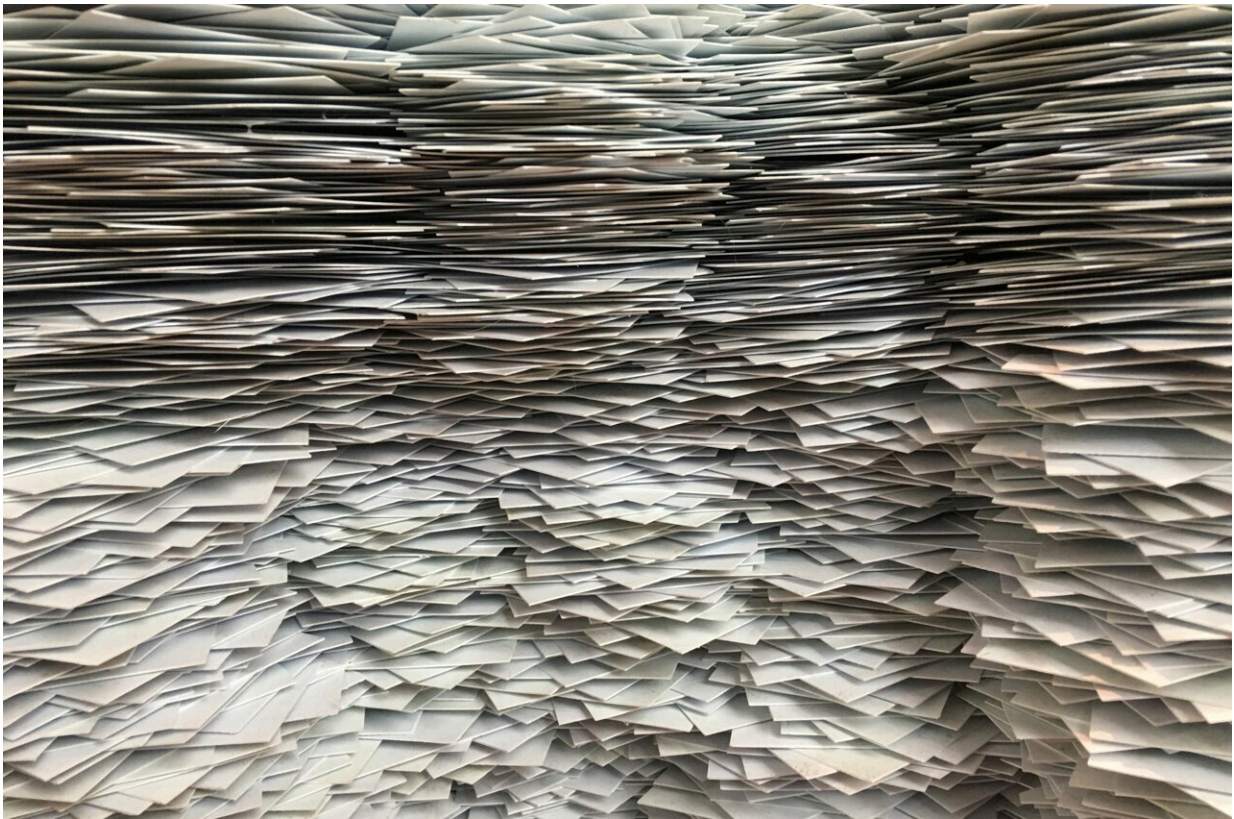


Comparing preprints and their finalized publications during the pandemic

February 1 2022



Credit: Christa Dodoo, Unsplash (CC 0, creativecommons.org/publicdomain/zero/1.0/)

Preprinting, the sharing of freely available manuscripts prior to peer-review, has been on the rise in the biosciences since 2013 and

experienced a surge during the COVID-19 pandemic, expediting the dissemination of timely research. But how do preprints relate to the final peer-reviewed papers? Two new studies publishing in the open access journal *PLOS Biology* February 1st took different approaches to explore how preprints posted on bioRxiv and medRxiv compare with their published versions.

One study, led by Dr. Jonathon Coates of Queen Mary University of London, manually compared over 180 preprints to their published versions in the first 4 months of the COVID-19 pandemic. The other study, led by Mr. David Nicholson of University of Pennsylvania's Perelman School of Medicine, used machine learning and textual analytics to explore the relationships between nearly 18,000 bioRxiv preprints and their published version.

Concerns over the quality of preprints have existed since the emergence of preprinting in the sciences. As Coates notes, "Approximately 40% of the early COVID-19 research was first shared as a [preprint](#) and these were used in policy and public health decisions. Therefore, knowing the quality of these preprints is vital in having trust in science at a time when many are attempting to erode that trust". Analysis of public scientific preprint repositories also has the potential to illuminate many previously hidden details of the peer-review process.

Coates and his colleagues compared all the COVID-19 preprints posted and published within the first 4 months of the pandemic and found that over 83% of COVID and 93% of non-COVID-related life sciences articles do not change from their preprint to final published versions.

Comparing the entire bioRxiv corpus to eventually published versions, Nicholson and colleagues found that many differences appear to occur from typesetting and the addition of supplementary materials; there were only modest changes in the linguistic characteristics of most manuscripts

during the peer-review and [publication process](#).

Furthermore, Nicholson and their team created a website that uses their [machine learning](#) tool to recommend potential journals that publish linguistically similar articles that can be found at <https://greenelab.github.io/preprint-similarity-search/>.

Dr. Casey Greene of the University of Colorado School of Medicine, a co-author on the Nicholson et al. study, adds, "Collectively, our studies both provide evidence supporting the reliability and use of preprints both during a global pandemic and for general scientific outputs. Examining preprint-publication pairs provides an opportunity to study the process of peer review and taken together our results should provoke a rethinking of the role and prominence of peer-[review](#) in the current publication system."

Coates adds, "With such a large proportion of early COVID-19 literature shared as non-peer reviewed preprints it is essential to know if those studies are reliable or not. By manually comparing the preprints to their peer reviewed, published, versions we show that over 83% of COVID-19 and 93% of non-COVID preprints are reliable and trustworthy."

More information: Brierley L, Nanni F, Polka JK, Dey G, Pálffy M, Fraser N, et al. (2022) Tracking changes between preprint posting and journal publication during a pandemic. *PLoS Biol* 20(2): e3001285. doi.org/10.1371/journal.pbio.3001285

Nicholson DN, Rubinetti V, Hu D, Thielk M, Hunter LE, Greene CS (2022) Examining linguistic shifts between preprints and publications. *PLoS Biol* 20(2): e3001470. doi.org/10.1371/journal.pbio.3001470

Provided by Public Library of Science

Citation: Comparing preprints and their finalized publications during the pandemic (2022, February 1) retrieved 20 June 2024 from <https://phys.org/news/2022-02-preprints-pandemic.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.