



## Retractions and post-retraction citations in the COVID-19 infodemic: is Academia spreading misinformation?

### *Retratações e citações pós-retratação na infodemia de COVID-19: a Academia está espalhando desinformação?*

Karen Santos-d'Amorim <sup>a</sup> 

Rinaldo Ribeiro de Melo <sup>a</sup> 

Raimundo Nonato Macedo dos Santos <sup>a,\*</sup> 

**ABSTRACT:** The speed in producing information and the rush to publish scientific articles on COVID-19 in several knowledge areas have resulted in what is known as an infodemic also in the scientific field, potentially producing inaccurate information and sources of misinformation at scholarly communication. This has led to some articles being retracted or withdrawn due to unintentional errors or deliberate misconduct, but they continue to be cited. This article (i) gives an overview of the COVID-19 retracted articles and preprints, and (ii) analyses a set of post-retraction citations in the context of the COVID-19 infodemic. We analyzed 56 retracted articles and preprints by using the list available in the section on “retracted coronavirus (COVID-19) papers” in the Retraction Watch (RW) webpage. We found that 64.3% of these retractions were articles published in journals, 33.9% were uploaded in preprints servers, and 1.8% conference papers. We also analyzed 162 eligible articles out of 612 records identified by using the Google Scholar search engine. This research found that an article from The Lancet continued to be cited even after being retracted. In this case, we identified 214 post-retraction citations, of which 38% were negative (n=81), 32% were neutral (n=69), and 30% were positive citations (n=64).

**Keywords:** Retracted articles; Retracted preprints; Post-retraction citations; Misinformation; Infodemic.

**RESUMO:** A velocidade da produção de informações e a pressa em publicar artigos científicos sobre a COVID-19 em diversas áreas do conhecimento resultaram no que se denomina infodemia também no campo científico, potencialmente produzindo informações imprecisas e fontes de desinformação na comunicação científica. Isso fez com que alguns artigos fossem retratados ou despublicados devido a erros não intencionais ou má conduta deliberada, mas eles continuam sendo citados. Este artigo (i) oferece uma visão geral dos artigos e *preprints* retratados no âmbito da pesquisa sobre COVID-19 e (ii) analisa um conjunto de citações pós-retratação no contexto da infodemia da COVID-19. Foram analisados 56 artigos e *preprints* retratados utilizando a lista disponível na seção “artigos sobre coronavírus retratados (COVID-19)” na página *Retraction Watch* (RW). Encontrou-se que 64,3% dessas retratações se referem a artigos publicados em periódicos, 33,9% foram carregadas em servidores de *preprints* e 1,8% artigos de conferência. Também foram analisados 162 artigos elegíveis dos 612 registros identificados usando o mecanismo de busca do *Google Scholar*. A pesquisa constatou que um artigo do *The Lancet* continuou a ser citado mesmo após a retratação. Nesse caso, foram identificadas 214 citações pós-retratação, sendo 38% negativas (n= 81), 32% neutras (n= 69) e 30% positivas (n= 64).

**Palavras-chave:** Artigos retratados; Preprints retratados; Citações pós-retratações; Desinformação; Infodemia.

<sup>a</sup> Programa de Pós-graduação em Ciência da Informação, Departamento de Ciência da Informação, Universidade Federal de Pernambuco, Recife, PE, Brasil.

\* Correspondência para/Correspondence to: Raimundo Nonato Macedo dos Santos. E-mail: raimundo.macedo@ufpe.br. Endereço: Programa de Pós-graduação em Ciência da Informação, Departamento de Ciência da Informação, Universidade Federal de Pernambuco, Av. da Arquitetura, S/N, Cidade Universitária, Recife, PE 50740-550, Brasil.

Recebido em/Received: 01/02/2021; Aprovado em/Approved: 21/05/2021.

Artigo publicado em acesso aberto sob licença [CC BY 4.0 Internacional](https://creativecommons.org/licenses/by/4.0/) 

## INTRODUCTORY REMARKS

In the ongoing pandemic of the coronavirus disease 2019 (COVID-19), the need for fast responses from the scientific community to the catastrophic effects of the virus led researchers and publishers to speed up scientific publications and peer-review processes (Else 2020; Palayew et al. 2020). As of this writing, the World Health Organization (WHO) COVID-19 database<sup>1</sup> had already screened 247,100 records, including scientific articles, preprints, and reports around the world from various knowledge areas. The speed of the information production on the pandemic has resulted in an informational epidemic, known as infodemic, potentially producing inaccurate information and sources of misinformation at scholarly communication (Zarocostas 2020).

Furthermore, this rush could lead to problems in public health (Steen 2011), as well as define and redefine some countries' research agenda and investment. In this context, some articles have already been corrected, retracted, or withdrawn. If compared to the number of publications, the number of retractions is not substantial; nonetheless, because many of the retractions are widely reported in the media, it raises questions of trust in science.

In this scenario, one of the concerns in *Academia* is that despite being retracted due to research misconduct, some of these articles continue to be cited and in some cases in a positive light (Bar-Ilan, Halevi 2017; Teixeira da Silva, Bornemann-Cimenti 2017; Teixeira da Silva, Dobránszki 2017; Schneider et al. 2020; Santos-d'Amorim et al. 2021, in press). Indeed, the effect of persistent misinformation in scholarly communication through post-retraction citations should be a matter of concern. Considering that scientific knowledge must be built on solid foundations and considering Bornmann and Daniel's (2008) statement that, “[s]cientists have complex citing motives which, depending on the intellectual and practical environment, are variously socially constructed”, it is important to understand the citing behavior in short-term post-retraction citations in the current context (Moed et al. 1985; Bornmann, Daniel 2008).

Thus, this study addresses (i) an overview of the retracted COVID-19 articles and preprints and (ii) the post-retraction citations in the context of the COVID-19 infodemic.

---

<sup>1</sup> Available at: <https://search.bvsalud.org/global-literature-on-novel-coronavirus-2019-ncov/> (accessed Apr 14, 2021).

