The Ethics of Data Sharing

MODERATOR:

Jennifer N Workman
Editorial Client Manager
J&J Editorial, LLC
Cary, North Carolina

Jennifer Lee Mahar
Managing Editor
Origin Editorial, Massachusetts

SPEAKERS:
Shelley Stall
Sr. Director, Data Leadership
American Geophysical Union
Washington, DC

Matthew Cannon
Head of Open Research
Taylor & Francis Group
Abingdon, United Kingdom

Trevor Lane
Education Consultant, AsiaEdit
Council Member, Committee on
   Publication Ethics

REPORTER:
Simona Fernandes
Quality Lead
Enago, Mumbai

COPE CORE PRACTICE
Data and reproducibility

" Journals should include policies on **data availability** and encourage the use of **reporting guidelines** and **registration** of clinical trials and other study designs according to standard practice in their discipline "

Figure. One of COPE's core principles.

Promoting data sharing among the scientific community is important; it helps in the advancement of science in small increments rather than through single blockbuster studies. Data sharing ensures replicability and thereby helps confirm a study's findings. It accelerates the time taken to progress from one breakthrough to the next and reduces the time and costs required to gain confidence about a particular discovery. However, many authors are reluctant to share their data, and editors, publishers, societies, and individual journals need to be able to access these data to promote transparency and fair practices in their publications.

The panel comprised Shelley Stall, Matthew Cannon, and Trevor Lane. Using cases and polls, these experts shared examples of the ethics-related muddle one often finds themselves in when it comes to data sharing and responsibility. The importance of verifying data and investigating them in case of ethical issues was emphasized. The session proceedings will particularly benefit individuals in scholarly publishing who would like to learn more about editor responsibilities, investigations when ethical issues arise, and the verification process.

Trevor Lane opened this session by discussing data problems, the practices the Committee on Publication Ethics (COPE)[1] follows, data stewardship, and responsible data sharing and shared 2 interactive cases with the participants. Data fabrication, falsification, and plagiarism are the 3 primary areas of ethical misconduct. The Cooperation & Liaison between Universities & Editors,[2] COPE, and Responsibilities of Publishers, Agencies, Institutions, and Researchers in Protecting the Integrity of the Research Record[3] guidelines are important guidelines to debunk fabrication, falsification, and plagiarism. Questionable research practices (unauthorized data use, data censoring, fishing, 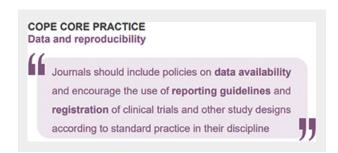hacking, etc.) and questionable publishing practices (image distortion, data misinterpretation, selective publication, salami slicing, etc.) are some of the commonly encountered data problems.

Authors need to exhibit data stewardship by adhering to laws and regulations, respecting the study subjects' consent and approval, and by themselves adhering to all ethics principles. Trevor discussed one of COPE's core practices and how these are required to attain the highest standards in publication ethics (Figure).

Next, Shelley Stall talked about the work she does at American Geophysical Union's (AGU's) Data Leadership Program. She highlighted that "data should be as open as possible, as closed as necessary." Until 2019, data were required to be cited in the paper so they could be preserved in a trusted repository with a proper identification tag. Although this seems easy to say, it is really difficult to put into practice, often requiring data stewards, and that's where ethics comes in. To substantiate her points, Stall presented a few cases to the participants via a Zoom poll.

Just to give an example, she discussed a particular scenario wherein the authors drafted a paper with links to the software used in the study. No discrepancy was highlighted during the review or even when the paper was published. However, another research group, when reviewing this paper while conducting their own study, identified an error in the software and contacted AGU. The session participants were allowed to answer a poll on what would be the best thing to do in such a situation—whether the paper should be retracted, the error ignored (just a software after all!), or the author contacted, so as to get an idea about what really went wrong. Most participants agreed that the author should be contacted, and that is what AGU did. The second research team was extremely particular that this error should be corrected because this software was a really important one and was actively being used in the research community, and future research would be affected if this error remained uncorrected. AGU did facilitate the discussion between the original authors and the second research team; the

CONTINUED

discussion was cordial, and the authors identified the error and corrected it via an erratum. The best part about this was that the paper was not retracted and the research findings published previously were in no way affected; moreover, the authors were pleased to have received the feedback. This is why data sharing is so important, and this case is a fantastic example of how data transparency can be fostered in the scientific community. Having a good ethics policy in place will also support reproducibility of research; hence, asking the right questions and making sure data are shareable is critical.

Finally, Matthew Cannon, who has been working in open research, discussed how his organization, Taylor & Francis, is trying to use the data sharing policies used in the sciences and applying those to the fields of arts and the humanities. In furtherance of Stall's session and cases, Cannon presented a few cases and discussed the best practices in terms of data sharing. A particularly interesting case he presented was of a patient who granted the author permission to publish data and a code to enable the creation of a 3D model of their brain (this would mean that anyone with access to a 3D printer would be able to print the model!). Editorial checks performed prior to publication led to the author being queried if the patient had allowed them to "just publish" their data or to "use the data in other ways as well without any further consent being required"; the author then confirmed that the patient had granted permission to "just publish" their data. The participants and Cannon reached a consensus that "it would be unfair if the patient was asked to consider all potential commercial and other reuses of their brain scan." The data and code were created to aid research; hence, the author agreed to restrict file access to bonafide researchers so as to protect the patient's rights.

The effort put into research comes to fruition when it is published and becomes available to the scientific community; however, publishing one's work following the required ethics is a challenge. This session greatly contributed toward spreading awareness among the researchers and the scientific community regarding publication ethics and how it can diminish misconduct in research. All-in-all, this was an extremely informative session and laid emphasis on why ethical standards are required in scholarly publishing—to ensure high-quality scientific publications and public trust in scientific findings, and so that people receive credit for their work and ideas.

References and Links

1. https://publicationethics.org/resources/guidelines-new/cooperation-between-research-institutions-and-journals-research-integrity
2. Wager E, Kleinert S. on behalf of the CLUE Working Group. Cooperation & Liaison between Universities & Editors (CLUE): recommendations on best practice. Res Integr Peer Rev. 2021; 6(1):1–2. https://doi.org/10.1186/s41073-021-00109-3
3. https://publicationethics.org/files/RePAIR%20Consensus%20Guidelines%20v2.pdf