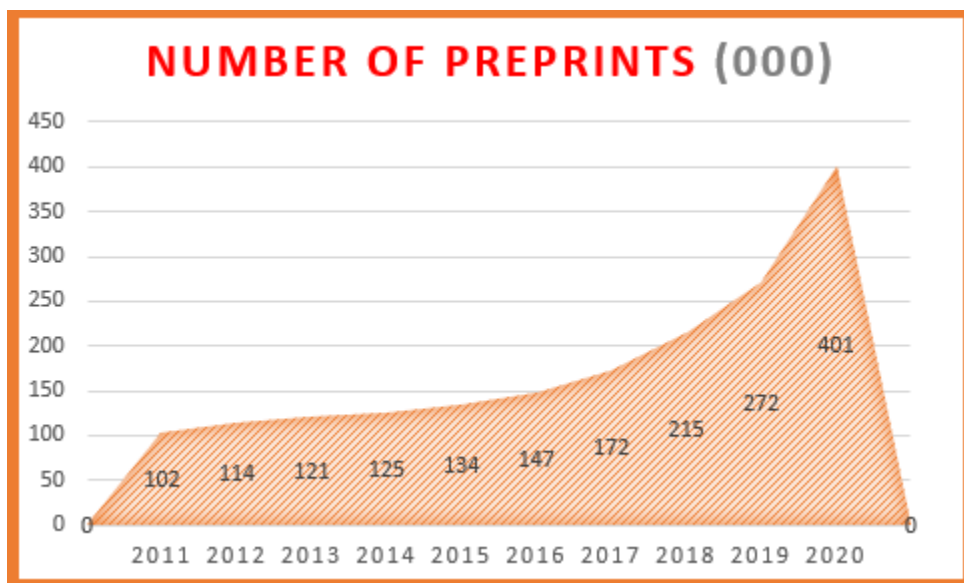**APPENDIX B: PREPRINTS**

"A growing problem in the scholarly and scientific community is a population of opportunists who try to use preprint servers as a place to post crackpot pseudo-science and misleading public health information, all under the flag of scholarly 'publishing.'" (Anderson, R)

According to a post on the NISO website (National Information Standards Organization) there is no one agreed upon definition of preprint (Polka). Below is the Open Science definition. "Preprint" and "pre-print" seem to be used interchangeably.

Pre-print - A manuscript draft that has not yet been subject to formal peer review, distributed to receive early feedback on research from peers. (Open Research Glossary)

 Dimensions is the best place to search for preprints. They "index directly from arXiv as well as a number of other whitelisted servers and indirectly via Crossref.(email Dimensions) Dimensions preprints increased almost 300% in the past ten years with a 47% increase between 2019 and 2020. The increase in medical documents grew 100% in the same time period (searched 1/5/2021).

There are now many sources of preprints and one preprint may appear in more than one preprint dataset. Kent Anderson suggests that this results in overcounting of preprints, using a Covid-19 preprint, "Covid-19 outbreak prediction with machine learning" as his example. (Anderson, K.) Other concerns are raised over the lack of accountability (Anderson, R.) and the fluidity of a preprint, and how it should be cited (Hunter, Kleshchevich, & Rosenblum).

EXAMPLE: A preprint that appeared multiple times and then was published in a journal

The following preprint appears 15 times in Dimensions and once as a retraction for being out of scope. Google Scholar also referenced the 15 versions.

☑ Show abstract   Sort by: Relevance⌄

Title, Author(s), Bibliographic reference - About the metrics

COVID-19 Outbreak Prediction with Machine Learning

Sina F. Ardabili, Amir Mosavi, Pedram Ghamisi, Filip Ferdinand, Annamaria R. Varkonyi-Koczy, Uwe Reuter, Timon Rabcz...

2020, Preprints.org - Preprint

Several outbreak prediction models for COVID-19 are being used by officials around the world to make informed-decisions and enforce relevant control measures. Among the standard models for COVID-19 gl... more

Citations 1   Altmetric 7   ☐ Open Access   ≡+ Add to Library   ⓘ Add to ORCID

COVID-19 Outbreak Prediction with Machine Learning

Sina F. Ardabili, Amir MOSAVI, Pedram Ghamisi, Filip Ferdinand, Annamaria R. Varkonyi-Koczy, Uwe Reuter, Timon Rabc...

2020, medRxiv - Preprint

Abstract Several outbreak prediction models for COVID-19 are being used by officials around the world to make informed-decisions and enforce relevant control measures. Among the standard models for CO... more

Citations 2   Altmetric 7   ☐ Open Access   ≡+ Add to Library   ⓘ Add to ORCID

COVID-19 Outbreak Prediction with Machine Learning

Sina F Ardabili, Amir Mosavi, Pedram Ghamisi, Filip Ferdinand, Annamaria R Varkonyi-Koczy, Uwe Reuter, Timon Rabczu...

2020, Research Square - Preprint

<p>Several outbreak prediction models for COVID-19 are being used by officials around the world to make informed-decisions and enforce relevant control measures. Among the stand... more

☐ View PDF   ≡+ Add to Library   ⓘ Add to ORCID

EXAMPLE 2: This document went from a preprint to a journal article, published in October (identified through Web of Science)

algorithms   MDPI

Article
COVID-19 Outbreak Prediction with Machine Learning
.

This article in the journal *Algorithms* published by MDPI ties in another aspect of the new normal of Scholarly publishing. In 2014, MDPI was on the predatory publishing list and now it is one of the top publishers of open access journals, with 221 titles on DOAJ as of 4 January 2021.

CONCLUSION: Preprints do not have direct impacts on journal metrics. Citations to preprints do not carry over to the journal versions. Both Web of Science and Scopus do have options for showing preprint citations.

This article in Web of Science shows no citations. Performing a **Cited Reference Search** includes citations to the preprint versions. Scopus does not include the article, but has a feature called **Secondary Sources** that list citations from Scopus articles to articles not in Scopus, including citations to preprints.

SOURCES

Anderson, K. (4 Jan 2021). Covid-19 preprint counts are inflated. *The Geyser, accessed at* https://thegeyser.substack.com/p/covid-19-preprint-counts-are-inflated

Anderson, R. (14 Dec 2020). Journalism, preprint servers and the truth: Allocating accountability. *The Scholarly Kitchen*, accessed at https://scholarlykitchen.sspnet.org/2020/12/14/journalism-preprint-servers-and-the-truth-allocating-accountability/

Dimensions Support Team (24 Jan 2021) email

Hunter, S.I., Kleshchevich, I., & Rosenblum, B. (18 Sep 2020). What's wrong with preprint citations, accessed at https://scholarlykitchen.sspnet.org/2020/09/18/guest-post-whats-wrong-with-preprint-citations/

Polka, J (12, 2019). Why "What is a preprint?" is the wrong question. NISO, accessed at https://www.niso.org/niso-io/2019/12/why-what-preprint-wrong-question

Sample preprint sources:

arXiv (1991) - Cornell university: https://arxiv.org/help/submit

medRxiv (2019) Yale university, BMJ, Cold Spring Harbor Laboratory https://www.medrxiv.org/submit-a-manuscript

OSF   is a platform from the Center for Open Science for preprint services and it also provides a searchable database for preprints. We introduced COS in Ruth's Rankings 45.