## BACKGROUND

From January 2020 until the date of writing this article, the COVID-19 disease (SARS-CoV-2) has already decimated about 2,970,389 people across the globe<sup>2</sup>. In addition to the catastrophic effects caused by this disease, the world is also facing an infodemic<sup>3</sup>. According to the Oxford English Dictionary, infodemic is “a proliferation of diverse, often unsubstantiated information relating to a crisis, controversy, or event, which disseminates rapidly and uncontrollably through news, online, and social media, and is regarded as intensifying public speculation or anxiety” (Oxford... 2020). Still according to the Oxford English Dictionary, the first entry for the word infodemic dates from 2003 and refers to a publication in the *Washington Post* “by David J. Rothkopf to describe the explosion of information (and misinformation) associated with the SARS epidemic of 2003<sup>4</sup>”.

In this arena, researchers are committed to solving the most diverse problems caused by the new coronavirus. In addition to the research developed in the different sciences fields and the medical sphere, researchers from social, cognitive, computational, and information sciences have dedicated themselves to understanding informational problems associated with the social media dilemma and its potential of undermining the effort made by institutions that struggle to control the destructive effects of the virus on society (Marin 2020; Pennycook et al. 2020; Quinn, Fazel, Peters 2020; Stephens 2020).

On this matter, Bramstedt (2020, p. 804) states that “[r]esearch normally occurs at the speed of a marathon, but during a pandemic, the pace is more like a sprint”. An example of this is illustrated by the COVID-19 Primer database (<https://covid19primer.com/>). Through its daily mapping of journals indexed by PubMed and preprints uploaded at bioRxiv, medRxiv, and arXiv servers, it offers a comprehensive view on the COVID-19 research. Figure 1 shows the increase in publications in the period between Jan 21, 2020 and Jan 15, 2021.

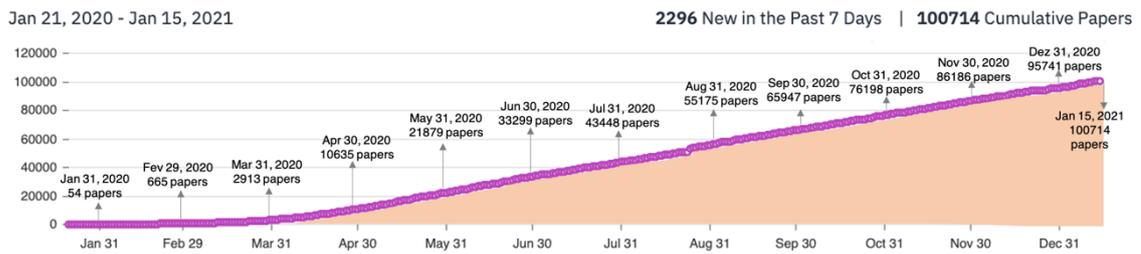
---

<sup>2</sup> According the COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University, available at <https://bit.ly/3aaJecn> (accessed Apr 14, 2021).

<sup>3</sup> As pointed out by Dr. Tedros Adhanom Ghebreyesus, World Health Organization Director-General, by saying that “we’re not just fighting an epidemic; we’re fighting an infodemic.”, at Munich Security Conference (Feb 15, 2020). Available at: <https://www.who.int/dg/speeches/detail/munich-security-conference> (accessed Jan 6, 2021).

<sup>4</sup> Available at: <https://www.oed.com/view/Entry/88407009> (accessed Jan 7, 2021).

**Figure 1.** Research publications on COVID-19.



Source: Adapted from COVID-19 Primer database (accessed Jan 15, 2021).

With this in mind, we can say that this “sprint” to publish research results has led to an infodemic, allowing for research with potentially flawed methodologies, misconduct, and misinformation in general, to go public.

With regard to misinformation<sup>5</sup>, the definitions in literature are countless (see, for example, Fox 1983; Floridi 2005; Karlova, Lee 2012; Fallis 2014). Swire-Thompson and Lazer (2020, p. 434) defined health misinformation as “information that is contrary to the epistemic consensus of the scientific community regarding a phenomenon”. In this research, we show that retracted articles and preprints are potential misinformation sources because they publish results based on flawed research. Therefore, in this context, is *Academia* contributing to generating infodemic and misinformation? In this article, the concept of misinformation (i.e., misleading, or inaccurate information) is used in its broader sense because the focus is not on ‘intent’, given that articles are retracted or withdrawn due to both intentional misconduct and unintentional errors<sup>6</sup>.

## DATA COLLECTION AND METHODS

For the first research objective, the data were recovered from the “retracted coronavirus (COVID-19) papers” section on Retraction Watch (RW) webpage (<https://retractionwatch.com/retracted-coronavirus-covid-19-papers>) accessed on 3<sup>rd</sup> January 2021. Retractions due to journal error, retracted and reinstated and

<sup>5</sup> For disinformation and malinformation disambiguations, see Fallis (2009, 2014), Karlova, Fisher (2013) and Wardle and Derakhshan (2018).

<sup>6</sup> Research misconduct occurs when there is “an intentional (or deliberate) deviation from accepted norms of scientific behavior” while “[d]eviations that are unintended (or accidental) are regarded as honest error, not misconduct” (Resnik, Stewart 2011 p. 2).

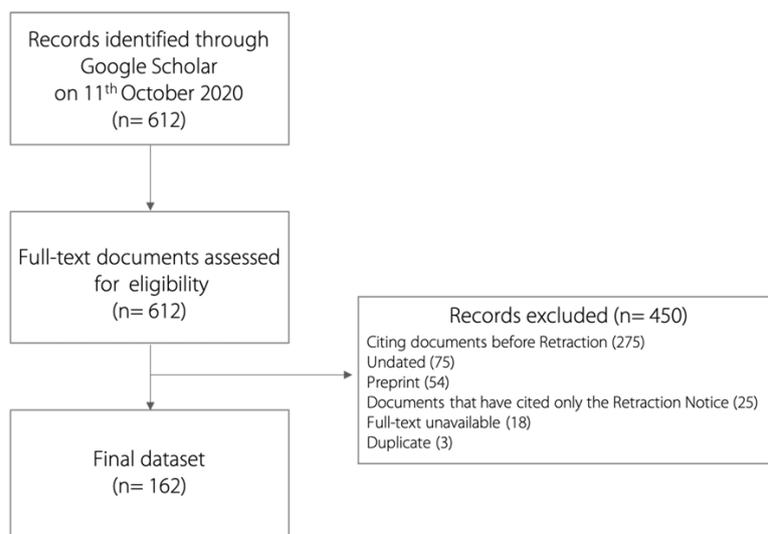
expressions of concern were not considered. Thus, 60 documents composed the corpus, and the following information was extracted and organized for analysis:

- 1) document typology;
- 2) reasons for retraction;
- 3) science field;
- 4) journal name and Journal Impact Factor (JIF), and
- 5) author's affiliation (Country).

