Who’s Guarding the Gate? The Reach of Prereviewed Emerging Science and Implications for Family Medicine Education

LaKesha N. Anderson, PhD; Christy J.W. Ledford, PhD

BACKGROUND AND OBJECTIVES: Online publication of prereviewed manuscripts disseminates research simultaneously to scientists, clinicians, and patients, enabling the media and public to act as scientific reviewers for studies that are not yet endorsed by the scientific and clinical community. This study describes the reach of prereview literature and frames it within the pursuit to teach evidence-based medicine.

METHODS: In this deductive content analysis, the primary unit of analysis was the individual preprint manuscript submitted to the medRxiv preprint server during the emergence of the COVID-19 pandemic. The coding scheme included study design, negative or positive findings, dissemination status (whether it was withdrawn from the server or eventually published), and three levels of reach: user engagement, news media coverage, and social media engagement.

RESULTS: Prereviewed manuscripts describe a variety of study methods. Dissemination status was significantly related to abstract views, manuscript views, news coverage, and social media exposure. Studies with negative findings had higher counts of abstract views, manuscript views, and news coverage, but no significant relationships were detected.

CONCLUSIONS: Results demonstrate that not only are scientists publishing negative findings, but that those studies reach a wide audience. Notably, eventually-withdrawn manuscripts, potentially containing incomplete or uncertain science, is reaching the public domain. Increasingly, family physicians will need to critically appraise emerging literature before it is peer reviewed, whether they encounter it in their own searches or when a patient presents information they found before an appointment.

(Fam Med. 2021;53(8):670-5.)

The practice of evidence-based medicine is a critical skill taught in family medicine.\textsuperscript{1,2} We teach family medicine trainees how to ask focused clinical questions, search for evidence that addresses those questions, appraise that evidence, and produce an evidence-based answer.\textsuperscript{3} Historically, this process has focused on evidence found in peer-reviewed literature. However, in 2020, as scientists and clinicians worked together to study the emerging threat of COVID-19, prereview literature entered both the public consciousness and medical literature. Preprints became available to clinicians via traditionally credible evidentiary sources such as the National Library of Medicine’s PubMed index, which, in June 2020, launched a preprint pilot program aimed at providing early access to COVID-19 research.\textsuperscript{4,5} Much of the early discussion on COVID-19 transmission was informed by preprint articles rather than peer-reviewed research.\textsuperscript{6} Prereview literature was also cited as evidence in decision-making in health care policy and practice. In June 2020, the Food and Drug Administration released new guidelines on the use of chloroquine phosphate and hydroxychloroquine sulfate, citing eight articles that question the effectiveness of these drugs in treating COVID-19.\textsuperscript{7} Of these eight articles, five were preprints published on the publicly available medRxiv preprint server.

Peer review serves a gatekeeper function,\textsuperscript{8} ensuring both quality and integrity as academic scholarship enters the evidence base. The review process acts as a sieve, separating reliable, valid, and rigorous science from rushed, possibly less accurate, and potentially harmful science. Peer review is the greatest factor in determining research quality and...
trustworthiness. Peer-reviewed publications have easily recognized gatekeeper roles. Editors and reviewers work with authors to improve a manuscript before eventually endorsing its quality. Without a positive endorsement from these peer reviewers, an article is not viewed as suitable for publication.

A different mechanism of dissemination, the preprint, seeks to increase the visibility of research via shortcut, by bypassing the peer-review process. The term preprint is different than online ahead-of-print or online first articles, which have completed the process of peer review. It is unclear why preprint servers have framed articles this way instead of as prereview, which is a more accurate term. No inherent gatekeeper ensures these articles meet scientific standards before they are disseminated to audiences. Rather, these articles are checked for basic content and organization, and likelihood to do harm. The medRxiv website describes “a basic screening process for offensive and/or non-scientific content and for material that might pose a health risk.” The website also recommends that preprints not be reported in the news media as established information, and that individuals reporting preprint results do so with the caveat that these reports were not evaluated by the medical community and may contain errors. Table 1 explains terms used for preprints and ahead-of-print peer-reviewed literature along with benefits and possible consequences of each manuscript type.

