

# Accessing Non-Peer-Reviewed Preprint Literature from Various Servers: An Essential Library Resource to Support Vaccine Development During A Global Pandemic

Sharon Glass, MLIS

SLA Annual  
Conference 2021



## BACKGROUND

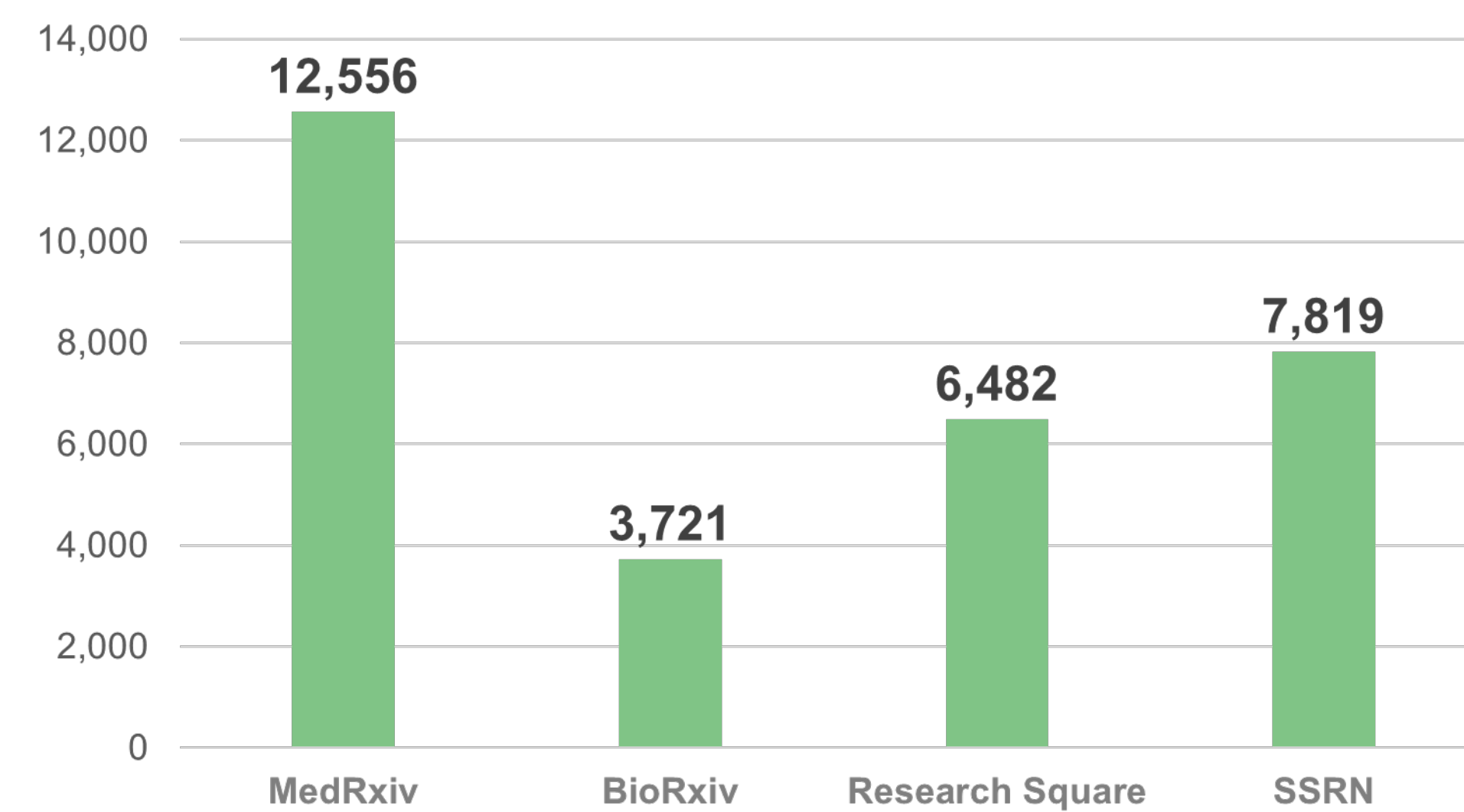
Preprints are becoming not only more prevalent, but more important in infectious disease research. Accessing these types of studies may be less straightforward than searching published literature. Corporate libraries typically depend on peer-reviewed published literature when conducting literature searches. During the Covid-19 pandemic, new information has been changing rapidly and there is significant lag time for new manuscripts to become fully published journal articles. Vaccine developers have unprecedentedly come to rely on non-peer-reviewed preprint studies to stay abreast of new data.

