

2024 Annual Ethics Cases

The Committee on Scientific Conduct and Ethics (CSCE) has prepared three cases for 2024 that deal with some important topics relating to research with human subjects and using artificial intelligence in research. These include:

Case 1: IRB Protocol Deviation

Case 2: Using AI to Write a Manuscript

Case 3: Using AI to Analyze Research Data

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 #	Case Study	Page #
1	IRB Protocol Deviation	2
2	Using AI to Write a Manuscript	4
3	Using AI to Analyze Research Data	6-7

Please note that the CSCE recommends using a gender-neutral honorific title (Mx., pronounced “mix” or “mux”) and “they/them/their” pronouns in the cases, to try to reduce bias in the case discussions. We encourage facilitators to take advantage of negative comments from participants about this practice to promote a positive discussion about reducing genderism (bias resulting from a gender binary view) and unconscious bias in science.

IRB Protocol Deviation (Case #1)

Mx. Fox is a research nurse at the NIH who recently started working with Dr. Bear on an IRB-approved Phase III clinical trial comparing three different FDA-approved medications for treating mild-to-moderate depression. The study has exclusion criteria pertaining to various health measures, such as blood pressure, kidney and liver function, depression score (based on two metrics), and body mass index (BMI). Participants' BMI must not be greater than 30 kg/m². One day, Mx. Fox was reviewing the records of new patients on the study and noticed that Dr. Bear had enrolled a patient with a BMI of 31, which is a protocol deviation. Mx. Fox asked Dr. Bear about this, but they shrugged and told Mx. Fox not to worry about it because, in their professional judgment, the patient was healthy enough to participate in the study. Not wanting to cause any trouble, Mx. Fox tried to forget about the incident, but Dr. Bear did the same thing the following week. This time when Mx. Fox asked about the deviation, Dr. Bear became angry, grabbed Mx. Fox's wrist and told Mx. Fox that it is none of their business. During lunch at the cafeteria, Mx. Fox told Mx. Badger, another research nurse, what Dr. Bear had done. Mx. Badger responded: "You better get used to it. Dr. Bear does not tolerate people questioning their judgment."

Questions for Case #1 discussion

1. Should Mx. Fox follow the advice of Mx. Badger to "get used" to Dr. Bear's behavior? Why/Why not?
2. What are the potential consequences of not addressing the protocol deviation?
3. What resources are available to support Mx. Fox in navigating this situation?
4. Is Mx. Fox being disloyal to the research team? How do they balance staff loyalty with ethical responsibilities to study participants and the scientific community?
5. What steps can be taken to ensure the safety and well-being of participants enrolled in the study and ensure the validity and reliability of the data collected?
6. How might this case impact the trust and confidence of participants in clinical research at the NIH?

[End of case study #1]

Please take the survey by either clicking on the link below or scanning the QR code on your handheld device: <https://www.surveymonkey.com/r/JTDK6JN>



Using AI to Write a Manuscript (Case #2)

Dr. Blue is principal investigator at the NIH who specializes in cancer genotyping. A prestigious review journal has asked Dr. Blue to write an article reviewing the current state of the field. Dr. Blue is very busy with clinical, research, and administrative responsibilities, so they ask Dr. Green, a postdoctoral fellow working in the lab, to write the review. Without telling Dr. Blue, Dr. Green uses an artificial intelligence (AI) tool to summarize the literature on this topic and generate references. Dr. Blue reads the review and congratulates Dr. Green on a job well done. They submit the solicited review to the journal. The article lists Drs. Blue and Green as authors but does not acknowledge the use of the AI in preparing the article. Two months after publication, an anonymous critique of the article, appearing in a post-publication peer review blog, claims that two of the citations in the article are fake. The editors of the review journal inform Dr. Blue about this and ask them to submit a correction. Dr. Blue meets with Dr. Green about the issue and asks how the problem occurred. Dr. Green admits to using an AI tool to help write the article and says the tool must have made the mistakes. Dr. Blue is furious at Dr. Green for using this tool without consulting with them first. They both carefully examine the references and verify that the two references mentioned by the critic are indeed fake. They also discover that three additional references are inaccurate, three are irrelevant, and two sentences in the article are copied word-for-word from another article without quotation marks or attribution.

