

# Managing Information From Preprints: The Scholarly Record and the Public Need for Information (Especially During a Pandemic)

## MODERATOR:

**Patricia K Baskin**Executive Editor, *Neurology Journals*American Academy of Neurology  
Minneapolis, Minnesota

## SPEAKERS:

**John R Inglis**Cold Spring Harbor Laboratory  
Cold Spring Harbor, New York**Bruce T Rosenblum**Inera  
Somerville, Massachusetts**Iratxe Puebla**ASAPBio  
San Francisco, California

## REPORTER:

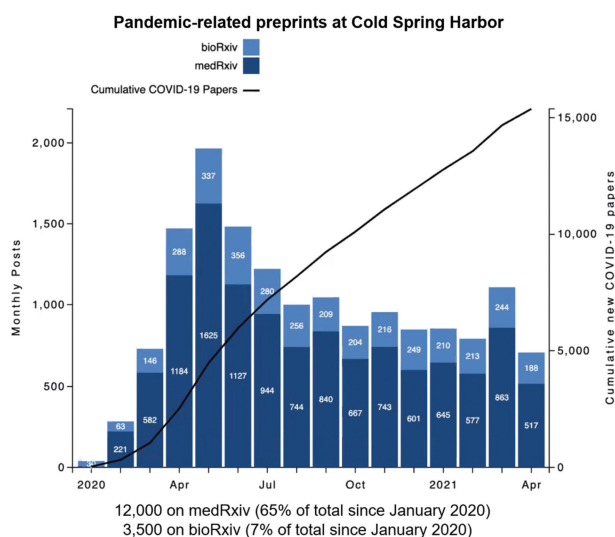
**Tony Alves**

Hopdale, Massachusetts

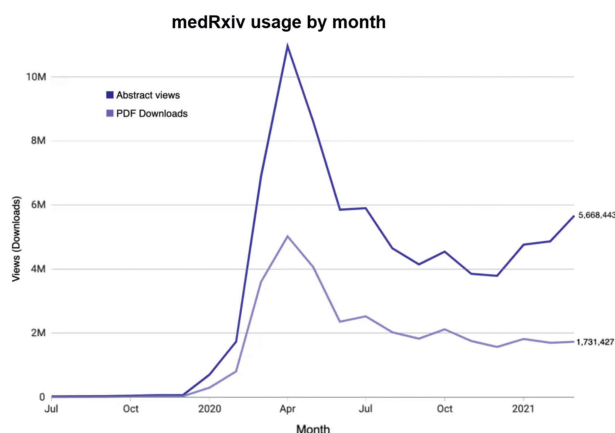
The session entitled "Managing Information from Preprints: The Scholarly Record and the Public Need for Information (Especially During a Pandemic)" was held on May 3, 2021 at the Council of Science Editors Annual Meeting. Moderated by Patty Baskin, Executive Editor of the *Neurology*<sup>®</sup> Journals, the session featured presentations from John Inglis, Co-founder of bioRxiv and medRxiv, and Executive Director of Cold Spring Harbor Laboratory Press; Iratxe Puebla, Associate Director of ASAPbio; and Bruce Rosenblum, Vice President of Content and Workflow Solutions at Inera, an Atypion Company.

During the COVID-19 pandemic, the usage and popularity of preprint servers has grown. Preprints have been common in disciplines such as physics, astronomy, and mathematics for decades, via the arXiv<sup>1</sup> preprint server, but they are a somewhat new phenomenon in biology and medicine. Preprints allow for an early look at new research, often before authors submit their manuscripts to scholarly journals for evaluation and peer review. Because readers have access to research before it is peer reviewed, confusion and misconceptions around preprints may arise, particularly among the general public.

