

# Responsible Conduct of Research

## Data

Susan Porter

Clinical Associate Professor,  
Department of Pathology & Laboratory Medicine  
Associate Dean, Professional Development  
Faculty of Graduate Studies

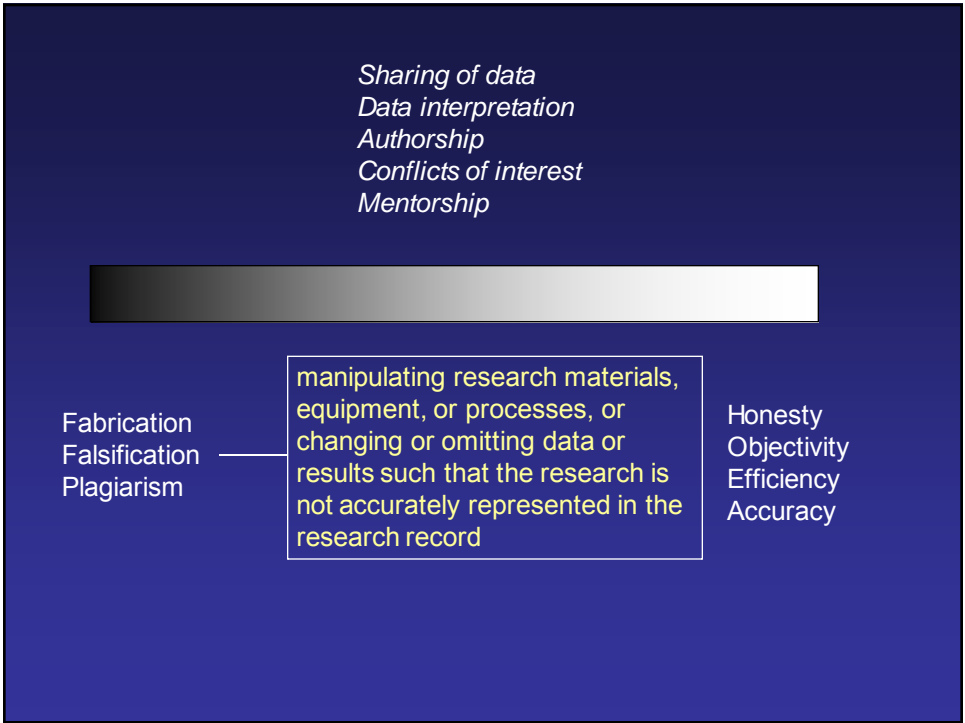


# Responsible Conduct of Research

## Data

Obtaining  
Recording  
Analyzing  
Reporting  
Ownership  
Sharing





*Sharing of data*  
*Data interpretation*  
*Authorship*  
*Conflicts of interest*  
*Mentorship*



Fabrication  
Falsification  
Plagiarism

appropriation of another person's  
ideas, processes, results, or  
words without giving appropriate credit

using inappropriate references (eg, not  
referencing original paper describing  
results)

Honesty  
Objectivity  
Efficiency  
Accuracy

## Obtaining Data

### Ensure

- experimental design is free from bias
- sample (if applicable) is appropriate
- conducted carefully
- proper controls are included

“ The first principle is that you must not fool yourself, and you are the  
easiest person to fool...I'm talking about a specific, extra type of integrity  
that is not lying, but bending over backwards to show how you're maybe  
wrong” - Richard Feynman, Nobel Laureate

If in doubt – ask (mentor, colleague), tell (publication)

## Obtaining Data

- external or administrative

### Ensure

- privacy and confidentiality is respected as appropriate
- awareness of any potential bias, limitations or assumptions in the data
- all aspects of the research question(s) have been thoroughly thought through
- sensitivity to organization and research team dynamics, good communication
- consideration of feedback to community group if appropriate

## Recording Data

Data books should allow:

1. You and others to easily know
  - WHY you did the experiment
  - WHEN you did the experiment
  - HOW you did the experiment
  - with what reagents you did the experiment
  - your results
  - your conclusions
  - (your interpretation)
  - what you're doing next
2. Others to repeat what you did (several years later)

## Recording Data

Therefore, notebooks should:

1. Be (bound) lab notebooks with numbered pages
2. Be in pen, not pencil – legibly, neatly
3. Be recorded as soon as possible after data collected
4. Include all raw data
5. Include calculations
6. Have errors corrected without erasing or obscuring original recording
7. Include correspondence and conversations relating to experiments
8. Include tangible data or information or its location

