Challenges to Large-scale Scientific Knowledge Sharing During the COVID-19 Pandemic

AUTHOR 1*, Institution
AUTHOR 2, Institution, State

The COVID-19 pandemic poses major challenges to researchers using existing cyberinfrastructures to collaborate and share COVID-19 scientific knowledge. This paper enriches the existing literature by investigating the following challenges: (1) the mass influx of publications imposes a cognitive overload on COVID-19 researchers, (2) increasing speed of publications requires additional efforts to preserve peer-evaluation quality, (3) decentralization of research communities imposes impediments to informing the larger community about inappropriate results/interpretations in papers, (4) rigidity of papers imposes barriers to updating prior findings, and (5) difficulties with establishing common ground make interdisciplinary research more challenging. We manually reviewed approximately 1,300 tweets about large-scale collaborative research problems. We conducted a thematic analysis of about 10% of these tweets through a grounded theory approach and affinity mapping. We call on the CHI community to design scalable solutions for rapid knowledge sharing that help researchers collaboratively summarize, organize, share, and peer-review/improve their knowledge.

CCS Concepts: • **Human-centered computing** → **Social content sharing**; *Collaborative content creation*; *Social networks*; **Computer supported cooperative work**; *Social media*.

Additional Key Words and Phrases: E-science, E-research, CSCW, Collaborative Research, COVID-19, Knowledge Graph, Knowledge Sharing

ACM Reference Format:
Author 1 and Author 2. 2018. Challenges to Large-scale Scientific Knowledge Sharing During the COVID-19 Pandemic. In *Woodstock '18: ACM Symposium on Neural Gaze Detection, June 03–05, 2018, Woodstock, NY*. ACM, New York, NY, USA, 46 pages. https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

There have been numerous studies on scientific data collection, sharing, filtering, analysis, and maintenance in the domains of E-science and CSCW. However, the design community has paid less attention to the dissemination, distillation, organization, and improvement of the result of analyzing, interpreting, and learning through the data, i.e., scientific knowledge. While traditional methods, such as knowledge sharing through journal publications, and classification of publications in knowledge repositories have been relatively efficient in past years, in the wake of the pandemic, new issues have arisen. In this study, we review the recent literature and analyze tweets by COVID-19 researchers explaining the difficulties they experience in scientific knowledge sharing, evaluating, and improving due to the nature of expedited research during the pandemic.

Within the pandemic, there is also an infodemic [125]. Misinformation about COVID-19 has spread across media platforms [32, 92], incorporated numerous topics [17] and exacerbated public

*Note 1

Woodstock '18, June 03–05, 2018, Woodstock, NY

ACM ISBN 978-1-4503-XXXX-X/18/06…$15.00
https://doi.org/10.1145/1122445.1122456

panic [35]. While all of this is bad, there exists another underlying cause of the infodemic, taking root in the place where the pursuit of novelty and rigorous evaluation is supposed to prevent misinformation, the inability to effectively share scientific knowledge during the pandemic.

Due to the massive influx of COVID-19 related knowledge, researchers are struggling to find enough time to understand the contributions and limitations of findings [16]. The number of publications is unprecedented, with the number of academic papers related to COVID-19 being approximately 12 times more than for the MERS outbreak and more than 30 times the SARS outbreak [49].

In a recent study, it was shown that the majority of COVID-19 related papers (80%) are from medical science and only 3% from computer science, psychology, and engineering together [48]. This distribution in the number of published papers indicates that there is not a robust multidisciplinary approach to fight against the pandemic.

Knowledge sharing has been hindered by a paywall of publications, with roughly 20% of publications costing money to access. To make things worse, the number of paywalled papers increasing faster than open-access ones [114], and expected to rise to as high as 50% [16].

Preprints (papers published before peer review) accelerate the knowledge sharing process [57] and have been widely adopted as a way to do so during the pandemic. Glasziou et al. [46] reported that views and downloads of preprints had increased 100 fold. While preprints have been found useful, they are not an alternative to peer-reviewed publications [71]. Embracing of preprints can be justified by the fact that around 6/24/2020, more than 283 papers appeared in preprints in comparison to about 261 papers in journals [68]. Many novel technologies were introduced to collect and centralize publishing to facilitate searching and finding papers. Additionally, some NLP based systems were deployed to retrieve information, classify and summarize them from those repositories. However, there is a doubt whether those technologies have succeeded to address problems of medical researcher [100].

Parallel to preprint repositories, researchers have used social media (e.g. Twitter) to disseminate information faster. It might have been overlooked so far, but the lack of communication among researchers and sharing research that has been failed may cause repetitive research. Researchers have also utilized messaging technologies like Slack to communicate with each other and get informed about the state of the are research [68].

Providing researchers with technologies supporting efficient knowledge sharing and trustworthy research is crucial to solving the current crisis.

2 BACKGROUND

The Web has drastically changed scientific knowledge generation, evaluation, dissemination, and improvement. This large-scale virtual lab has helped researchers, engineers, and technicians to asynchronously connect and collaborate to share, analyze, and interpret data, information, and knowledge. Social media, collaborative editing tools, cloud computing, and scalable data processing units, accessible data warehouses, cheap sensory data collectors, crowdsourcing data collection and labeling, and fast searchability of a plethora of research papers including pre-prints have significantly increased the speed and the volume of knowledge generation and dissemination. In addition, citizen science has enabled non-scientists to engage in the process of data collection, filtering, assessment, and labeling, which has significantly facilitated transdisciplinary knowledge coproduction. These new technologies and methods have been studied in multiple disciplines including cyberinfrastructure, e-Science, citizen science, and transdisciplinarity. Pacheco et al. [88] classified these disciplinary fields under the term "digital science" and defined it as "a system that yields common (digital) spaces for knowledge coproduction based on open access and connectivity" (p. 378).

A large spectrum of people like researchers, citizen scientists, students, academic staff benefits from digital science. They collaborate and share information through connected technologies such as cloud computing. A common knowledge space is required to collaborate interdisciplinarily. [88, p. 5,6] Cyberinfrastructure (CI) was coined to give a prospect of how computing technologies can be exploited to facilitate academic research or researchers (i.e., scientists, engineers, humanists) [6]. National Science Foundation (NSF) defines CI as a union of technology, interdisciplinary groups, and new educational and workforce enterprises. In other words, CI is simply a combination of shared "computing systems, data, information resources, networking, digitally enabled-sensors, instruments, virtual organizations, and observatories, along with an interoperable suite of software services and tool" among interdisciplinary research groups who generate, establish and utilize "transformative approaches to scientific and engineering discovery and learning [28]" [88, p. 2].

In European countries, CI corresponds t the notions of e-science (eS) and e-infrastructure. E-science is defined as "global collaboration in key areas of science and the next generation of infrastructure that will enable it. [52]"

Some times e-infrastructure is used instead of CI or to refer to the infrastructure part of CI. In this sense, e-science might be considered as a scientific method which facilitates the acquiring the latest digital platforms known as e-infrastructures [88, p. 2,3].

In the UK the realm of e-Science is so widely spread that not includes natural and physical sciences but also the social sciences and humanities. As a result, the term e-Science is being substituted by 'e-Research.' Consequently, Digital Humanities and e-Social Science have attracted more attention and financial support. e-Research can be seen as a collection of computing infrastructures and applications aimed to support distributed and multidisciplinary research collaboration. [56]

Several digital science challenges and trends have been investigated in the literature Pacheco et al. [88]:

- **Social and psychological aspects: Science of Team Science, Citizen Science and Transdisciplinarity**: concerns roles, goals, norms, incentives, conflicts, attitudes, training, project management, ethics, equity, actions, and reactions within research communities and those between researchers and non-researchers who directly or indirectly engage in research projects.
  - *Science of Team Science*: Hall et al. [50] comprehensively reviewed the empirical evidence and research gaps on collaboration in science including formation, composition, performance of scientific research teams, within and between group interactions, institutional influences, and interdisciplinary research.
  - *Citizen science*: Bonney et al. [15] gives an overview of the application, requirements, and difficulties with citizen science and suggests next steps and strategies to follow. They exemplify the dataset of crowdsourced observations of more than five million birds peer month, which has been the source of analysis in more than 100 research papers so far. Carlson et al. [18] explains the coral reef data collection and analysis by 90 Earthwatch volunteers who made 6,000 measurements from 2000 to 2008 in Jamaica, as one of the successful applications of citizen science to enhance zoological research and policy making. Crabbe [30] discusses the case of Anchorage Coastal Beluga Survey conducted by trained volunteers between 2008 and 2011. Through 444 hours of observation with 507 beluga whale sightings, they collected the data regarding the presence, characteristics, and behavior of beluga whales.
  - *Transdisciplinary research*: Pohl et al. [96] describe how transdisciplinary research can address wicked societal problems through engaging private and public sectors, representatives of the society, and researchers from multi-disciplines. In the same textbook [42],

Defila and di Giulio [34], through analyzing survey results, shed light on how inter- and transdisciplinary teams frame problems based on a common ground, to reach out to a consensus in their collaboration. They also discuss "content-rich moderation" and the importance of balancing group and individual work to achieve the best outcomes.

- **Technological aspects: E-Science and Cyberinfrastructure**:
  - *Cloud and grid computing*: Lee [73] review portability, interoperability, and scalability of cloud computing; and suggest a trajectory of best practices and standards for development, deployment, and fundamental research agendas. Nabrzyski et al. [82] provides a detailed description of the grid computing applications and technologies with a special focus on resource management.
  - *CSCW and e-research*: The affordances provided by CSCW technologies have enabled scientific collaboration to take place in large-scale, distributed, and international contexts. One of the prominent examples is the emergence of virtual labs and remote experimentations [3]. *Crowdsourcing scientific data management* is another application of CSCW in e-research. Lee [73] present Super-Ego as a framework for location information privacy management in ubiquitous environments and through a 2-week study show that their automated methods can predict privacy preferences for 80% and semi-automated methods for up to 90% of users. Law et al. [72] proposes strategies and circumstances of using crowdsourcing for generating, processing, or analyzing research data. Wilkinson et al. [127] define and promote the FAIR guiding principles to improve management and reusability of scientific data. O'Grady et al. [86] shed light on the issues with sparse sensor networks and suggest leveraging the power of citizen scientists as data collection relays. This trend of research has contributed to a greater interrelation between CSCW and e-science and e-research. Designers seeking to address e-science problems through CSCW must be able to integrate the dimensions of science, sociality, and technology and develope a socio-technical perspective and approach [56].
- **Political aspects: Science and Technology Management**: König and Gorman [65] study how public research funding agencies such as NSF and ERC identify, promote, and constrain interdisciplinary research and how organizational constraints affect them. Huutoniemi and Rafols [55] provides a meta-analysis of the notion of interdisciplinarity in research, how to evaluate its value and significance, and tools that help with this evaluation. Roemer and Borchardt [102] discusses concerns and controversies about measuring scholarly impact of publications using bibliometrics and different alternative, and arguably more robust objective metrics to measure quality and impact of scientific publications. Walker [118] compares national, transnational, and international science; the fact that science is being perceived as national and its implications.

2.1 Scientific knowledge sharing during the COVID-19 pandemic

The rapid spread of the coronavirus was met with an increasing number of publications and scientific knowledge. However, the existing infrastructures for knowledge sharing failed to cope with these abrupt changes. In this section, we investigate the literature about knowledge sharing problems during the pandemic and evolved solutions to address those issues.

2.2 Repetitions in COVID-19 Literature

Papes and Ozimec [91] complains about the copious number of repetitive papers and review letters and call for a solution for the upcoming similar crisis. Majority of early COVID-19 articles provide no new information on the disease [36]. In a recent study, a total number of 14719 papers about COVID-19 were collected from Google Scholar, Springerlink, and PubMed. By using Mendeley

Desktop they recognized 8499 were duplicate studies [117]. In another study, 10 paper out of 157 collected papers were identified as repetitive papers [106]. Which might be because of the poor communication among researchers and being informed about ongoing research by other communities which highlights the importance of having robust social networks.

2.2.1 Use of Social Media for knowledge Sharing. Twitter with more than 330 million users [109], turned out to be a useful platform for sharing health care information [126]. Twitter was used extensively to share information during the pandemic to the extent that there is a new COVID-19 related Tweet every 45 milliseconds [2]. Between January 16 and March 15 there were 2, 792, 513 Tweets and 18, 168, 161 retweets about COVID-19 [109], and by March 30, 2020, the number of tweets using #COVID-19 reached 386, 686, 511 [104]. About 6, 162 COVID-19 literature was mentioned in 1.4 million tweets. It was shown that about 68.1% of COVID-19 publications were mentioned by Twitter users at least once. Papers were buzzing within the first couple of days after their publication [39]. For instance, a paper titled "uncanny similarity of unique inserts in the 2019-nCoV spike protein to HIV-1 gp120 and Gag" with 21.360 mentions was among the top ten controversial papers on twitter which was retracted later [40]. As long as very few percent of annually published two million academic papers, tend to be cited and employed by other researchers, discussing papers on Twitter enhances the probability of papers to be cited and employed [85]. Currently, there are 27 active Twitter-based journal clubs that use Twitter to discuss and learn about important publishings. Recently the discussions are often centered around COVID-19. After a discussion, reviews are archived to be searched and retrieved later.

A shortcoming of scientific communication on Twitter is that tweets are evanescent. Additionally, scrolling through tweets after the journal club is over, may not be a convenient and efficient way to grasp the ideas and information discussed live [110]. Twitter is considered a useful platform for the rapid dissemination of information. However, to overcome the flood of information it is suggested to limit the use of Twitter and only following verified sources [104]. Twitter, like other social networks, is not immune to misinformation and myths [87, 109]. One study showed 24.8% of COVID-19 related tweets incorporated misinformation, and 17.4% included unverifiable information[66]. It was shown that unreliable information spread at the same rate as reliable ones on Twitter [21] which makes Twitter vulnerable to misinformation and hence may limit its use by researchers.