Each document was manually checked and after this process, we detected that two articles included in the corpus were retracted due to editor error; and two articles have already been corrected and republished. Thus, we disregarded four articles at this stage. For this reason, 56 documents composed the final corpus investigated here.

For the second research objective, Figure 2 summarizes the data collection process. First, we retrieved all documents that cited Mehra, Desai, Ruschitzka, and Patel<sup>7</sup> post-retraction citations by using the Google Scholar search engine on October 11, 2020 and found 612 records.

**Figure 2.** Data collection process.



Source: research data (2020).

<sup>7</sup> Article published in *The Lancet* on May 22, 2020 and retracted on June 4, 2020 - “Hydroxychloroquine or chloroquine with or without a macrolide for treatment of COVID-19: a multinational registry analysis”.

Second, for the sake of accuracy, we manually checked the receiving date of each article to separate the ones that were submitted before and after the retraction notice. Citing articles before retraction ( $n= 275$ ), undated articles ( $n= 75$ ), preprints ( $n= 54$ ), documents that have cited only the Retraction Notice ( $n= 25$ ), full-text unavailable ( $n= 18$ ), and duplicate articles ( $n= 3$ ) were excluded from the final dataset. Finally, we analyzed the citation context in 162 articles.

We analyze the context of the citations in the text, considering that the “citation context is defined by the words which are located around a specific citation” (Bornmann, Haunschild, Hug 2018). The classification of each citation was conducted using the categorization proposed by Bar-Ilan and Halevi (2017), that classify post retraction citations as:

1. *Positive* A positive citation indicates that the retracted article was cited as legitimate prior work and its findings used to corroborate the author/s current study.
2. *Negative* A negative citations indicates that the authors mentioned the retracted article as such and its findings inappropriate.
3. *Neutral* A neutral citation indicates that the retracted article was mentioned as a publication that appears in the literature and does not include judgement on its validity”.

## RESULTS AND DISCUSSION

This section is presented into two parts. Part I gives an overview on the retracted COVID-19 articles and preprints, and Part II presents an analysis of the post-retraction citations in the context of the COVID-19 infodemic.

### Part I: an overview of the COVID-19 retracted articles and preprints

Within the scope of the COVID-19 research, Bramstedt (2020) found that 19 articles and 14 preprints were retracted or withdrawn up to July 2020. As of January 3, 2021, the Retraction Watch list had already traced 60 retracted documents. To offer an overview of the COVID-19 retracted articles and preprints, in addition to investigating the reasons for retraction, we also analyzed some characteristics of these documents, as detailed in Table 1.

**Table 1.** General characteristics of retractions.

Characteristic	Retractions (n)	Share (%)
<b>Document typology</b>		
Article	36	64.3
Conference Paper	1	1.8
Preprint	19	33.9
<b>Authorship</b>		
Single-authored	9	16.1
Co-authored	3	5.3
Multi-authored	42	75
N/A	2	3.6
<b>Journal Impact Factor <sup>a</sup></b>		
> 8	7	19.4
5-8	3	8.3
2-5	11	30.6
< 2	5	13.9
Without Journal Impact Factor (JIF)	10	27.8
<b>Preprint servers</b>		
medRxiv	9	47.4
SSRN: Social Science Resource Network	6	31.6
bioRxiv	4	21.0

N/A = not available

<sup>a</sup> Journal Impact Factor was manually checked in the 2019 Journal Citation Reports released by Clarivate Analytics in 2020 at the Journal Citation Reports (JCR).

According to the data collected, we observed that 64.3% of these retractions were from articles published in journals and 33.9% were uploaded in preprint servers. Of all the retracted articles, 30.6% were published in journals with Impact Factor (JIF) between 2 to 5. After checking each retraction notice, of the 32 retracted documents (articles and preprints), we were only able to identify 11 reasons for the retractions (Table 2); in 24 of them (17 articles, 6 preprints, and one conference paper) the reasons were not disclosed.

Of the 17 articles with undisclosed reasons, 13 are part of the Elsevier conglomerate, which is reason for concern considering that they do not provide clear information regarding the retraction guidelines as recommended by the Committee on Publication Ethics (COPE) (Wager 2009). Their standard statement on the 13 cases is “[t]his article has been withdrawn at the request of the authors and the editors. The Publisher apologizes for any inconvenience this may cause”.

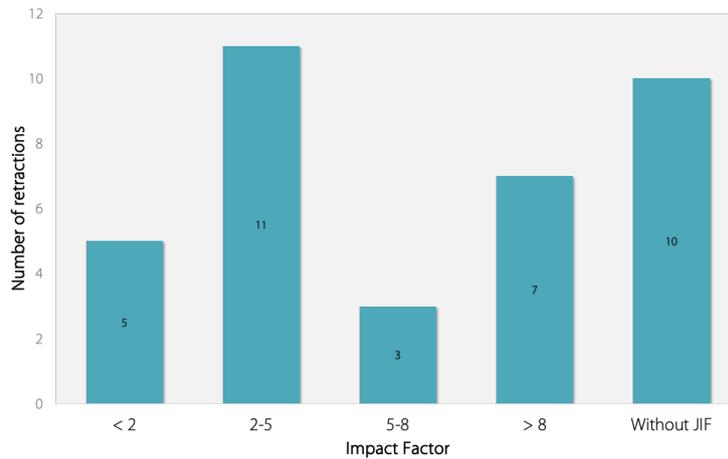
**Table 2.** General reasons for article retraction/withdrawn.

