Scientific misconduct continues to be a serious and ongoing problem in the biomedical research community

Since 1994, there has been an average of two misconduct cases that have been examined by Inquiry, and in some cases, Investigation Committees in the NIH Intramural Research Program EACH year.

Recently a finding of misconduct in the extramural community resulted in a 366-day Federal prison term for a scientist because his actions led to loss of government funds, obstruction of justice, and abuse of a position of trust. The sentenced scientist had the following explanation for his actions:

"First, I believed that because the research questions I had framed were legitimate and worthy of study, it was okay to misrepresent "minor" pieces of data to increase the odds that the grant would be awarded to UVM and the work I proposed could be done. Second, the structure at UVM created pressures which I should have, but was not able to, stand up to. Being an academic in a medical school setting, I saw my job and my laboratory as expendable if I were not able to produce. Many aspects of my laboratory, including salaries of the technicians and lab workers, depended on my ability to obtain grants for the university. I convinced myself that the responsibility I felt for these individuals, the stress associated with that responsibility, and my passion and personal ambition justified "cutting corners". Third, I cannot deny that I was also motivated by my own desire to advance as a respected scientist because I wanted to be recognized as an important contributor in a field I was committed to."

Underlying this case was the issue of data management and the detection by one of the scientist's staff of inappropriate data management. He admitted to destruction of electronic evidence of his falsifications and fabrications, among other things. Scientific misconduct is detrimental to all parties involved. Everyone in a lab has a responsibility to be informed and vigilant about appropriate data management to prevent instances of scientific misconduct.

Several of the following cases are based on actual misconduct cases.

Case 1 - Data Management of Computer-generated Files

Case 2 – Handling of Images and Graphs

Case 3 – Appropriate Use of Statistics

Case 4 – Appropriate Sources of Data and Decision to Publish

Case 5 – Handling of Clinical Data

Comments and Guidelines derived from the Cases

Case 1 - Data Management of Computer-generated Files

Dr. Fred has been at the NIH for three years and is anxious because he has not published a paper and wants to begin looking for a job. He was finally able to purify all of the mutant genes needed for his analysis and recently presented a draft of a manuscript to Dr. Wilma, his mentor. Dr. Wilma found the data interesting but wanted to see the original scans for Table 1, which supposedly were carried out in February 2006. Dr. Fred could not produce a copy of the original scans because after he received a warning that his folder was full, he inadvertently deleted the data on the lab computer associated with the spectrophotometer. As a result, he had only an Excel table with the numbers he had written down from the plots produced by the lab computer. Dr. Wilma was able to obtain a backup copy of all the February scans that had been backed up on the institute's server, but none of these files contained data corresponding to the numbers in the Excel file. Since he could not produce the missing data, Dr. Fred carefully repeated the experiments and showed Dr. Wilma that all numbers were consistent with the original Table 1. Two pages from Dr. Fred's lab notebook covering February 2006 are attached.

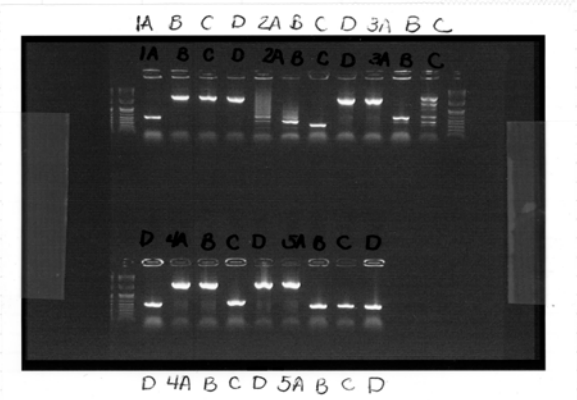
Does this case represent scientific misconduct?

- Are there any problems regarding data management, and if so, what are they?
- Who is at fault?
- How could this case have been prevented?
- Can you show all of the primary data for each experiment that you performed a year ago?
- Could someone reproduce the details of your experiments from your lab notebook?
- What are the elements of a good lab notebook?

ligation

insert	3
vector	1
5X	2
ligase	1
	3
	10

checked clones



purified :

1B
1C
2D
3A
4A
4B
5A

1B	5.4	10.6	57.2		
1C	5.6	11.2	62.7		
2D	5.9	50.3	296.8		
3A	6.1	1.1	6.7		
4A	6.2	9.0	55.8		
4B	5.5	9.0	49.5		
5A	5.7	101.1	576.3		

