Unidentified retracted articles on publishers' websites, bibliographic databases, academic social networks, and the Sci-Hub black open access website: a problem that should no longer be ignored

Journal:	*BMJ*
Manuscript ID	BMJ-2022-071385
Article Type:	Analysis
Date Submitted by the Author:	12-Apr-2022
Complete List of Authors:	Boudry, Christophe; UNICAEN; PSL Research University Howard, Katherine; Deakin University Mouriaux, Frederic; CHU Rennes Department of Ophthalmology; Université de Rennes 1
Keywords:	Information management, Ethics, medical, Guideline adherence, Information technology

Unidentified retracted articles on publishers' websites, bibliographic databases, academic social networks, and the Sci-Hub black open access website: a problem that should no longer be ignored

Christophe Boudry (0000-0002-8730-8731) [1,2]
Katherine Howard (0000-0002-9327-8989)[3]
Frederic Mouriaux (0000-0001-9183-4235) [4]

[1] Normandie Univ, UNICAEN, Média Normandie. 14032 Caen cedex, France
[2] URFIST, Ecole Nationale des Chartes, PSL Research University. 75005 Paris, France
[3] Deakin University, Melbourne, Australia
[4] INSERM, Univ Rennes, CHU Rennes, Department of Ophthalmology, CLCC Eugène Marquis, COSS [(Chemistry Oncogenesis Stress Signaling)] – UMR_S 1242, 35000 Rennes, France

Correspondence to:
Christophe Boudry
Email: christophe.boudry@chartes.psl.eu
Phone: (+33) 1.43.26.85.22

KEY MESSAGES

- **The identification of retracted articles is essential to avoid the propagation of scientific errors caused by citing retracted articles**
- **Publishers' websites and sites hosting references and/or full-text articles insufficiently identify retracted articles, most notably Google Scholar and Sci-Hub**
- **Solutions exist to overcome this worrying problem. It is a collective responsibility that requires a real commitment on the part of all actors involved**

Contributors and sources

CB and FM have published articles related to bibliometrics in the past several years, especially in ophthalmology. They have also studied the difficulty involved in accessing scientific literature around the world with KH, a specialist in information science. The idea for this article came from FM who had used a retracted article in his bibliography. Zotero had identified this retracted article and FM was able to identify it. However, after discussion with colleagues, no one was careful not to cite retracted articles. So CB and FM, in collaboration with KH, wanted to know the extent to which retraction information is made available across article available on publishers' websites and sites hosting references and/or full-text articles. CB and FM drafted the manuscript; CB and FM extracted data. CB analyzed data; and CB, KH and FM revised the manuscript critically for important intellectual content. All authors approved the final version of the manuscript. CB is the guarantor of the article.

Acknowledgements

The authors thank Retraction Watch for their assistance in providing the authors with free access to the Retraction Watch Database. Thanks to Sebastien Kerever for constructive discussions related to this article.

Patient involvement
No patients were involved in this work.

Conflicts of Interest
We have read and understood BMJ policy on declaration of interests and we have no conflicts of interest to declare.

Christophe Boudry and colleagues call for better identification of retracted articles on publishers' websites and sites hosting references and/or full-text articles.

Article retraction

Article retraction, which mainly results from research misconduct or errors [1], is described by the Committee on Publication Ethics (COPE) as 'a mechanism for correcting the literature and alerting readers to articles that contain such seriously flawed or erroneous content or data that their findings and conclusions cannot be relied upon' [2]. The main reasons for retraction are inadvertent errors or mistakes, non-replicable findings, research misconduct, and redundant or duplicate publication [3]. Although retractions are still rare, around 0.5 retractions per 100,000 articles published [4–6]), numerous papers have highlighted an increasing rate of retractions over time (it has doubled in the last ten years) [4,7–9], and the field of medicine is no exception [5] (see Box 1). It is very difficult to know if this increase is a result of a higher number of suspicious papers due to a rise in fraud and error [10] or if the scientific community is improving its ability to detect and report them, reflecting a more self-monitoring community [4,7,8,11,12].

Box 1: Characterisation of retractions in biological and medical sciences

The retraction rate per 10,000 publications was 0.38 in 1985 and it rose to 2.03 in 2000 with a peak of 5.95 in 2014. In recent years the rate has decreased because there will likely be more retractions in the future among papers published in recent years [5]. The maximum number of retractions occured within 1 year of publication, then the number decreased as the years went by. It took 3.8 years (on average) for a publication to be retracted [5].