YouTube has a wide spectrum of users from ordinary people and patients to healthcare staff. It has about 2 billion users [10]. YouTube with Coefficient of relative amplification of 0.1 cuts unreliable information out over the time [21]. However useful videos are overwhelmed by deceiving videos based on average view counts, view ratios, and "Video Power Indices" and only about 37.5% of COVID-19 related videos are useful and 10.4% of English videos have misleading content [5]. Another study showed that 27.5% of most-viewed English videos on YouTube contained unauthentic or misleading information, attaining over 62 million views [74]. This might be because uploaded videos on YouTube are not peer-reviewed. For instance there was a video on Youtube claiming finding an effective drug against COVID-19. Pharma Society tweeted (11 RT, 12 L): "Response to Claims of a COVID-19 Cure by Biobert Research group"Mr. Kijumbi, let the clinical trials and experiments on your drugs talk for you not the YouTube videos" [115]."

Wikipedia has been used for collaborative knowledge sharing. In the initial stages of the pandemic, more than 4, 500 new Wikipedia pages on COVID-19 were created and have attained almost 250M views. Content generation is continuing on WikiProject COVID-19 [22]. COVID-19 articles are usually cited in Wikipedia in advance to their official publication, because of early access versions of articles [22].

Although social media platforms have been used to alleviate knowledge sharing and scientific communication, lack of peer-review and dissemination misinformation and disinformation on those platforms [132] may hinder their utility.

2.2.2 Preprints and Repository Systems. After the pandemic began, preprint servers took a larger role in knowledge dissemination as the research community was ever more in need for sharing findings and results. Because fast dissemination is critical, but the traditional publication takes time, more researchers are utilizing the preprint servers, and in some cases, advertising their papers via social media, e.g., Twitter. Preprint servers like arXiv, Research Square, PrePrints.Org, OSFPreprints, bioRxiv, medRxiv have been used to quickly publish findings [57]. The number of posted papers in medRxiv and the page view has been largely increased [41].

Search and Repository tools have also been proposed to facilitate finding new publishing. CORD-19 dataset deployed a set of about 59, 000 papers including coronavirus literature from as early as the 1950s [16]. However, by one an estimate only 40% of the papers in CORD-19 are even directly related to COVID-19, and about 19, 000 papers in the dataset lack the full-text [16] and only 13% were deemed genuinely significant to the medical knowledge on COVID-19 [29]. CoAID [33] is another dataset focusing on news and social networks posts including verified true and false information to help differentiate fake and real news for covid2019 hashtag. They showed about 17% of the news was fake [33]. LitCovid is another literature hub for tracking up-to-date scientific data [20], WHO introduced its own database [124]. COVID-END is a database maintained by McMaster University [129].

Extracting valuable and meaningful knowledge from such a large set of literature is critical for decision making during a pandemic [51]. To this end, machine learning and NLP techniques were exploited to distill information from the ocean of publications. CAiRE-COVID [111] is a system that incorporates natural language processing (NLP) question answering (QA) to find related results based on a keyword and also provides a short summary of the papers [111].EVIDENCEMINER [119] is a web application for textual data exploration for life sciences. Users' query is received as a natural language statement and EVIDENCEMINER returns textual data from a background corpora (i.e. CORD-19) automatically. Covidex [133] is a similar online service that employs deep learning and neural architectures for ranking the papers recovered from CORD-19 and incorporates BioBERT to highlight the sentences in the returned results. SciSight is a dataminer service based on CORD-19. It shows related literature to facilitate filter search results and draws connections within the literature on browseable maps [16].

However, multiple concerns and criticisms are raised regarding the extensive usage of preprint servers even before the pandemic, one of which is the absence of a "peer-review" process. This is why the total number of peer-revied publications is roughly three times more than preprints [54]. If the lack of a peer-review process were to result in unsound science, it is far more consequential for the COVID-19 research to directly touch on a matter of life and death. As far as preprints are not peer-reviewed they might be prone to include nonfactual findings and misinformation and readers must self-assess reliability of information. Misinformation can costs lives and exacerbate the spread of COVID-19 [43, 83]. As a solution, fact-checking websites (e.g.MythBusters [123]) and crowdsourced peer-reviewing are proposed [103].

Numerous platforms are deployed to employ crowdsourced peer-reviewing. Science Rapid PREreview is a system for fast review of preprints related to COVID-19 [58]. Reviews on this web system are crowd-sourced and can be anonymous[107]. PubPeer is another service for post-publication peer review [71]. Additionally, to inform the research community about retracted papers Retractionwatch was launched as a web system aimed to make a list of retracted papers [37]. According to this website, more than 30 papers about COVID-19 were retracted, being 1410 times

cited totally. Each paper retracted because of misconduct costs the National Institutes of Health (NIH) about 425, 000US$ [37].

Preprints are not meant to be a substitution of peer-reviewed publication, they are only hoped to reduce the reviewing delay [71]. Researchers highlight the quality of papers, not the quantity, many papers have been shared on social media or as preprints— which are not peer-reviewed and may contain non-factual and misleading information [63, 68, 130].

2.2.3 Multidisciplinary Research During the Pandemic. The uneven distribution of COVID-19 related papers among different disciplines [48] may indicate the lack of interdisciplinary collaborative research. As long as fighting against the pandemic is a multifaceted task and involves researchers of a variety of disciplines, ignoring researchers of the fields outside medical science has raised some concerns. Miranda van Tilburg tweeted (6 RT, 68 L): "Sad to notice the pediatric research agenda around #COVID19 excludes psychosocial aspects of the #pandemic and #lockdown. We cannot view health as purely biologic. Currently the approach to this virus is behavioral (eg, social distancing) [112]."

2.3 CSCW and knowledge management

The history of knowledge management can be divided broadly into two generations: repositories and expertise or knowledge sharing. The former is associated with data and information flows while the latter deals with the transmission of knowledge through interpersonal communication channels [CoP , knowledge embedded within practice] [1].

Having mutual knowledge between collaborating researchers is integral to their succession in coordinating actions in order to reach their research goals. Mutual knowledge is the knowledge that communicating parties share and are aware that they share. However, for dispersed researchers whose main form of communicating and collaborating is through technological spaces, it is more challenging to maintain mutual knowledge. When relaying knowledge to their peers, researchers often experience what is called "uncertainties". This is defined as "the gaps in information needed for decision-making during the process of designing and executing a research project" [72]. Five identified challenges in maintaining mutual knowledge are as follows: "failure to communicate and retain contextual information, unevenly distributed information, difficulty communicating and understanding the salience of information, differences in speed of access to information, and difficulty interpreting the meaning of silence" [31].

Socioality and its connection to technology and science can be examined through the lens of communities of practice (CoP). CoP is a concept developed to investigate the ongoing practices and interactions within a community as the community's means of developing knowledge and expertise within a particular domain [122]. A scientific community can be examined as a CoP, which usually will be made up of groups of knowledge workers. Each group will focus on a particular problem or topic [60]. When compared to CoP in corporate contexts, scientific communities are more heterogeneous and its groups will operate with greater independence [38].

3 METHODS

We analyzed 130 tweets which was about 10% of the total tweets we manually collected from COVID-19 researchers discussing problems, or topics related to problems, with large large-scale collaborative research. The tweets we selected for our analysis were chosen based on their relevance to our research question: What are the challenges of large-scale scientific knowledge sharing and collaboration in the context of the pandemic?

We followed a grounded theory [45] approach to develop concepts and categories which informed our theoretical sampling. As we developed and refined these concepts and categories, we

simultaneously consulted relevant literature to compare our interpretations to existing theories. To facilitate our analysis, we used open and axial coding alongside an iterative construction and reconstruction of an affinity diagram.

3.1 Data collection

An initial phase of data collection allowed us to create a preliminary set of problem categories to refine our search. These initial categories were substantiated and expanded by a review of the pertinent literature [4, 16, 90]. In our initial literature review, we identified a few themes of challenges and met with four senior epidimiology PIs at R1 schools in the U.S. who had recently published a significant number of research papers on the COVID-19 pandemic. These researchers validated the themes we had identified and provided us with accounts of experiences they have had with these issues, including their frustrations with the magnitude of incoming research papers they have to review. They also introduced us to new themes and keywords to include in our research data collection process. Based on the keywords we learned from these PIs and the literature, we initiated our tweet search and expanded the list of keywords as we found them in tweets. We searched for tweets that included content related to our problem categories:

- Time-consuming to self-judge many rapidly released papers leading to cognitive overload;
- Difficult to efficiently inform the research community about inappropriate results/findings;
- Repetition among studies;
- Published papers are static and do not get updated dynamically;
- Difficulties in multidisciplinary research collaboration;
- Flawed preprint papers are shared without peer-review process;
- Marketing and subscription fees of journals leading to inaccessibily to knowledge;
- Personal ambitions and career incentives.

We followed two approaches to locating tweets:

(1) Keyword search for all tweets posted during the COVID-19 pandemic;
(2) Keyword search filtered to only include tweets made by a list of prominent COVID-19 researcher profiles we had followed.

For our second approach to tweet searches, we followed Twitter accounts of COVID-19 researchers from multiple disciplines including epidemiology, statistics, science communication, virology, and biostatistics. The researchers that we followed were usually verified by Twitter (indicated by a blue checkmark icon) or had their academic/research title(s) and affiliated institution(s) listed in their profile bio. These researchers had titles such as *Professor*, *Researcher*, *Scientist*, *Epidemiologist*, and *Director*. We also continuously followed new accounts to expand our author sources for collected data. During the retrieval period (May 1, 2020 – August 19, 2020), we used Twitter's search function and entered different keywords suggesting a problem (e.g., *difficulty*, *challenge*, *barrier*, *jargon*) combined with keywords suggesting collaborative, COVID-19 research contexts (e.g., *COVID-19*, *collaborative*, *research*, *multidisciplinary*, *across disciplines*. All the data we collected from Twitter is public and includes both original tweets and reply tweets in Twitter threads. We compared new data to our existing thematic analyses and iteratively constructed and reconstructed a large affinity diagram to aid in data and code organization and interpretation. All the tweets and their classifications from our affinity diagram are available in the appendix.

3.2 Thematic analysis

Our thematic analysis was facilitated by iterative phases of coding and affinity mapping. Through this process we were able to determine which tweets to analyze based on how well they addressed

	Title	Post	Sequences	Causes / Conditions	Themes
1	Natalie E. Dean tweets about the challenges she's encountered while reading through scientific papers	"Most common issues I see in scientific writing: • Doesn't explain why the problem is important. • No clear takeaway message. (How would the reader explain it to a colleague?) • Sentences too long/complex.	• Scientists produce papers with poor scientific writing > Papers are less useful to and less understood by other scientists • 💡Researchers are eager to share their insights with the research community > Researchers write conclusions that don't discuss limitations	• 💡Researchers often write without considering their audience • 💡Researchers assume that scientific papers should or must be written in complex, wordy, and jargon-laden prose in order to be published and/or accepted in	Scientific writing and discourse Info overload
2	Kareem points the need to show both sides to the argument using literature review to come to a consensus.	"In the age of covid, I've seen a few literature reviews that amount to "a list of all the papers that say I'm right". Gentle reminder: the goal of a literature review is to determine the state of the scientific CONSENSUS through a BALANCED survey of ALL available evidence."	• Incomplete/unbalanced literature reviews > Papers published that exhibit selection and confirmation bias • 💡Preprint papers are published without peer review > Scientists are able to conduct cherry picked literature reviews that reinforce their	• 💡Covid-19 research is being published through preprint papers which lack peer review • 💡Scientists are under pressure to publish Covid-19 research with strong and definite conclusions/implications	No peer review Criticism of claims
3	Steven Nono tweets about the importance of science communication with so many papers being published	This pandemic has proven to me that researchers have to spend less time in lab and more engagement in #SciComm. Great COVID-19 research is happening but it's being lost in the sea of journal publication. We need more front-facing researchers communicating results + implications.	• Important Covid-19 research is conducted > Research is "buried" in the influx of new journal articles	• Massive influx of literature being published on Covid-19 • Papers are published and their significance is not explicit / communicated • 💡Researchers don't have means of easily surfacing most important and relevant	Info overload Scientific writing and discourse Limits of journals
4 1	Ellie Murray tweets about how to communicate with other scientists from other disciplines to collaborate	This pandemic has scientists & scholars from a multitude of fields all working towards a common goal, which is good! But it means when we write our research findings, we can no longer assume shared context or jargon among our readers.	• Media misinterprets science > scientists reinforce misconceptions > leading to further misunderstanding among non-specialists • Researchers lacking key context/expertise > Researchers fail to understand significance (or lack thereof) of research finding	• Researchers communicating with other researchers in different disciplines • Researchers lacking key context/expertise • Researchers across disciplines have the same goal	Scientific writing and discourse Disciplinarities and expertise Criticism of claims

Fig. 1. Our open coding of tweets using Airtable

our research question. We also considered whether a tweet addressed a specific problem and provided context.

To begin our thematic analysis we first looked at the causes and conditions of each tweet. We then used an open coding approach to code terms to specific phrases within these causes and condition statements.

These general themes emerged in our initial coding phase:

- Scientific writing and discourse
- Information overload
- Lack of peer-review
- Criticism of claims made by researchers
- Limits of journals
- Disciplinaries and expertise
- Knowledge access
- Credibility of researchers
- Roles/audience (public versus science)

After labelling tweets with associated codes, we categorized grouped codes using axial coding. These categories were then used to restructure our affinity diagram. Through an iterative theoretical sampling approach, we were able to redevelop, refine, and validate our concepts and categories.