2-17

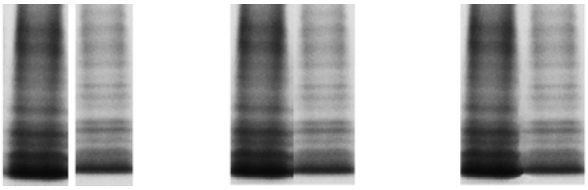
1B	5.4	110.1	594.5		
1C	5.6	10.7	59.9		
2D	5.9	53.4	315.1		
3A	6.1	0.9	5.5		
4A	6.2	8.9	55.2		
4B	5.5	9.2	50.6		
5A	5.7	102.8	586.0		

1B	6.4	9.5	60.8		
1C	6.2	10.9	67.6		
2D	5.9	54.0	318.6		
3A	6.1	0.8	4.9		
4A	6.3	9.1	57.3		
4B	6.2	8.9	55.2		
5A	5.8	105.6	612.5		

Case 2 – Handling of Images and Graphs

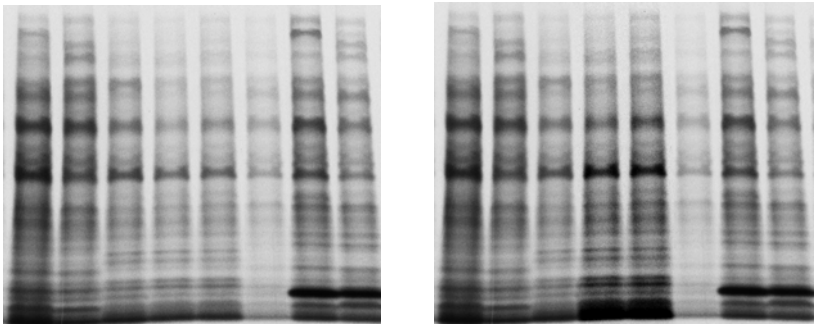
Dr. Green is preparing her manuscript for submission to a high-impact journal and is trying to decide the most effective way to present her gel and image data. Colleagues tell her that the data need to be "impressive" and "clean." She comes to you for advice about which of the following versions would be the most effective presentation of her data.

A. Spliced lanes from different parts of the same gel.



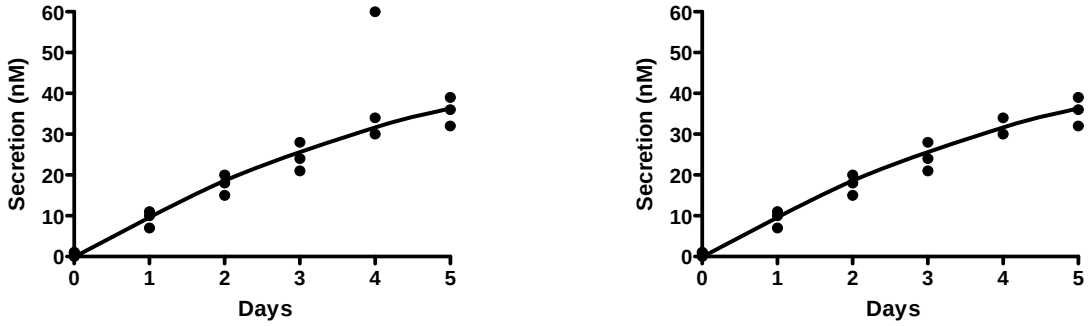
Originals Joined Blurred

B. Lanes enhanced for emphasis.



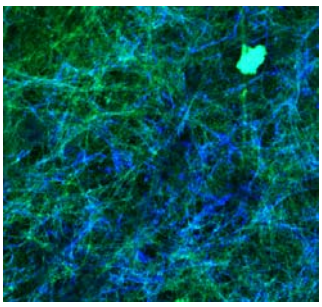
Original Two lanes enhanced

C. Deletion of outlying point.

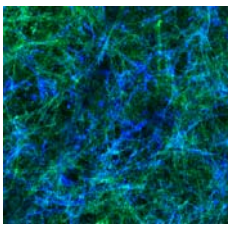


Original data and graph Anomalous point deleted

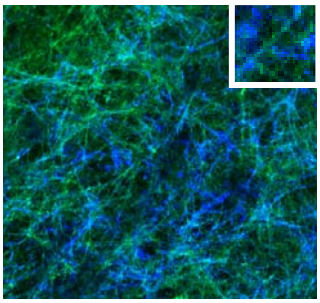
D. Cropping and cosmetic fixes.