In 2020, the most common reason for retraction was scientific misconduct (62.3%), followed by error(s) in the manuscript (37.4%) and issues with the journal or publisher (19.4). Nearly 70% of retracted articles had more than 1 reason for retraction [6].

Multiplicity of coexisting versions of an article online, article retraction and scientific error propagation

Articles can be hosted on different websites: preprint servers, publishers' websites, bibliographic databases, academic social networks or, illegal black open access websites such as Sci-Hub. Because copies of the indexed references or available full texts may be unidentified retracted versions of an article, retraction information may thereby not be updated on some of these websites [13–17]. Furthermore, retractions may sometimes occur years after publication of articles. So, if sites hosting references and/or full texts do not periodically check their database of articles, unidentified retracted articles will appear in their database. In these cases, authors may obviously not be aware that they are reading and citing a retracted article [18]. Thus, they may continue to cite retracted articles without reference to the retraction [13–15,19–22] and, unknowingly participate in scientific error propagation (see Box 2).

Box 2: Post-retraction citations and error propagation

Numerous articles have studied post-retraction citations in Medicine showing that retraction had little or no impact on the total number of citations. In 2008, Redman et al. showed that some papers in Medicine may have high late post-retraction citations, i.e., cited a total of 225 times occurring 4 or more years after retraction [23]. For example, in cancer research, Bozzo et al. showed that retracted papers were on average cited 45 times with a range of 0-742 [24]. Another example is that in retracted papers reporting a radiology-imaging diagnostic method, Bolboaca et al. showed that the number of post-retraction citations was higher than the number of citations before retraction in 30 of 54 cases (56%) [3]. This phenomenon involves the most cited articles and the most prominent journals. In such cases, highly cited articles can accumulate almost three-fold more citations after retraction than before retraction [20], with sometimes more than half of the citations lacking a reference to the retraction [25]. Moreover, a study published in 2010 in the high profile journal Cell [26], continued to be widely cited in 2016 [22] despite its retraction in 2014.

Citations of unidentified retracted articles in the field of medicine may have potentially serious consequences for patient health, especially when clinical papers are based on findings from retracted papers [4,11,15,18,27].

To avoid the propagation of scientific errors by citing retracted articles, the identification of retracted articles is essential. Thus, retracted articles should be unmistakably identified as such in all online sources. The purpose of this analysis is to determine if this identification is correctly carried out on the various sites hosting articles, i.e., on publishers' websites, on preprint servers (such as bioRxiv or medRxiv), on sites hosting references and/or full-text articles; bibliographic databases (such as Scopus, Web of Science (WoS) and Google Scholar), academic social networks (such as ResearchGate); and finally, on the illegal black open access website Sci-Hub.

Publishers' websites and sites hosting references and/or full-text articles insufficiently identify retracted articles

Information on the retraction of articles should in principle appear first on the publishers' websites. However publishers do not always follow the specific recommendations from COPE related to article retraction as described since 2011 [19]. Consequently, past literature has shown that flagging retractions on publishers' websites are insufficient [19,28–30]. This is confirmed by data presented in Table 1 which shows that, to date, some publishers are still not identifying retractions correctly on their websites, with 20.8% of references and 18.1% of full texts of retracted articles still unidentified as retracted. This shows that there has been no significant improvement on publishers' websites, while electronic journals can now easily and

reliably annotate retracted articles compared to paper journals, for example [19]. It is a real problem because references and full texts of scientific articles are commonly reached by researchers using publishers' websites [14]. Reasons why publishers are inconsistent in how they report retractions and fail to alert readers to retractions seem unclear: perhaps fear for their reputation or ignorance of the seriousness of the problem of retractions? If articles are not corrected on publisher' websites, others sites hosting references and/or full-texts articles can collect these uncorrected references and full texts.

Preprint servers are now numerous [31] and the number of preprints in the life sciences has grown dramatically over the past few years [32]. One study has shown that, among the 30 preprints linked to retracted journal publications present on Research Square, bioRxiv, and medRxiv, 16 (53%) included an indication of the retraction on the preprint [17]. This confirms that these three preprint servers had not implemented any monitoring mechanisms to alert readers to retractions [21].

Subscription-based bibliographic databases such as Scopus and the WoS do not always clearly and consistently display retractions [29], and have been singled out as failing to correctly identify of retracted articles references [19,29,33]. The subscription-based bibliographic databases Scopus and the WoS, still poorly identify retracted articles references: 54.6% and 41.9% respectively (Table 1). This confirms results recently published by Suelzer et al. [33] : 50% for Scopus and 40.6% for the WoS.

