Preprints and Data Repositories
Best Practices for Citation

Peter J. Olson, ELS
Sheridan Journal Services
Waterbury, VT
Preprints

What is a preprint?
A preprint is an unpublished report of research that is posted online to encourage public discussion, feedback, and data sharing. Preprints are hosted on an online server and have not been peer reviewed.
Identifying a preprint as such in text makes it clear to the reader that a reference being used to support a particular statement has not been peer-reviewed.

Identifying a preprint as such in the reference list distinguishes it from an accepted article and makes it clear that it is not a published-ahead-of-print article.
Updating preprint references

Publish-ahead-of print (PAP) articles that are cited as references can often be updated before publication of the article in which they are cited:

However, the amount of time that elapses between the initial posting of a preprint and its publication as a peer-reviewed article may not allow for an update:

From Stack Exchange. Is it necessary to include and update the status of a paper on arXiv once it has been accepted for publication in a journal? https://academia.stackexchange.com/questions/8621/is-it-necessary-to-include-and-update-the-status-of-a-paper-on-arxiv-once-it-has. Accessed Nov 6, 2018.
Data repositories

What is a data repository?

A data repository (a.k.a “data warehouse” or “data mart”) is an online, publically accessible library of a specific data population that is used to support current research and aid in the development and execution of future research.
Data repository citations: title page footnote

Data repository citations: data and software availability statement

**Cell Reports**

Large-Scale Human Dendritic Cell Differentiation Revealing Notch-Dependent Lineage Bifurcation and Heterogeneity

**DATA AND SOFTWARE AVAILABILITY**

The accession numbers for the raw single cell RNA-sequencing data reported in this paper are GEO: GSE108546 and GSE115405.

**NOTE** The “data and software availability statement” can appear in the front matter, back matter, and/or full text HTML version of the article.

Certain data repositories will not provide working links for newly deposited data until the article in which the data are cited has been published in some form (namely, publish-ahead-of-print). As such, the author may need to request a functional link from the repository before final publication.

Querying the author

Sample query

All cited data must be deposited in a publicly accessible database before publication. For any data cited in the text and/or the data deposition footnote, please provide complete reference information for the applicable database(s). If any data mentioned the data deposition footnote are not also cited in the text, please (a) insert a citation at an appropriate point in the text and (b) provide a complete reference for the applicable database.

Critical components

Ask the author for:

• A description of the data
• The name of the database
• The accession number(s)
• Confirmation that the data have been deposited (for newly deposited data)

NOTE Consider crafting a standard (or “standing”) query to ensure that the same message is always delivered to your authors.
Food for thought...

**Preprints**

- Should preprint references be identified as such, or is a DOI sufficient?
- If a referenced preprint is eventually published as a peer-reviewed article, should the reference be updated to match the article?
- Is citing your own preprint a form of self-citation?

*NOTE* Authors often refer to preprints as “working papers”; make sure to distinguish these from “manuscript in preparation” citations that do not have a DOI.

*NOTE* Preprint ≠ published ahead of print.

**Data repositories**

- Do data repositories need to appear in the reference list, or is a footnote/link sufficient?
- Are specific accession numbers being cited? If so, is a separate reference required for each accession number, regardless of how many are cited?

*NOTE* Distinguish previously deposited data from newly deposited data by including an “accessed” or “deposited” date, respectively, in the reference list.
Further Reading

What are preprints, and how do they benefit authors?

Preprints and citations: should non-peer reviewed material be included in article references?

Plan for new medical preprint server receives a mixed response

What’s up with data citations?

Thank you!

Peter J. Olson, ELS
Sheridan Journal Services
Waterbury, VT
peter.olson@sheridan.com