4 ANALYSIS AND RESULTS

Online communities of researchers emerge out of a need to collectively summarize, share, and discuss recently published studies (i.e., to quickly and efficiently share scientific knowledge). This section focuses particularly on the inefficiencies and unsuitability of current information organization and filtration mechanisms of these platforms for the scientific community. We also highlight the consequences of not having an appropriate infrastructure in place for scientists during the COVID-19 pandemic and the necessity for it.

One example of Twitter being used to summarize and disseminate recent scientific research can be seen in a tweet by Professor Francois Balloux (3, 000 RT, 5, 100 L): "Over the last weeks, a substantial amount of new evidence has become available about immunity to #SARSCoV2. This information can be difficult to process and integrate. As such I felt it may be helpful to write a thread to summarise the information and provide some context (1/13) [7]. "Understanding copious

amounts of new literature regarding COVID-19 literature is timeconsuming and tiresome to do by oneself.

While Balloux acknowledges these difficulties, he also provides insight into an important solution - the power of crowds. By summarizing the literature and sharing it with others to more efficiently learn the novelty of the findings, he has saved other researchers more time than he spent in his efforts.

Much of the struggle COVID-19 researchers are facing with effectively sharing knowledge could be alleviated with more awareness of the issues. Through our analysis of tweets and recently published articles, we identified the following themes of challenges to the large-scale scientific knowledge sharing during the COVID-19 pandemic:

- Shortcomings of scientific knowledge sharing on Twitter;
- Preserving quality of evaluation while increasing speed;
- Difficulties in efficiently informing research community about inappropriate results/interpretations in papers;
- Reliance on bibliometrics and altmetrics;
- Repetition in Published Papers;
- Rigidity of papers;
- Difficulties in multidisciplinary research.

Throughout the rest of this section we will discuss our analysis of each of these challenges in details. We refer to all authors of tweets by the specific profile name they used on Twitter at the time we collected their tweet. Additionally, we will be including the number of retweets(RT) and likes(L) as part of tweet descriptions.

4.1 Shortcomings of Scientific Knowledge Sharing on Twitter

In lack of a platform tailored for and populated by scientists for collective knowledge sharing, researchers have turned to Twitter as a very popular social media platform for research communication. Researchers have turned to Twitter as a very popular social media platform for research communication [23, 59, 116]. However, this platform lacks key design affordances needed to address the particular needs of users engaged in scientific knowledge sharing. As evinced by Professor Francois Balloux's tweet in , Twitter is used by researchers to quickly and efficiently share knowledge, but the dissemination of misinformation and an experience that is not tailored to research communication has left researchers using the platform looking for more support. PTPeters tweeted (1 RT, 1 L) "The word limit, and fleeting nature of Tweets, makes this format best for a simple IT KSM: Knowledge Sharing Network. ... Twitter is not good format for debate... [98]."

Lacassin et al. [70] call the reliability and legitimacy of knowledge sharing on Twitter into question, due to the lack of peer-review of the content. Additionally, they claim that some scientists may be reluctant to share novel findings or knowledge on Twitter because there is a chance of publishing the data or analysis by other researchers rather than the original author. It can be dangerous to both the public and research community when unreliable content is widely disseminated on Twitter. Dr. Krutika Kuppalli describes this situation in Figure 2 referring to a tweet made by Michael Levitt, with $2K$ RT, $4.5K$ L.

Pershad et al. [93] discuss the drawbacks of knowledge sharing on Twitter in medical sciences. One issue is that the limited number of characters in tweets restrains researchers to briefly convey sophisticated research, which may result in miscommunications and negatively impact scientific communication. The second problem is the time required to go through all tweets to self-judge and filter the necessary information. For instance, 80% of tweets from reliable professionals are posted

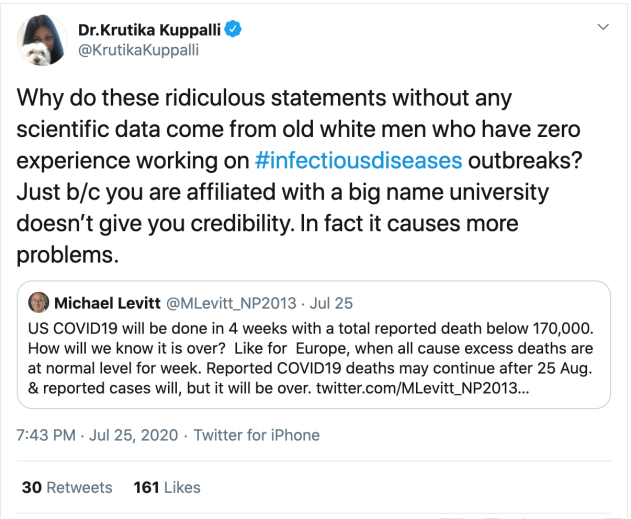


Fig. 2. Twitter used to disseminate inaccurate information by academics with institutional credibility [69].

during the workday, from 9 a.m. to 5 p.m. that may deprive health care workers to allocate their time to patient care routines.

Researchers are using Twitter to communicate science. While some take their papers to social media to share, others use the platform to actively challenge studies that they found misleading, as an unofficial form of quality control [16]. In a conversation among researchers on Twitter, the context of the conversation can get confusing. A researcher may have a diverse set of followers from many backgrounds: other researchers (possibly from different disciplines), laypeople, politicians, etc. This can have negative ramifications on how a message is framed and received, since different audiences can react differently to the same message. Kareem Carr tweeted (156 RT, 486 L) "One big problem is there are at least four possible ways for a scientist to have a conversation on Twitter...Talking to other scientists as a scientist...Talking to other scientists as a member of the public...Talking to the public as a scientist...Talking to the public as a member of the public...These conversations are happening all at once and it can change from sentence to sentence. Scientists are generally not good about specifying the role we're playing in public discussions. This can lead to terrible fights [19]."

4.2 Preserving Quality of Evaluation While Increasing Speed

Rapid dissemination of the COVID-19 research findings is important to facilitate the acceleration of the research. Dissemination of fast and rapidly reviewed papers makes researchers self-judge and screen most of the studies on their own, which is extremely time-consuming. While fast dissemination of factual information is beneficial, misinformation can grow fear and panic or cause lethal movements in society [77]. Palayew et al. [90] stresses the importance and the need for novel solutions to guarantee sustainable trust in the large-scale scientific publishing process during a pandemic. They conducted single-query searches on PubMed and analyzed data for COVID-19 and Ebola-related papers. The results showed that COVID-19 publications had a median acceptance time of 6 days, while it was about 15 days for Ebola. The fast dissemination of information is needed to keep up with a global pandemic but has cast doubt on the quality of publications and the probability of misinformation [90]. Some medical journals have drastically reduced publication

time [54, 68]. Peer-review delay has been reduced by 49%, or 57 days on average (even higher for some journals like Archives of Virology) [54]. Fourteen medical journals that published most COVID-19 papers reduced the average time taken from submission to publication to about 60 days [16], which may adversely affect the quality of published papers.

Lonni Besançon investigated 12, 682 publications related to COVID-19 on PubMed. Their study revealed that about 8% of papers were published in the day that they were submitted. What can be inferred is that the peer-review process might have been rushed for some papers. Editorial conflicts of interest were recognized in 43% of these papers. Such a short acceptance interval raises concerns about the integrity and clearness of the peer-review process [12]. Some concerns regarding quick peer-review surround the ethics of research being done. Mamas tweeted (74 RT, 145 L): "Pandemic vs paperdemic, the risks of speed science and research ethics as exemplified by #covid19 Sometimes we take at face value the honesty of researchers and veracity of data [76]." Mamas' concerns are noteworthy given recent scandals surrounding data that could have cost lives, namely Surgisphere's mishandling of Hydroxychloroquine data [108]. Sapan Desai, the founder of Surgisphere who has published two papers in the most reputable medical journals, is currently being sued in two different medical mal-practiced cases and had incentive to promote work using data from his company, and yet the Lancet and New England Journal of Medicine both published his work in fields completely unrelateed from his expertise, largely due to a lack of adequate peer-review [95].

Falsified data and misleading findings in publications may also stain the reputation of journals and decrease public trust in scientific findings. Gregory Makles tweeted (0 RT, 0 L): "NEJM is suddenly reliable again? Have they investigated their latest peer review incredible failure? [75]."

Speedy peer-review is not the only gateway for improper research to disseminate throughout the scientific community. Preprints, who publish work with no peer-review process also pose a threat. Block tweeted (0 RT, 0 L): "...I'm seeing a ton of COVID-19 pre-print research coming out, though that poses issues of its own because of lack of peer review Block [14]." Preprints have certainly been a useful resource for scientists to communicate with one another efficiently, but at what cost?

4.3 Difficulties in Efficiently Informing Research Community About Inappropriate Results/interpretations in Papers

Researchers have expressed concerns about the lack of an efficient mechanism to communicate their criticisms. After analyzing recent studies, they lack a unified platform to document their criticisms. Other researchers who come across the same studies may not be aware of those criticisms if they do not have contact with previous readers or forgot what has been critiqued.

We have noticed many instances that researchers express their frustrations in informing the research community when they identify issues in analysis, methods, results, or interpretations reported in a paper. An example of researchers using Tweets to inform the research community about retracting a paper is shown in Figure 3. Other examples of advocating on social media for retraction on Twitter is shown in Figure 4. The use of Twitter for science communication has been very valuable, however for reasons we discuss in subsection 4.1, system design tailored to science communication are needed.

In a recent study, it was found that the researchers' followers on Twitter are relatively low and only 25% of researchers have more than 429 followers[25]. Another study discusses that nearly 40% of scientists' followers are nonscholars [27]. Also, it was shown that only 1.1% of Web of Science researchers in the field of biomedical and health sciences use twitter. And the use of twitter by scholars is mostly (40%) confined to the USA and the UK [26].

Fig. 3. An example of researchers using blog posts, Tweets, and any medium they can to inform the research community about retracting a paper [13]

Fig. 4. Examples of advocating for retraction on Twitter [101, 121]

4.4 Reliance on Bibliometrics and Altmetrics Exacerbates the Issue

Bibliometrics is a set of methods for estimating scholarly impact which might generate conflicts and concerns [102]. Internet-based impact metrics known as alternative metrics or altmetrics

were proposed to mitigate those problems. There is an ongoing challenge to find and develop new methods for tracking scholarly impact [102]. Merely relying on citation count as a metric to measure the *quality* of research papers has exacerbated the difficulties bound up with efficiently informing the research community of inappropriate results and/or interpretations.

It is shown that the number of citation and publications are not a reliable gauge for scholars' talent and scientific merits [105]. NatureNews tweeted (0 RT, 0 L): "...Bibliometrics alone clearly do not offer a reliable way to gauge the quality of research [24]." Joe Kraus tweeted (0 RT, 0 L): "...Citation data does not provide reliable evidence of quality of research [67]." This heuristic can be misleading, especially, in case of retractions. We observed that retracted papers continue to receive citations. As an example, on 3/20/2020, a paper titled, "Hydroxychloroquine and azithromycin as a treatment of COVID-19: results of an open-label non-randomized clinical trial" was published by International Journal of Antimicrobial Agents [44]. The manuscript appeared on medRxiv on 3/16/2020. It was also submitted to the International Journal of Antimicrobial Agents on 3/16/2020. The paper was accepted for publication by the journal on 3/17/2020, and was published on 3/20/2020. This paper raised many concerns by multiple scientists on Twitter and ten days after publication, on 3/30/2020, another article titled, "A Rush to Judgment? Rapid Reporting and Dissemination of Results and Its Consequences Regarding the Use of Hydroxychloroquine for COVID-19" [62] was published by Annals of Internal Medicine, retracting Gautret et al. [44] and pointing out various flaws in it. However, as of 09/02/2020, while Gautret et al. [44] has received 2,304 citations on Google Scholar, Kim et al. [62] has received only 97. We hypothesized that most of the papers citing Gautret et al. [44] may have criticized or retracted it. However, reviewing a random sample of 100 papers out of the 2,304 that were citing the study, showed that 82 out of 100 papers were using the results or reporting follow-up studies, which indicates inefficiency of:

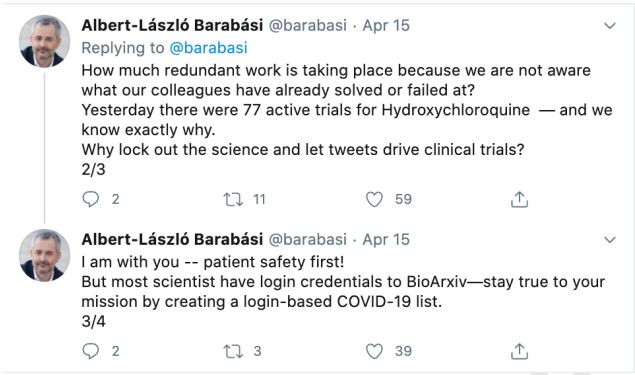
- Evaluation of publications merely based on their number of citations;
- Informing the research community about inappropriate results/interpretations in papers.

As another example, a paper titled "Cardiovascular Disease, Drug Therapy, and Mortality in Covid-19" was retracted on 4/6/2020 and had 358 citations however by 9/13/2020 the number of citation had reached 452. We hypothesized that it might be because the upcoming papers are retracting the original paper. We randomly went through 25 of those papers and to our surprise, 18 of those papers were really exploiting the findings of the retracted paper. To be able to compare the results we checked a third paper titled "SARS-CoV-2 infects T lymphocytes through its spike protein-mediated membrane fusion" was published in Nature Cellular & Molecular Immunology on 4/7/2020 and retracted about three months later on 7/10/2020 [120]. By 9/13/2020 this paper had 152 citations. We randomly picked 10 out of 40 papers citing that paper, which were published after the retraction date. We observed that only one paper was retracting the original paper. This might be due to prolonged delay before retracting.

