2025 Annual Ethics Cases: Good Recordkeeping, Responsible AI Use, and Laboratory Safety

The Committee on Scientific Conduct and Ethics (CSCE) has prepared three cases for 2025 that deal with some important topics relating to good recordkeeping and reproducibility, use of AI in writing official and clinical documents and maintaining laboratory safety. These include:

Case 1: Recordkeeping, Reproducibility, and Authorship

Case 2: Responsible use of AI in Preparing IRB Submissions

Case 3: Laboratory Safety as a Shared Responsibility

Since it may not be possible to cover all three cases in the allotted time, we suggest that facilitators cover the cases that meet the needs and interests of their audience.

CONTENTS

Case 1: Recordkeeping, Reproducibility, and Authorship

After five years of work on a project in an NIH research group, a postdoctoral fellow ("postdoc") wrote a manuscript to report the exciting discovery of a novel human cancer protein. The plan was to submit the manuscript to a high-visibility journal with the postdoc as first author and with the postdoc's Principal Investigator (PI) and a clinical collaborator as co-authors. With a job opportunity in hand, the postdoc departed the NIH in a rush, leaving the draft manuscript with the PI to submit.

The PI appointed a postbaccalaureate fellow ("postbac") to the research group shortly after the postdoc left, with the intent of continuing the postdoc's project. The PI asked the postbac to reproduce one of the key experiments from the draft manuscript. Unfortunately, examination of the postdoc's incomplete electronic lab notebook and clinical files made it difficult for the postbac to understand some of the details from the initial experiments, and the methods section of the manuscript was quite brief. The postbac, who had never performed research full-time previously, asked the PI what to do. The PI, looking through the lab notebooks more closely, found the level of information entered to be minimal and wondered if there was supporting information stored elsewhere in the lab. Although the PI had felt confident about providing proper instruction in good recordkeeping practices to lab members, the PI did not actually review the records regularly and in this case, regretted not having reviewed them prior to the postdoc's departure. The PI emailed the postdoc to ask for any possible additional information, but the postdoc replied that everything necessary had been entered into the electronic lab notebooks that each lab member had been assigned in June 2024. The postdoc noted that there were some paper notes from before 2024, along with more recent scrap paper in a drawer in the lab that might provide some additional insight, but that the postdoc did not think those rose to the level of inclusion in the electronic lab notebook.

The postbac found the paper notes and, together with the PI, pieced together what had been done experimentally. The postbac repeated the experiment and found results that were similar to those in the manuscript, but with some key differences. The PI, concerned about the reproducibility of the data, asked the postbac to repeat the other experiments as well. The PI, now less confident in the data generated by the postdoc, closely examined the manuscript figures, found some suspicious duplications within a key image and feared that the data might be unreliable. Over several months, the postbac was able to successfully recreate all the postdoc's experiments. While the main scientific story had not changed, the PI decided to replace the figures from the original manuscript with those generated by the postbac, as they had clearer supporting details recorded in an electronic lab notebook. The postbac then helped to edit the manuscript, rewriting some sections. The postdoc, who was concerned that there had not yet been an update about the manuscript being submitted, emailed the PI to ask if it was going to be submitted soon. The PI had doubts about whether the postdoc should still be first author (or even an author at all) anymore because the updated manuscript did not contain any of the postdoc's original data, but the

PI did not want to upset the postdoc, so the PI replied that the group was still working on the edits to the manuscript.

Questions for Case #1 discussion (with facilitator notes):

1. Why is it important to keep highly detailed records of laboratory and clinical research data and analyses – which should have been electronic at the NIH since June 2024?

   Detailed research records allow scientists to understand how experiments were conducted and to replicate them accurately. Maintaining excellent research records is essential not only for experimental reproducibility but also for assisting new group members in completing projects that rely on data initially gathered by former group members who have since moved on to the next stage of their careers. Good record keeping may help defend against false allegations of potential research misconduct or provide evidence to support one's claim to warrant authorship in an authorship conflict resolution process. For a more detailed discussion on the requirements and advantages of good record keeping please refer to the "Scientific record keeping" chapter of the Guidelines and Policies for the Conduct of Research document.

2. Is it acceptable for a fellow to decide that some notes do not need to be included in the official lab notebook? How detailed does a lab notebook need to be?

   Each group member must exercise careful judgment about what should be included in the official lab notebook, as the experimentalist is best equipped to determine which details are crucial for accurate replication. Every scientist should view the lab notebook from the perspective of someone unfamiliar with the experiment and consider whether there is sufficient information for that person to replicate the experiment accurately.