Google Scholar is considered as the most comprehensive source of scientific information [34] but has no formal monitoring mechanism to alert readers to retractions. Google Scholar fails in identify 66.3% of retracted article references (Table 1). Recently, Teixeira et al. [35] showed that, in *NEJM* [36] and *Lancet* [37], high-profile publications related to COVID-19 were present in 19 and 59 versions in Google Scholar, respectively, without any indication that the articles had been retracted.

Academic social networks such as ResearchGate, which is mostly oriented towards the science, technology, and medical fields, claims to have 135 million free articles available on its site, mainly indexed by researchers in their profiles. This would make it one of the largest scientific article databases in the world [38]. ResearchGate has implemented monitoring mechanisms to alert readers of retractions for references to articles hosted on its site, but monitoring systems have not been implemented for full texts that are uploaded by researchers. There was no information on retractions in 22.9% of the references hosted by ResearchGate, which represents the lowest percentage observed. In contrast 81.8% of full text retracted articles available on this site were unidentified as retracted articles (Table 1).

The black open access website Sci-Hub allows subscriptions to be bypassed to recover full text 'paywalled articles' and is massively used by researchers [39], especially in low income countries having less efficient institutional access [40]. However, Sci-Hub had not

implemented any monitoring mechanisms to alert readers to retractions. It is very worrying to note that more than 71% of full texts downloadable on Sci-Hub were unidentified retracted articles (Table 1). This considerably damages the usefulness of this site. Google Scholar and Sci-Hub had not implemented any monitoring mechanisms to alert readers to retractions, which explains their failure to correctly identify most of the retracted articles.

Table 1: Number and percentage of references and full texts unidentified as retracted on publishers' websites, bibliographic databases, academic social networks, legal open access, and black open access

A search for retracted articles was done on PubMed on April 8th 2021. PubMed was used because it is considered by the ICMJE as the authoritative source for information about retractions. PubMed uses information published by journals (e.g., retraction notices) to identify retractions. A total of 8559 articles were retrieved from which 500 were randomly selected. For this, a random sequence of integers between 1 and 8559 was generated using Random.org and attributed to the 8559 articles. An ascending classification was performed based on these random integers, and only the first 500 articles were considered for analysis. A check for the presence of any information allowing identification of retraction in references and full-text articles was performed for each of these 500 articles and on each of the sites or services studied. Digital Object Identifiers (DOI) or titles of articles were used as queries. All searches were conducted between April and July 2021. Full-text paywalled articles on publishers' websites were searched using the institutional access of our university. Searches using Sci-Hub were performed in France using non-university internet access. No university or institution affiliated with the authors of this article were therefore involved in downloading articles via Sci-Hub.

"-" : unavailable

		Number of articles with references available (identified or unidentified as retracted) (%)	Number of references unidentified as retracted (%)	Number of articles with full texts available (identified or unidentified as retracted) (%)	Number of full texts unidentified as retracted (%)	Number of articles with both references and full text available (identified or unidentified as retracted) (%)	Number of references and full texts unidentified as retracted (%)
Publisher	**Publishers' websites**	476 (95.2)	99 (20.8)	403 (84.4)	73 (18.1)	403 (84.4)	47 (11.7)
Subscription-based bibliographic databases	**Scopus**	403 (80.6)	220 (54.6)	-	-	-	-
	WoS	420 (84)	176 (41.9)	-	-	-	-
Free bibliographic database	**Google Scholar**	499 (99.8)	331 (66.3)	-	-	-	-
Academic social network	**ResearchGate**	497 (99.4)	114 (22.9)	176 (35.4)	144 (81.8)	176 (35.4)	47 (26.7)
Black open access (illegal)	**Sci-Hub**	437 (87.4)	330 (75.5)	435 (99.5)	310 (71.3)	435 (99.5)	244 (61)

What solutions should be considered and possibly implemented?

Data shown in this analysis show that the identification of retracted articles on publishers' websites, and on different sites such as preprint servers, bibliographic databases, academic social networks, and Sci-Hub is far from satisfactory and that they participate in scientific error propagation. What should be done to overcome this problem?