Reason for retraction/withdrawn	Number of articles (n= 36)
Reason not disclosed	17
Duplicate publication	4
Plagiarism	3
Raw data not made available to the authors/editor/referee/auditor	3
Unreliable results	2
Authors' desire to perform additional experiments	1
Authorship dispute: author does not satisfy the authorship criteria	1
Erroneous data	1
Flawed methodology	1
Lack of consent for an institutional review board (IRB)	1
“Serious scientific fraud”	1
“The wrong paper has been published due to some technical glitch. The information pertaining in this paper is misleading the readers and creating massive conflicts amid the scientific community”	1

Besides the 17 incidences where no reason for the retraction was given, we found recurrent cases of duplicate publication – papers that have already been published in part or in full elsewhere – ( $n= 4$ ), plagiarism ( $n= 3$ ), raw data not made available to the authors/editor/referee/auditor ( $n= 3$ ), and unreliable results ( $n= 2$ ). The generic terms used in the retraction notes, such as, “serious scientific fraud” and “technical glitch”, did not allow us to conduct a proper analysis.

Regarding the Journal Impact Factor (JIF), Figure 3 demonstrates that out of the 36 retractions, 14% ( $n= 5$ ) of the articles had been published with a JIF under two, 31% ( $n= 11$ ) between 2-5, 8% ( $n= 3$ ) between 5-8, and 19% ( $n= 7$ ) with a JIF higher than 8. We also note that 28 % ( $n= 10$ ) of the retractions were from journals that were not indexed in the 2019 Journal Citation Report (JCR). These results differ from those found by Fang and Casadevall (2011) when investigating the correlation between Journal Impact Factor and retractions – what they called “retraction index” – where they suggest that “the probability that an article published in a higher-impact journal will be retracted is higher than that for an article published in a lower-impact journal” (p. 3856). However, the limited analysis corpus precludes drawing more precise conclusions.

**Figure 3.** Journal Impact Factor distribution among the retracted articles.

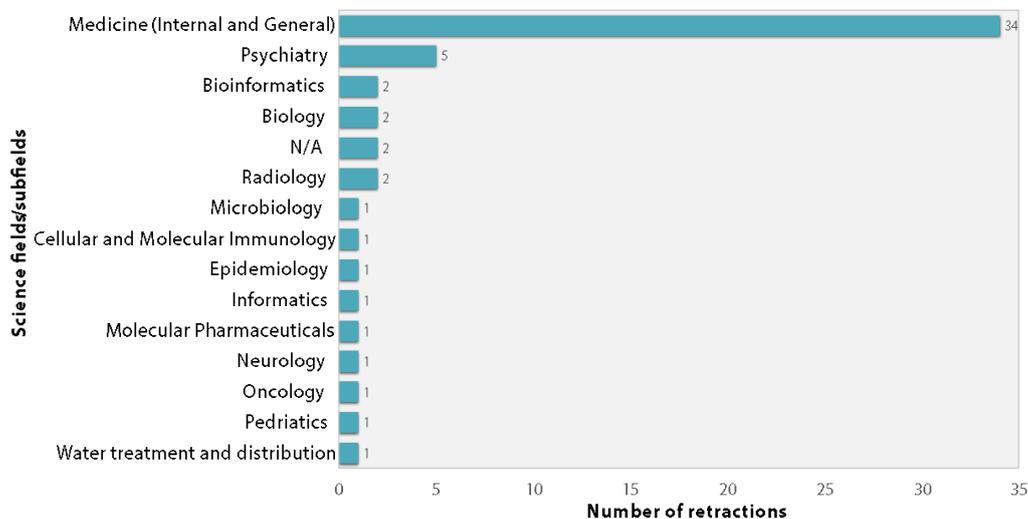


Source: research data (2021).

By using the 2019 Journal Impact Factor (JIF) released by the 2020 JCR, the highest JIF found were from the New England Journal of Medicine (JIF = 74.699 /  $n= 1$ ), The Lancet (JIF= 60.390 /  $n= 2$ ), the Annals of Internal Medicine (JIF= 21.317 /  $n= 1$ ), and the American Journal of Respiratory and Critical Care Medicine (JIF= 17.452 /  $n= 1$ ).

Figure 4 shows the stratification by science field. Considering that COVID-19 is an infectious disease and that the scientific production on this matter is mostly concentrated around Medicine, our findings were not surprising: Medicine with the most retractions ( $n= 34$ ), followed by Psychiatry ( $n= 5$ ), and other specialties (subfields) of Medicine, such as Radiology ( $n= 1$ ) Neurology ( $n= 1$ ), Oncology ( $n= 1$ ), and Pediatrics ( $n= 1$ ).

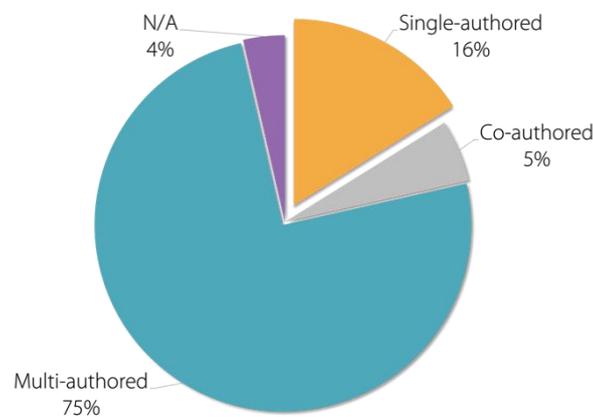
**Figure 4.** Retractions (articles and preprints) by science fields.



Source: research data (2021).

Figure 5 shows that articles written by three or more authors were responsible for 75% (n= 42) of the retractions, followed by the ones with one author, 16% (n= 9), and two authors 5% (n=3). Considering that 86% of these retractions were in the medical literature, including specialties, we can point out that even in normal circumstances (i.e., not in a pandemic situation), a growth in the number of authors per article can be observed in this field. This is explained by the complexity of multi-disciplinary and multi-institutional collaboration (An et al. 2020). In 4% (n= 2) of the cases, it was not possible to have access to the documents.