## Recording Data

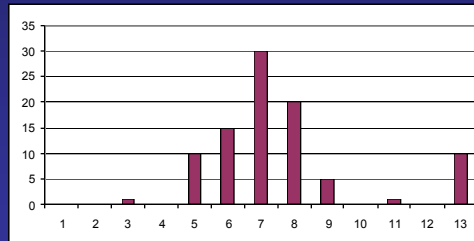
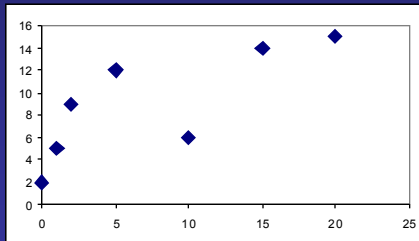
Electronic notebooks or data collection:

1. Control entry and alteration access: authorized individuals only
2. Ensure information trail of who and when accessed
3. Ensure proper backup

## Analyzing and Reporting Data

### Data trimming

- removing data points to fit a desired or more statistically pleasing result



## Analyzing and Reporting Data

### Outliers

- if possible, determine likely cause
- decide how to deal with these beforehand if possible
- use statistical methods to assess
- disclose basis for dropping or modifying data

## Analyzing and Reporting Data

### Data enhancing

“We should conspire to end the fetish of the perfect image.”

“Beautification is a form of misrepresentation.”

## Analyzing and Reporting Data

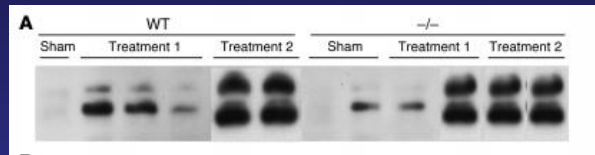
### Image data

from *Nature* and *JCI*:

- List all image acquisition tools, image processing software, and key settings and manipulations
- Don't combine images gathered at different times or from different locations unless stated. If juxtaposing images is essential, clearly demarcate borders (eg, black line between lanes).
- Avoid touch-up tools
- Don't remove “important” data by cropping
- Avoid high contrast settings (have grey background)
- Only adjust brightness, contrast, etc equally across image
- Contrast or settings should not be adjusted so that data disappear.
- Do not reuse previously published figures even if cropped/rotated

## Analyzing and Reporting Data

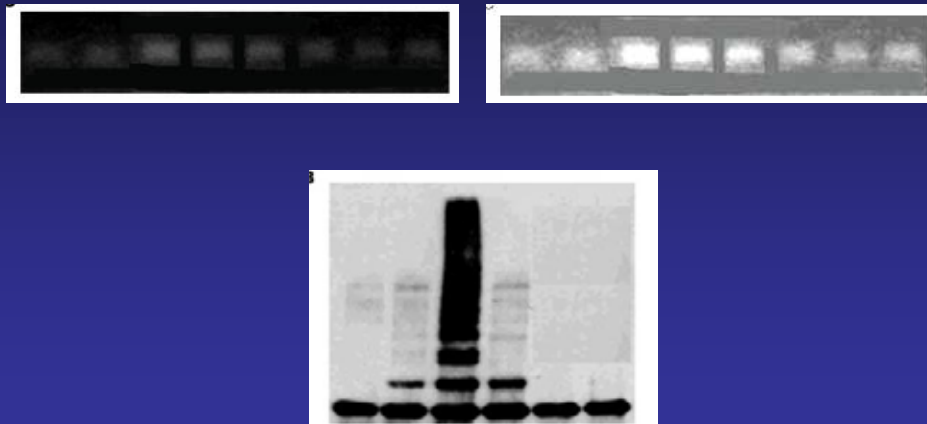
### Image data



- Label graphic data clearly
- Keep files well organized
- Take high-resolution pictures from start

## Analyzing and Reporting Data

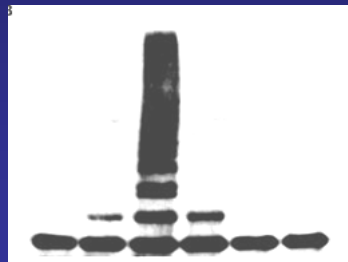
### Image data





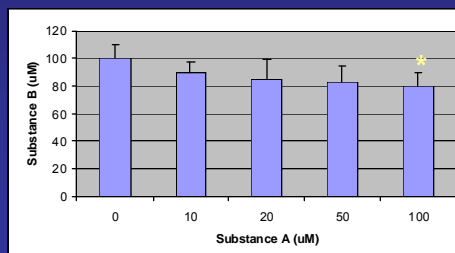
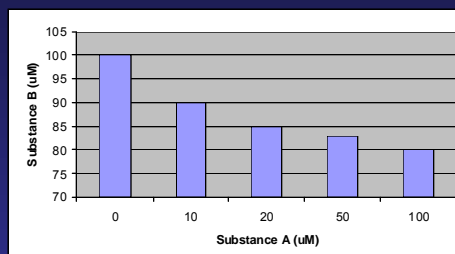
## Analyzing and Reporting Data