From the standpoint of a publisher

First of all, publishers should, as suggested by the COPE retraction guidelines for retractions [2], clearly identify retracted articles. This may seem trivial and obvious, but 20.8% of references to retracted articles are still unidentified as retracted on publishers' websites. Therefore, improvements in adherence to COPE retraction guidelines are urgently needed [11]. Furthermore, as suggested by Suelzer et al. [33], systematically adding the prefix of 'retracted' to the title of retracted articles should be considered. This would have the advantage of allowing researchers to better identify retracted articles. Secondly, journals should avoid publishing articles citing retracted articles without reference to the retraction. In order to achieve this objective, in 2006 the members of the International Committee of Medical Journal Editors (ICMJE) changed the Uniform Requirements for Manuscripts to specify the author's responsibility to check their manuscripts for references to retracted articles [41,42]: 'Authors are responsible for checking that none of the references cite retracted articles except in the context of referring to the retraction'. Unfortunately, these recommendations are not necessarily known and applied by the authors, so do not prevent them from citing retracted articles without checking for retraction. Thus, publishers could require authors, through the 'instructions for authors' submission guidelines, to attest that they have checked their submitted manuscripts' references and that no retracted articles are included unwittingly [33]. However, to the best of our knowledge, the vast majority of publishers have not included such statements in their submission guidelines. As an example, one exception to this result found to date is the Methodological Expectations of Cochrane Interventions Reviews (MECIR) Manual [43], which states that retraction statements and errata must be examined prior to including any study in the review or meta-analysis. It is worth noting that Cochrane Standard does not forbid the inclusion of retracted articles, but it states that "care should be taken to ensure that this information [details of the retraction or errata] is retrieved in all database searches ...". To help authors to identify retracted articles in their reference list, the publishers could create a web-based program that would check the manuscript's reference list against a master list of retracted articles and send a report to the manuscript's corresponding author during the submission process [41]. Achieving this does not appear to be out of range, especially since this master list of retracted article already exists, i.e., the Retraction Watch DataBase (RWDB) which was specifically developed to identify retracted articles [44]. The

RWDB was able to identify 477 retracted articles among the 500 (96.4%) included in this analysis.

Reviewing is the last possible means of detecting retracted publications in reference lists of submitted articles. However, even though the reviewers and editors involved are experienced researchers, they depend on the same information sources as the authors [14]. Moreover, the reviewing process is sufficiently time-consuming for reviewers as it is. Having reviewers undertake this task could dissuade some of them and, in the end, make it more difficult for editors to find reviewers for their manuscripts. In our opinion, for this reason this solution would be difficult to apply. Finally, CrossMark [45] could also be worth exploring. This system allows publishers to update versions of articles, including retracted versions, on the condition that the publisher is a member CrossRef (Crossref is the official Digital Object Identifier (DOI) Registration Agency of the International DOI Foundation). To date, CrossMark has updated 130 000 articles, of which nearly 5 000 are retracted articles and, ensures that researchers have access to the latest version of articles [46].

From the standpoint of sites hosting references and/or full-text articles

Globally, sites hosting references and/or full-text articles included in this analysis insufficiently identify retracted articles. However, there are no insurmountable difficulties (financial, technical, etc.) of any kind that can explain this. It seems that this problem is not considered important enough to be recognized by these sites, and therefore has remained largely underestimated for some time. Hence, as a priority for sites hosting articles, if they lack a formal mechanism for alerting readers to retraction of an article (e.g., Google Scholar and Sci-Hub) they should set up a system to identify retracted articles over time to keep their article databases up to date. A routine check using the CrossRef database, would allow identification and flagging of at least 5 000 retracted articles, which would represent significant progress. Such monitoring is very simple to achieve through an Application Programming Interface (API) that is made available. Bibliographic databases such as Scopus and WoS should also try to improve retraction flagging on their sites by developing specific applications or calling upon external partners (e.g., RWDB). These sites could potentially add a mechanism whereby researchers, authors, research support staff, and librarians are able to flag retracted articles in much the same way as they already are able to request author mergers and/or attribution to authors where the attribution is missing.

From the standpoint of researchers

To prevent themselves from citing retracted articles, researchers should not add retracted articles to their stored articles, whether in paper or digital format, when using reference manager software. To do this, they can query the Retraction Watch DataBase article by article

each time they store an article. This is obviously unrealistic because it is a very tedious process that few researchers would accept to implement in their daily practice. Moreover, as retractions sometimes occur years after an article is published, researchers should periodically check the articles they have saved and delete retracted articles if necessary. Reference manager software such as Zotero and EndNote could be a way to avoid using retracted articles [14]. Contrary to other reference manager software services (e.g., Mendeley), these softwares had built-in capabilities for monitoring retractions by checking databases for documents that have been retracted. This feature was developed in partnership with the Retraction Watch in 2019 for Zotero and in 2021 for EndNote. As an example, retracted publications stored in Zotero, present in the RWDB, are flagged in the article list, and a warning informs the user that the stored article is retracted (Figure 1). In addition, each time the software is run, all articles are checked for retractions in the RWDB allowing an automatic update over time.