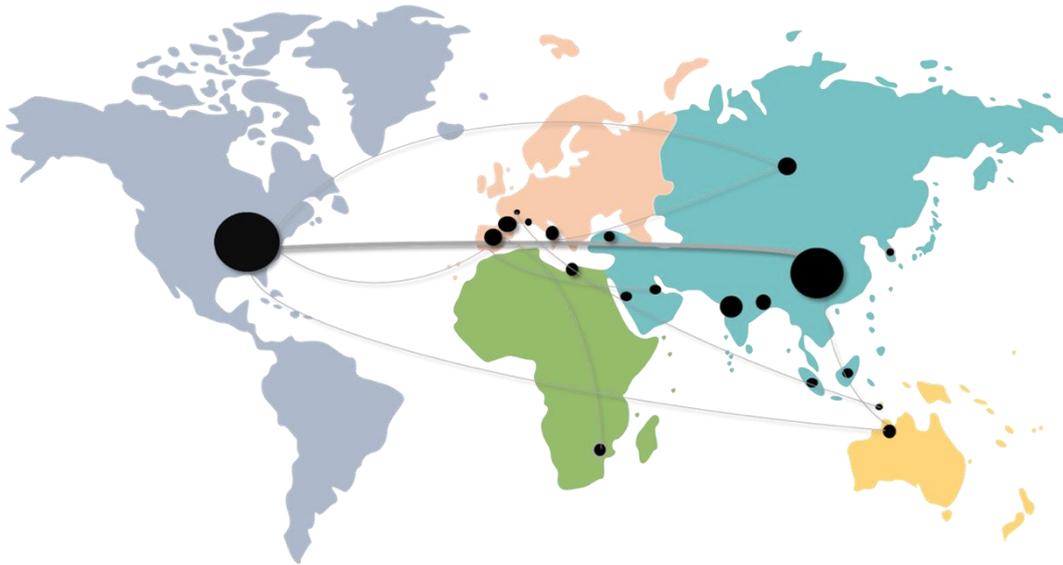
**Figure 5.** Percentage of retractions by the size of authorship team.



Source: research data (2021).

In addition to the identification of the percentage of retractions by the number of authors in each article, the author's affiliation was also considered in order to identify institutional collaboration between countries around the globe. We identified that, among the co-authored and multi-authored articles, 11% (n= 5) were from authors of the same institution and 81% (n= 38) from different institutions; 8% (n= 4) of the articles were not available for consultation. The highest incidence of interactions among countries was between China-USA (n= 4). Other interactions were between Italy-USA-Russia; Italy-USA; Australia-USA; Spain-Mozambique; Belgium-France-Switzerland-East Timor; and USA-Switzerland, among others, as shown in Figure 6.

**Figure 6.** Collaborative countries on COVID-19 retracted articles.



Source: research data (2021).

Regarding preprints, because they offer some advantages, such as accessibility, faster dissemination, easy feedback (Vlasschaert, Topf, Hiremath 2020), or credit for priority in science discovery (Merton 1957), at first glance, some researchers have been opting to upload their manuscripts on preprints servers. Although they are documents that have not been subject to peer scrutiny, preprints have been an important subject of discussion, the reason why we included them in this study. Table 3 summarizes the reasons for retraction/withdrawal of the preprints analyzed. We found eight different reasons.

**Table 3.** General reasons for preprints retraction/withdrawn.

Reason for retraction/withdrawn	Number of preprints (n= 19)
Reason not disclosed	6
Authors' desire to update their dataset	3
Lack of full consent for publication by all authors	2
Lack of consent for an institutional review board (IRB)	2
Controversies about hydroxychloroquine	2
Concerns about research findings	2
Authors' desire to perform additional experiments	1
Privacy issues	1

By examining these documents, we found that, as in the case of the research articles, the preprints with non-disclosed reasons (n= 6) are the most frequent, followed by

“the authors' wish to update their database” reason ( $n= 3$ ). Unlike the research articles, we found no retractions due to plagiarism or duplicate publication in preprints.

In a comment for *Retraction Watch*, Richard Sever, co-founder of bioRxiv and medRxiv preprint servers, clarifies that unless a preprint contains fraud or some ethical or legal violation, “[a] preprint is usually withdrawn only at an authors’ request. A preprint server will not typically withdraw a flawed preprint against the wishes of the author since the server made no claim to have peer-reviewed and certified the scientific content in the first place” (Retraction... 2019 n.p.).

However, to curb the advancement of publication of potentially bad research, preprint servers have been taking measures such as in the case of bioRxiv. As reported by Kwon (2020), “[t]he biomedical repository would no longer accept manuscripts making predictions about treatments for COVID-19 solely on the basis of computational work”, since, according to Richard Sever, co-founder of bioRxiv and medRxiv, numerous speculative papers have been published using computational models.

## **Part II: post-retraction citations in the COVID-19 infodemic context**

The example of post-retraction citations we chose to analyze is from the article “Hydroxychloroquine or chloroquine with or without a macrolide for treatment of COVID-19: a multinational registry analysis”, published in the journal *The Lancet* on May 22, 2020. The results disclosed in this paper led the WHO to temporarily suspend the tests with these drugs (Offord 2020). It is very important to understand what happens when retracted articles are cited and whether the short-term post-retraction citations in research publications on the COVID-19 pandemic suffer the effect of misinformation.

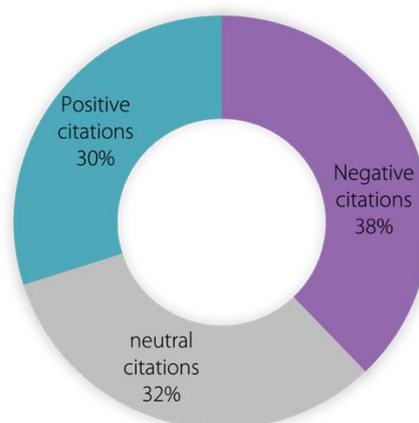
Known as “The Surgisphere Scandal” (Offord 2020; Piller 2020), at the request of three of the four authors, this article was retracted on June 4, 2020, ([http://doi.org/10.1016/S0140-6736\(20\)31180-6](http://doi.org/10.1016/S0140-6736(20)31180-6)) because they could no longer guarantee the veracity of the primary sources of the data used to support the research. According to the retraction notice, the authors justified their action by stating that one of the co-authors, who was also the founder of the company (Surgisphere Corporation, Chicago, IL), claimed not to be able to transfer the complete database, which would allow an independent review to ensure the integrity of the primary sources and the replicability of the analysis (Mehra, Ruschitzka, Patel 2020).