## METHODS

- Novavax' Library Services uses an information aggregator to search for published literature through Medline™, Embase™, and Biosis™. Although there is some coverage of preprint literature through Medline, it is limited and therefore site visits to each individual preprint server have been necessary to obtain a more comprehensive list of studies. Without sophisticated search engines on many of these sites, we had to adjust to a more manual way of seeking out relevant studies. The library developed keyword search strategies specifically developed for each of the main preprint servers used: MedRxiv™, BioRxiv™, Research Square™, and SSRN™. The keyword search strategy was a title search using the terms Covid-19 or SARS-COV-2 with no date restraints. The search was conducted weekly.
- The same keyword search strategy was conducted in the Medline database, selecting 'preprint' as the document type.
- Lastly, a PubMed™ search using the same criteria was conducted.

## RESULTS

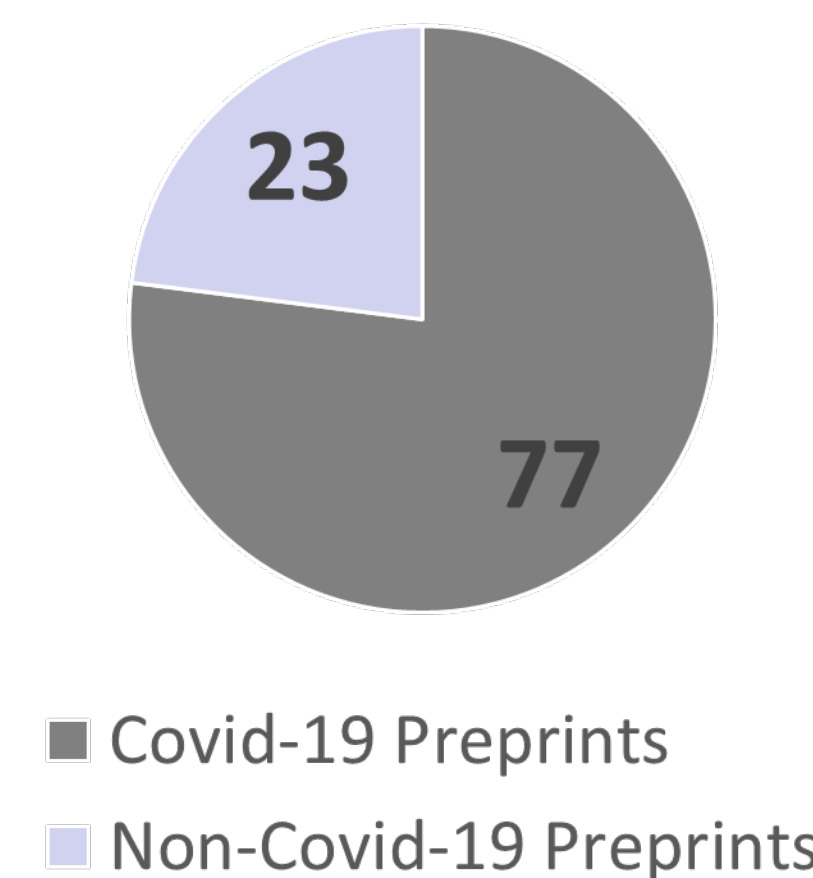
**Fig 1. Total Number of Covid-19 Articles by Individual Preprint Server through May 2021\***



\* Total numbers reflect a title search for Covid-19 or SARS-COV-2 with no date restraints.