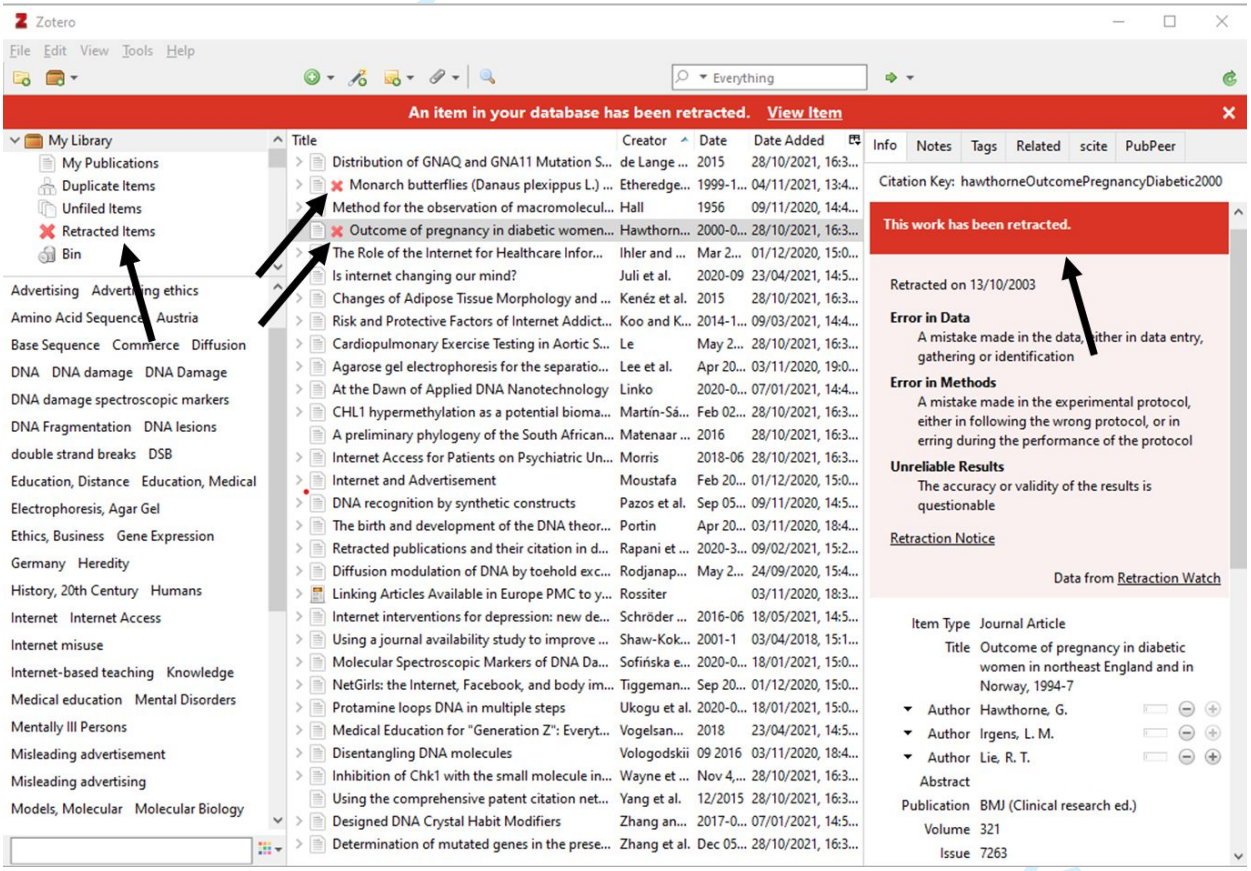

Fig 1 | Display of a retracted article in Zotero

Using Zotero or EndNote, and therefore the RWDB, would result in identifying 477 retracted articles among the 500 (96.4%) in this analysis. This clearly indicates that free and simple-to-implement solutions exist that can dramatically limit the storage, and consequently the citing, of retracted articles by researchers. As said by Suelzer [25], the failure to use such tools 'is a disservice to the readers and researchers', and 'inappropriate citations of retracted articles are

hard to excuse'. Finally, as suggested by Frampton [11], the responsibility to ensure that all copies of the article documents that are available are clearly marked as retracted could be assumed by authors once a decision has been made to retract an article. This should apply to copies of the article on all sites hosting references and/or full-text articles, including illegal sites such as Sci-Hub. One should recognize, however, that in the latter case, the illegal nature of these sites could make this task difficult or even unrealistic.

