Pharmacological interventions for the COVID-19 pandemic and the use of preprint articles: The good, the bad and the ugly

Bryan Nicolalde
Universidad San Francisco de Quito https://orcid.org/0000-0002-7043-5515

Diego Añazco
Universidad San Francisco de Quito https://orcid.org/0000-0003-1829-7001

Mariam Mushtaq
Universidad San Francisco de Quito https://orcid.org/0000-0001-5929-3184

Isabel Espinosa
Universidad San Francisco de Quito

Jimena Gimenez
Universidad San Francisco de Quito

Enrique Teran (eteran@usfq.edu.ec)
Universidad San Francisco de Quito https://orcid.org/0000-0001-6979-5655

Research Article

Keywords: preprints, COVID-19, medRxiv, bioRxiv, Pubmed

DOI: https://doi.org/10.21203/rs.3.rs-34689/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

**Background:** Preprints are preliminary reports that have not been peer-reviewed. On Dec 2019, a novel coronavirus appeared in China, and since then, scientific production, including preprints, has drastically increased. In this study, we intend to evaluate how often preprints regarding pharmacological interventions against COVID-19 were cited, in spite of the fact that some of these preprints remained unpublished.

**Methods:** We conducted a search on medRxiv and bioRxiv to identify preprints related to pharmacological interventions against SARS-CoV-2 from Jan 1st to Mar 31, 2020. We gathered metadata on included preprints and identified if they had been published in a peer-reviewed journal. We performed Mann-Whitney U tests to evaluate if published articles had differences in citation numbers or usage, as defined by PDF downloads and abstract views, when compared to preprints that were not published.

**Results:** Our sample included 97 preprints, of which only 14 were published on peer-reviewed journals and 83 remained unpublished. The most common study designs we found among preprints were basic science research and case series. Published articles had a significantly higher number of citations and metrics (PDF and abstract downloads) when compared to unpublished preprints.

**Conclusions:** The use of preprints during this pandemic has been higher than in previous outbreaks, however, the publication rate in peer-reviewed journals in our sample was low. Preprints should be used as a mean to display preliminary data rapidly in order to obtain feedback by the scientific community, or to guide further research. However, due to the lack of peer-review, and potentially flawed data analysis, preprints alone should not be used to guide clinical practice, as the risk of unwarranted modifications to management is concerning.

**Background**

A preprint is a preliminary report that is shared publicly before it has been peer-reviewed. Most preprints have a digital object identifier (DOI) so they can be cited in other research articles (1). There are multiple servers that host preprints, notably including bioRxiv and medRxiv, operated by Cold Spring Harbor Laboratory. BioRxiv is a free server for unpublished manuscripts related to biological sciences, while medRxiv focuses on medical, clinical and health science issues (1).

The usage of information from preprints has been controversial. Advocates of preprints claim that this medium can accelerate access to science findings and improve quality of published works by permitting faster feedback of the work by the scientific community before publishing. As well, they argue that the audience for preprints is larger, because many articles published in peer-reviewed journals do not have open access (2).

Those who are against the use of preprints state that many of these investigations may be flawed due to the lack of a peer-review process, and may generate unwarranted modifications to clinical practice that
might harm patients (3). Therefore, readers should be aware that articles on preprint servers are not final versions and might contain errors, and the information reported in preprints have not been reviewed or endorsed by any means by the scientific or medical community (1,2,4).

Coronavirus disease 19 (COVID-19) started on December 2019 in Wuhan, China and the World Health Organization (WHO) declared it as a pandemic on March 12, 2020 (5). Due to the drastic increase in scientific production on this novel disease, peer-reviewed journals have been overwhelmed, and in response, they have shortened their usual processes to accelerate the time it takes to publish an article (4). Nonetheless, due to the sanitary emergency, authors have been urged to share relevant information as soon as possible, thus, an increasing amount of studies have been posted as preprints (6).

In this study, we intend to evaluate how often preprints on pharmacological interventions against COVID-19 were cited, in spite of the fact that some of these preprints remained unpublished.

**Methods**

We conducted a search to identify preprints on pharmacological interventions for COVID-19 posted in medRxiv and bioRxiv from January 1st to March 31, 2020. The keywords used were based on prior published reviews of treatment (7,8), and we decided to include case series reporting pharmacological interventions (Figure 1) (9).