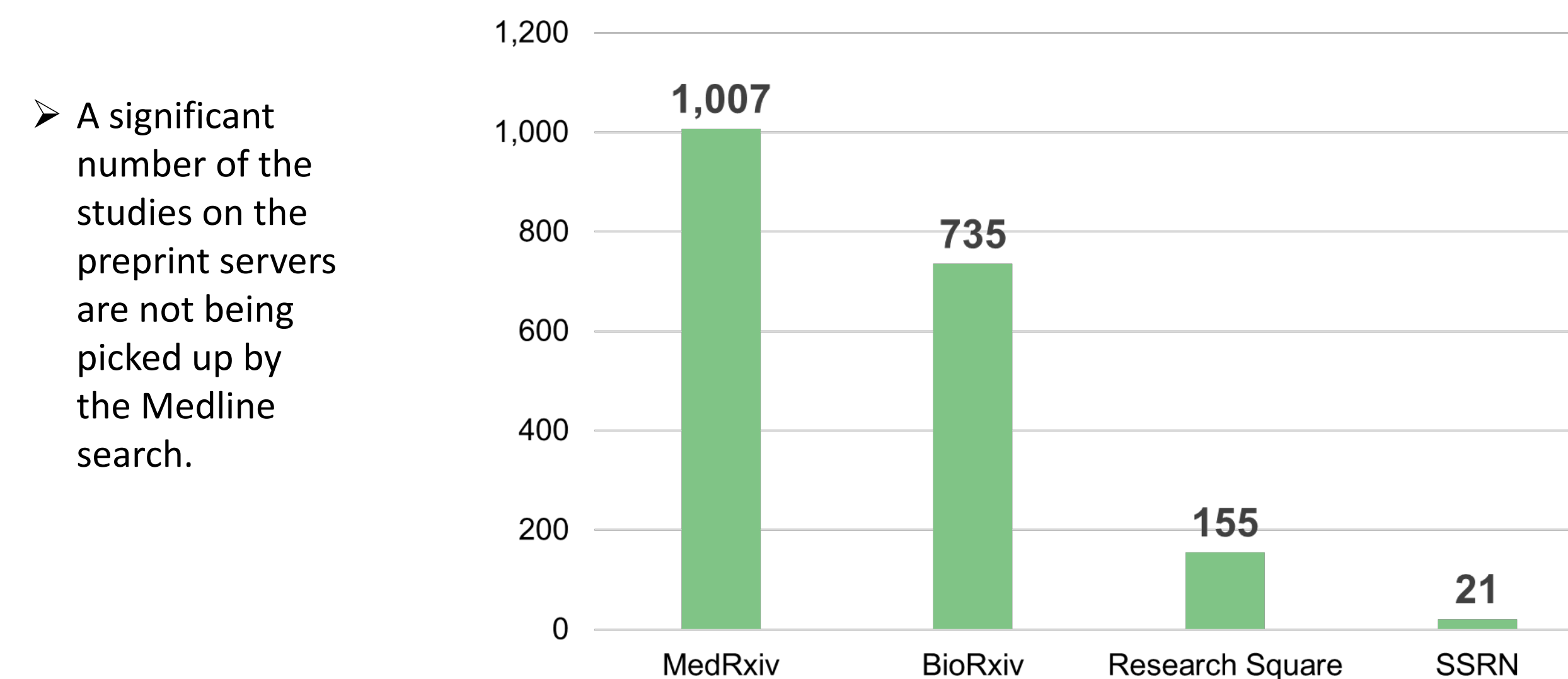
- The search strategies created for each individual preprint server site which were conducted weekly yielded a unique set of results. We found very little overlap in preprints and concluded that most studies were not posted to more than one preprint server.

**Fig 2. Percent of Preprint Articles Devoted to Covid-19 on MedRxiv through May 2021**



- MedRxiv was established in 2019. A title search for Covid-19 or SARS-COV-2 revealed that since its inception about 77% of its content has been devoted to Covid-19 related research.
- One study showed that 10% of all preprints submitted in 2020 were related to Covid-19 research.<sup>1</sup>
- A study estimated that between 17% and 30% of all Covid-19 research papers submitted in 2020 were in preprint form.<sup>1</sup>
- The average preprint takes about 160 days to publish to a journal.<sup>2</sup>