3. Is it unusual for a new group member to be unable to replicate previous experimental results?

   This response will vary based on individual experiences in attempting or observing others' attempts to replicate experimental results. It will likely depend on many factors, including but not limited to the quality of the training and recordkeeping within the research group, the seniority and experience of the scientists, the complexity of the experiments, and the reliability of the reagents and instruments.

4. Should the postdoc be the first author or a co-author on the new version of the paper?

   Because the postdoc had written the original version of the manuscript, based on NIH General Guidelines for Authorship Contributions, drafting the manuscript warrants first authorship. However, if the manuscript no longer contains any original data from the postdoc and if the postbac had rewritten the manuscript so substantially that in essence, an entirely new manuscript had been drafted, then the PI may consider the manuscript to be a new paper. If that were the PI's decision, it would be likely to be a contentious decision, as the postdoc had presumably conceptualized many of the experiments and therefore merited authorship with original ideas, planning, developing protocols and reagents, and other intellectual contributions to the project. The postbac could deserve first authorship if the postbac's contribution surpassed that of the postdoc's. Alternatively, the postdoc and the postbac could be co-first authors on the manuscript based on the extent of their contribution and the authorship guidelines. All co-authors must take responsibility for the full manuscript and agree to its submission, which can be a challenge if there are fundamental disagreements about authorship.

5. Could the postdoc's low quality of recordkeeping lead to suspicions of research misconduct? How should the PI respond in this case?

   The low quality of the postdoc's recordkeeping is unlikely to directly lead to suspicions of research misconduct, unless it appeared that the experiments performed by the postdoc could not actually have been conducted as reported. However, when faced with suspicious data, the PI is more likely to be reassured of the quality of the results and integrity of the experimentalist when the work has been responsibly documented. On the other hand, if the PI confirmed the presence of "suspicious duplications within a key image," a consultation with the NIH Agency Intramural Research Integrity Officer (AIRIO) would be warranted.

6. Was it fair for the PI to ask a new lab member, especially a postbac, reproduce experiments in a manuscript that is ready to be submitted to a high visibility journal? What kind of pressure does this situation put on the postbac?

   While any new group member will feel pressure to replicate the results of a previous group member, more junior group members may feel increased pressure, especially if they are just learning a new technique for the replication effort. In addition to the typical pressure of wanting to appear competent and

not wanting to contradict an established group member, the fact that the paper is ready to be submitted to a journal only creates additional perceived consequences for the postbac if the results are not consistent with those in the draft manuscript. In this particular case, some of this pressure could have been alleviated if the postbac had robust research records to refer to when starting the work.

7. What could each person have done differently to prevent the problems that occurred in this case?

   The PI could have kept a closer eye on the recordkeeping in the research group and asked a different group member, perhaps a staff member, to repeat the experiments. The PI should also have consulted appropriate resources regarding authorship and provided transparent updates to the postdoc about the status of the manuscript, including any authorship changes. The postdoc should have kept better research records, in a single, agreed-upon location, and asked questions if unsure about best practices; the postdoc should also have considered the importance of avoiding potential misconduct in the preparation of manuscript figures. The PI and the postdoc should have reviewed the postdoc's records before departure. This is particularly important since most journals currently require the authors to submit the underlying raw data, including an exhaustive methods and materials section in the supplemental information section. This is especially relevant in the context of the recent policies on open science and data sharing. The postbac could express concerns to the PI about the pressure of replicating these experiments and consider asking if there is another, more appropriate group member to perform the replication efforts.

   The postdoc could have asked to take a copy of all data relevant to the manuscript before leaving, in case questions came up before submission or during revisions. In that case, the postdoc would have needed to complete an NIH Form 3000 at least 45 days before departure.

   Please refer to the NIH IRP Authorship Conflict Resolution policy for more information about authorship and conflict resolution.

[End of case study #1]

Please take the survey by either clicking on the link below or scanning the QR code on your hand-held device: https://www.surveymonkey.com/r/M68XD9V

