



Cornell University®

**2019 Winter  
Responsible Conduct of Research (RCR) Symposium:  
Data Acquisition and Management**

**January 17, 2019**

---

## CASE STUDIES

---

### Case 1: Managing Research Data

(Adapted from:

[www.onlineethics.org/Resources/TeachingTools/Modules/19237/resethpages/advislab.aspx](http://www.onlineethics.org/Resources/TeachingTools/Modules/19237/resethpages/advislab.aspx))

You are a graduate student and for two years have been working on kidney function in Professor Carr's group. Professor Carr has pioneered methods to understand kidney development in mice by reconstructing a time-line of kidney development from high-resolution fluorescent microscopy. Several years before you joined the lab, Carr published a paper describing a computer program for generating 3-D images of developing kidneys from multiple image files. The paper describes the program in reasonable detail but does not include the actual computer code.

You have developed a new analysis method and computer program that appears to give significant differences from results in previous work. You would like to figure out the cause of these differences. You ask Prof. Carr if you could run your images using the older program. Prof. Carr replies that the older program was stored on a computer that failed several years ago and is thus no longer available.

#### **Discussion Questions for the Facilitators:**

- How does this general scenario apply to your own type of research?
- What data management practices could have prevented this scenario from occurring?
- What can/should you do? What, if any, ambiguities do you face? What are the potential consequences of taking action, or not taking action?
- What are the potential consequences for Prof. Carr? Can he continue to cite the older paper? Should he consider retracting it?

## Case 2: Data Sharing

(Adapted from: <https://ori.hhs.gov/case-two-data-sharing-fever>)

Mary admires the NIH-funded work of her postdoctoral advisor, Henryk, who pioneers research on alternative treatments for fever due to infectious diseases. Mary is one of many co-workers who has assisted Henryk in compiling the most comprehensive database ever assembled, tracking many different infectious agents, species of animals, and different interventions and their outcomes. The database contains all the raw data as well as comparative analyses and summaries.

Henryk's interpretation of this rich dataset suggests that some "alternative medicines" are highly effective in certain species, but have no therapeutic value in others. He is completing his analysis and interpretation, and is preparing a manuscript for submission. Mary will be a co-author because of her part in collecting data for the study.

Mary has recently started an Assistant Professor position and wants to build on her postdoctoral work. She asks Henryk for permission to use the dataset to develop her own project. However, she plans to use a different methodology for analysis and interpretation of the dataset to address a different aspect of the outcomes of treatment. She plans to use her analyses to develop her first grant proposal to submit to the NIH.

Henryk is unwilling to share the entire dataset prior to publishing his interpretation of these data. However, Mary has access to the database as part of her current project, and therefore she decides that it is ethical for her to look more closely at the data.

Mary spends quite a lot of time looking at the data and Henryk's analysis, and realizes that he has excluded specific datapoints that impact his interpretation. Henryk's draft manuscript carefully justifies the exclusion of these data in the methods section so that there is no issue with data falsification.

Mary realizes that with her methodology and by including all the data points, that she comes to a very different interpretation than Henryk's, which could lead to an entirely new understanding of therapies to treat fever. Mary is excited about her upcoming grant proposal, but is concerned about how to broach the discussion of her use of the data with Henryk.

### **Discussion Questions:**

- How does this general scenario apply to your own type of research?
- Who owns the database at this point: Henryk? The institution? NIH? The public?
- Is it ethical for Mary to go ahead and use her analysis for her grant proposal?
- Is it ethical for Mary to go ahead and publish her own analysis before Henryk?
- Must Henryk share his database with Mary before publication? After publication? Must he share it with others, outside his lab, and if so, when?
- Suppose that after a year that Henryk has still not published his analysis of the data. Mary's first grant submission was rejected, in part because the reviewers wanted her analysis to first be peer-reviewed and published. The deadline for Mary to submit a revised grant proposal is approaching, and she really wants to get her paper submitted. How might Mary respond?
- Mary and Henryk have different views on whether or not to use all the data points in their respective analyses. Is this a legitimate case of scientific disagreement, or not?