The problem with retracted papers and how often they are cited is summarized in a tweet by Andrew R. Timming (34 RT, 334 L): "Why the h-index is a poor performance metric: I could write a bunch of egregiously bad papers and everyone would cite them by pointing out all the flaws. Plus, it punishes single authored papers and artificially inflates the worth of co-authored ones. Am I wrong? [113]"

4.5 Repetition in Published Papers

To allow researchers to keep up with the massive and rapid dissemination of academic papers, it is important that the redundancy of content is as minimal as possible. With no great way to signal to an entire community what has already been done, we are in danger of wasting resources on what is already known (see Figure 6 and Figure 7). Take the hydroxychloroquine controversy where an



tweet.png

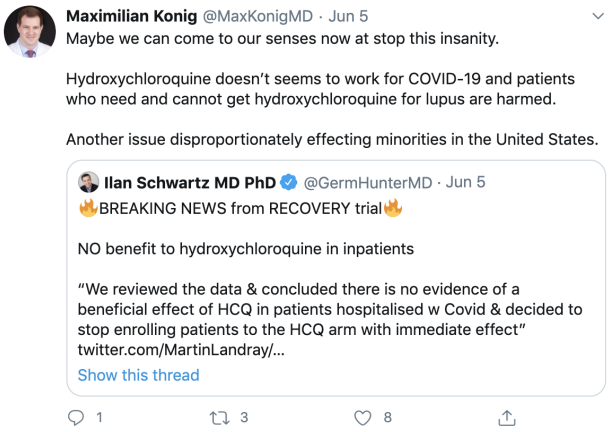
Fig. 5. Repetition in published papers [8]

Fig. 6. Repetition in published papers [94]

excess of evidence came out against its efficacy, possibly wasting valuable time, energy and risk in the process. In Figure 5 Barbási tweets about the issue of redundancy as it pertains to active trials regarding hydroxychloroquine, stating that there was a whopping total of 77 active trials within the previous day. Barbási goes as far as to place the blame on "...lock[ing] out the science...", and letting twitter drive research directions (see subsection 4.1 for more analysis on why twitter is not the optimal place for scientific dissemination).

Richard Poynder tweeted: “The number of publications related to COVID-19 is growing exponentially. Many articles are being hastily and non-critically published, contain repetitive or inaccurate information, illogical and non-evidence based recommendations, and are highly biased [97](5 RT, 12 L)."

Dmitri Zaitsev shared their hopes for reducing redundancy and the problems it creates: “When this pandemic subsides we could use this situation as a motive to reduce redundancy in publication, to modernize the way data is presented, and to avoid the same issues when the new pandemic arrives.- Any details about how that would happen? :) [131](0 RT, 1 L).”



tweet.png

Fig. 7. Repetition in published papers [64]

4.6 Rigidity of Papers

When a paper is published, its content stays static. Review papers of the publications of the previous day or week have gained significant attention, but they often miss some of the main points or misinterpret findings. As new investigations are made, there is no way to update these papers. Egon Willigh tweeted "I am really frustrated with the fact that our publications are static objects. why can't we improve them over time?... [128]."

This may be especially problematic when the number of citations is considered to be the primary metric to assess the quality of a publication and whether to pay attention to it. It would take a fair amount of time for new publications to get a comparable number of citations (Mathew's effect [11]) to grab attention and make the outdated papers obsolete, similar to the problems discussed in subsection 4.4. George Musser tweeted (2400 RT, 3200 L): "Google Scholar seems to be altering scholarly citation patterns. Citations are getting more concentrated: the same few papers get cited over and over, @jevinwest has found. People lazily cite whatever papers the search engine ranks highly [81]."

Jeffrey Barrett tweeted about desiring the capability of editing papers that have content that is redundant and not relevant in the near future (4 RT, 36 L): "I don't know who needs to hear this, but if you're writing a paper about Covid-19, just delete the sentence that states there is a pandemic and provides the current (already out of date) estimate of the number of global cases and deaths. Just. Delete. It. [9]." His sentiment echoes the claims from subsection 4.5, that too much space in COVID-19 literature is used to repeat the same things to the same same audience.

To mitigate this problem a system called Manubot has been suggested. Manubot is an open collaborative authoring platform providing handy version control and multiple access levels. Casey Greene tweeted, asking about its usage (1 RT, 13 L): "What if you're writing a collaborative review paper using manubot, and you use its capabilities to update from external data sources each night? [47]." While Manubot works to alleviate some of the problems, using it requires some computational skils (e.g version control workflows and GitHub). Also, discussions between writers are not tracked using Git [53]. As a result, collaborative authoring in many disciplines including medical science and epidemiology might be not very efficient using the platform [53].

4.7 Difficulties in Multidisciplinary Research

Large-scale collaboration across different disciplines presents obstacles in reaching shared research goals due to differences in terminology and context within each discipline. Many researchers have taken to Twitter, open conversations about this insufficiency.

Ellie Murray tweets about how to communicate with other scientists from other disciplines to collaborate (58 RT, 221 L): "This pandemic has scientists & scholars from a multitude of fields all working towards a common goal, which is good! But it means when we write our research findings, we can no longer assume shared context or jargon among our readers [80]."

"There's been a fair amount of discussion of the dangers of pre-prints being misinterpreted by the public or journalists. But, we don't talk about a potentially more dangerous issue: scientists & scholars in other disciplines getting the wrong message b/c they lack key context [79]."

Murray continues in another tweet (2 RT, 36 L): "Second, since combating the pandemic is our shared goal, any failure to communicate between scientists & scholars creates a barrier to achieving that goal, and thus has a cost in human lives lost [78]."

Another tweet, by Dr. Angela Rasmussen, discusses the difficulty of communicating across disciplines because of different definitions and contexts to terms used in different disciplines (9 RT, 53 L): "I think the real problem has been communication across disciplines. The terminology is confusing and means different things to different people, which is why I agree with SaskiaPopescu that we need to reconsider the language we use to discuss airborne' transmission [99]."

5 DISCUSSION AND DESIGN IMPLICATIONS

We found that researchers are having a difficult time keeping up with the massive influx of papers. This is due in part to the use of inefficient methods of organization and filtration of information that do not work at scale. Cognitive overload is being imposed on COVID-19 researchers as they struggle to stay aware of what has been studied.

Readers familiar with Natural Language Processing (NLP) may feel that the question of its efficiency in facilitating scientific knowledge summarization must be addressed. Many applications using NLP algorithms have been developed to summarize the newly published papers and reduce the overload on scientists who must review hundreds of newly published papers every day. However, researchers have raised concerns over NLP-based tools promoting retracted papers or papers with overstated findings [16]. One reason for this issue is the fact that the content of about 20% of subscribers-only journals papers is not free to download (which has a forecasted trend of a 50% increase) and therefore, cannot be reviewed by NLP-based tools [16]. Even among free publications, many of their papers are in a PDF format, which are not considered to be FAIR (Findable, Accessible, Interoperable, Reusable) and therefore cannot be processed by NLP algorithms [84]. To train machine learning models, researchers rely on preprint databases such as arXiv.org to be able to retrieve the content of a large number of papers for free [16].

In addition, reviewing the papers can be its own challenge. Scientific jargon is often complex and difficult to understand even by human standards. The terminology found in peer-reviewed literature is not always present in the mainstream text where the model training is performed. This problem is more common when the literature is domain-specific (outside of the training domain). This is present in the CORD-19 dataset (and others), leading to adverse effects on the fine-tuning of the text summarization models [61]. Some scholars also doubt the efficacy of automated procedures to help the research community. In our informal interviews with COVID-19 PIs, two of them told us "automated text summarization is very helpful for use-cases like distilling product reviews on Amazon.com, but research papers are totally different. If understanding the analysis or findings in

research studies was that easy, our undergraduate students would be able to do it and we would not need to spend so much time reviewing papers on our own."

COVID-19 researchers' need for a faster mechanism to learn about new studies on one hand, and their frustration with the NLP-based text summarization services, on the other hand, necessitates design and development of an online crowdsourced research hub, where researchers collaboratively summarize newly published literature and continuously update it as new findings are published. Every day COVID-19 researchers and their students study and summarize newly published research papers. They present their summaries in focus groups, classes, and journal clubs to help each other cope with the accelerated rate of publications. For the same reason, some research groups have also focused on publishing review papers of studies published in the past few days to help a larger community of researchers. However, focus groups and journal clubs are only helpful for small groups of people. Review papers are also rigid (as discussed in subsection 4.6) and any misinterpretation or missing findings of a study would not be corrected until new review papers make the old ones completely obsolete. The time and energy used by the existing research community could be more efficiently utilized if an online crowdsourcing research hub existed which would enable researchers to collaboratively summarize, discuss, comment, and criticize research publications.

In the remainder of this section, we discuss and ideate the necessary features of such a research hub to mitigate the issues discussed in the analysis and results section. We follow the same order of classifications of difficulties based on the themes identified from the analyzed tweets.

5.1 Difficulties in efficiently informing research community about inappropriate results/interpretations in papers

Our analysis of tweets indicates that many COVID-19 researchers experience difficulties to find a tribune to let other researchers hear their voice. They have expressed a need for a large-scale online research hub only for certified researchers, where they can easily access:

- Each other's names and validated credentials;
- Published papers with their number of citations;
- Comments about each publication and related discussions;
- Classified questions and answers about each topic;
- Direct contact channel to other researchers for follow-up questions/comments.

5.2 Reliance on bibliometrics and altmetrics

We found many tweets indicating the researchers' reliance on citation count to assess the quality of research papers has mislead many, especially in the case of retracted papers. While algorithms such as PageRank [89] have been proven helpful for website searching, similar citation-based ranking does not necessarily guarantee the quality of research papers. While altmetrics also consider mentions of a paper on tweets, blogposts, and other online social networks, it is still insufficient. A mere citation does not indicate anything about the quality of a study. In the research context, a citation should be assigned a positive, neutral, or negative weight. Then a weighted summation of the citations could be produced and it would be a more accurate metric for evaluation of research papers.

- **Positive**: if a study is cited to be admired/replicated/continued/rederived in another paper.
- **Neutral**: if a study is cited to be reviewed/meta-analyzed/classified in another paper.
- **Negative**: if a study is cited to be retracted/contradicted/corrected in another paper.

5.3 Difficulties in multidisciplinary research

The COVID-19 pandemic has made the need for multidisciplinary research more prominent than ever. However, COVID-19 researchers' tweets indicate that they need a more efficient communication method between researchers from different disciplines. They need a common ground around a shared set of vocabulary that can be easily deciphered by any researcher, regardless of their discipline and specialty area. To fulfill this objective the large-scale online research hub should provide a new search UI to researchers where they can easily find the term/concept they need to understand and facilitate back-tracking prerequisite concepts to be learned to understand the searched query.

Conventionally, scientific jargon or words are highly interdependent to each other where understanding one requires understanding others and to understand the others, one should read a graduate level textbook. This makes building a common ground between researchers extremely difficult. To solve this issue, a collaborative prerequisite linking mechanism is necessary where researchers link every concept/term/jargon to its prerequisites and expand upon them. This should be an asynchronous collaborative process because the scale of the terminology in different disciplines is growing fast and providing data to such a search tool requires crowdsourcing and collaboration between many scientists and possibly their students. The tool should enable researchers to start from a term/concept and back-track the prerequisite knowledge required to understand that term/concept recursively. Compared to the conventional method of studying a textbook to learn a new field's jargon, this tool enables researchers to only focus on the concepts/terms necessary to efficiently learn what they need without getting distracted by other concepts or discussions around them.

6 CONCLUSION

REFERENCES

[1] Mark S Ackerman, Juri Dachtera, Volkmar Pipek, and Volker Wulf. 2013. Sharing knowledge and expertise: The CSCW view of knowledge management. *Computer Supported Cooperative Work (CSCW)* 22, 4-6 (2013), 531–573.

[2] Eimear Lambe Alex Josephson. 2020. *Brand communications in time of crisis.* https://blog.twitter.com/en_us/topics/company/2020/Brand-communications-in-time-of-crisis.html

[3] Gustavo R Alves, Manuel G Gericota, Juarez B Silva, and João Bosco Alves. 2007. Large and small scale networks of remote labs: a survey. *Advances on remote laboratories and e-learning experiences* 15 (2007).

[4] Matt Apuzzo and David D Kirkpatrick. 2020. Covid-19 changed how the world does science, together. *New York Times* (2020).

[5] Omer Atac, Yunus Ozalp, Rifat Kurnaz, Osman Murat Guler, Meliksah Inamlik, and Osman Hayran. 2020. YouTube as an information source during the Coronavirus disease (COVID-19) pandemic. *medRxiv* (2020).

[6] Daniel Atkins. 2003. *Revolutionizing science and engineering through cyberinfrastructure: Report of the National Science Foundation blue-ribbon advisory panel on cyberinfrastructure.* Technical Report.

[7] Prof Francois Balloux. 2020. Over the last weeks, a substantial amount of new evidence has become available about immunity to #SARSCoV2. This information can be difficult to process and integrate. As such I felt it may be helpful to write a thread to summarise the information and provide some context.(1/13). https://twitter.com/ballouxfrancois/status/1277885591376535553

[8] Albert-Laszla Barabasi. 2020. How much redundant work is taking place because we are not aware what our colleagues have already solved or failed at? Yesterday there were 77 active trials for Hydroxychloroquine - and we know exactly why. Why lock out the science and let tweets drive clinical trials? 2/3. https://twitter.com/barabasi/status/1250500822003322881

[9] Jeffrey Barrett. 2020. I don't know who needs to hear this, but if you're writing a paper about Covid-19, just delete the sentence that states there is a pandemic and provides the current (already out of date) estimate of the number of global cases and deaths. Just. Delete. It. https://twitter.com/jcbarret/status/1298574234680283136

[10] Bibhusan Basnet, Suraj Bhattarai, Amit Khanal, Manisha Upadhyay, and Ashma Baruwal. 2019. Quality of YouTube patient information on prostate cancer screening. In *Baylor University Medical Center Proceedings*, Vol. 32. Taylor & Francis, 361–363.