Chloroquine and Hydroxychloroquine were persistently used as a political device, being promoted as “miraculous” drugs. Both have long been used in the treatment for malaria and autoimmune diseases and presented *in vitro* potential effects in the treatment and prevention of infections caused by the new coronavirus. Thus, they were seen as promising drugs and as the fastest route in the treatment of COVID-19 infection (Lei et al. 2020; Mohamed, Rezaei 2020). However, their inefficacy for the treatment of COVID-19 has been confirmed (Cavalcanti et al. 2020).

In summary, the Mehra, Desai, Ruschitzka, and Patel retracted article, highlighted the non-evidence of Chloroquine and Hydroxychloroquine effectiveness, as well as associated their use with the risk of ventricular arrhythmias and the greater risk of in-hospital death for patients with COVID-19.

Nonetheless, our investigation found that this article continued to be cited even after the retraction. We found 214 post-retraction citations in 162 articles, including multiple citations in some of them. By analyzing the citation using the Bar-Ilan and Halevi (2017) categorization, we found 38% negative citations ( $n= 81$ ), 32% neutral citations ( $n= 69$ ), and 30% positive citations ( $n= 64$ ), as shown in Figure 7.

**Figure 7.** Categorization of the post-retraction citations in the Mehra, Desai, Ruschitzka, and Patel retracted article.



Source: research data (2020).

In its concept, neutral citations do not include judgment of value. They act as information devices that indicate the existence of that article in literature, not contributing effectively to the misinformation effect. They are probably cited for social reasons, such as for having been published in a prestigious journal (see Shadish et al.

1995; Bornmann, Daniel 2008). The positive citations are still a matter of concern since they consider the work to be legitimate and are used to corroborate the cited study.

## CONCLUDING REMARKS

For the sake of investigating the main characteristics of the retracted COVID-19 articles and preprints and examine the nature of a set of post-retraction citations in the context of the COVID-19 infodemic, by using a paper as a case study, we set out a methodological strategy based on information metrics studies and the in-text citation analysis technique. To enhance the understanding of the object, we analyze the reasons for retraction and the following characteristics present in the research corpus: document typology, science field, journal name, and Journal Impact Factor (JIF), author's affiliation (Country), and their respective collaborations through authorship analysis.

In summary, we analyzed 56 retracted articles and preprints under the COVID-19 research and found that 64.3% (n= 36) of these retractions were articles published in journals, 33.9% (n= 19) were uploaded in preprints servers, and 1.8% (n= 1) conference paper. Our analysis identified the prevalence of retraction in multi-authored articles characterized by the complexity of multidisciplinary and multi-institutional collaborations mainly in the field of medicine and its specialties. Based on the verifications presented here, we can conclude that this informational epidemic, known as infodemic, is also present in the scientific sphere.

Concerning about the reasons for retraction, the lack of transparency in the retraction notices is still a matter of concern. In 17 of the 36 analyzed articles, the reasons that led to the retraction are unknown, characterized by the text "Reason not disclosed". The same occurs for the preprints, with the recurrence of the term "reason not disclosed" in 6 of the 19 preprints retracted from the analyzed corpus.

We also analyzed 214 post-retraction citations in 162 articles. Of these, we found 81 negative citations (38%), 69 neutral citations (32%), and 64 positive citations (30%). The prevalence of the negative and neutral post-retraction citations can indicate that *Academia* is aware of the dynamics of scientific publications, they are not spreading misinformation, and in a sense, they are shielded from political interferences.

In conclusion, we emphasize that retraction is an efficient mechanism of the scientific record correction that flags flaws in the published research - due to unintentional errors or deliberate misconduct. It is also noteworthy that the rate of articles retracted can also serve as an indicator for measuring the misinformation phenomenon in science, enabling one to assess trends and mitigate current problems and at future scenarios. It is understood that given the need for fast responses from the scientific community to the catastrophic effects caused by the coronavirus disease 2019 (COVID-19), more efficient and effective peer review processes can prevent faulty articles from being published and avoiding, *ab origine*, the spread of misinformation.

## FUNDING

This work was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) - Finance Code 001, and also by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

## DATA AVAILABILITY

Our full dataset is available on Zenodo repository and can be downloaded on demand at <https://doi.org/10.5281/zenodo.4695045>

## REFERENCES

AN, Julie Y., et al., 2020. Authorship growth in contemporary medical literature. *Sage open medicine* [online]. 2020. Vol. 8, p. 1-10. [Accessed 17 January 2021]. DOI 10.1177/2050312120915399. Available from: <https://journals.sagepub.com/doi/10.1177/2050312120915399>

BAR-ILAN, Judit and HALEVI, Gali, 2017. Post retraction citations in context: a case study. *Scientometrics* [online]. 2017. Vol. 113, p. 547-565. [Accessed 3 September 2020]. DOI 10.1007/s11192-017-2242-0. Available from: <https://link.springer.com/article/10.1007/s11192-017-2242-0>

BORNMANN, Lutz and DANIEL, Hans-Dieter, 2008. What do citation counts measure? A review of studies on citing behavior. *Journal of Documentation* [online]. 2008. Vol. 64, no. 1, p. 45-80. [Accessed 15 September 2020]. DOI 10.1108/00220410810844150. Available from: <https://doi.org/10.1108/00220410810844150>

BORNMANN, Lutz, HAUNSCHILD, Robin and HUG, Sven E., 2018. Visualizing the context of citations referencing papers published by Eugene Garfield: a new type of keyword co-occurrence analysis. *Scientometrics* [online]. 2018. Vol. 144, p. 427-437. [Accessed 15 September 2020]. DOI 10.1007/s11192-017-2591-8. Available from: <https://doi.org/10.1007/s11192-017-2591-8>

BRAMSTEDT, Katarina, 2020. The carnage of substandard research during the COVID-19 pandemic: a call for quality. *Journal of Medical Ethics* [online]. 2020. Vol. 46, p. 803–807. [Accessed 3 January 2021]. DOI 10.1136/medethics-2020-106494. Available from: <https://jme.bmj.com/content/46/12/803>

CAVALCANTI, Alexandre B., et al., 2020. Hydroxychloroquine with or without Azithromycin in Mild-to-Moderate Covid-19. *The New England Journal of Medicine* [online]. 2020. Vol. 383, p. 2041-2052. [Accessed 31 January 2021]. DOI 10.1056/NEJMoa2019014. Available from: <https://www.nejm.org/doi/full/10.1056/NEJMoa2019014>