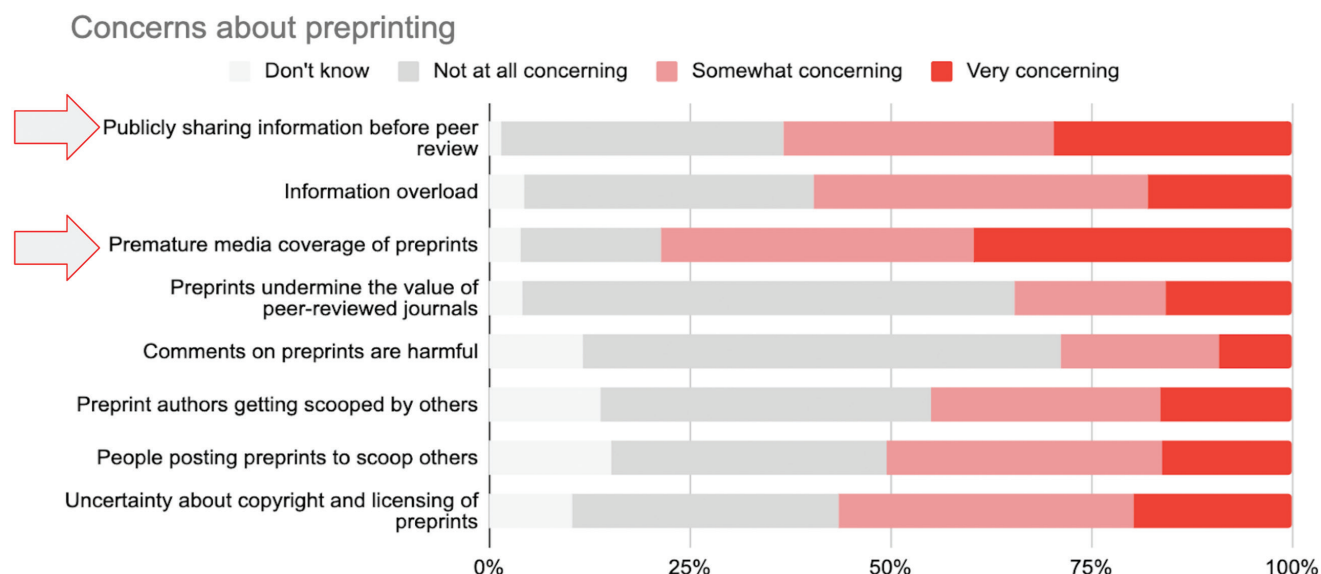
John Inglis presented first and provided a comprehensive overview of how preprints are transforming research communications, using bioRxiv<sup>2</sup> and medRxiv<sup>3</sup> from Cold Spring Harbor Laboratory (CSHL) as examples. Inglis spoke



of the enthusiasm that has arisen because of the speed that preprints offer, exposing useful research in a matter of days rather than months. Preprints also offer more control for authors—they are free to read and submit, and they



CONTINUED



#biopreprints2020 survey n=512

<https://asapbio.org/biopreprints2020-survey-initial-results>

make possible a dialogue between authors and readers. Inglis noted that CSHL's preprint servers are not a product; rather, they are a service from an academic institution with a long history of innovation in science communication and education. The preprints are not peer reviewed, but they are screened, and they can be revised at any time until they are accepted by a journal.

Since January 2020, there has been a deluge of pandemic-related preprints, with 16,000 on medRxiv and bioRxiv alone. In one article, "The first 12 months of COVID-19: a timeline of immunological insights," published in *Nature Reviews Immunology*,<sup>4</sup> of 168 citations, 12 were to preprints, and 42 to journal articles with preprint versions. COVID-19-related preprints posted in 2020 were published in journals faster than non-COVID-19-related preprints, and citation of medRxiv preprints has skyrocketed.

The adoption of preprints in biomedicine has been criticized by some because it brings attention to unvalidated science. Some preprint authors have made exaggerated or mistaken claims that have been amplified by the media and adopted by politicians. Inglis reviewed the precautions

taken at bioRxiv and medRxiv to reduce mistrust. In-house content specialists and independent principal investigators approve all posted manuscripts. Submissions are declined if they aren't research papers, don't have conflict of interest and ethics approval statements, report unregistered clinical trials, encourage self-medication, or promote conspiracy theories. It is interesting to note that only 33 (0.2%) of the pandemic-related preprints have been withdrawn. Retraction Watch's database lists more than 100 retracted journal papers on COVID-19.