Box 3: Retracted papers and COVID-19

The COVID-19 crisis has been the source of a significant number of article retractions [11,47,48], with high retraction rates compared to some related research fields for example up to four times higher compared to other infectious diseases such as HIV, H1N1 or Ebola [49]. Such retractions also include high-profile papers [50]. Several retracted studies have resulted from accelerated publication of papers, resulting in the appearance of expedited peer-review processes [48,51,52]. As shown by Teixeira da Silva et al. and Anderson et al., the high speed of publication is believed to lead to less rigorous peer-review [35] and may have led to duplicate articles being published [48].

In conclusion, scientific error propagation, by citing retracted articles without mentioning retractation, may have potential consequences for the quality of research, and in medicine for patient health. Unfortunately, as shown in this analysis, the persistence of unidentified retracted articles is still very high on sites hosting references and/or full-text articles. However, it seems that the solutions mentioned in this article have so far not been sufficiently taken into account by the entire scientific community. Given that retractions are relatively uncommon, they are probably still considered by the vast majority of the scientific community as an epiphenomenon. This has greatly limited awareness of the deleterious consequences of persistent unidentified retracted articles on almost all of the sites used by researchers for documentation. It is also possible that the resolution of this problem, dependent on multiple actors (publishers, sites hosting references and/or full-text articles, researchers), has experienced an effect of dilution of their individual responsibility, limiting the motivation of each to take action [53]. This sentiment is likely accentuated by the fact that the COPE guidelines do not provide any specific recommendations for tracking and deleting unlabelled retracted articles present on a multiplicity of websites and sites hosting references and/or full-text articles once a decision has been made to retract an article.

The COVID-19 crisis has, along with other consequences, placed the issue of scientific retractions at the forefront of scientific and public debate (see Box 3). We certainly hope that this will rapidly lead to heightened awareness of the importance of treating article retractions

appropriately. As pointed out in this article, solutions exist to overcome this worrying problem. It is a collective responsibility that requires a real commitment on the part of all actors involved in scientific publication.

References

1 Fang FC, Steen RG, Casadevall A. Misconduct accounts for the majority of retracted scientific publications. *PNAS* 2012;**109**:17028–33. doi:10.1073/pnas.1212247109

2 Retraction guidelines. COPE: Committee on Publication Ethics. https://publicationethics.org/files/cope-retraction-guidelines-v2.pdf (accessed 20 Sep 2021).

3 Bolboaca SD, Buhai D-V, Aluas M, *et al.* Post retraction citations among manuscripts reporting a radiology-imaging diagnostic method. *PLoS One* 2019;**14**:e0217918. doi:10.1371/journal.pone.0217918

4 Brainard J. Rethinking retractions. *Science* 2018;**362**:390–3. doi:10.1126/science.362.6413.390

5 Bhatt B. A multi-perspective analysis of retractions in life sciences. *Scientometrics* 2021;**126**:4039–54. doi:10.1007/s11192-021-03907-0

6 Gaudino M, Robinson NB, Audisio K, *et al.* Trends and Characteristics of Retracted Articles in the Biomedical Literature, 1971 to 2020. *JAMA Internal Medicine* 2021;**181**:1118–21. doi:10.1001/jamainternmed.2021.1807

7 Van Noorden R. Science publishing: The trouble with retractions. *Nature* 2011;**478**:26–8. doi:10.1038/478026a

8 Fanelli D. Why Growing Retractions Are (Mostly) a Good Sign. *PLOS Medicine* 2013;**10**:e1001563. doi:10.1371/journal.pmed.1001563

9 Grieneisen ML, Zhang M. A Comprehensive Survey of Retracted Articles from the Scholarly Literature. *PLOS ONE* 2012;**7**:e44118. doi:10.1371/journal.pone.0044118

10 Steen RG, Casadevall A, Fang FC. Why Has the Number of Scientific Retractions Increased? *PLOS ONE* 2013;**8**:e68397. doi:10.1371/journal.pone.0068397

11 Frampton G, Woods L, Scott DA. Inconsistent and incomplete retraction of published research: A cross-sectional study on Covid-19 retractions and recommendations to mitigate risks for research, policy and practice. *PLOS ONE* 2021;**16**:e0258935. doi:10.1371/journal.pone.0258935