Questions for Case #2 discussion (with facilitator notes)

1. When Dr. Blue and Dr. Green submit their correction to the journal, should they also address the inaccurate and irrelevant references and the copied sentences and acknowledge the use of the AI tool?
2. Should they explain how the problem occurred, i.e., that the AI tool made the mistakes?
3. Should they retract the article?
4. Did they commit research misconduct, i.e., plagiarism?
5. What are the responsibilities of authors when using AI tools to review the literature?

[End of case study #2]

Please take the survey by either clicking on the link below or scanning the QR code on your handheld device: <https://www.surveymonkey.com/r/JTDK6JN>



Using AI to Analyze Data (Case #3)

Dr. Falcon, a postdoc in Dr. Hawk's research group, has struggled to analyze health survey and genomic data from a longitudinal NIH intramural research study with 10,000 human participants. Dr. Falcon wonders if they might be able to use artificial intelligence (AI) tools to help analyze the data. Dr. Falcon has an account for an NIH ChatGPT platform, but this version of ChatGPT does not have the functionality needed for this data analysis, so they sign up for a personal account with a commercial AI platform, HotBot1, which is able to analyze data from publicly accessible health databases that is similar to the IRP study data. Dr. Falcon uses HotBot1 to search for statistical relationships among dozens of variables from the public databases; however, Dr. Falcon soon realizes that to make significant progress, they need to supplement the publicly available data with additional, more detailed data. Fortunately, the IRP study includes the data needed to improve the analysis and HotBot1 allows users to upload data to the platform.

Dr. Falcon de-identifies the intramural study data so it includes no names or personal identifiers and uploads them to HotBot1. After several weeks of work, Dr. Falcon has some promising results, including a genetic association that could have important public health implications. Although the analysis appears to misrepresent findings for an underrepresented minority cohort of the data, Dr. Falcon is confident that the rest of the analysis is completely reliable. Dr. Falcon shares the results of this work with Dr. Hawk at their next regularly scheduled meeting and tells Dr. Hawk how HotBot1 was used to analyze both the public and intramural datasets together. While Dr. Hawk is not very familiar with AI tools, Dr. Hawk is excited about the new findings. They quickly draft a manuscript reporting the results of their data analysis and submit it for publication clearance review in their IC.

Questions for Case #3 discussion (with facilitator notes)

1. Has Dr. Falcon done anything wrong? If so, what actions should be taken to mitigate any mistakes?
2. Were the steps that Dr. Falcon took to protect NIH data sufficient? Has Dr. Falcon committed a data breach incident that should be reported?
3. How can scientists balance the need to develop their research program quickly with their lack of formal education in emerging technologies?
4. How could HotBot1 have made an error in analyzing the underrepresented minority cohort of the population? What are the implications of using the entire dataset despite the concerns? How could this problem have been anticipated or prevented?
5. In your opinion, is Dr. Hawk appropriately overseeing the research of Dr. Falcon? Should Dr. Hawk have been informed by Dr. Falcon that they were embarking on this exploratory path? Should Dr. Hawk delve more deeply into the work that Dr. Falcon did using HotBot1, or is it acceptable for Dr. Hawk to trust Dr. Falcon without independently verifying any of the analyses?

6. Is it ever acceptable to use personal credentials instead of official credentials to set up an account using an NIH computer to analyze data? If so, under what circumstances?
7. More generally, what types of AI tools are permissible to use in your research?

[End of case study #3]

Please take the survey by either clicking on the link below or scanning the QR code on your handheld device: <https://www.surveymonkey.com/r/JTDK6JN>