ELSE, Holly, 2020. How a torrent of COVID science changed research publishing - in seven charts. *Nature* [online]. 2020. Vol. 588, p. 553. [Accessed 3 January 2021]. DOI 10.1038/d41586-020-03564-y. Available from: <https://www.nature.com/articles/d41586-020-03564-y>

FALLIS, Don, 2009. A conceptual analysis of disinformation. In: *iCONFERENCE* [online]. Chapel Hill: University of North Carolina. 2009. [Accessed 17 January 2021]. Available from: <http://hdl.handle.net/2142/15205>

FALLIS, Don, 2014. A Functional Analysis of Disinformation. In: *iCONFERENCE* [online]. Berlin: University Humboldt. 2014. [Accessed 17 January 2021]. Available from: <http://hdl.handle.net/2142/47258>

FANG, Ferric and CASADEVALL, Arturo, 2011. Retracted Science and the Retraction Index. *Infection and Immunity* [online]. 2011. Vol. 79, no. 10, p. 3855-3859. [Accessed 15 December 2020]. DOI 10.1128/IAI.05661-11. Available from: <https://iai.asm.org/content/79/10/3855>

FLORIDI, Luciano, 2005. Is Semantic Information Meaningful Data? *Philosophy and Phenomenological Research* [online]. 2005. Vol. 70, no. 2, p. 351-370. [Accessed 17 January 2021]. DOI 10.1111/j.1933-1592.2005.tb00531.x. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1933-1592.2005.tb00531.x>

FOX, Christopher John, 1983. *Information and misinformation: An investigation of the notions of information, misinformation, informing, and misinforming*. Westport, CT: Greenwood Press, 1983.

KARLOVA, Natascha and FISHER, Karen, 2013. A social diffusion model of misinformation and disinformation for understanding human information behavior. *Information Research* [online]. 2013. Vol. 18, no. 1. [Accessed 17 January 2021]. Available from: <http://www.informationr.net/ir/18-1/paper573.html#.YASiauhKJIU>

KARLOVA, Natascha and LEE, Jin Ha, 2012. Notes from the underground city of disinformation: A conceptual investigation. *Proceedings of the American Society for Information Science and Technology* [online]. Vol. 48, no. 1, p. 01-09, 2012. [Accessed 17 January 2021]. DOI 10.1002/meet.2011.14504801133. Available from: <https://asistdl.onlinelibrary.wiley.com/doi/full/10.1002/meet.2011.14504801133>

KWON, Diana, 2020. How swamped preprint servers are blocking bad coronavirus research. *Nature* [online]. 2020. Vol. 581, p. 130-131. [Accessed 3 January 2021]. DOI 10.1038/d41586-020-01394-6. Available from: <https://www.nature.com/articles/d41586-020-01394-6>

LEI, Zi-Ning, et al., 2020. Chloroquine and hydroxychloroquine in the treatment of malaria and repurposing in treating COVID-19. *Pharmacology & Therapeutics* [online]. 2020. Vol. 2016, p. 1-13. [Accessed 25 October 2020]. DOI 10.1016/j.pharmthera.2020.107672. Available from: <https://doi.org/10.1016/j.pharmthera.2020.107672>

MARIN, Lavinia, 2020. Three contextual dimensions of information on social media: lessons learned from the COVID-19 infodemic. *Ethics and Information Technology* [online]. 2020. [Accessed 5 January 2020]. DOI 10.1007/s10676-020-09550-2. Available from: <https://link.springer.com/article/10.1007/s10676-020-09550-2#citeas>

MEHRA, Mandeep R., RUSCHITZKA, Frank and PATEL, Amit N., 2020. Retraction—Hydroxychloroquine or chloroquine with or without a macrolide for treatment of COVID-19: a multinational registry analysis. *The Lancet* [online]. 2020. Vol. 395, no. 10240, p. 1820. [Accessed 5 September 2020]. DOI 10.1016/S0140-6736(20)31324-6. Available from: [https://doi.org/10.1016/S0140-6736\(20\)31324-6](https://doi.org/10.1016/S0140-6736(20)31324-6)

MERTON, Robert King, 1957. Priorities in Scientific Discovery: A Chapter in the Sociology of Science. *American Sociological Review* [online]. 1957. Vol. 22, no. 6, p. 635-659. [Accessed 10 January 2021]. DOI 10.2307/2089193. Available from: <https://www.jstor.org/stable/2089193?seq=1>

MOED, Henk F., et al., 1985. The use of bibliometric data for the measurement of university research performance, *Research Policy* [online]. 1985. Vol. 14, no. 3, p. 131-149. [Accessed 28 September 2020]. DOI 10.1016/0048-7333(85)90012-5. Available from: [https://doi.org/10.1016/0048-7333\(85\)90012-5](https://doi.org/10.1016/0048-7333(85)90012-5)

MOHAMED, Kwathar and REZAEI, Nima, 2020. COVID-19 pandemic is not the time of trial and error. *The American Journal of Emergency Medicine* [online]. 2020. [Accessed 20 November 2020]. DOI 10.1016/j.ajem.2020.09.020. Available from: <https://doi.org/10.1016/j.ajem.2020.09.020>

OFFORD, Catherine, 2020. The Surgisphere Scandal: What Went Wrong? *The Scientist* [online]. 1 October 2020. [Accessed 15 October 2020]. Available from: <https://www.the-scientist.com/features/the-surgisphere-scandal-what-went-wrong--67955>

OXFORD English Dictionary, 2020. Infodemic. *OED Third Edition* [online]. April 2020. [Accessed 15 December 2020]. Available from: <https://www.oed.com/view/Entry/88407009>

PALAYEW, Adam, et al., 2020. Pandemic publishing poses a new COVID-19 challenge. *Nature Human Behaviour* [online]. 2020. Vol. 4, p. 666-669. [Accessed 28 September 2020]. DOI 10.1038/s41562-020-0911-0. Available from: <https://doi.org/10.1038/s41562-020-0911-0>

PENNYCOOK, Gordon, et al., 2020. Fighting COVID-19 Misinformation on Social Media: Experimental Evidence for a Scalable Accuracy-Nudge Intervention. *Psychological Science* [online]. Vol. 31, no. 7, p. 770-780. [Accessed 11 November 2020]. Available from: <https://doi.org/10.1177/0956797620939054>

PILLER, Charles, 2020. Who's to blame? These three scientists are at the heart of the Surgisphere COVID-19 scandal. *Science* [online]. 8 June 2020. [Accessed 6 October

2020]. Available from: <https://www.sciencemag.org/news/2020/06/whos-blame-these-three-scientists-are-heart-surgisphere-covid-19-scandal>

SANTOS-D'AMORIM, Karen, et al., 2021. Retratos, despublicados e ainda citados: análise das citações recebidas pós-retratação em artigos de pesquisadores brasileiros. *Informação & Sociedade*, Vol. 31, no. 1, 2021. In press.