For each preprint, we gathered metadata including the DOI, title, authors, date of first publication on the server, and metrics (PDF downloads and abstract reads). To identify manuscripts that were already published on a journal, we used links to the published versions available in MedRxiv or bioRxiv, and we conducted an individual search for each preprint on Google Scholar. Then, we obtained data including the new DOI, journal of publication, final title, list of authors, date of submission, acceptance, and publication if available, for all the peer-reviewed published versions. Also, we obtained the average time to acceptance in those journals before the COVID-19 crisis. Finally, the number of citations of each article was obtained from Google Scholar.

*Statistical analysis*

To compare the number of citations, we performed a Mann-Whitney U test on IBM SPSS Statistics for Windows, version 25 (IBM Corp., Armonk, N.Y., USA).

**Results**

From a total of 97 preprints, 14 were published on a peer-reviewed journal, and 83 remained unpublished up to May 16, 2020. The publication rate in this sample was only 14.4%. Considering the date in which preprints were posted for the first time on the servers and the date in which we collected data (May 16, 2020), a minimum time of 46 days, and a maximum of 109 days (median: 65 days) passed.
Figure 2 shows that the most common study types in preprints were basic science research (n= 39; 40.2%) and case series (n=34; 35.1%). Among the preprints that were eventually published, we identified basic science research (n=7; 50%), case series (n=6; 42.9%) and an interventional study (n=1; 7.1%).

Preprints that were eventually published had a significantly higher number of abstract (p=0.01) and PDF (p=0.04) downloads, when compared to preprints that remained unpublished (Table 1).

We identified preprints that had no citations at all (n=20; 20.6%), and we excluded them for our analysis. As well, we identified a preprint within the published group that was an outlier (article by Gautret et al.), which had 760 cites, thus, we also excluded this article for our analysis. After excluding this data, preprints that were eventually published on a journal had a significantly higher number of citations (median:13; IQR:6-34) than preprints that were not published (median:4 ; IQR:2-8 ; p=0.012).

Among the 14 manuscripts that were eventually published on a journal, the time that elapsed since date of posting on a preprint server to date of publication on a peer-reviewed journal ranged from 0 to 63 days (mean: 33.9 days). As well, the time from submission to a journal to acceptance in our sample ranged from 1 to 52 days (median: 20). In 2019, before the COVID-19 crisis, the time from submission to acceptance in these group of journals ranged from 27 to 193 days (median: 56) (Table 2).

Interestingly, we found that half (n=7) of the preprints that were published had modifications on the title or the result section, six articles (42.9%) had modifications in the methods, and five articles (35.7%) had modifications made to the discussion section. The list of authors only changed in two articles (14.3%).

**Discussion**

Preprints intend to accelerate the access to preliminary data for the scientific community, mainly to receive rapid feedback prior to entering a peer-review process, which is needed for publication in the majority of indexed journals (10). We found that preprints that were eventually published had significantly more abstract and PDF downloads than preprints that were not published, which reflects that presumably preprints with a higher quality were more often used by researchers.

Thus, medRxiv and bioRxiv, widely-known preprint servers (11), have a disclaimer on their homepage that states: “these are preliminary reports that have not been peer-reviewed. They should not be regarded as conclusive, guide clinical practice/health-related behavior, or be reported in news media as established information”.

We found that in spite of the fact that the time to acceptance in peer-reviewed journals during the COVID-19 crisis was almost four time shorter than the time to acceptance in 2019 for the same journals, only 14.4% of the preprints in our sample were published. Furthermore, in the same period, i.e. January to March 2020, we identified 1340 articles in PubMed related to therapeutic interventions against COVID-19, which is consistent with the idea that most of the preprints did not meet the requirements to be published.
Interestingly, a recent preprint estimated that the average time to publication in journals during the COVID-19 pandemic is approximately two months (4,12).

We found that half of the preprints that were subsequently published had significant modifications in the result section, which suggests that preprints can change importantly after peer-review, raising concerns on the possibility of significant errors in the data analysis of preprints that are not peer-reviewed and published, as previously reported (11,13).