12 Steen RG. Retractions in the scientific literature: is the incidence of research fraud increasing? *Journal of Medical Ethics* 2011;**37**:249–53. doi:10.1136/jme.2010.040923

13 da Silva JAT, Bornemann-Cimenti H. Why do some retracted papers continue to be cited? *Scientometrics* 2017;**110**:365–70. doi:10.1007/s11192-016-2178-9

14 Bornemann-Cimenti H, Szilagyi IS, Sandner-Kiesling A. Perpetuation of Retracted Publications Using the Example of the Scott S. Reuben Case: Incidences, Reasons and

Possible Improvements. *Sci Eng Ethics* 2016;**22**:1063–72. doi:10.1007/s11948-015-9680-y

15 Schneider J, Ye D, Hill AM, *et al.* Continued post-retraction citation of a fraudulent clinical trial report, 11 years after it was retracted for falsifying data. *Scientometrics* 2020;**125**:2877–913. doi:10.1007/s11192-020-03631-1

16 Davis PM. The persistence of error: a study of retracted articles on the Internet and in personal libraries. *J Med Libr Assoc* 2012;**100**:184–9. doi:10.3163/1536-5050.100.3.008

17 Avissar-Whiting M. Downstream retraction of preprinted research in the life and medical sciences. 2022. doi:10.31222/osf.io/xdekq

18 Steen RG. Retractions in the medical literature: how many patients are put at risk by flawed research? *Journal of Medical Ethics* 2011;**37**:688–92. doi:10.1136/jme.2011.043133

19 Wright K, McDaid C. Reporting of article retractions in bibliographic databases and online journals. *J Med Libr Assoc* 2011;**99**:164–7. doi:10.3163/1536-5050.99.2.010

20 da Silva JAT, Dobranszki J. Highly cited retracted papers. *Scientometrics* 2017;**110**:1653–61. doi:10.1007/s11192-016-2227-4

21 Bar-Ilan J, Halevi G. Temporal characteristics of retracted articles. *Scientometrics* 2018;**116**:1771–83. doi:10.1007/s11192-018-2802-y

22 Bar-Ilan J, Halevi G. Post retraction citations in context: a case study. *Scientometrics* 2017;**113**:547–65. doi:10.1007/s11192-017-2242-0

23 Redman BK, Yarandi HN, Merz JF. Empirical developments in retraction. *Journal of Medical Ethics* 2008;**34**:807–9. doi:10.1136/jme.2007.023069

24 Bozzo A, Bali K, Evaniew N, *et al.* Retractions in cancer research: a systematic survey. *Res Integr Peer Rev* 2017;**2**:5. doi:10.1186/s41073-017-0031-1

25 Piller C. Disgraced COVID-19 studies are still routinely cited. *Science* 2021;**371**:331–2. doi:10.1126/science.371.6527.331

26 Donmez G, Wang D, Cohen DE, *et al.* RETRACTED: SIRT1 suppresses beta-amyloid production by activating the alpha-secretase gene ADAM10. *Cell* 2010;**142**:320–32. doi:10.1016/j.cell.2010.06.020

27 van der Vet PE, Nijveen H. Propagation of errors in citation networks: a study involving the entire citation network of a widely cited paper published in, and later retracted from, the journal Nature. *Research Integrity and Peer Review* 2016;**1**:3. doi:10/gkhpzr

28 Decullier E, Huot L, Samson G, *et al.* Visibility of retractions: a cross-sectional one-year study. *BMC Research Notes* 2013;**6**:238. doi:10.1186/1756-0500-6-238

29 Bakker C, Riegelman A. Retracted Publications in Mental Health Literature: Discovery across Bibliographic Platforms. *Journal of Librarianship and Scholarly Communication* 2018;**6**. doi:10.7710/2162-3309.2199

30 Dal-Ré R, Ayuso C. For how long and with what relevance do genetics articles retracted due to research misconduct remain active in the scientific literature. *Account Res* 2021;**28**:280–96. doi:10.1080/08989621.2020.1835479

31 Kirkham JJ, Penfold NC, Murphy F, *et al.* Systematic examination of preprint platforms for use in the medical and biomedical sciences setting. *BMJ Open* 2020;**10**:e041849. doi:10.1136/bmjopen-2020-041849

32 Biology preprints over time. ASAPbio. https://asapbio.org/preprint-info/biology-preprints-over-time (accessed 28 Feb 2022).

