



## NIH Preprint Pilot – Q3 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 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 March 9, 2021, the pilot has been running for 9 months (3 quarters).

**Preliminary Results** – as of the end of the third quarter of the pilot:

- Nearly 2,000 preprint records have been added to PMC and made discoverable in PubMed, up from approximately 1,500 at the end of Q2.
- Preprint records have been viewed more than 1 million times in PMC. Associated preprint metadata and abstract records have been viewed approximately 800,000 times in PubMed.
- Approximately 40% of the preprint records made available in NLM databases are matched to a published journal article.
- Preprints included in the pilot have been cited approximately 10,000 times.
- Quarterly growth continues in the number of preprint records made available under some type of Creative Commons (CC) license, including NIH-recommended CC-BY and CC0 (public domain) licenses.

### Q3 Implementation Details

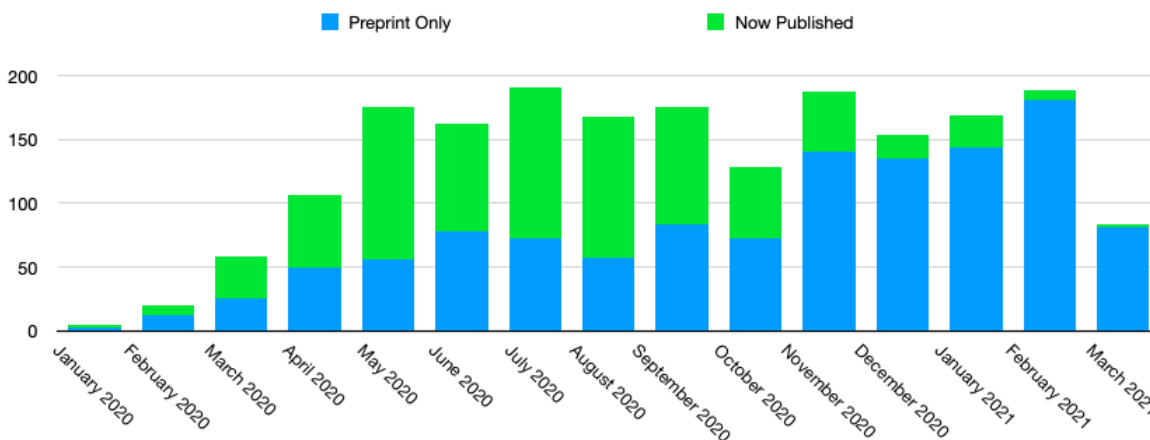
- **Scope:** Preprints acknowledging NIH support and relating to COVID-19.
- **Preprint Server Sources:** arXiv, bioRxiv, ChemRxiv, medRxiv, Research Square, and SSRN.
- **Workflows:** New preprint records uploaded on a weekly basis. Checks for updated preprint and journal article versions run 1-2 times a week.
- **User Interface:** Preprint records include preprint banners and citations, links to preprint servers via DOI and LinkOut, links to additional preprint versions, as well as resulting journal articles (when available). Improvements to search results display continue to be explored.
- **NEW! Automated Checks:** To ensure continued trust through responsible stewardship of the pilot, NLM automated its checks for withdrawn preprints during Q3.

### Outcomes

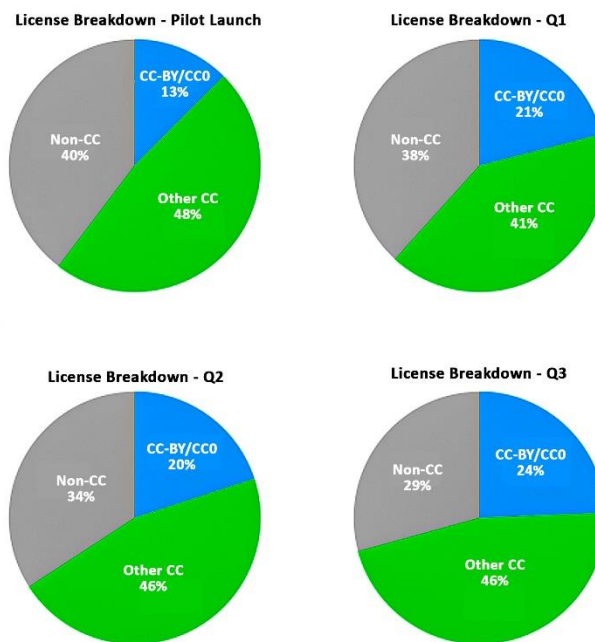
- **Faster dissemination** – Approximately 40% of preprint records in PMC have been linked to a published journal article (up slightly from around 35% at the end of the second quarter). Preprints posted earlier in 2020 were more likely to have been published by the end of Q3.

Based on data from March 2021:

- 58% of the preprints in PMC posted between Jan 2020 – June 2020 have been subsequently published and linked to a journal article.
- 60% of preprints posted between July 2020 – Sept 2020 have been published.
- 26% of preprints posted between Oct 2020 – Dec 2020 have been published.
- 3% of preprints posted between Jan 2021 – Mar 9, 2021 have been published.



- **Increased discoverability** –Traffic to preprints in PMC and PubMed comes primarily from third-party search engines, direct search in PubMed, or social media links. The availability of full-text XML in PMC plays a key role in search engine discovery. Preprints frequently appear in PubMed trending articles.
- **Awareness of NIH-recommended preprint practices** –The share of preprint records posted under an NIH-recommended CC-BY or CC0 license has increased each quarter of the pilot (see figure to the right). Additionally, conflict of interest statements are included in more than half of the full-text preprints in PMC, and other open science indicators (e.g., supplementary material sharing, data availability statements, data and software links) appear regularly.



### Communications & Engagement on Preprint Best Practices

- A [New England Journal of Medicine perspectives piece](#) identified the NIH Preprint Pilot as one way the NIH has advanced transparency and accelerated research during the pandemic.
- Updates about the NIH Preprint Pilot were presented on panels organized by the Friends of the National Library of Medicine (January) and [NISO Plus](#) (February).
- NLM is engaging with stakeholders, including repositories, on managing article versions and maintaining the currency of the scientific record, including withdrawals.