In comparison to our findings, in previous infectious outbreaks such as Zika or Ebola, the publication rate for preprints on a peer-reviewed journal was of around 60% and 48%, respectively (14). However, only 174 and 75 preprints were posted during the Zika (Nov 2015 to Aug 2017) and the Ebola (May 2014 to Jany 2016) outbreaks, respectively (14). Until May 30, 2020, 3544 preprints about COVID-19 have been posted on medRxiv and 842 preprints have been posted on bioRxiv, evidencing the drastic increase in preprint production during the COVID-19 pandemic.

To our concern, we found that during the COVID-19 pandemic multiple preprints have been used in the development of clinical guidelines (15), public health policies, and the development of scientific articles (7,13,16).

Nonetheless, some preprints might contain useful information, for example, a study showed that the infectivity index, R0, calculated using data available on preprints was not different to the one estimated in peer-reviewed articles (16), and preprints on the viral sequence and structure have allowed for early investigation of potential therapeutic options or vaccines (4,17). However, preprints should be used responsibly, as they contain preliminary information that needs to be confirmed through a peer-review process.

Lack of a peer-review process in preprints may be an important implication, due to the fact that the basic screening process employed by preprint servers may not be enough to avoid the dissemination of flawed information (18). For example, a preprint that was posted on bioRxiv suggested significant molecular similarities between SARS-CoV-2 and HIV (19). Even though this preprint was later withdrawn, by the time that happened, it had already sparked controversy and conspiracy theories.

In spite of the fact that peer-review intents to be an exhaustive and thorough process that improves the quality of a manuscript, articles published on a peer-reviewed journal should not be taken as non-refutable knowledge. To illustrate this, a couple of peer-reviewed articles have been recently withdrawn from two prestigious journals due to significant concerns on primary data validity (20,21).

We acknowledge that our study has limitations

Main limitations to our study include the fact that we only included preprints on pharmacological interventions against COVID-19, and that we only used medRxiv and bioRxiv as preprint servers to obtain our sample. However, due to the large sample of preprints we included in our study, and the low publication rate we identified, a significant difference in other aspects seems unlikely.
Conclusion

Preprints should be used as mean of obtaining rapid feedback from the scientific community before it goes through a peer-review process. Also, preprints could be used as preliminary data to inform further research. However, preprints alone should not be used to guide clinical decision making, as data has not been peer-reviewed and may be flawed, resulting in unwarranted modifications to patient management.

Declarations

Ethics approval and consent to participate

Does not apply

Consent for publication

Does not apply

Availability of data and materials

The datasets generated and/or analyzed during the current study are available in the [Open Science Framework] repository, [https://doi.org/10.17605/OSF.IO/VJ27U]

Competing interest

None of the authors have something to declare

Funding

No extramural funding

Authors contribution

BN, DA, MM, IE, JG, and ET designed the study and collected the data. MM and IE mainly analyzed the data, processed the results, and developed the figures and tables. DA and ET designed the methods. BN and ET contributed to the discussion and conclusions. JG contributed to the introduction. All the authors contributed to the development of the original manuscript. All the authors agree with the final version of the manuscript and their contributions.

Acknowledgements

Does not apply

References


Tables

Table 1. Characteristics (median, IQR, and range) of metrics of the preprints on pharmacological therapies for COVID-19 posted on medRxiv or bioRxiv between January to March, 2020.

<table>
<thead>
<tr>
<th>Metrics of preprint server, n</th>
<th>Published preprints (n=14)</th>
<th>Unpublished preprints (n=83)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDF n</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median (IQR)</td>
<td>8287 (5887-19300.5)</td>
<td>4211 (1895-7247)</td>
<td>p=0.011</td>
</tr>
<tr>
<td>Range (Min-Max)</td>
<td>631-148827</td>
<td>670-87505</td>
<td></td>
</tr>
<tr>
<td>Abstract</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median (IQR)</td>
<td>17081 (9472-22362.5)</td>
<td>7100 (3595-17087)</td>
<td>p=0.04</td>
</tr>
<tr>
<td>Range (Min-Max)</td>
<td>1577-279632</td>
<td>516-115738</td>
<td></td>
</tr>
</tbody>
</table>

IQR=interquartile range.