The use of preprints has steadily grown from their start in the 1960s when the National Institutes of Health (NIH) mailed preprint manuscripts to biologists. Today, more than 60 preprint servers are available around the world, including many that focus on science and medicine, such as medRxiv and F1000Research. In 2017, NIH began encouraging researchers to include draft preprints in their grant applications. Proponents argue that prereviewed research articles advance science, pointing to faster dissemination of scientific findings, opportunities to receive rapid feedback, and benefits to younger scholars to make their research more visible. As an example of how quickly preprints can be disseminated, preprints studies exploring the Zika and Ebola outbreaks were posted an average of 100 days prior to their peer-reviewed journal publication. In both science and medicine, where clinicians comprise the primary readership, quicker access to research findings is an attractive option.

Preprint dissemination may also overcome publication bias, where articles reporting negative findings are less likely to be published. Research finds that manuscripts reporting negative results are less likely to be published than those reporting positive results. Studies are most likely to be published when the results are statistically significant, confirm a priori beliefs, or are somehow shocking. However, publication bias may lead to unreliable evidence available for decision-making, which may hinder informed decision-making by clinicians and patients.

As family medicine trainees and physicians increasingly encounter preprints in PubMed searches and clinical conversations, we need to develop a better understanding of the phenomenon. The current study describes the reach and potential impact of prereview literature, framing the analysis within the context of teaching evidence-based medicine.

Method
This study is a content analysis of preprint manuscripts available on medRxiv that were written during the emergence of the COVID-19 pandemic. Contemporary preprint articles proliferated throughout the COVID-19 pandemic. The two most popular preprint servers for these articles were bioRxiv and medRxiv. Combined, these servers offered more than 14,700 articles related to the pandemic as of February 27, 2021.

This study describes the reach of preprint studies as well as two specific characteristics of these preprint studies: study design and publication bias, referred to as either positive

| Table 1: Definitions, Descriptions, Benefits, Concerns of Preprint Literature |
|-----------------------------|-----------------------------|
| **Term and Synonyms**       |                             |
| Preprint                    |                             |
| Prereview literature        |                             |

**Definition—How It’s Different Than Traditional Literature**

Manuscripts that have not been peer reviewed

**Potential Benefits**

- Quicker dissemination to science and health care communities
- Fast advice and comments from a wider audience
- Greater visibility for younger scholars

**Concerns**

- Lack the quality control that comes with peer review
- May contain methodological errors or present poorly-written findings that can be easily misinterpreted
- May provide information that has not been critically evaluated by appropriate experts, including peers and scientists
reach: user engagement, news media coverage, and social media engagement. Counts are provided in the metrics section of each medRxiv article. User engagement was operationalized as the number of total abstract views and total full manuscript (PDF) views. News media coverage was operationalized as the number of news stories generated about the manuscript and number of news outlets providing coverage of the manuscript. Social media engagement was operationalized as the number of unique tweets about a manuscript, number of Twitter users engaged with a manuscript, and the estimated Twitter audience-reach of a manuscript.

We coded three additional variables specific to the science of the manuscript: research questions, study design, and study findings (ie, did the treatment work?). A copy of the coding instrument can be obtained by contacting the corresponding author (L.N.A.).

A research assistant coded all abstracts in the sample for the objective items. We coded all full manuscripts for the study findings variable. When author coders disagreed on whether the study presented negative or positive findings, the manuscript PDF was reviewed until codes were agreed upon.

Results

The 39 articles reviewed included three types of research questions: tests of drugs as increased susceptibility markers, tests of prophylactic use of drugs, and tests of drug efficacy. Article methodologies included systematic reviews (including reviews of registered ongoing trials), computer or mathematical modeling, retrospective chart reviews (including case reports, case series, and observational cohorts), and efficacy trials. Manuscripts in the data set also included trial protocols that had not been completed. Table 2 presents study methods and findings grouped by treatment studied.

To understand the potential additional reach of prereview manuscripts, we used analyses of variance to test two independent variables onto reach outcomes. First, we tested how dissemination status (published in peer-review literature, published only in medRxiv, or published then withdrawn from medRxiv) was related to the reach of the manuscripts. Dissemination status was significantly related to abstract views, $F(2, 36) = 3.84, P = .03$; to manuscript views, $F(2, 36) = 3.98, P = .03$; to news coverage, $F(2, 36) = 4.14, P = .02$; and to social media exposure, $F(2, 36) = 7.78, P = .002$. Table 3 presents reach counts by category. Manuscripts that were published in peer-reviewed journals had significantly more abstract views, manuscript views, and news coverage than those that remained only on the medRxiv preprint server. However, manuscripts that were eventually withdrawn from the medRxiv server had significantly higher social media exposure that manuscripts that were eventually published or remained on the medRxiv server alone.