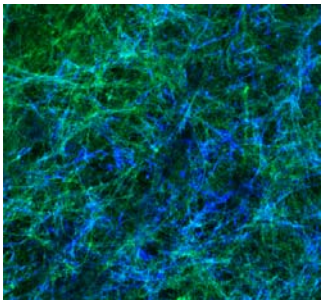
Original with “junk”



Cropped



Hidden by inset



Fixed using Photoshop

- What is your advice?
- Which changes are acceptable, borderline, or inappropriate?
- How do you choose the findings that you actually publish?

Dr. Green raises the issue that she has had two experiments that “worked,” i.e. both showed that stress increases synthesis of the protein stimulin considerably more than it increases borin. However, an earlier experiment had shown the opposite.

- What should you advise her?
- How do you decide which of multiple varying results or experiments to trust?

Case 3 – Appropriate Use of Statistics

After years of research, Dr. Venable, a gerontologist, had developed a hypothesis that dementia is strongly correlated with fish consumption. He designed a prospective, longitudinal clinical study to test this hypothesis. The study was to follow the health histories of two groups of patients over the age of 40, one of which ate fish at least three times a week, and the other essentially never ate fish. The groups were balanced for gender, race/ethnicity, and socioeconomic status, and chosen to exclude such confounding factors as smoking, substance abuse, and the use of dietary supplements. He submitted a proposal to the NIA for funding which included only one scientific aim: to evaluate the relationship between fish consumption and dementia. The only specified primary outcome measures involved assessments of cognitive function. The study section was enthusiastic about the proposal, but insisted, especially given the expense of the project, that he track the incidence not only of dementia but also of several other common disorders manifest in the geriatric population. Although Dr. Venable had no reason to believe that any of these other conditions were affected by fish consumption, he did modify his study and the revised application was funded upon resubmission.

- Was it appropriate for Dr. Venable to alter his study design in response to the study section's recommendations?
- Given that he agreed to their suggestions, what changes, if any, should Dr. Venable make to the study's scientific aims, specified outcome measures, and statistical analysis plan?

Dr. Venable's study reached its first analytic time point after five years. Using a standard statistical package, his staff calculated the following P values for positive or negative association of fish consumption with the 15 conditions evaluated:

Disease Entity	**P Value**
Dementia	0.013
Leukemia	0.78
Lymphoma	0.86
Colon Cancer	0.87
Breast Cancer	0.46
Lung Cancer	0.23
Suicide	0.0087
Hypertension	0.93
Obesity	0.67
Osteoporosis	0.18
Arthritis	0.16
Glaucoma	0.61
Nephritis	0.50
Cardiovascular Disease	0.23
Diabetes	0.51

Around this time, Dr. Venable accepted an invitation to be a keynote speaker at an international gerontology conference and met with his lab a week prior to the meeting to discuss the presentation of their long-awaited results. Dr. Gray, an assistant professor who had collected many of the case medical records, said, "The timing of this meeting is great. We'll be able to report strong evidence for a positive correlation of fish consumption with both dementia and suicide."

"Wait a minute," said Dr. Oldham, a first-year postdoctoral fellow who had performed the statistical analyses. "The P values I gave you have not been corrected for the multiple hypotheses tested, without which they cannot be validly interpreted. For example, the Bonferroni correction requires that a P value of 0.05 be divided by 15, i.e. 0.0033, for any of these results to be considered statistically significant. I'm afraid you can't report any of the clinical findings from this study as being statistically significant."

- Is Dr. Oldham correct in insisting that the P values be corrected for multiple comparisons?

"Look," said Dr. Venable, "I don't know about suicide, but my original hypothesis concerned only dementia. I set out to prove a correlation between fish consumption and dementia, and I proved it. I want to report at least that at the meeting. Including the other clinical outcomes was based on the recommendation of the study section's comments and there was no a priori hypothesis regarding their relationship to fish consumption. We can discuss the statistical analyses of the other data when we prepare the results for publication, but for now, please go back and re-analyze the data focusing on the findings for dementia so that I can show an impressive slide at the conference."

- Is Dr. Venable making a reasonable request?
- Should Dr. Oldham re-analyze the data without correcting for multiple comparisons, on the grounds that Dr. Venable's original specific aim involved only dementia?
- Should she try other less stringent correction algorithms for multiple comparisons until she finds one that yields a significant P value for the dementia correlation?