Iratxe Puebla spoke second, and provided a strong endorsement of preprints, pointing out that ASAPbio<sup>5</sup> supports preprint adoption in the life sciences. ASAPbio is a "scientist-driven nonprofit working to drive open and innovative communication in the life sciences." They promote both the use of preprints and transparent peer review. Puebla's presentation was called "Building Trust in Preprints: Opportunities Ahead," and she focused on how preprint metadata practices can enable transparency and trust. Reporting on a set of recommendations by stakeholders, she first talked about how it should be clear what version of the work is being read or cited, and that readers should always be pointed to the newest version of the preprint, recommending persistent identifiers for each version. The published journal article, if there is one, should always be linked to the preprint once it is available. Puebla recommended that the metadata for withdrawn or removed preprints clearly identify that it was withdrawn or removed, and that it is desirable to include the reason for the withdrawal. With regard to reviews of preprints, review services can register Crossref DOIs for the reviews,

#### Example of a "good" preprint citation

##### ➤ From the *AMA Manual of Style*:

1. Bloss CS, Wineinger NE, Peters M, et al. A prospective randomized trial examining health care utilization in individuals using multiple smartphone-enabled biosensors. bioRxiv. Preprint posted online October 28, 2015. doi:10.1101/029983

## CONTINUED

and preprint servers could also use Crossref Event Data to surface commentaries and reviews for readers.

Puebla, like Inglis, addressed concerns about preprints, such as public sharing of research before peer review, and premature media coverage of preprints. These concerns and others can be handled through transparency in labeling and reporting, and Puebla outlined ASAPbio's guiding principles<sup>6</sup> for preprint servers on preprint labelling. Finally, Puebla discussed a few platforms that provide peer review of preprints, such as Review Commons,<sup>7</sup> PREREVIEW,<sup>8</sup> and the overlay journal *Rapid Reviews: COVID-19*.<sup>9</sup>

The final speaker was Bruce Rosenblum, an expert on scholarly citations. His talk was entitled "Challenges Citing Preprints and How to Tackle Them." He pointed out several challenges with preprint citation. First, preprint servers don't always identify contents as preprints. Second, recommended citations may be incomplete. Third, author citations are frequently incomplete. Finally, incomplete metadata makes citation completion and verification challenging.

Rosenblum ran through various examples of confusing citations, including a National Geographic article that cited a preprint as "New research" without specifying that the research had been posted as a preprint and was not yet peer reviewed. He pointed out how different citation formats don't include clear guidelines on how to handle preprint citations, and that DOIs are often missing. He provided a lot of good advice, recommending that journal style guides, reference management software, and author instructions be updated to include elements such as preprint server name, preprint title, a preprint indicator, and preprint DOI.

Other issues covered by Rosenblum include the lack of consistent linking between preprints and their published articles, inconsistent Crossref query results, posting the same preprint on multiple servers, and missing "withdrawn" status for withdrawn and removed preprints. One particularly

problematic situation is when a preprint is replaced by the published journal article. This is a problem because if an author has cited a preprint, and the information in the published article has materially changed since it was cited as a preprint, then this will undermine the chain of citation.

Rosenblum finished his presentation with a list of recommended practices such as: 1) Clearly identify non-peer reviewed content at all stages; 2) update citation style guides; 3) ensure preprints are treated as unique citation types, are promptly updated with publication status, never replaced by a published journal version or deposited to Crossref as a journal article; and 4) automatically link published articles with their preprints and identify withdrawn preprints. He also recommended 3 articles.<sup>10–12</sup>

## References and Links

1. <https://arxiv.org/>
2. <https://www.biorxiv.org/>
3. <https://www.medrxiv.org/>
4. Carvalho T, Krammer F, Iwasaki, A. The first 12 months of COVID-19: a timeline of immunological insights. *Nat Rev Immunol* 2021;21:245–256. <https://doi.org/10.1038/s41577-021-00522-1>
5. <https://asapbio.org/>
6. <https://asapbio.org/public>
7. <https://www.reviewcommons.org/>
8. <https://www.prereview.org/>
9. <https://rapidreviewscovid19.mitpress.mit.edu/>
10. Hunter SI, Kleshchevich I, Rosenblum B. Guest post—what's wrong with preprint citations? The Scholarly Kitchen. <https://scholarlykitchen.sspnet.org/2020/09/18/guest-post-whats-wrong-with-preprint-citations/>
11. Beck J, Ferguson CA, Funk K, Hanson B, Harrison M, Ide-Smith M, Lammey R, Levchenko M, Mendonça A, Parking M, et al. Building trust in preprints: recommendations for servers and other stakeholders. [accessed June 1, 2021]. <https://doi.org/10.31219/osf.io/8dn4w>
12. Patel, J. Opinion: preprints in the public eye. *The Scientist*. March 18, 2021. <https://www.the-scientist.com/news-opinion/opinion-preprints-in-the-public-eye-68563>