[11] Jöran Beel and Bela Gipp. 2009. Google Scholar's ranking algorithm: an introductory overview. In *Proceedings of the 12th international conference on scientometrics and informetrics (ISSI'09)*, Vol. 1. Rio de Janeiro (Brazil), 230–241.

[12] Lonni Besançon, Nathan Peiffer-Smadja, Corentin Segalas, Haiting Jiang, Paola Masuzzo, Cooper A Smout, Maxime Deforet, and Clémence Leyrat. 2020. Open Science Saves Lives: Lessons from the COVID-19 Pandemic. *BioRxiv* (2020).

[13] Elisabeth Bik. 2020. Blog post describing some problems with the Gautret et al. paper about Hydroxychloroquine and Azithromycin treatment of COVID-19 infectionshttps://t.co/Ez5dnm8hoM. https://twitter.com/MicrobiomDigest/status/1242556354398531584

[14] Jonathan Block. 2020. In that case, what do you propose can be done now to enhance sharing in the research community? I'm seeing a ton of COVID-19 pre-print research coming out, though that poses issues of its own because of lack of peer review. https://twitter.com/jblockhcare/status/1266505013775474688

[15] Rick Bonney, Jennifer L Shirk, Tina B Phillips, Andrea Wiggins, Heidi L Ballard, Abraham J Miller-Rushing, and Julia K Parrish. 2014. Next steps for citizen science. *Science* 343, 6178 (2014), 1436–1437.

[16] Jeffrey Brainard. 2020. *Scientists are drowning in COVID-19 papers. Can new tools keep them afloat?* https://www.sciencemag.org/news/2020/05/scientists-are-drowning-covid-19-papers-can-new-tools-keep-them-afloat

[17] J Scott Brennen, Felix Simon, Philip N Howard, and Rasmus Kleis Nielsen. 2020. Types, sources, and claims of Covid-19 misinformation. *Reuters Institute* 7 (2020), 3–1.

[18] Barbara Svarny Carlson, Christy Sims, and Sylvia Brunner. 2015. Cook Inlet Beluga Whale, Delphinapterus leucas, observations near Anchorage, Alaska between 2008 and 2011: Results from a citizen scientist project. *Marine Fisheries Review* 77, 2 (2015), 115–130.

[19] Kareem Carr. 2020. DURING THE PANDEMIC, we scientists made a lot of science communication mistakes. One big problem is there are at least four possible ways for a scientist to have a conversation on Twitter. Let me explain ... 1/. https://twitter.com/kareem_carr/status/1282108190583410688

[20] Qingyu Chen, Alexis Allot, and Zhiyong Lu. 2020. Keep up with the latest coronavirus research. *Natur* 579, 7798 (2020), 193–193.

[21] Matteo Cinelli, Walter Quattrociocchi, Alessandro Galeazzi, Carlo Michele Valensise, Emanuele Brugnoli, Ana Lucia Schmidt, Paola Zola, Fabiana Zollo, and Antonio Scala. 2020. The covid-19 social media infodemic. *arXiv preprint arXiv:2003.05004* (2020).

[22] Giovanni Colavizza. 2020. COVID-19 research in Wikipedia. *bioRxiv* (2020).

[23] Kimberley Collins, David Shiffman, and Jenny Rock. 2016. How are scientists using social media in the workplace? *PloS one* 11, 10 (2016), e0162680.

[24] Nature News& Comment. 2010. david_colquhoun noahWG Bibliometrics alone clearly do not offer a reliable way to gauge the quality of research. https://twitter.com/NatureNews/status/20099932509704192

[25] Rodrigo Costas and Márcia R Ferreira. 2020. A Comparison of the Citing, Publishing, and Tweeting Activity of Scholars on Web of Science. In *Evaluative Informetrics: The Art of Metrics-Based Research Assessment*. Springer, 261–285.

[26] Rodrigo Costas, Philippe Mongeon, Márcia R Ferreira, Jeroen van Honk, and Thomas Franssen. 2020. Large-scale identification and characterization of scholars on Twitter. *Quantitative Science Studies* (2020), 1–21.

[27] Isabelle M Côté and Emily S Darling. 2018. Scientists on Twitter: Preaching to the choir or singing from the rooftops? *Facets* 3, 1 (2018), 682–694.

[28] Cyberinfrastructure Council. 2007. Cyberinfrastructure vision for 21st century discovery. *Arlington, VA: National Science Foundation). http://www. nsf. gov/pubs/2007/nsf0728/nsf0728. pdf* (2007).

[29] Stephen Covey. [n.d.]. Mercy of C0VID-19 Research. ([n. d.]).

[30] M James C Crabbe. 2012. From citizen science to policy development on the coral reefs of Jamaica. *International Journal of Zoology* 2012 (2012).

[31] Catherine Durnell Cramton. 2001. The mutual knowledge problem and its consequences for dispersed collaboration. *Organization science* 12, 3 (2001), 346–371.

[32] Jose Yunam Cuan-Baltazar, Maria José Muñoz-Perez, Carolina Robledo-Vega, Maria Fernanda Pérez-Zepeda, and Elena Soto-Vega. 2020. Misinformation of COVID-19 on the internet: infodemiology study. *JMIR public health and surveillance* 6, 2 (2020), e18444.

[33] Limeng Cui and Dongwon Lee. 2020. CoAID: COVID-19 Healthcare Misinformation Dataset. *arXiv preprint arXiv:2006.00885* (2020).

[34] Rico Defila and Antonietta di Giulio. 2017. Managing consensus in inter-and transdisciplinary teams: Tasks and expertise. In *The Oxford handbook of interdisciplinarity*.

[35] Anneliese Depoux, Sam Martin, Emilie Karafillakis, Raman Preet, Annelies Wilder-Smith, and Heidi Larson. 2020. The pandemic of social media panic travels faster than the COVID-19 outbreak.

[36] Nicola Di Girolamo and Reint Meursinge Reynders. 2020. Characteristics of scientific articles on COVID-19 published during the initial 3 months of the pandemic. *Scientometrics* (2020), 1–18.

[37] Ricardo Jorge Dinis-Oliveira. 2020. COVID-19 research: pandemic versus "paperdemic", integrity, values and risks of the "speed science". *Forensic Sciences Research* (2020), 1–14.

[38] Eckehard Doerry, Sarah A Douglas, Arthur E Kirkpatrick, and Monte Westerfield. 1997. Participatory design for widely-distributed scientific communities. In *Proc. of the Third Conference on Human Factors and the Web, Denver, CO.*

[39] Gunther Eysenbach. 2011. Can tweets predict citations? Metrics of social impact based on Twitter and correlation with traditional metrics of scientific impact. *Journal of medical Internet research* 13, 4 (2011), e123.

[40] Zhichao Fang and Rodrigo Costas. 2020. Tracking the Twitter attention around the research efforts on the COVID-19 pandemic. *arXiv preprint arXiv:2006.05783* (2020).

[41] Justin Fox. 2020. *A Pandemic Moves Peer Review to Twitter.* https://www.bloomberg.com/opinion/articles/2020-05-05/coronavirus-research-moves-faster-than-medical-journals

[42] Robert Frodeman, Julie Thompson Klein, and Roberto Carlos Dos Santos Pacheco. 2017. *The Oxford handbook of interdisciplinarity.* Oxford University Press.

[43] Laurie Garrett. 2020. COVID-19: the medium is the message. *The lancet* 395, 10228 (2020), 942–943.

[44] Philippe Gautret, Jean-Christophe Lagier, Philippe Parola, Line Meddeb, Morgane Mailhe, Barbara Doudier, Johan Courjon, Valérie Giordanengo, Vera Esteves Vieira, Hervé Tissot Dupont, et al. 2020. Hydroxychloroquine and azithromycin as a treatment of COVID-19: results of an open-label non-randomized clinical trial. *International journal of antimicrobial agents* (2020), 105949.

[45] Barney G Glaser. 1998. *Doing grounded theory: Issues and discussions.* Sociology Press.

[46] Paul P Glasziou, Sharon Sanders, and Tammy Hoffmann. 2020. Waste in covid-19 research.

[47] Casey Greene. 2020. What if you're writing a collaborative review paper using manubot, and you use its capabilities to update from external data sources each night? #AskingForAFriend https://t.co/NheVDeLtxu https://t.co/v7QYaCsGn9. https://twitter.com/GreeneScientist/status/1298575744113750016

[48] Milad Haghani and Michiel CJ Bliemer. 2020. Covid-19 pandemic and the unprecedented mobilisation of scholarly efforts prompted by a health crisis: Scientometric comparisons across SARS, MERS and 2019-nCov literature. *arXiv preprint arXiv:2006.00674* (2020).

[49] Milad Haghani, Michiel CJ Bliemer, Floris Goerlandt, and Jie Li. 2020. The scientific literature on Coronaviruses, COVID-19 and its associated safety-related research dimensions: A scientometric analysis and scoping review. *Safety Science* (2020), 104806.

[50] Kara L Hall, Amanda L Vogel, Grace C Huang, Katrina J Serrano, Elise L Rice, Sophia P Tsakraklides, and Stephen M Fiore. 2018. The science of team science: A review of the empirical evidence and research gaps on collaboration in science. *American Psychologist* 73, 4 (2018), 532.

[51] Elizabeth M Hechenbleikner, Daniel V Samarov, and Ed Lin. 2020. Data explosion during COVID-19: A call for collaboration with the tech industry & data scrutiny. *EClinicalMedicine* (2020).

[52] Tony Hey and Anne E Trefethen. 2002. The UK e-science core programme and the grid. In *International Conference on Computational Science.* Springer, 3–21.

[53] Daniel S Himmelstein, Vincent Rubinetti, David R Slochower, Dongbo Hu, Venkat S Malladi, Casey S Greene, and Anthony Gitter. 2019. Open collaborative writing with Manubot. *PLoS computational biology* 15, 6 (2019), e1007128.

[54] Serge PJM Horbach. 2020. Pandemic Publishing: Medical journals strongly speed up their publication process for Covid-19. *Quantitative Science Studies* Just Accepted (2020), 1–16.

[55] Katri Huutoniemi and Ismael Rafols. 2016. Interdisciplinarity in research evaluation. (2016).

[56] Marina Jirotka, Charlotte P Lee, and Gary M Olson. 2013. Supporting scientific collaboration: Methods, tools and concepts. *Computer Supported Cooperative Work (CSCW)* 22, 4-6 (2013), 667–715.

[57] Michael A Johansson, Nicholas G Reich, Lauren Ancel Meyers, and Marc Lipsitch. 2018. Preprints: An underutilized mechanism to accelerate outbreak science. *PLoS medicine* 15, 4 (2018), e1002549.

[58] Michael A Johansson and Daniela Saderi. 2020. Open peer-review platform for COVID-19 preprints. *Natur* 579, 7797 (2020), 29–29.

[59] Qing Ke, Yong-Yeol Ahn, and Cassidy R Sugimoto. 2017. A systematic identification and analysis of scientists on Twitter. *PloS one* 12, 4 (2017), e0175368.

[60] Andrea Kienle and Martin Wessner. 2005. Principles for cultivating scientific communities of practice. In *Communities and Technologies 2005.* Springer, 283–299.

[61] Virapat Kieuvongngam, Bowen Tan, and Yiming Niu. 2020. Automatic Text Summarization of COVID-19 Medical Research Articles using BERT and GPT-2. *arXiv preprint arXiv:2006.01997* (2020).

[62] Alfred HJ Kim, Jeffrey A Sparks, Jean W Liew, Michael S Putman, Francis Berenbaum, Alí Duarte-García, Elizabeth R Graef, Peter Korsten, Sebastian E Sattui, Emily Sirotich, et al. 2020. A rush to judgment? Rapid reporting and dissemination of results and its consequences regarding the use of hydroxychloroquine for COVID-19.

[63] Silvana Konermann. 2020. Just checked their results. The similarity is spurious. Out of 4 inserts they identify between NCov and SARS, 2 are found in bat coronavirus. Of the remaining two, only one is most similar to HIV, and is so short (6 AA) that the similarity is not higher than chance given database. https://twitter.com/SKonermann/status/1223344537772290048?s=20

[64] Maximilian Konig. 2020. Maybe we can come to our senses now at stop this insanity. https://twitter.com/MaxKonigMD/status/1268957290935267328

[65] Thomas König and Michael E Gorman. 2016. The challenge of funding interdisciplinary research: A look inside public research funding agencies. (2016).

[66] Ramez Kouzy, Joseph Abi Jaoude, Afif Kraitem, Molly B El Alam, Basil Karam, Elio Adib, Jabra Zarka, Cindy Traboulsi, Elie W Akl, and Khalil Baddour. 2020. Coronavirus goes viral: quantifying the COVID-19 misinformation epidemic on Twitter. *Cureus* 12, 3 (2020).

[67] Joe Kraus. 2013. Need moar #altmetrics MT ColizScott JudithFreedman sglassmeyer Citation data does not provide reliable evidence of quality of research. https://twitter.com/OAJoe/status/397801608131510272

[68] Kai Kupferschmidt. 2020. *A completely new culture of doing research.' Coronavirus outbreak changes how scientists communicate.* https://www.sciencemag.org/news/2020/02/completely-new-culture-doing-research-coronavirus-outbreak-changes-how-scientists

[69] Dr.Krutika Kuppalli. 2020. Why do these ridiculous statements without any scientific data come from old white men who have zero experience working on #infectiousdiseases outbreaks? Just b/c you are affiliated with a big name university doesn't give you credibility. In fact it causes more problems. https://t.co/k5HtXIKfHU. https://twitter.com/KrutikaKuppalli/status/1287171595941773312

[70] Robin Lacassin, Maud Devès, Stephen P Hicks, Jean-Paul Ampuero, Rémy Bossu, Lucile Bruhat, Desianto F Wibisono, Laure Fallou, Eric J Fielding, Alice-Agnes Gabriel, et al. 2020. Rapid collaborative knowledge building via Twitter after significant geohazard events. *Geoscience Communication* 3, 1 (2020), 129–146.