In the second set of ANOVAs, we tested findings type (positive or negative findings) by the reach variables. This test excludes studies that were withdrawn since we were unable to code for study findings. The test also excludes one comparative effectiveness study that determined positive findings for remdesivir and negative findings for hydroxychloroquine. Table 3 presents reach counts by category. Studies with negative findings had higher counts of abstract views, manuscript views, and news coverage, but no significant relationships were detected.

Discussion

We historically teach evidence-based medicine as the “critical appraisal of medical research literature.”

Many of our evidence-based medicine teaching practices rely on the understanding that medical research literature is peer-reviewed. Findings here indicate that learners are likely to encounter prereview medical research, via their own literature searches or via conversations with
patients. As family medicine educators, we need to equip learners with the skills to critically appraise research that may not have been peer reviewed. While steps are being taken to teach medical students about COVID-19 misinformation, these efforts often focus heavily on responding to media inaccuracies. Results demonstrate that, not only are scientists publishing negative findings, but also that those studies reach a wide audience. In fact, readers engaged with preprint studies with negative findings more than with preprint studies reporting positive findings. Two unique circumstances may influence this relationship. First, the credibility of speakers is important when communicating complex science. The wider conversation about COVID-19 and its potential treatments has included political figures whose positions afford them credibility, from emergency responders to political leaders. Sensational headlines related to these political figures may have increased scientists’ motivations to share not only the “right” treatment options but also the “wrong” ones, or to correct public miscommunication about treatment options. Second, the emergent and collegial nature of scientists searching for COVID-19 treatments may motivate scientists to disseminate findings to their fellow scientists to use in their own experiments. Additionally, scientists’ desire for a treatment may lead to flawed studies making it to the public domain.

It is interesting to note the number of negative findings on the preprint server. Authors reporting negative findings often face more critical comments about their research design, and clinical trials reporting negative findings take longer to publish and have a lower publication rate overall. However, as Oldehinkel states, researchers who report negative findings “should be praised...”

Table 2: Type of Findings by COVID-19 Treatment

<table>
<thead>
<tr>
<th>Study methods</th>
<th>Hydroxychloroquine n=27</th>
<th>Remdesivir n=4</th>
<th>Studies Both Treatments n=8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Efficacy trial</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Computer modelling</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Systematic review or meta-analysis</td>
<td>3</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Retrospective chart review</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protocol description</td>
<td>3</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Secondary data analysis</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Bench science</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Findings</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>7</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Negative</td>
<td>14</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Other</td>
<td>4</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Manuscript later withdrawn</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Manuscript later published</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 3: Relationships Between Dissemination Status and Manuscript Reach, and Findings Type and Manuscript Reach

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Dependent Variables: Reach (Mean, SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Abstract Views</td>
</tr>
<tr>
<td>Dissemination Statusa</td>
<td>Published (n=5)</td>
</tr>
<tr>
<td></td>
<td>medRxiv only (n=32)</td>
</tr>
<tr>
<td></td>
<td>Withdrawn (n=2)</td>
</tr>
<tr>
<td>Findings Typeb</td>
<td>Positive (n=14)</td>
</tr>
<tr>
<td></td>
<td>Negative (n=16)</td>
</tr>
</tbody>
</table>

a Dissemination status was significantly related to abstract views, F(2, 36)=3.84, P=.03; to manuscript views, F(2, 36)=3.98, P=.03; to news coverage, F(2, 36)=4.14, P=.02; and social media exposure, F(2, 36)=7.78, P=.002.
b No significant differences by findings type.
for their resistance to the temptation to give their work a positive gloss.23 The need for rapid research during a pandemic may help explain the dissemination of negative findings on the medRxiv preprint server. It is possible that negative findings were only reported because of scientists’ desire to combat misinformation or conflicting ideas about COVID-19 treatments. This finding presents an opportunity for family medicine educators to teach trainees about the value of reading studies with both positive and negative findings and considering the implications of these studies for their practice.