33 Suelzer EM, Deal J, Hanus K, *et al.* Challenges in Identifying the Retracted Status of an Article. *JAMA Netw Open* 2021;**4**:e2115648. doi:10.1001/jamanetworkopen.2021.15648

34 Halevi G, Moed H, Bar-Ilan J. Suitability of Google Scholar as a source of scientific information and as a source of data for scientific evaluation—Review of the Literature. *Journal of Informetrics* 2017;**11**:823–34. doi:10.1016/j.joi.2017.06.005

35 Teixeira da Silva JA, Bornemann-Cimenti H, Tsigaris P. Optimizing peer review to minimize the risk of retracting COVID-19-related literature. *Med Health Care and Philos* 2021;**24**:21–6. doi:10.1007/s11019-020-09990-z

36 Mehra MR, Desai SS, Kuy S, *et al.* RETRACTED: Cardiovascular Disease, Drug Therapy, and Mortality in Covid-19. *N Engl J Med* 2020;**382**:e102. doi:10.1056/NEJMoa2007621

37 Mehra MR, Desai SS, Ruschitzka F, *et al.* RETRACTED: Hydroxychloroquine or chloroquine with or without a macrolide for treatment of COVID-19: a multinational registry analysis. *The Lancet* 2020;**0**. doi:10.1016/S0140-6736(20)31180-6

38 Gusenbauer M. Google Scholar to overshadow them all? Comparing the sizes of 12 academic search engines and bibliographic databases. *Scientometrics* 2019;**118**:177–214. doi:10.1007/s11192-018-2958-5

39 Bohannon J. Who's downloading pirated papers? Everyone. *Science* 2016;**352**:508–12. doi:10.1126/science.352.6285.508

40 Boudry C, Alvarez-Muñoz P, Arencibia-Jorge R, *et al.* Worldwide inequality in access to full text scientific articles: the example of ophthalmology. *PeerJ* 2019;**7**:e7850. doi:10.7717/peerj.7850

41 Sox HC, Rennie D. Research misconduct, retraction, and cleansing the medical literature: lessons from the Poehlman case. *Ann Intern Med* 2006;**144**:609–13. doi:10/ghzwh5

42 ICMJE | Recommendations | Preparing a Manuscript for Submission to a Medical Journal. http://www.icmje.org/recommendations/browse/manuscript-preparation/preparing-for-submission.html#g (accessed 16 Mar 2022).

43 Higgins J, Lasserson T, Chandler J, *et al.* Methodological Expectations of Cochrane Intervention Reviews. London: : Cochrane 2022. https://community.cochrane.org/mecir-manual (accessed 27 Mar 2022).

44 Retraction Watch Database. http://retractiondatabase.org/RetractionSearch.aspx? (accessed 27 May 2021).

45 FabienneMichaud. Crossmark. Crossref. https://www.crossref.org/services/crossmark/ (accessed 7 Oct 2021).

46 Clark R. You are Crossref. Crossref. https://www.crossref.org/ (accessed 7 Oct 2021).

47 Cortegiani A, Catalisano G, Ippolito M, *et al.* Retracted papers on SARS-CoV-2 and COVID-19. *Br J Anaesth* 2021;**126**:e155–6. doi:10.1016/j.bja.2021.01.008

48 Anderson C, Nugent K, Peterson C. Academic Journal Retractions and the COVID-19 Pandemic. *J Prim Care Community Health* 2021;**12**:21501327211015590. doi:10.1177/21501327211015592

49 Yeo-Teh NSL, Tang BL. An alarming retraction rate for scientific publications on Coronavirus Disease 2019 (COVID-19). *Account Res* 2021;**28**:47–53. doi:10.1080/08989621.2020.1782203

50 Two elite medical journals retract coronavirus papers over data integrity questions. https://www.science.org/content/article/two-elite-medical-journals-retract-coronavirus-papers-over-data-integrity-questions (accessed 28 Feb 2022).

51 El-Menyar A, Mekkodathil A, Asim M, *et al.* Publications and retracted articles of COVID-19 pharmacotherapy-related research: A systematic review. *Sci Prog* 2021;**104**:368504211016936. doi:10.1177/00368504211016936

52 Watson C. Rise of the preprint: how rapid data sharing during COVID-19 has changed science forever. *Nature Medicine* 2022;**28**:2–5. doi:10.1038/s41591-021-01654-6

53 Suelzer EM, Deal J, Hanus KL, *et al.* Assessment of Citations of the Retracted Article by Wakefield et al With Fraudulent Claims of an Association Between Vaccination and Autism. *JAMA Network Open* 2019;**2**:e1915552. doi:10.1001/jamanetworkopen.2019.15552