[71] Allison Langham-Putrow and Amy Riegelman. 2019. Discovery and scholarly communication aspects of preprints: Sources for online information. *College & Research Libraries News* 80, 9 (2019), 506.

[72] Edith Law, Krzysztof Z Gajos, Andrea Wiggins, Mary L Gray, and Alex Williams. 2017. Crowdsourcing as a tool for research: Implications of uncertainty. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. 1544–1561.

[73] Craig A Lee. 2010. A perspective on scientific cloud computing. In *Proceedings of the 19th ACM International Symposium on High Performance Distributed Computing*. 451–459.

[74] Heidi Oi-Yee Li, Adrian Bailey, David Huynh, and James Chan. 2020. YouTube as a source of information on COVID-19: a pandemic of misinformation? *BMJ Global Health* 5, 5 (2020), e002604.

[75] Gregory Makles. 2020. NEJM is suddenly reliable again? Have they investigated their latest peer review incredible failure?https://t.co/vFId6bTqRC. https://twitter.com/GregoryMakles/status/1301567570781392900

[76] Mamas Mamas. 2020. Pandemic vs paperdemic, the risks of speed science and research ethics as exemplified by #covid19 https://t.co/0Qm7uzNcqm Sometimes we take at face value the honesty of researchers and veracity of data. Recent events question this vass_vassiliou CMichaelGibson HeartOTXHeartMD pic.twitter.com/h4V8hjMz1X. https://twitter.com/mmamas1973/status/1272405779996446722

[77] Simon Scarr Manas Sharma and Kate Kelland. 2020. *Speed Science The risks of swiftly spreading coronavirus research.* https://graphics.reuters.com/CHINA-HEALTH-RESEARCH/0100B5ES3MG/index.html

[78] Ellie Murray. 2020. Second, since combating the pandemic is our shared goal, any failure to communicate between scientists & scholars creates a barrier to achieving that goal, and thus has a cost in human lives lost.4/6. https://twitter.com/EpiEllie/status/1262221352544722945

[79] Ellie Murray. 2020. There's been a fair amount of discussion of the dangers of pre-prints being misinterpreted by the public or journalists. But, we don't talk about a potentially more dangerous issue: scientists & scholars in other disciplines getting the wrong message b/c they lack key context. https://twitter.com/EpiEllie/status/1262221350250459136

[80] Ellie Murray. 2020. This pandemic has scientists & scholars from a multitude of fields all working towards a common goal, which is good!But it means when we write our research findings, we can no longer assume shared context or jargon among our readers. A thread of musings 1/6. https://twitter.com/EpiEllie/status/1262221347767402497

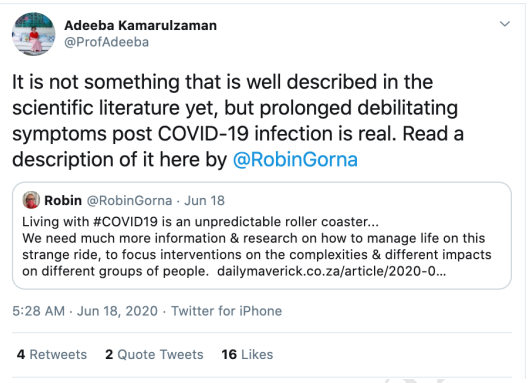
[81] George Musser. 2019. . https://twitter.com/gmusser/status/1170466414345375746

[82] Jarek Nabrzyski, Jennifer M Schopf, and Jan Weglarz. 2012. *Grid resource management: state of the art and future trends*. Vol. 64. Springer Science & Business Media.

[83] Scott Neuman. 2020. *Man Dies, Woman Hospitalized After Taking Form Of Chloroquine To Prevent COVID-19.* https://www.npr.org/sections/coronavirus-live-updates/2020/03/24/820512107/man-dies-woman-hospitalized-after-taking-form-of-chloroquine-to-prevent-covid-19?t=1594377313159

[84] Allard Oelen, Mohamad Yaser Jaradeh, Markus Stocker, and Sören Auer. 2020. Generate FAIR Literature Surveys with Scholarly Knowledge Graphs. *arXiv preprint arXiv:2006.01747* (2020).
[85] Avital Y O'Glasser, Rebecca C Jaffe, and Michelle Brooks. 2020. To tweet or not to tweet, that is the question. In *Seminars in Nephrology*, Vol. 40. Elsevier, 249–263.
[86] Michael J O'Grady, Conor Muldoon, Dominic Carr, Jie Wan, Barnard Kroon, and Gregory MP O'Hare. 2016. Intelligent sensing for citizen science. *Mobile Networks and Applications* 21, 2 (2016), 375–385.
[87] Sunday Oluwafemi Oyeyemi, Elia Gabarron, and Rolf Wynn. 2014. Ebola, Twitter, and misinformation: a dangerous combination? *Bmj* 349 (2014), g6178.
[88] Roberto CS Pacheco, Everton R Nascimento, and Rosina O Weber. 2018. Digital science: cyberinfrastructure, e-Science and citizen science. In *Knowledge Management in Digital Change*. Springer, 377–388.
[89] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. 1999. *The PageRank citation ranking: Bringing order to the web.* Technical Report. Stanford InfoLab.
[90] Adam Palayew, Ole Norgaard, Kelly Safreed-Harmon, Tue Helms Andersen, Lauge Neimann Rasmussen, and Jeffrey V Lazarus. 2020. Pandemic publishing poses a new COVID-19 challenge. *Nature Human Behaviour* 4, 7 (2020), 666–669.
[91] Dino Papes and Elizabeta Ozimec. 2020. Redundancy in reporting on COVID-19. *European Journal of Clinical Investigation* (2020), e13257.
[92] Gordon Pennycook, Jonathon McPhetres, Yunhao Zhang, Jackson G Lu, and David G Rand. 2020. Fighting COVID-19 misinformation on social media: experimental evidence for a scalable accuracy-nudge intervention. *Psychological science* 31, 7 (2020), 770–780.
[93] Yash Pershad, Patrick T Hangge, Hassan Albadawi, and Rahmi Oklu. 2018. Social medicine: Twitter in healthcare. *Journal of clinical medicine* 7, 6 (2018), 121.
[94] Philippe Philippe. 2020. #hydroxychloroquine inutile... again & again & again !!! https://twitter.com/Philipp31561573/status/1271553622007058433
[95] C Piller. 2020. Who's to blame? These three scientists are at the heart of the Surgisphere COVID-19 scandal. *Science Magazine* (2020).
[96] Christian Pohl, Bernhard Truffer, and Gertrude Hirsch Hadorn. 2017. Addressing wicked problems through transdisciplinary research. *The Oxford handbook of interdisciplinarity* (2017), 319–331.
[97] Richard Poynder. 2020. "The number of publications related to COVID-19 is growing exponentially. Many articles are being hastily and non-critically published, contain repetitive or inaccurate information, illogical and non-evidence based recommendations, and are highly biased." https://t.co/7kkYLVIRcU. https://twitter.com/RickyPo/status/1257180666669731847
[98] PTPeters. 2020. The word limit, and fleeting nature of Tweets, makes this formate best for a simple IT KSM: Knowledge Sharing Network. Fortunately Twitter is not good format for debate: as it is wasted of anyway. Do not argue w people on Covid - or ANYTHING for that matter. https://twitter.com/Paulp6363/status/1245801892917194752
[99] Dr. Angela Rasmussen. 2020. I think the real problem has been communication across disciplines. The terminology is confusing and means different things to different people, which is why I agree with @SaskiaPopescu that we need to reconsider the language we use to discuss "airborne" transmission. 2/2. https://twitter.com/angie_rasmussen/status/1282421963122634752
[100] Philip Resnik, Katherine E Goodman, and Mike Moran. 2020. Developing a Curated Topic Model for COVID-19 Medical Research Literature. (2020).
[101] Carlos del Rio. 2020. Hydroxychloroquine and azithromycin as a treatment of COVID-19: results of an open-label non-randomized clinical trial - Time to take this study down sciencedirect it was wrong topublish it, too much damage already. Take this study down! https://t.co/tDjaKagtKb. https://twitter.com/CarlosdelRio7/status/1248081364215308292
[102] Robin Chin Roemer and Rachel Borchardt. 2012. From bibliometrics to altmetrics: A changing scholarly landscape. *College & Research Libraries News* 73, 10 (2012), 596–600.
[103] Kevin Roitero, Michael Soprano, Beatrice Portelli, Damiano Spina, Vincenzo Della Mea, Giuseppe Serra, Stefano Mizzaro, and Gianluca Demartini. 2020. The COVID-19 Infodemic: Can the Crowd Judge Recent Misinformation Objectively? *arXiv preprint arXiv:2008.05701* (2020).
[104] Hans Rosenberg, Shahbaz Syed, and Salim Rezaie. 2020. The Twitter pandemic: The critical role of Twitter in the dissemination of medical information and misinformation during the COVID-19 pandemic. *Canadian Journal of Emergency Medicine* (2020), 1–4.
[105] Giancarlo Ruocco, Cinzia Daraio, Viola Folli, and Marco Leonetti. 2017. Bibliometric indicators: the origin of their log-normal distribution and why they are not a reliable proxy for an individual scholar's talent. *Palgrave Communications* 3, 1 (2017), 1–8.
[106] Anwar Santoso, Raymond Pranata, Arief Wibowo, Makhyan Jibril Al-Farabi, Ian Huang, and Budhi Antariksa. 2020. Cardiac injury is associated with mortality and critically ill pneumonia in COVID-19: a meta-analysis. *The American*

Journal of Emergency Medicine (2020).
[107] Simon Schwab and Leonhard Held. 2020. Science after Covid-19 Faster, better, stronger? *Significance* 17, 4 (2020), 8–9.
[108] Kelly Servick and Martin Enserink. 2020. The pandemic's first major research scandal erupts.
[109] Lisa Singh, Shweta Bansal, Leticia Bode, Ceren Budak, Guangqing Chi, Kornraphop Kawintiranon, Colton Padden, Rebecca Vanarsdall, Emily Vraga, and Yanchen Wang. 2020. A first look at COVID-19 information and misinformation sharing on Twitter. *arXiv preprint arXiv:2003.13907* (2020).
[110] Sinead Stoneman and Swapnil Hiremath. 2020. Twitter-based journal clubs: bringing critical appraisal to the social table. In *Seminars in nephrology*, Vol. 40. Elsevier, 264–272.
[111] Dan Su, Yan Xu, Tiezheng Yu, Farhad Bin Siddique, Elham J Barezi, and Pascale Fung. 2020. CAiRE-COVID: A Question Answering and Multi-Document Summarization System for COVID-19 Research. *arXiv preprint arXiv:2005.03975* (2020).
[112] Miranda van Tilburg. 2020. Sad to notice the pediatric research agenda around #COVID19 excludes psychosocial aspects of the #pandemic and #lockdown. We cannot view health as purely biologic. Currently the approach to this virus is behavioral (eg, social distancing). https://t.co/UK6KVzkfJk. https://twitter.com/DrvanTilburg/status/1261973218522533888
[113] Andrew R. Timming. 2020. Why the h-index is a poor performance metric: I could write a bunch of egregiously bad papers and everyone would cite them by pointing out all the flaws. Plus, it punishes single authored papers and artificially inflates the worth of co-authored ones. Am I wrong? https://twitter.com/timminglab/status/1258571672522637312
[114] Daniel Torres-Salinas, Nicolas Robinson-Garcia, and Pedro A Castillo-Valdivieso. 2020. Open Access and Altmetrics in the pandemic age: Forescast analysis on COVID-19 related literature. *BioRxiv* (2020).
[115] Pharma Society Ug. 2020. Response to Claims of a COVID-19 Cure by Biobert Research group "Mr. Kijumbi, let the clinical trials and experiments on your drugs talk for you not the YouTube videos" Full statement: https://t.co/TG614oUETr#COVID19UG #StaySafeUG. https://twitter.com/PharmacistsUg/status/1248944049559670785
[116] Laura Van Eperen and Francesco M Marincola. 2011. How scientists use social media to communicate their research.
[117] Apostolos Vantarakis, Ioanna Chatziprodromidou, and Thomas Apostolou. 2020. COVID-19 and Environmental factors. A PRISMA-compliant systematic review. *medRxiv* (2020).
[118] Mark Walker. 2012. The'national'in international and transnational science. *The British Journal for the History of Science* (2012), 359–376.
[119] Xuan Wang, Weili Liu, Aabhas Chauhan, Yingjun Guan, and Jiawei Han. 2020. Automatic Textual Evidence Mining in COVID-19 Literature. *arXiv preprint arXiv:2004.12563* (2020).
[120] Xinling Wang, Wei Xu, Gaowei Hu, Shuai Xia, Zhiping Sun, Zezhong Liu, Youhua Xie, Rong Zhang, Shibo Jiang, and Lu Lu. 2020. SARS-CoV-2 infects T lymphocytes through its spike protein-mediated membrane fusion. *Cellular & Molecular Immunology* (2020), 1–3.
[121] Retraction Watch. 2020. Hydroxychlorine-COVID-19 study did not meet publishing society's "expected standard" https://t.co/LSKO0GnSo9 pic.twitter.com/MR0dWs0lAn. https://twitter.com/RetractionWatch/status/1247183601944408065
[122] Etienne Wenger, Richard Arnold McDermott, and William Snyder. 2002. *Cultivating communities of practice: A guide to managing knowledge*. Harvard Business Press.
[123] WHO. [n.d.]. *WHO:Mythbuster*. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public/myth-busters
[124] WHO. 2020. *COVID-19 Global literature on coronavirus disease*. https://search.bvsalud.org/global-literature-on-novel-coronavirus-2019-ncov/
[125] WHO. 2020. *Munich Security Conference*. https://www.who.int/dg/speeches/detail/munich-security-conference
[126] Justin Wilford, Kathryn Osann, and Lari Wenzel. 2018. Social Media Use Among Parents of Young Childhood Cancer Survivors. *Journal of Oncology Navigation & Survivorship* 9, 1 (2018).
[127] Mark D Wilkinson, Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, Jan-Willem Boiten, Luiz Bonino da Silva Santos, Philip E Bourne, et al. 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data* 3, 1 (2016), 1–9.
[128] Egon Willigh gen. 2018. I am really frustrated with the fact that our publications are static objects... why can't we improve them over time? Fix a typo here, some character casing there... just the typography, not the content... sigh #onlyHuman. https://twitter.com/egonwillighagen/status/972908516901089280
[129] Www.mcmasterhealthforum.org. [n.d.]. COVID-END. https://www.mcmasterforum.org/networks/covid-end
[130] Chuan Xiao, Xiaojun Li, Shuying Liu, Yongming Sang, Shou-Jiang Gao, and Feng Gao. 2020. HIV-1 did not contribute to the 2019-nCoV genome. *Emerging microbes & infections* 9, 1 (2020), 378–381.