Dr. Oldham provides Dr. Venable with slides showing a significant relationship between fish consumption and dementia, which he proudly presents at the conference.

- Does Dr. Venable have an obligation to mention the other clinical outcomes included in his study, and whether and how he corrected for multiple comparisons?
- Does he have similar or different obligations for disclosure in the peer-reviewed publication?
- To what extent does this case raise issues of honest disagreement over statistical methods, rather than issues of scientific misconduct?

Case 4 – Appropriate Sources of Data and Decision to Publish

Joe Smith is a graduate student working on an M.P.H. degree with Dr. Sampler, a famous epidemiologist with an appointment at the NIH and an adjunct professorship at Jensen University. His research includes epidemiological studies of human exposure to air-borne asbestos in dust clouds that result from building demolitions. At one well-publicized destruction of a large stadium, thousands of people witnessed the event. Because of the light prevailing winds that day, the resulting dust cloud slowly drifted south over a neighborhood of nearly 4,000 residents while to the north a similar neighborhood, being upwind, remained dust-free. Mr. Smith proposed to Dr. Sampler that he analyze the exposure data derived from this one event as his master's thesis research, and Dr. Sampler agreed.

After several months, Mr. Smith came to Dr. Sampler with the preliminary results of his research: in the six months following the demolition, there was a 2-fold higher incidence of school absences among children aged 5 to 8 on the downwind (exposed) side of the old stadium compared to those on the upwind (unexposed) side. Dr. Sampler asked Mr. Smith if there was any way to confirm that the school absences were related to upper-respiratory illnesses. Mr. Smith contacted Stella Seller, a friend who works in the purchasing department of CTS, the largest retail drugstore chain in that area. Mr. Smith asked Ms. Seller if she could review the chain's records of sales of children's nasal decongestants in the months immediately prior to and following the demolition, to see if the sales at upwind and downwind stores were different. Ms. Seller provided him with a print-out that showed month-to-month changes in sales volumes for the three top-selling decongestants, for six upwind and five downwind stores. For corporate confidentiality reasons, Ms. Seller provided only aggregate sales data expressed as percent change in sales from the same period a year earlier. Nevertheless, the sales data showed a statistically significant increase in decongestant sales at the downwind stores in the three months following the demolition.

Mr. Smith proposes to Dr. Sampler that he begin writing up his thesis, so that he can graduate in June and enter medical school in the fall. He also presents Dr. Sampler with a rough draft of a manuscript he is thinking of submitting to a journal. Dr. Sampler agrees that Mr. Smith can write the thesis, but says that a publication would be premature before the study can be confirmed by analyses of similar data from other demolitions that Dr. Sampler and colleagues at Jensen University are working on. Their study will not be completed for another 2-3 years however. When Mr. Smith tells Ms. Seller about the analysis and his plans to publish the work, she tells him that she should be a co-author on the paper and that a lawyer at CTS will need to review the paper prior to publication.

- Should Mr. Smith complete his thesis even if the paper cannot be written at this time?
- Is Dr. Sampler justified in requiring that Mr. Smith wait two or three years for confirmatory data from other demolitions before publication? Does the length of time (2-3 years) he might have to wait influence the decision?

- Was it appropriate for Ms. Seller to provide data to Mr. Smith? Should Ms. Seller be a co-author on the paper?
- What rights does CTS have with respect to publication of the paper?

Case 5 – Handling of Clinical Data

Dr. Bob is a promising junior faculty member at Z University. His major clinical research project, funded by an NIH grant, is a prospective, longitudinal study of changes over time in plasma levels of protein X and their association with cardiovascular disease. Previous cross-sectional studies by others have suggested that protein X levels increase with age and are associated with increased risk of cardiovascular disease. Dr. Bob's is the first longitudinal study to address this issue. A successful study would be publishable in a high-impact journal and give a substantial boost to his achieving tenure.

Dr. Miriam, a resident at the Z University Medical School, approaches Dr. Bob for advice about a research career and he offers to let her help analyze data from the first 3 time points of his protein X study. She eagerly accepts this offer as an opportunity to gain research experience and perhaps co-authorship on a high-impact paper.

- When is it appropriate for Dr. Miriam to discuss her authorship status with Dr. Bob? Should she raise the issue now, before agreeing to analyze the data, or wait until after the results are known?

