart
Purpose: to restore mechanical function of damaged heart tissue

Image courtesy of Dr. Glenn R. Gaudette
Materials used in Experiment

SDMP #1 creating data

Photo courtesy of nlm.nih.gov

Photo courtesy of Dr. Glenn R. Gaudette
Animal Model for Experiment: Rat

SDMP #1: creating data

SDMP #2: File formats

Surgical log for #2345

LABORATORY NOTEBOOK
Collective data from experiment

-2 days: Incubate stem cells with markers

-1 day: Stem cells in solution with biological suture

0 day: #1 Surgery: infarct/delivery of stem cells to damaged heart tissue

7 day: #2 Surgery: examination, high speed imaging/LVPs, isolate heart and place it in freezer

8 days +: Section heart, tissues on slides, staining, images of tissues, tracking particles on heart

SDMP #1 Creating data
Timeline of Experiment

2 days before surgery

Incubate adult stem cells with markers (Q dots)

1 day before surgery

Place stem cells in solution—inject into tube with biological suture

Images courtesy of Dr. Glenn R. Gaudette
Timeline of Experiment
Day of Surgery

1. Document pre-op data
2. Operative procedure: open up thoracic cavity.
   a. Create a myocardial infarction
   b. Place biological suture with stem cells in area of infarct (damaged heart tissue) \(\rightarrow\) document operative data

Image courtesy of Dr. Glenn R. Gaudette
1. Thoracic cavity is reopened
2. Images of heart are collected with 2 cameras
3. Pressure transducer syncs with images — measures left ventricular pressure
4. Euthanize rat
5. Isolate heart and fix it in a fixative

Timeline of Experiment

7 Days Post Stem Cell Delivery: 2

DMP #1 Creating Data
DMP #2 Formats, Naming
DMP #3 Storage

Image courtesy of Dr. Glenn R. Gaudette
Timeline of Experiment

24 hours after placement of heart in freezer: Sectioning heart tissue

1. Cutting sections of heart and placing on slides. (Yield ~200 slides per experiment)

2. Tissues are kept in 3 freezers.
Timeline of Experiment
8+ Days: (up to “a couple” months):

1. Some of the slides are stained immediately
2. Other slides stained from a day to a couple of months after that
3. Stains: some are stained with trichrome, some other stains for specific markers

~200 slides

Image courtesy of Dr. Glenn R. Gaudette
Timeline of Experiment

Unspecified but after several slides are stained: Taking microscopic images of stained slides

- Take images on lab’s epifluorescent microscope

If happy with staining, use confocal microscope to take better quality images

- SDMP #1 Creating data
- SDMP #2 File formats Naming conventions
- SDMP #3 Storage Security
Team looks at the data that they acquired and use home-grown software to track particles on the surface of the heart to see how far and how fast those particles are moving.
## File Formats

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Images</td>
<td>?</td>
</tr>
<tr>
<td>Left ventricular</td>
<td>?</td>
</tr>
<tr>
<td>pressure measurements</td>
<td></td>
</tr>
<tr>
<td>Home made software</td>
<td>MATLAB or C</td>
</tr>
<tr>
<td>Histology sections</td>
<td>Slides—file name based on stain—example .act is actinin stain</td>
</tr>
<tr>
<td>Contextual</td>
<td>Paper lab notebook, animal log</td>
</tr>
</tbody>
</table>

Directory that links data sets together: Excel spread sheet
Analyzing the data

“There could easily be up to 10 people involved in data analysis and we have not yet found a good way to link all the data.”

Dr. Glenn R. Gaudette, PI
<table>
<thead>
<tr>
<th>Data Set</th>
<th>Storage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optical images taken during Surgery: (~10,000 images) 1000 images for each data set</td>
<td>Hard drive acquisition computer &gt; Drobo backup&gt;hard drive of network computer backed up by institution</td>
</tr>
<tr>
<td>Left ventricular pressures (numeric) correlating with specific images</td>
<td>Same as above</td>
</tr>
<tr>
<td>Tissue sections</td>
<td>Slide boxes—could be in any of 3 or 4 freezers</td>
</tr>
<tr>
<td>Software</td>
<td>?</td>
</tr>
<tr>
<td>Images from different stained tissue after second surgery</td>
<td>Drobo &gt; DVD backups</td>
</tr>
<tr>
<td>Contextual data</td>
<td>Paper lab notebook (lab, PI’s office), surgical log (with animal)</td>
</tr>
</tbody>
</table>
What are the policies for reuse of data?

What is the process for gaining access to data?

What is the long term plan for preservation and maintenance of the data?
Case B Analysis using Simplified Data Management Plan

1. Types of data: Images of heart, LVP measurements, histology slides (tissues), images of slides, software

2. Contextual details (metadata) needed to make data meaningful to others:
   - experiment #
   - dates of experimental activities
   - stem cell line
   - details about animal (species, age, identifier)
   - area of infarct
   - area where stem cells implanted
   - type of stain used
   - instrumentation

3. Storage, Backup, and Security:
   - Hard drive of acquisition computer (initial)
   - Drobo
   - networked hard drive (backed up by institution)
   - DVDs
   - Slide boxes in freezers
3. (continued):
• Lab notebooks/animal surgical logs (some backup when data is transcribed from animal surgical log to paper lab notebook)

4. Provisions for Protection/Privacy:
• Files are not password protected
• Paper lab notebooks are kept in lab/older ones in PI’s office (key card access to these rooms)

5. Policies for re-use:
• Not addressed – need to ask researcher (possibilities: reuse with permission of PI, institutional policies)

6. Policies for access and sharing:
• Not addressed – need to ask researcher (possibilities: PI will make accessible after paper publication, will share them immediately, may depend on funding agency reqs.)
Case B Analysis using Simplified Data Management Plan

7. Plan for archiving and preservation of access
   • Not addressed — ask researcher (possibilities: convert files in proprietary formats to generic formats, appraisal of data and data versions, decision on where data should be archived (IR or DR)


Imperial College London, Department of Engineering. N.D. “Definition of biomedical engineering.” Accessed March 20, 2013 at http://www3.imperial.ac.uk/pls/portallive/docs/1/51182.PDF