kamarulzaman tweet.png

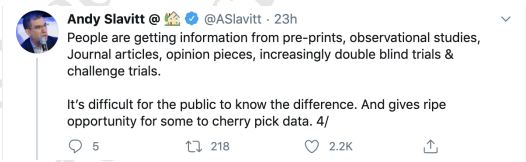
Fig. 8. Adeeba Kamarulzaman tweets about COVID-19 findings that are not yet widely talked about in scientific literature

[131] Dmitri Zaitsev. 2020. "When this pandemic subsides we could use this situation as a motive to reduce redundancy in publication, to modernize the way data is presented, and to avoid the same issues when the new pandemic arrives."- Any details about how that would happen? :). https://twitter.com/Dmitri145/status/1258350943927537664
[132] John Zarocostas. 2020. How to fight an infodemic. *The Lancet* 395, 10225 (2020), 676.
[133] Edwin Zhang, Nikhil Gupta, Rodrigo Nogueira, Kyunghyun Cho, and Jimmy Lin. 2020. Rapidly deploying a neural search engine for the covid-19 open research dataset: Preliminary thoughts and lessons learned. *arXiv preprint arXiv:2004.05125* (2020).

A TWEETS

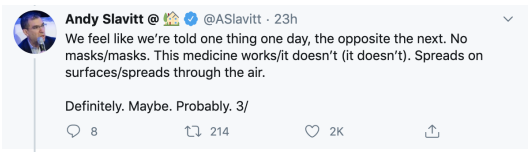
sim tweet.png

Fig. 9. Alice Sim tweets about fueling the infodemic with misinterpreted terminology



slavitt tweet.png

Fig. 10. Andy Slavitt tweets about the variety of sources of information and the public being unaware of the differences



slavitt tweet2.png

Fig. 11. Andy Slavitt tweets about confusion over scientists advice

rasmussen tweet.png

Fig. 12. Dr.Angela Rasmussen tweets about the miscommunication across disciplines

rasmussen tweet2.png

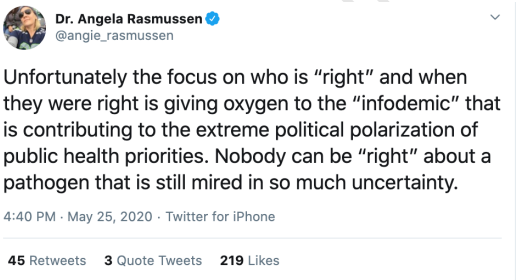


Fig. 13. Dr.Angela Rasmussen tweets about the focus on validity of statements

rasmussen tweet3.png

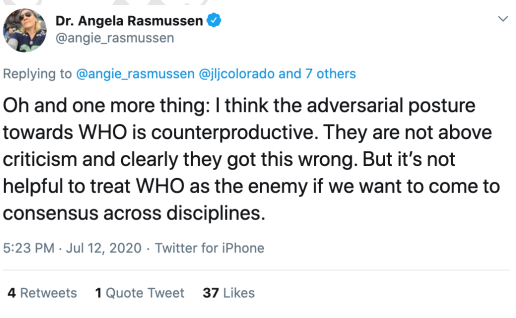
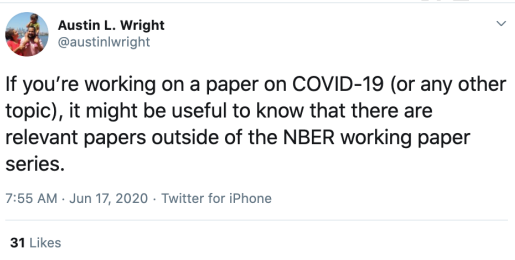


Fig. 14. Dr.Angela Rasmussen tweets how painting WHO as the enemy is counterproductive in terms of coming to a consensus

rasmussen tweet4.png

Fig. 15. Dr.Angela Rasmussen tweets about the harm of others not sharing data and being unable to make her own analyses

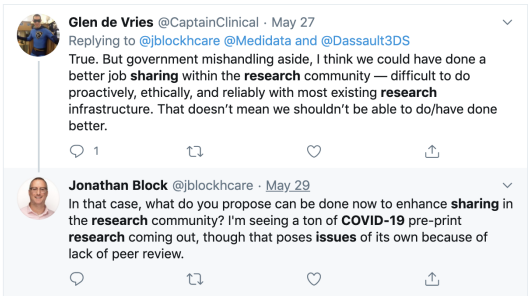


wright tweet.png

Fig. 16. Austin L. Wright tweets about using a variety of resources for finding papers

mukherjee tweet.png

Fig. 17. Bhramar Mukherjee tweets about the increasing number of papers



and vries tweet.png

Fig. 18. Jonathan Block replies to tweet by Glen de Vries about sharing research within the research community

michael gibson tweet.png

Fig. 19. C. Michael Gibson tweets about surgisphere (a health analytics company) lying about collaborating with NHS Scotland

rivers tweet.png

Fig. 20. Caitlin Rivers tweets the original WHO report on COVID-19 modes of transmission

rivers tweet2.png

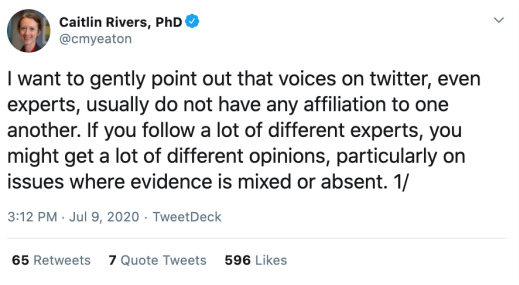


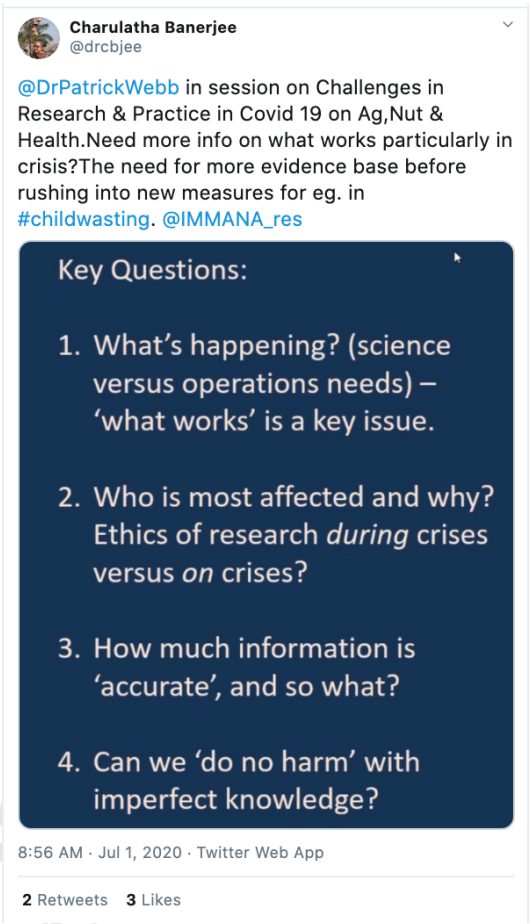
Fig. 21. Dr.Caitlin Rivers tweets about the confusion led by many different expert opinions being shared on Twitter

rivers tweet2 contd.png

Fig. 22. Dr.Caitlin Rivers tweets about the confusion led by many different expert opinions being shared on Twitter

chen and arun divakaruni tweet.png

Fig. 23. Caroline Chen tweets about being clear and defining terminology when speaking to an audience as scientists and Arun Divakaruni replies in agreement

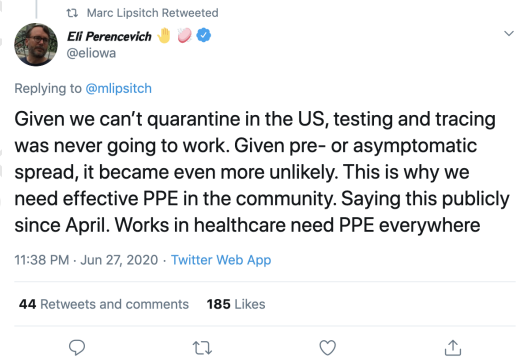


banerjee tweet.png

Fig. 24. Charulatha Banerjee tweets about the need for more evidence base for publishing on COVID-19

willighagen tweet.png

Fig. 25. Egon Willighagen tweets about the inaccessibility to papers and their cited papers



perencevich tweet.png

Fig. 26. Eli Perencevich tweets about how he's been advocating for PPE since April

murray tweet.png

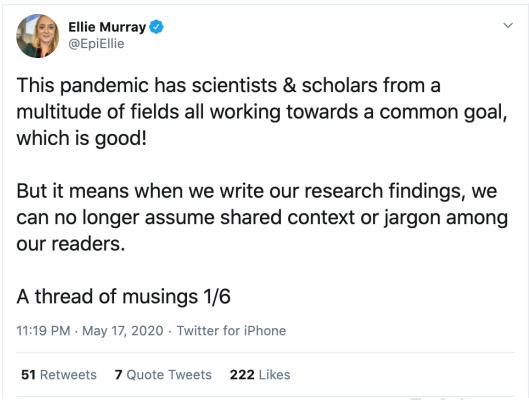
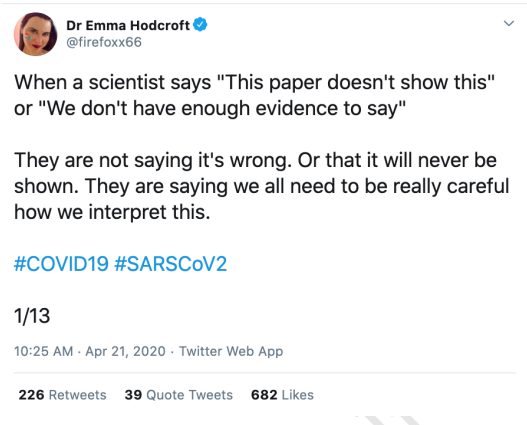


Fig. 27. Ellie Murray tweets about precautions when sharing research findings

murray tweet contd.png

Fig. 28. Ellie Murray tweets about precautions when sharing research findings

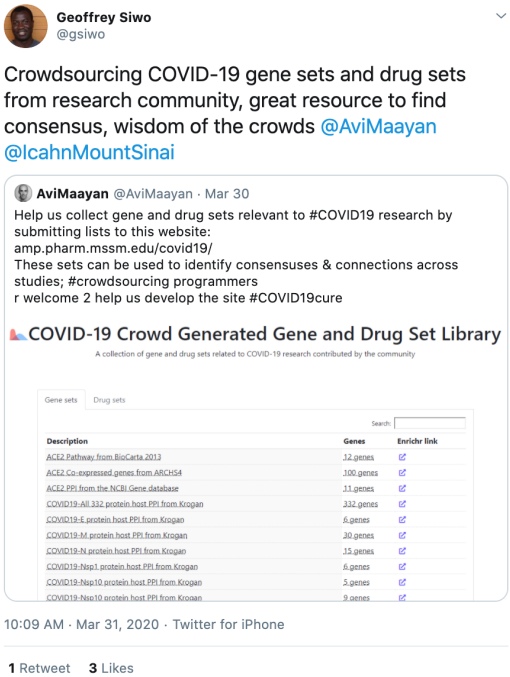


hodcroft tweet.png

Fig. 29. Dr. Emma Hodcroft tweets that when a scientist points out that a paper doesn't have enough evidence, they are not saying it is wrong

