



NIH Preprint Pilot – Q2 Status Update

On June 9, 2020, NLM launched a pilot project designed to test the viability of making full-text preprints resulting from NIH-funded research searchable in PubMed Central (PMC), with an accompanying citation discoverable in PubMed. Including preprints in PMC and PubMed is a way to enhance their discoverability.

Phase 1

Phase 1 of the NIH Preprint Pilot is focused on preprints resulting from NIH-funded COVID-19 research and curated for inclusion the [NIH iSearch COVID-19 Portfolio](#). As of December 9, the pilot had been running for 6 months (2 quarters).

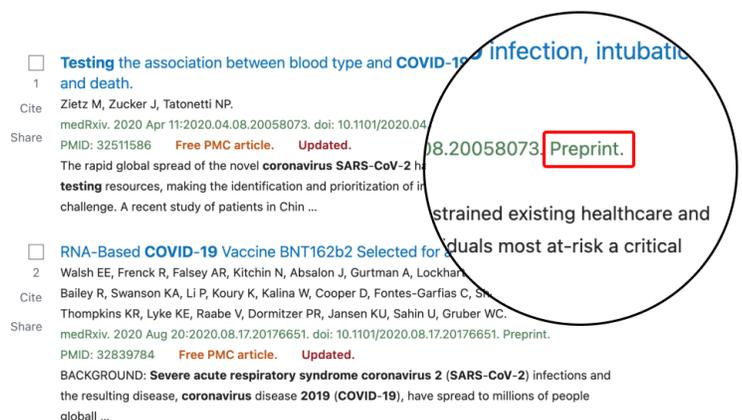
Preliminary Results – as of the end of the second quarter of the pilot:

- Nearly 1,500 preprint records have been added to PMC and made discoverable in PubMed.
- Preprint records have been viewed more than 800k times in PMC. Associated preprint metadata and abstract records have been viewed approximately 500k times in PubMed.
- Some 35% of the preprint records made available in NLM databases are matched to a published journal article.
- About a quarter of the preprint records have more than one version available in a given preprint server.
- Nearly two-thirds of the preprint records are made available under some type of Creative Commons (CC) license; 18% had a CC-BY or CC0 (public domain) license.

Q2 Implementation Details

- **Scope:** Preprints acknowledging NIH support and relating to COVID-19.
- **Preprint Server Sources:** arXiv, bioRxiv, ChemRxiv, medRxiv, Research Square, and SSRN.
- **Workflows:** New preprint records uploaded on a weekly basis. Curation and updating of preprint records has been largely automated.
- **User Interface:** Preprint banners and citations, links to preprint servers via DOI and LinkOut, links to additional preprint versions, as well as resulting journal articles. “Preprint” flag was added to preprint search results in PubMed.

Preprint Indicators in PubMed Search Results Display



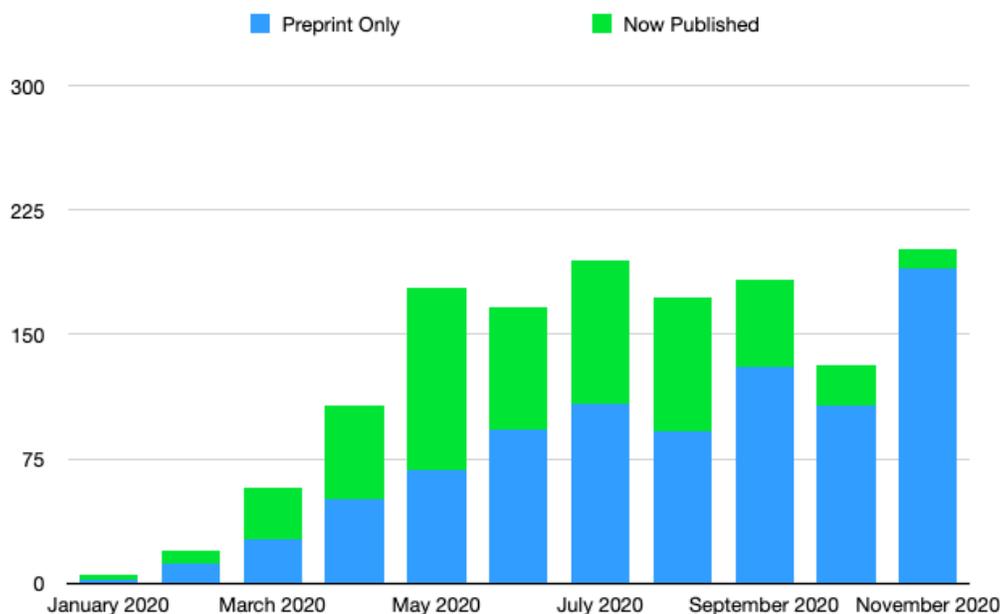
1 **Testing the association between blood type and COVID-19 infection, intubation and death.**
 Cite Zietz M, Zucker J, Tatonetti NP.
 medRxiv. 2020 Apr 11:2020.04.08.20058073. doi: 10.1101/2020.04.08.20058073. **Preprint.**
 Share PMID: 32511586 **Free PMC article. Updated.**

2 **RNA-Based COVID-19 Vaccine BNT162b2 Selected for Phase 1 Testing**
 Cite Walsh EE, Frencik R, Falsey AR, Kitchin N, Absalon J, Gurtman A, Lockhart S, Bailey R, Swanson KA, Li P, Koury K, Kalina W, Cooper D, Fontes-Garfias C, et al.
 medRxiv. 2020 Aug 20:2020.08.17.20176651. doi: 10.1101/2020.08.17.20176651. **Preprint.**
 Share PMID: 32839784 **Free PMC article. Updated.**

BACKGROUND: Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infections and the resulting disease, coronavirus disease 2019 (COVID-19), have spread to millions of people globally ...

Outcomes

- **Faster dissemination** – Approximately 35% of preprints records in PMC have been linked to a published journal article (up from around 18% at the end of the first quarter). Earlier posted preprints were more likely to have been published by the end of Q2. Based on data from January 2021:
 - 54% of the preprints in PMC posted between Jan 2020 – June 2020 have been subsequently published and linked to a journal article.
 - 45% of preprints posted between July 2020 – Sept 2020 have been published.
 - 14% of preprints posted between Oct 2020 – Dec 9, 2020 have been published.



- **Increased discoverability** – Comparing data on preprint record views in PMC to those in PubMed indicates that the availability of full-text XML may increase the discoverability of these records. Additionally, the most viewed records in PMC are generally those with Creative Commons licenses, which allow for full text archiving and indexing.
- **Awareness of NIH-recommended preprint practices** – The share of preprint records posted under an NIH-recommended CC-BY or CC0 license ticked up slightly since the pilot launch (13% of preprints posted Jan - June 2020 vs. 21% of preprints posted July - December 9, 2020).

Communications & Engagement on Preprint Best Practices

- Released the NIH Preprint Pilot: A Librarian Toolkit (<https://learn.nlm.nih.gov/documentation/training-packets/T000141112/>) that consolidates pilot overview documentation, educational materials, and related resources.
- Presented on NIH Preprint Pilot progress to date as part of a Network of the National Library of Medicine Resource Pick (September 2020) and at a workshop on changing publication practices during COVID-19 organized by the Friends of the National Library of Medicine (October 2020).
- Provided quarterly updates to the NLM Board of Regents Collections Working Group.

As the pilot continues, NLM will monitor preprint server compliance with the [pilot eligibility criteria](#).