**Fig 3. Total Number of Covid-19 Preprints Indexed in Medline by Server through May 2021\***

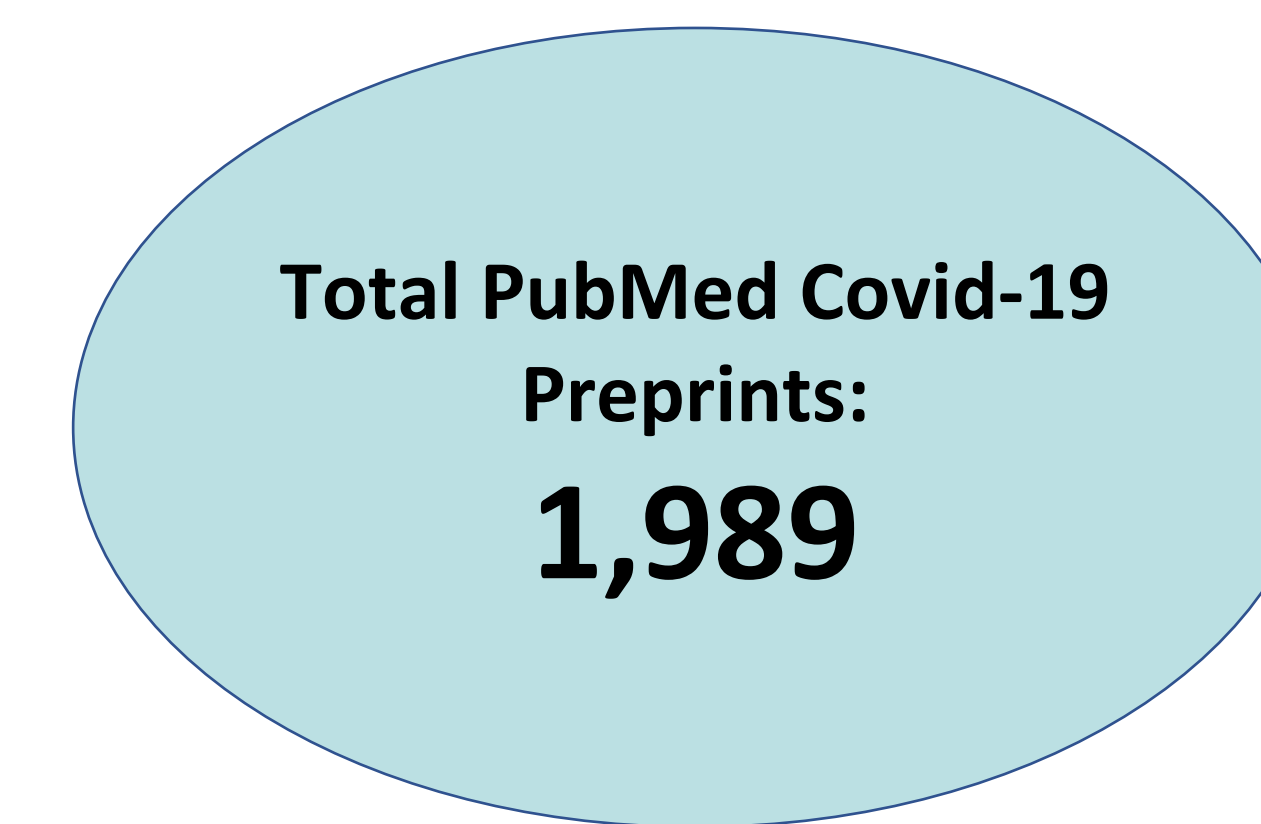


- A significant number of the studies on the preprint servers are not being picked up by the Medline search.

\* Total numbers reflect a title search for Covid-19 or SARS-COV-2 with no date restraints.

## RESULTS

**Fig 4. Total Number of Covid-19 Preprints in PubMed through May 2021**



- A PubMed search using the same criteria captured 1,989 Covid-19 preprints, far fewer than expected.
- Furthermore, Covid-19 related articles tagged as publication type 'preprint' accounted for 82% of the total preprints on PubMed.

## CONCLUSIONS

- Preprints are an important resource for vaccine developers to stay informed of new competitor data. Although time consuming, it is necessary for Library Services to perform varied, manual searches on preprint servers to capture relevant data.
- The Library has found that a weekly update that includes a combination of searching peer-reviewed published literature through an information aggregator including Medline, Embase, and Biosis, as well as searching the non-peer-reviewed literature on each individual preprint server is the best way to capture a comprehensive list of studies related to Covid-19 vaccines to ensure our pre-clinical, clinical, safety, regulatory, and medical affairs groups are up-to-date on relevant literature and newly conducted studies within a rapidly changing environment.
- Note: When sending any type of preprints to end-users, a disclaimer should always be included which indicates that the material is not peer-reviewed.

## REFERENCES

1. <https://www.nature.com/articles/d41586-020-03564-y>
2. [https://www.elsevier.com/\\_data/assets/pdf\\_file/0003/920721/ACAD\\_RI\\_SC\\_CS\\_Perspectives\\_-\\_Preprint-to-Publication.pdf](https://www.elsevier.com/_data/assets/pdf_file/0003/920721/ACAD_RI_SC_CS_Perspectives_-_Preprint-to-Publication.pdf)

### Trademark Companies:

- Medline (U.S. National Library of Medicine, U.S. NIH)
- Embase (Elsevier)
- Biosis (Clarivate Analytics)
- MedRxiv (Cold Spring Harbor Laboratory)
- BioRxiv (Cold Spring Harbor Laboratory)
- Research Square (Research Square AJE LLC)
- SSRN (Elsevier)
- PubMed (U.S. National Library of Medicine, U.S. NIH)