### Image data



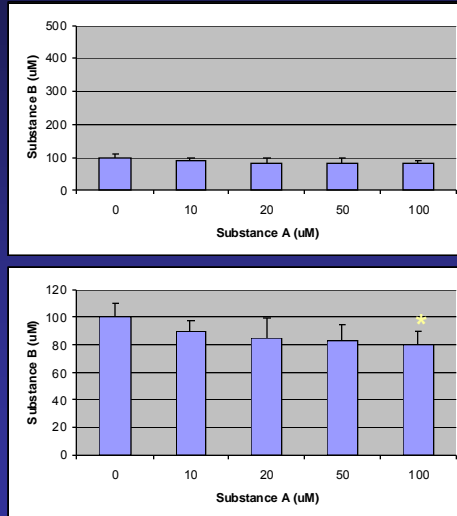
## Analyzing and Reporting Data

### Graphic data



## Analyzing and Reporting Data

### Graphic data



## Analyzing and Reporting Data

### Statistical analysis

“Lies, damned lies, and statistics”

## Analyzing and Reporting Data

### Statistical analysis

*American Statistical Association:*

- Use only appropriate methodologies suitable to the data
- Remain current
- Understand the theory, the data, and the methods
- Report assumptions
- Account for all data considered in a study
- Report the data cleaning and screening procedures used
- Address potential confounding variables
- Carefully consider statistical methods (or enlist statistical help) before research is undertaken

## Analyzing and Reporting Data

### Interpretation, Discussion

#### **Common errors (honest or otherwise)**

- *post hoc or cum hoc, ergo propter hoc* (association  $\neq$  cause)
- *a posteriori* hypothesis presented as *a priori*
- not taking into account all factors
- not taking into account all data (yours or others)
- misleading or incomplete description of methods or materials

## Analyzing and Reporting Data

### Interpretation, Discussion

Popovic M, Srangadhara MG, Read E and Gallo RC. *Science* 224:497 (1984)

“This cell line, termed HT...”

“The concentrated fluids [from patient T cell cultures] were first shown to contain particle-associated RT.”

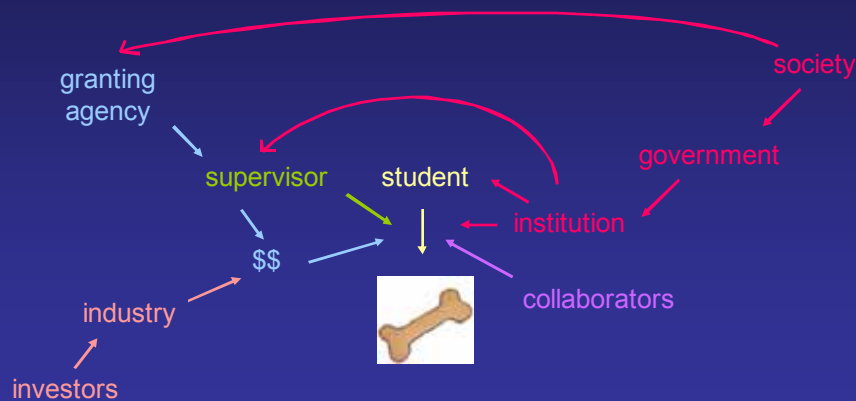
“Continuous HTLV-III production...”

“...suggest that HTLV-III and LAV may be different. However, it is possible that this is due to insufficient characterization of LAV because the virus has not yet been transmitted to a permanently growing cell line for pure isolation”

## Ownership of Data (& Intellectual Property)



## Ownership of Data (& Intellectual Property)



## Ownership of Data (& Intellectual Property)

### Law

- Research data usually belong to the institution or jointly by the institution and researcher(s)

### Norms

- Individual researchers are generally permitted to have a copy of the study data (after subject identifiers are removed)

## Ownership of Data (& Intellectual Property)

### Use of data



## Sharing of Data (& Materials)



“An author’s obligation is not only to release data and materials to enable others to verify or replicate published findings (as journals already implicitly or explicitly require) but also to provide them in a form on which other scientists can build with further research.”

- NATIONAL RESEARCH COUNCIL OF *THE NATIONAL ACADEMIES*

### UPSIDE

“the uniform principle for sharing integral data and materials expeditiously”

## Sharing of Data (& Materials)

### Corollary Principles & Recommendations

- In publication, include whatever data is necessary to support claims and enable others to replicate them. If impractical, they should be accessible by other means
- Use publicly accessible repositories, if available, by the time of publication.
- Anticipate which materials are likely to be requested and indicate how to obtain them.
- If integral material is patented, provide under a licence for research use.
- It is not acceptable for provider of material to demand an exclusive license to commercialize a new substance from the material or to require collaboration or coauthorship

## Sharing of Data (& Materials)

### Corollary Principles & Recommendations

- Transfer of requested materials or data should occur within 60 days.
- If this doesn't happen, it is acceptable for the requestor to contact the journal, then the author's institution or funder.
- Recipients should acknowledge the source in publications