Dr. Bob gives Dr. Miriam an Excel spreadsheet which he describes as containing all the relevant data from study subjects. Dr. Miriam performs a statistical analysis, but her results are not consistent with the hypothesis Dr. Bob wrote in his grant application. Protein X levels appear unchanged over time, and there is no association with cardiovascular risk. When Dr. Miriam presents her analysis to Dr. Bob, he is noncommittal. He says he will take the Excel spreadsheet home with him over the weekend to check her work. The next week, Dr. Bob returns the spreadsheet to Dr. Miriam, explaining that he has corrected a few mistaken data entries. He asks her to re-do the analysis.

- Is it appropriate for Dr. Bob to take the clinical data home with him?
- Would it make a difference whether or not the Excel spreadsheet contained personally identifiable information about the research subjects?

When Dr. Miriam reanalyzed the data, the hypothesis was confirmed. However, she was puzzled that correction of "a few mistaken data entries" would so substantially change the outcome of the analysis. She compared the "corrected" spreadsheet with the study's case report forms and found that the majority of data entries had been changed, always in the direction consistent with the hypothesis.

- Is it appropriate for Dr. Miriam to check the new spreadsheet against the case report forms?
- Should she have accepted Dr. Bob's corrections and confined herself to the re-analysis? Under what circumstances would one check a transcribed or secondary data set against the primary or source data?

When Dr. Miriam presented the data discrepancies to Dr. Bob and asked to see the original patient files, he brushed this off as unnecessary. He blamed the apparent discrepancies on his own ineptitude with Excel and on his use of imputed data (i.e., data entries derived from a statistical model, rather than actual measurements). Concerned about the situation, Dr. Miriam began reviewing patient records on her own without telling anyone. To her dismay, she found that many data entries in the spreadsheet had been changed from their true values, that some data entries did not correspond to actual measurements, and that some patients recorded as participating in the study did not actually exist.

- Is Dr. Bob's explanation of the data discrepancies justifiable? When is it appropriate to mix measurement-derived data with imputed data in the same data set?
- Is it appropriate in this context for Dr. Miriam to access patient records? Should she first have shared her concerns with someone in authority and gotten permission?
- Does this situation represent scientific misconduct? If so, what type of misconduct is it?

Dr. Miriam continued to work with Dr. Bob while she searched for a new mentor, but did not tell him of her findings. She did share her concerns with one of Dr. Bob's former fellows and with a collaborating faculty member in his department. Dr. Bob learned of Dr. Miriam's questioning of his scientific integrity and stopped working with her. In response, Dr. Miriam lodged a formal complaint of scientific misconduct against Dr. Bob with the university.

- Should Dr. Miriam have shared her concerns with others without first talking with Dr. Bob or lodging a formal complaint?
- What other steps could she have taken before lodging a complaint? When would have been the best time to lodge a formal complaint of scientific misconduct?

Dr. Bob was eventually convicted of scientific misconduct and resigned from his faculty position. NIH demanded repayment of the grant money that funded his study. Several patients who participated in the study felt exploited because they were exposed to risk without any balancing scientific gain.

- What factors might have motivated Dr. Bob to commit scientific misconduct?
- What obligation does Dr. Bob have toward the NIH? What ethical or legal obligations does he have toward the patients in his study?

Comments and Guidelines derived from the Cases

- The honest and accurate presentation of scientific findings is the most important thing a scientist can do. The illustrations must provide an accurate representation of the data obtained. The consequences of misrepresenting data far outweigh the short-term gains.

- Many recent cases of scientific misconduct in both the intramural and extramural programs involve inappropriate data manipulation using programs such as Photoshop, or inappropriate statistical analysis. As a result, journals now analyze images to detect inappropriate manipulations or send manuscripts out for separate statistical review.

- Changes in brightness, contrast, etc. should be applied simultaneously to all panels in a figure, including positive and negative controls. Attention must be paid to avoid saturating the brightest details and to avoid changing the relative brightness of different areas. Be aware that any change to an image has implications.

- Parts of images or graphical data should not be arbitrarily modified.

- Combining of gel lanes should be indicated by a break or line in the image.

- For safety, two copies/versions of data should be kept (original + figure version, two hard copies, hard copy + scan, computer file + backup, etc.).

- For digital images, the original data file must always be kept, with its original name (as recorded in a notebook); subsequent modified versions, and versions finalized for publication must be maintained as separate files.

- The scientific integrity and credibility of clinical trial data depend on a sound trial design with clearly identified primary and secondary endpoints and a description of statistical methods to be employed. This is a requirement for clinical studies under the jurisdiction of the FDA.