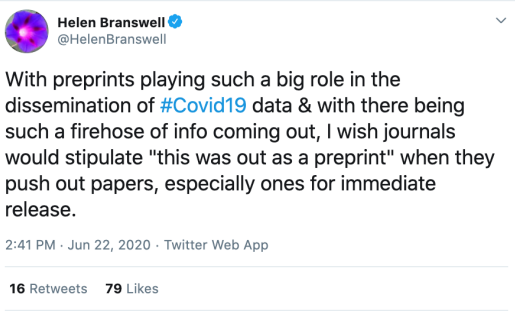
carson tweet.png

Fig. 30. Gail Carson tweets a call to action on responsible communication of scientific research



siwo tweet.png

Fig. 31. Geoffrey Siwo tweets about crowdsourced COVID-19 gene library that helps reach consensus



branswell tweet.png

Fig. 32. Helen Branswell tweets about how journals should be clear about using COVID-19 preprints

todaro tweet.png

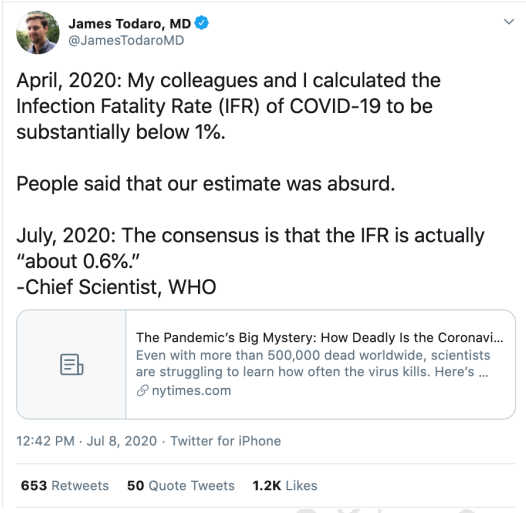


Fig. 33. James Todaros tweets about not finding a consensus on the IFR and current WHO estimates

pogue tweet.png

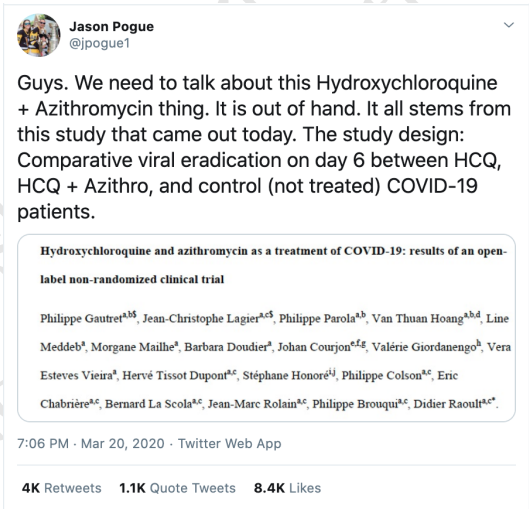


Fig. 34. James Pogue tweets about opening up a conversation on Hydroxychloroquine and Azithromycin

howard tweet.png

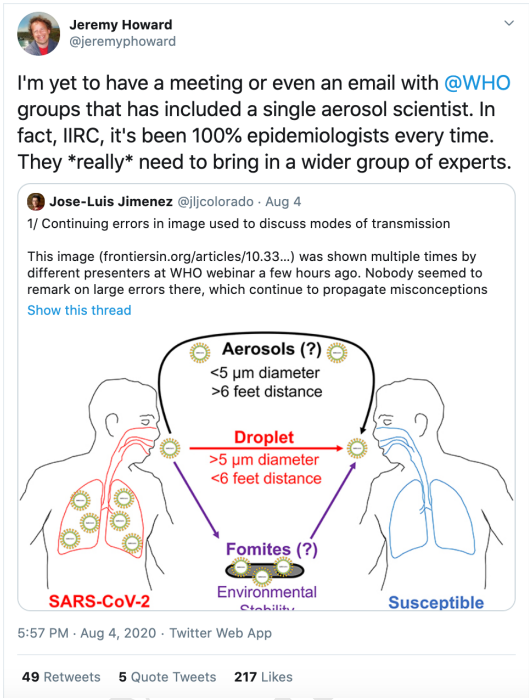
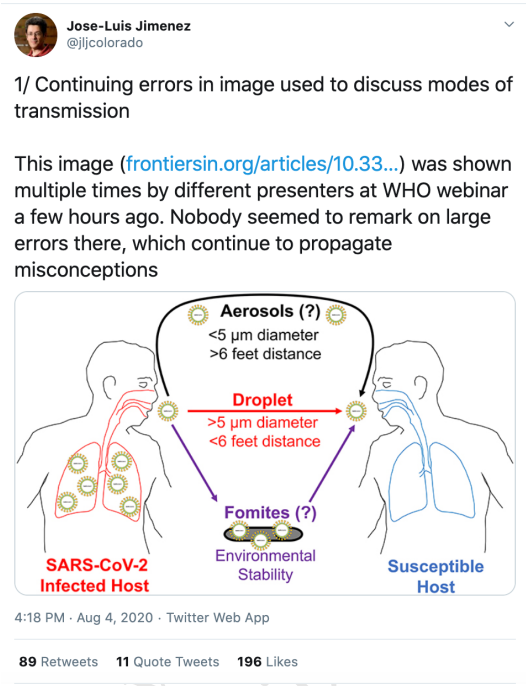
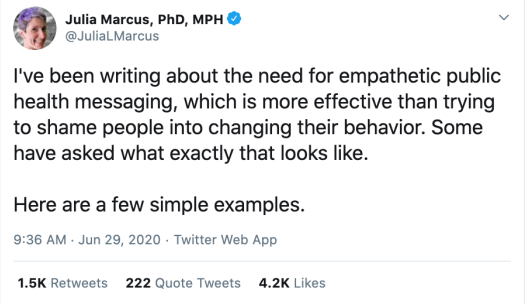


Fig. 35. Jeremy Howard tweets about the lack of variety in experts at meetings



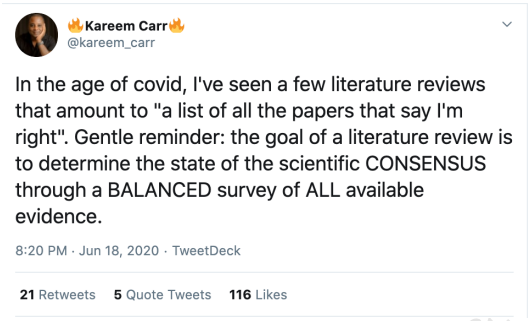
jimenez tweet.png

Fig. 36. Jose-Luis Jimenez tweets about errors in presentation slides talking about modes of transmission



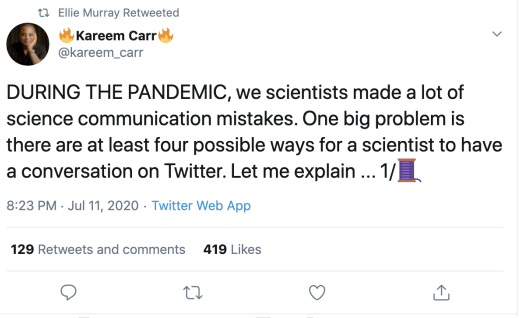
marcus tweet.png

Fig. 37. Julie Marcus tweets about how public messaging should be more empathetic



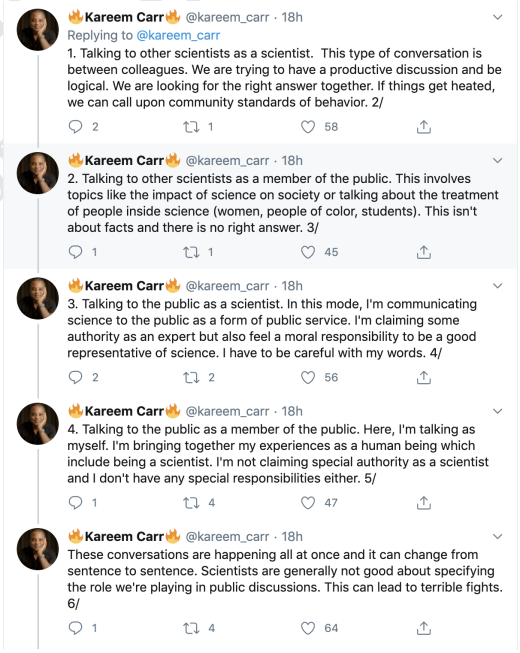
carr tweet.png

Fig. 38. Kareem Carr tweets about unbalanced literature reviews



carr tweet2.png

Fig. 39. Kareem Carr tweets about science communication mistakes made by scientists



carr tweet2 contd.png

Fig. 40. Kareem Carr tweets about science communication mistakes made by scientists

kuppalli tweet.png

Fig. 41. Krutika Kuppalli tweets about releasing data on media

tapia tweet.png

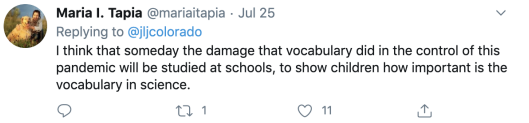


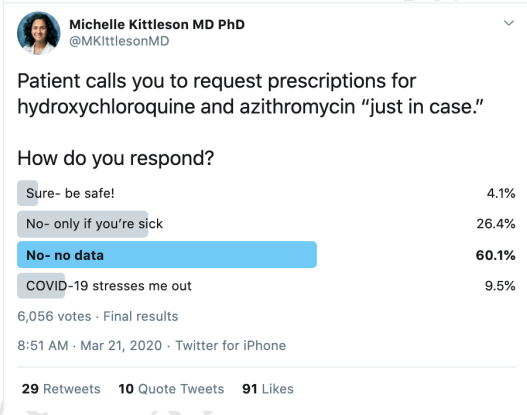
Fig. 42. Maria Tapia tweets about the importance of vocabulary in science

leefland tweet.png

Fig. 43. Mariska Leefland tweets about how many COVID papers are being published but they are making little policy difference

le tissier tweet.png

Fig. 44. Matt Le Tissier tweets about the contradictions in inscientist point of views on COVID-19



kittleson tweet.png

Fig. 45. Michelle Kittleson tweets a poll of how people respond to request for hydroxychloroquine

putman tweet.png

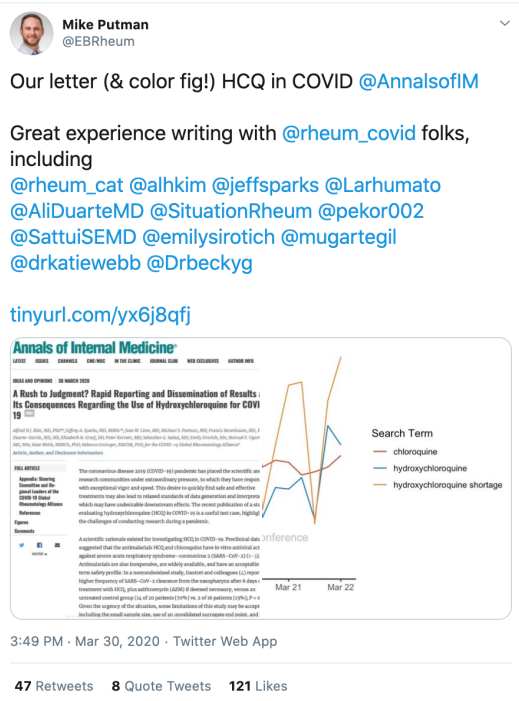


Fig. 46. Mike Putman tweets a letter they wrote on Hydroxychloroquine in COVID-19

dean tweet.png

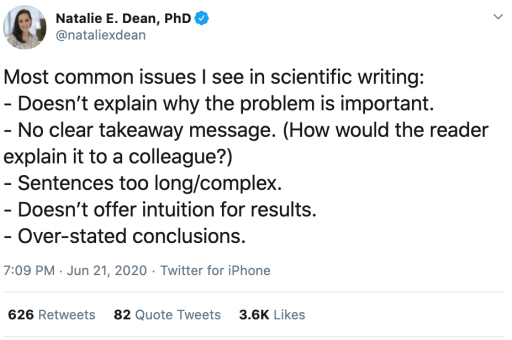
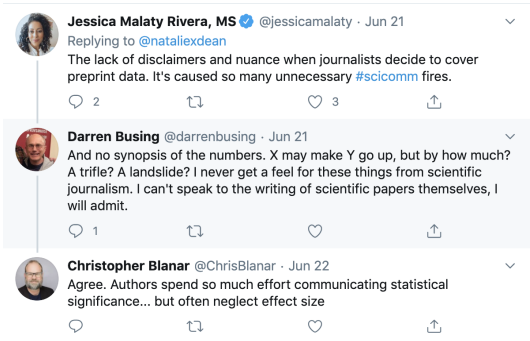


Fig. 47. Natalie Dean tweets about the issues seen in scientific writing

ferguson tweet.png

Fig. 48. Neil Ferguson tweets about focusing on uncertainties as scientists



tweet.png

Fig. 49. Jessica Malaty Rivera tweets about the coverage of preprints

rocha tweet.png

Fig. 50. Roberto Rocha has a twitter thread on the conversation between scientists on finding a consensus about COVID-19

popescu tweet.png

Fig. 51. Dr.Saskia Popescu tweets about evolving data

popescu tweet2.png

Fig. 52. Dr.Saskia Popescu tweets about how the term "airborne" is communicated by scientists

popescu tweet2 contd.png

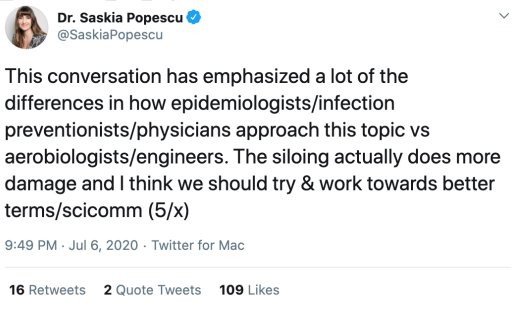


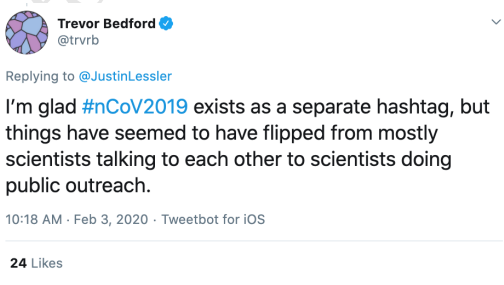
Fig. 53. Dr.Saskia Popescu tweets about how the term "airborne" is communicated by scientists

lab tweet.png

Fig. 54. Soragni:Lab tweets their thoughts on experts from unrelated fields with no relevant expertise contributing to COVID-19 research

nono tweet.png

Fig. 55. Steven Ñoño tweets about the lack of science communication among COVID-19 researchers



bedford tweet.png

Fig. 56. Trevor Bedford tweets about a Twitter space whose function is no longer about hosting conversations but is now about outreach