Case 2: Responsible use of AI in Preparing IRB Submissions

A busy, understaffed assistant clinical investigator decides to take advantage of the NIH's AI tool, ChIRP (Chatbot for the Intramural Research Program), to write initial drafts of documents for a submission to the Institutional Review Board (IRB), including the protocol and the informed consent form. The investigator shares these documents with the clinical research coordinator (CRC) and asks the CRC to review and edit them for IRB submission. The investigator informs the CRC that ChIRP wrote these initial drafts, but neither the investigator nor the CRC disclose the use of the AI the scientific review committee. The study is a Phase II trial of an off-label use of commercially available blood pressure medication to treat symptoms of an enlarged prostate. Two months after the study begins, a participant dies from renal failure. An investigation into the incident determines that participant's death was due to an adverse interaction between the study medication and an over-the-counter medication that the participant was taking but did not inform the investigator about. The adverse drug interaction was reported only in studies published in the last year, which was beyond the range of the AI's training data. Since the investigator relied solely on the AI to perform the literature review included in the protocol and other documents submitted to the IRB, the literature did not mention these drug interactions. Furthermore, the consent form did not include any questions regarding the concomitant use of other drugs while the patients were in the study trial. Following an investigation into the incident, the clinical investigator admits to using the AI to perform the literature review and write an initial draft of the protocol and informed consent form and takes full responsibility for this tragedy.

[Proceed to next page]

Questions for Case #2 discussion (with facilitator notes):

1. Is it ethically permissible to significantly rely on or even use AI to help write documents, such as the protocol or informed consent form, submitted for IRB review?

   AI large language models (LLMs) (such as ChIRP) are useful tools that can assist with writing and editing scientific articles, reports, memos, and other documents. With respect to IRB submissions, an LLM could be helpful in rewriting technical language used in consent documents to make it more understandable to research participants or in editing a protocol for clarity and organization. However, LLMs are susceptible to various problems, including fabrication of facts or citations, omission of important information, skewed representation of information, and potential plagiarism. Because of these problems, while it is acceptable to use AI systems to help write or edit documents submitted for IRB review, researchers should not fully rely on them.

2. What are the investigator's and CRC's responsibilities when using AI to help prepare an IRB submission?

   Carefully review the work of the AI, disclose substantial use of the AI, and take full responsibility for its output. The NIH IRB does not require disclosure for use of AI in preparation of protocols and other IRB documents, and disclosure requirements from ICs and Scientific Review Committees (SRC) may vary. However, the NIH Guidelines for the Conduct of Research (2025 revision) include guidance on responsible AI use. The guidelines define substantial use of AI. Using AI to write text is a substantial use. See, NIH Guidelines on the Conduct of Research, chapter on information technology:
   https://oir.nih.gov/sites/default/files/uploads/sourcebook/documents/ethical_conduct/guidelines-conduct_research.pdf

3. Should the investigator and CRC have disclosed the use of the AI to the scientific review committee?

   According to the NIH Guidelines for the Conduct of Research, if AI was used to substantially write these documents, then its use should have been disclosed to the SRC (if required by the IC). Informing the committee about the use of the AI could have helped them to be on the alert for potential problems with the submission, such as a literature review not being up to date. Although the investigator is ultimately responsible for disclosure, the CRC also has some responsibility to disclose, since the CRC assisted with the submission. The CRC should have reminded the investigator that the use of the AI needs to be disclosed.

4. What are the responsibilities of the scientific review committee and/or the IRB in this case?

The Scientific Review Committee is responsible for ensuring that the research is well-designed to address important scientific questions. The IRB is responsible for ensuring that the research complies with ethical and legal requirements. It is possible that the Scientific Review Committee failed to meet its responsibilities because the protocol did not address risks of the drug interactions that had been identified in the recent literature. If this information had been in the protocol, then the IRB would be able to make sure that it would also be included in the consent form and considered when reviewing the study.

5. What sorts of uses of AI are permissible and useful at NIH?

   To analyze data and draft, translate and edits documents. [Note: this could be a very general discussion that could go in various directions, depending on the audience.] Please refer to the guidance provided by NIH Office of Science Policy for more information - https://osp.od.nih.gov/policies/artificial-intelligence/

6. What types of uses of AI within NIH should be disclosed?

   Substantial uses. The guidelines define substantial use of AI. Using AI to write text or analyze data are substantial uses. See, NIH Guidelines on the Conduct of Research, chapter on information technology: https://oir.nih.gov/sites/default/files/uploads/sourcebook/documents/ethical_conduct/guidelines-conduct_research.pdf

7. Do you think people are hesitant to disclose the use of AI in drafting documents?

The Conduct of Research guidelines provide some suggested disclosure language – "*AI Usage Disclosure: This document was created with assistance from AI tools. The content has been reviewed, edited, and approved by all authors. For more information on the extent and nature of AI usage, please contact the author(s).*" [1]

[End of case study #2]

Please take the survey by either clicking on the link below or scanning the QR code on your hand-held device: https://www.surveymonkey.com/r/M68XD9V

[1] Why We Should Normalize Open Disclosure of AI Use. Retrieved from https://www.chronicle.com/article/why-we-should-normalize-open-disclosure-of-ai-use