An unexpected, yet important finding is the large social media reach of withdrawn manuscripts, which indicates that incomplete or uncertain science reached the public domain. We cannot determine if social media conversations about these withdrawn articles focused on a debate of the methods and/or findings or if these conversations promoted the findings. However, research shows that once information is in the public domain, it may be seen as fact by audiences who are unable, or unwilling, to critically evaluate the science. At minimum, weaker research can be overstated in the media.35 Without a traditional gatekeeper, it may be difficult to ensure the accuracy and safety of information focused on human science and/or medical treatments. This finding puts extra pressure on medical educators to teach evidence-based medicine, incorporating higher-order skills to interpret original research papers such as study design and statistical interpretation.36 Increasingly, family physicians will need to critically appraise emerging literature before it is peer reviewed, whether they encounter it in their own searches or when a patient presents information they found before an appointment.

Lastly, it is important to note that of the articles reviewed in this study, five went on to be published in peer-reviewed journals while two were retracted by the authors. It is unclear if preprint servers play a role in making accurate, rigorous science better. What is known is that two retracted studies made their way to the server, which indicates that potentially rushed, unreliable, and harmful science was made available to the public and to science and health professionals. In the traditional model of peer review, these manuscripts would have been submitted as manuscript drafts that a gatekeeper would have flagged for correction or improvement before entering the public domain. This finding is even more striking because our study discovered that manuscripts withdrawn from medRxiv had significantly greater reach than other manuscripts in our sample.

Three limitations inform interpretation of these findings. First, this study was bound in the context of COVID-19, a novel disease and the first pandemic in the social media age.37 The articles reviewed in this study represent only those found in the author-designated subject area of “infectious disease” and focused only on the potential COVID-19 treatments hydroxychloroquine and remdesivir. The decision to limit the treatments reviewed is a result of heavy news coverage of the two medications examined. This search limitation did reduce the studies identified for inclusion in this study. Second, during a pandemic, science moves quickly in a rush to determine effective treatments. This study was conducted as COVID-19 cases were rapidly increasing in the United States, and therefore presents a static picture of a dynamic information environment. Since data were collected, it is likely that abstracts were updated with newer versions, or even retracted, and it is very likely that engagement numbers have increased. Third, the intense public and media interest in COVID-19 may have generated greater interest in scientific studies like the ones analyzed in this study. This likely would not be the case for more routine scientific studies, such as ones focused on treatments for osteoarthritis. However, the findings of this study indicate what may happen when studies explore topics of high public interest.

Although limitations exist, this study indicates a number of opportunities for future research. Research shows that once an idea reaches the public domain, it may be seen as fact by audiences. Those with low science literacy may be more vulnerable to reading unreliable, inaccurate science. However, an inability to correctly understand or decipher information is not relegated to those with low science literacy. Once misinformation is presented, it can take multiple efforts to unlearn a falsehood as misinformation is often resistant to correction.38 Future research should look into the roles of literacy and memory in both interpreting and disseminating research.

More research is needed to develop and understand the theoretical implications of our findings. COVID-19 is the first pandemic to occur in the age of social media,37 and what separates it from past pandemics is the speed at which information about the disease spreads.39 Clinical practitioners and their patients are left to seek information from preprint servers, without a gatekeeper, such as an editor, to help them determine the best science to disseminate to others.

In the rapidly evolving COVID-19 crisis, preprint servers appear to have amplified emerging science, with both positive and negative findings. These same channels enabled science to enter the public domain before the science was ready for implementing in practice. Although the peer-review process is under heavy scrutiny, as concerns about data manipulation and questionable research practices abound, peer review still plays a critical role as gatekeeper to help readers determine and decipher quality science. Thus, family medicine educators need to modify how they teach evidence-based medicine in this evolving evidentiary landscape.
ACKNOWLEDGMENTS: The authors thank Amanda Crawford for her critical role in data collection.

CORRESPONDING AUTHOR: Address correspondence to Dr LisaKesh A. Anderson, Johns Hopkins University Advanced Academic Program, 1765 N Street, NW, Washington, DC 20036. 202-534-1003. lander73@jh.edu.

References