QUINN, Emma K., FAZEL, Sajjad S. and PETERS, Cheryl E., 2020. The Instagram Infodemic: Cobranding of Conspiracy Theories, Coronavirus Disease 2019 and Authority-Questioning Beliefs. *Cyberpsychology, Behavior, and Social Networking* [online]. 18 December 2020. [Accessed 28 September 2020]. DOI 10.1089/cyber.2020.0663. Available from: <http://doi.org/10.1089/cyber.2020.0663>

RESNIK, David B. and STEWART, Neal, 2012. Misconduct versus Honest Error and Scientific Disagreement. *Accountability in Research* [online]. 2012. Vol. 19, no. 1, p. 56–63. [Accessed 17 January 2021]. DOI 10.1080/08989621.2012.650948. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3443861/>

RETRACTION Watch database, 2019. Retracted coronavirus (COVID-19) papers. *Retraction Watch* [online]. 2019. [Accessed 15 December 2020]. Available from: <https://retractionwatch.com/retracted-coronavirus-covid-19-papers/>

SCHNEIDER, Jodi, et al., 2020. Continued post-retraction citation of a fraudulent clinical trial report, 11 years after it was retracted for falsifying data. *Scientometrics* [online]. 2020. Vol. 125, p. 2877-2913. [Accessed 10 November 2020]. DOI 10.1007/s11192-020-03631-1. Available from: <https://doi.org/10.1007/s11192-020-03631-1>

SHADISH, William R., et al., 1995. Author judgments about works they cite – three studies from psychology journals. *Social Studies of Science* [online]. 1995. Vol. 25, no. 3, p. 477-498. [Accessed 10 November 2020]. DOI 10.1177/030631295025003003. Available from: <https://doi.org/10.1177/030631295025003003>

STEEN, R. Grant, 2011. Retractions in the medical literature: how many patients are put at risk by flawed research? *Journal of Medical Ethics* [online]. 2020. Vol. 37, p. 688-692. [Accessed 10 November 2020]. DOI 10.1136/jme.2011.043133. Available from: <https://jme.bmj.com/content/37/11/688>

STEPHENS, Monica, 2020. A geospatial infodemic: Mapping Twitter conspiracy theories of COVID 19. *Dialogues in Human Geography* [online]. 2020. Vol. 10, no. 2, p. 276-281. [Accessed 22 December 2020]. DOI 10.1177/2043820620935683. Available from: <https://journals.sagepub.com/doi/full/10.1177/2043820620935683>

SWIRE-THOMPSON, Briony and LAZER, David, 2020. Public health and online misinformation: Challenges and recommendations. *Annual Review of Public Health* [online]. 2020. Vol. 41, no. 1, p. 433–451. [Accessed 17 January 2021]. DOI 10.1146/annurev-publhealth-040119-094127. Available from: <https://www.annualreviews.org/doi/10.1146/annurev-publhealth-040119-094127#>

TEIXEIRA DA SILVA, Jaime A. and BORNEMANN-CIMENTI, Helmar, 2017. Why do some retracted papers continue to be cited? *Scientometrics* [online]. 2017. Vol. 110, p. 365-370. [Accessed 10 November 2020]. DOI 10.1007/s11192-016-2178-9. Available from: <https://link.springer.com/article/10.1007%2Fs11192-016-2178-9>

TEIXEIRA DA SILVA, Jaime A. and DOBRÁNSZKI, Judit, 2017. Highly cited retracted papers. *Scientometrics* [online]. 2017. Vol. 110, p. 1653-1661. [Accessed 10 November 2020]. DOI 10.1007/s11192-016-2227-4. Available from: <https://link.springer.com/article/10.1007%2Fs11192-016-2227-4>

VLASSCHAERT, Caitlyn, TOPF, Joel M. and HIREMATH, Swapnil, 2020. Proliferation of Papers and Preprints During the Coronavirus Disease 2019 Pandemic: Progress or Problems With Peer Review? *Advances in chronic kidney disease* [online]. 2020. Vol. 27, no. 5, p. 418–426. [Accessed 10 November 2020]. DOI 10.1053/j.ackd.2020.08.003. Available from: [https://www.ackdjournal.org/article/S1548-5595\(20\)30119-1/fulltext#%20](https://www.ackdjournal.org/article/S1548-5595(20)30119-1/fulltext#%20)

WAGER, Elizabeth, 2009. Retractions: Guidance from the Committee on Publication Ethics (COPE). *Croatian Medical Journal* [online]. 2009. Vol. 50, no. 6, p. 532–535. [Accessed 5 January 2020]. DOI 10.3325/cmj.2009.50.532. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2802086/>

WARDLE, Claire and DERAKHSHAN, Hossein, 2018. Thinking about ‘information disorder’: formats of misinformation, disinformation, and mal-information. In: IRETON, Cherilyn and POSETTI, Julie. *Journalism, ‘fake news’ & disinformation*. Paris: UNESCO; 2018. p. 43-54. [Accessed 17 January 2021]. Available from: <https://bit.ly/2FW3Esb>

ZAROCOSTAS, John, 2020. How to fight an infodemic. *The Lancet* [online]. 2020. Vol. 395, no. 10225, p. 676. [Accessed 26 November 2020]. DOI 10.1016/S0140-6736(20)30461-X. Available from: [https://doi.org/10.1016/S0140-6736\(20\)30461-X](https://doi.org/10.1016/S0140-6736(20)30461-X)