Case 3: Laboratory Safety as a Shared Responsibility

Dr. C's pathology lab is being moved to a new building. Dr. C is not happy about this and complains about it to other members of the lab repeatedly. Dr. C is especially upset about the disruption in work that will occur and loss of time to perform important experiments needed to get a paper (returned for revisions) accepted in a high-impact journal, which will help with Dr. C's upcoming tenure review. The moving day arrives, and members of the lab begin unpacking boxes in the new location. A senior postdoc asks for a boxcutter to help open boxes. After looking for a boxcutter but failing to find one, the postdoc decides to use a scalpel to open boxes. On the fourth box, the scalpel slips and cuts the postdoc's thumb, which starts bleeding profusely. A technician quickly grabs some tissues and uses them to apply pressure to the thumb to control the bleeding. The technician and postdoc wash the thumb with water from the laboratory sink and pour some alcohol on it to sterilize the wound. The technician also creates a make-shift bandage for the thumb. Dr. C enters the room and says, "I heard someone cry out. Is everything OK in here?" The postdoc responds, "I just nicked my thumb while opening a box. We bandaged it and it is all good". Dr. C reminds the postdoc that, according to the annual safety training they both took recently, the postdoc should contact OMS to report it. The postdoc, who did not pay close attention to the training, looks at the pile of boxes remaining to be unpacked and decides not to report the incident to Occupational Health. Dr. C, who does not see the postdoc's bandaged thumb, returns to the office.

[Proceed to next page]

Questions for Case #3 discussion (with facilitator notes):

1. How could this accident have been avoided?

   By making sure that everyone had the proper equipment and training for the move (e.g., boxcutters), by not putting pressure on people to unpack quickly, and by promoting a safety culture.

2. What are some of the factors that contributed to this accident?

   Lack of a safety culture was a major factor in this case. As the PI, Dr. C is ultimately responsible for promoting a safety culture. However, everyone in the lab plays a role. The postdoc should have explained to the PI the seriousness of the cut. They all should have reported it to Occupational Health. Since they all took safety training, one wonders whether they took it seriously. Additional factors in this case include the lab environment more generally. The PI was facing pressure (e.g. paper review, tenure decision, moving the lab), which was sensed by others in the lab, leading them to also feel under pressure.

3. Why in your opinion didn't the postdoc tell Dr. C exactly what happened?

   It's likely that the postdoc was concerned that telling Dr. C what happened would make Dr. C feel even more stressed out. People in the lab should feel like they can openly discuss important scientific issues, including safety concerns and incidents. Note: this could lead into a discussion of the importance of having an open lab culture and factors that can help promote it or hinder it.

4. Was the postdoc responding to pressure from Dr. C, whether real or perceived?

   Probably. However, even though the postdoc was feeling some pressure, the postdoc should have been aware of the reporting responsibilities, especially since the postdoc was the senior person in the lab and had taken the refresher training recently. This could have been avoided especially after being reminded by Dr. C of the reporting requirement, if the postdoc had paid attention to the training.

5. Was it appropriate to not report this incident to Occupational Health?

   No, According to Occupational Health and Safety, "All work-related injuries and illnesses at the NIH must be reported to your local OMS." and "When a work-related injury or illness occurs it is important that we investigate why it happened so we can suggest changes to prevent a similar incident from happening again. Most injuries and illnesses are preventable. The purpose of our discussion is not to assign blame, but to identify contributing factors which

can then be controlled. Similar incidents can be avoided in the future by identifying the factors that led to the incident and then changing the conditions or actions."
https://ors.od.nih.gov/sr/dohs/HealthAndWellness/OccupationalMedical/Pages/Work-Related-Injuries-and-Illnesses.aspx

https://ors.od.nih.gov/sr/dohs/safety/incidents_accidents/Pages/Work-Related-Injury-and-Illness-Investigation.aspx"

If the technician came into contact with the postdoc's blood, this should also be reported. Even if the incident is minor, OMS should still be contacted to determine the appropriate response.

6. Who was responsible for reporting the incident to OMS?

   Each person in this story could have reported it. The postdoc was required to report the injury, the technician's role in providing first aid should have been reported, and Dr. C could have ensured that the lab staff visited OMS to be evaluated.

7. Was Dr. C promoting a safety culture in the lab?

   It's not easy to say without more information. Clearly, some people in the lab did not feel like they could share important information about an accident openly. However, as noted above, a safety culture is everyone's responsibility.

[End of case study #3]

Please take the survey by either clicking on the link below or scanning the QR code on your hand-held device: https://www.surveymonkey.com/r/M68XD9V