Table 2. Detail information of the 14 peer-review articles on pharmacological therapies for COVID-19 that originally were posted as preprint from January-March, 2020.
## Time to Acceptance in COVID-19

<table>
<thead>
<tr>
<th>Authors</th>
<th>Date of submission</th>
<th>Date of acceptance</th>
<th>Time to acceptance (days)</th>
<th>Journal</th>
<th>Mean of the time to acceptance in 2019 (days) per journal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gordon et al., 2020</td>
<td>23-03-2020</td>
<td>22-04-2020</td>
<td>30</td>
<td>Nature</td>
<td>193</td>
</tr>
<tr>
<td>Qian et al., 2020</td>
<td>04-03-2020</td>
<td>05-03-2020</td>
<td>1</td>
<td>International Journal of Medicine</td>
<td>28</td>
</tr>
<tr>
<td>Fan et al., 2020</td>
<td>Information not available</td>
<td>Information not available</td>
<td>x</td>
<td>Clinical Gastroenterology and Hepatology</td>
<td>28</td>
</tr>
<tr>
<td>Zhou et al., 2020</td>
<td>05-02-2020</td>
<td>02-03-2020</td>
<td>26</td>
<td>Cell Discovery</td>
<td>Information not available</td>
</tr>
<tr>
<td>Jin et al., 2020</td>
<td>09-02-2020</td>
<td>01-04-2020</td>
<td>52</td>
<td>Nature</td>
<td>193</td>
</tr>
<tr>
<td>Beck et al., 2020</td>
<td>22-02-2020</td>
<td>25-03-2020</td>
<td>32</td>
<td>Computational and Structural Biotechnology Journal</td>
<td>57</td>
</tr>
<tr>
<td>Hu et al., 2020</td>
<td>27-02-2020</td>
<td>03-03-2020</td>
<td>5</td>
<td>Science China Life Sciences</td>
<td>40</td>
</tr>
<tr>
<td>Yuan et al., 2020</td>
<td>12-02-2020</td>
<td>03-03-2020</td>
<td>20</td>
<td>PLOS ONE</td>
<td>157</td>
</tr>
<tr>
<td>Jeon et al., 2020</td>
<td>Information not available</td>
<td>Information not available</td>
<td>x</td>
<td>Antimicrobial Agents and Chemotherapy</td>
<td>47</td>
</tr>
<tr>
<td>Qiu et al., 2020</td>
<td>Information not available</td>
<td>Information not available</td>
<td>x</td>
<td>Journal of Medical Virology</td>
<td>Information not available</td>
</tr>
<tr>
<td>Zehender et al., 2020</td>
<td>16-03-2020</td>
<td>24-03-2020</td>
<td>8</td>
<td>Journal of Medical Virology</td>
<td>Information not available</td>
</tr>
<tr>
<td>Abbott et al., 2020</td>
<td>11-03-2020</td>
<td>13-04-2020</td>
<td>33</td>
<td>Cell</td>
<td>56</td>
</tr>
<tr>
<td>Gautret et al., 2020</td>
<td>16-03-2020</td>
<td>17-03-2020</td>
<td>1</td>
<td>International Journal of Antimicrobial Agents</td>
<td>33,6</td>
</tr>
</tbody>
</table>

**Median of time to acceptance** | 20 | Median of time to acceptance | 47

---

**Figures**
Search Strategy.

Preprint servers: medRxiv, bioRxiv

Search criteria:

Keywords: “chloroquine”, “interferon”, “hydroxychloroquine”, “azithromycin”, “favipiravir”, “lopinavir”, “remdesivir”, “tocilizumab”, “ivermectin”, “baloxavir”, “darunavir”, “camostat”, “ribavirin”

Published between January 1, 2020 and March 31, 2020

Results obtained by keywords:

Chloroquine: 103
“Interferon”, “Covid-19”: 102
Hydroxychloroquine: 29
Azithromycin: 52
Favipiravir: 14
Lopinavir: 70
Remdesivir: 59
Tocilizumab: 19
Ivermectin: 10
Baloxavir: 9
Darunavir: 14
Camostat: 8
Ribavirin: 53

Total: 542 preprints

Exclusion criteria

(Applied to abstract, or to full text, if information in abstract was insufficient)

1. Not related to SARS-CoV-2 infection
2. Did not report or analyzed data on pharmacological interventions
Figure 1

Search strategy: there were obtained 542 preprints. After removing duplicates and applying exclusion criteria, we included 97 preprints in the study.

Figure 2

Different types of designs were identified in preprints and the proportion of each among the total sample.