The relationship and incidence of three editorial notices in PubPeer: errata, expressions of concern and retractions

José Luis Ortega

Institute for Advanced Social Studies (IESA-CSIC), Córdoba, Spain

Joint Research Unit Knowledge Transfer and Innovation, (UCO-CSIC), Córdoba, Spain

jortega@iesa.csic.es ORCID: 0000-0001-9857-1511

Abstract

This article studies for the first time the incidence and relationship of three important relevant editorial notices: errata, expressions of concern and retractions. The journal club PubPeer was used to extract 39,449 research articles and the 2,308 errata, 189 expressions of concern and 1,531 retractions associated with these publications. The relationship, time delay and evolution of these publications were then compared, as was their incidence in journals and disciplines. The results show that the relationship between them is scant, the increase in these notices has slowed in proportion to the scientific literature and the time delay is more than three years. According to incidence, editorial notices are more frequent in journals specializing in biochemistry, medicine and multidisciplinarity—cancer journals release more errata. Research areas with more editorial notices are the life and health sciences. The research fields with the highest percentages of errata and retractions are *Biochemistry, Genetics and Molecular Biology* and *Immunology and Microbiology* are the research fields that show the highest percentages of errata and retractions. Finally, the paper discusses possible limitations of PubPeer as data source and how they could affect the results.

Key points

- The increase in editorial notices has slowed in proportion to the scientific literature with a time delay of more than three years.
- Editorial notices are more frequent in journals specializing in biochemistry, medicine and multidisciplinarity.
- PubPeer covers more publications from biochemistry, medicine and psychology than from the physical sciences.

Introduction

An essential element in the validity of scientific research is the need to disseminate and publish results to be approved and agreed by the scientific community. Traditionally, research journals have been the principal means to issue these results, a key agent in scientific communication and the main intermediary between scientists and their communities (Meadows, 1979). The leading role of the journal has contributed to the publication of research results being considered the main indicator by which to value the performance of the system (Gingras, 2016). Consequently, there is considerable pressure to publish results in top-tier journals (publish or perish), in some cases, leading to recurrent mistakes, plagiarism, data manipulation or a disrupted peer review process (Tijdink et al., 2014).

This ongoing rise in problem publications is causing editorial boards to carry out serious investigations into the honesty and soundness of already published papers (Brainard and You, 2018; Khadilkar, 2018). The image and prestige of journals could be at risk if they do not adopt measures against mistakes and unethical practices (Smith, 1997). Errata, corrections, addenda, expressions of concern and retractions are some editorial actions taken to correct misconduct and clarify dubious results. Analysis of these types of document would help explore how these errors and unethical practices are contained by publishers and authors, and provide insights into how these practices are being created and eradicated.

Another factor that has favoured the proliferation of these editorial actions is the appearance of web sites focused on detecting and reporting bad academic practices. Retraction Watch (retractionwatch.com) and PubPeer (pubpeer.com) are platforms that discuss and report the soundness of research studies (Knoepfler, 2015). The pressure of these sites is forcing journals and editors to open investigations and to require authors to explain their methods and results.

However, although these editorial notices are increasingly common and are being registered by different sources (Retraction Watch, PubMed, etc.), there is a gap in the literature on the manner in which they are quantitatively distributed and the relationship between them. The reason could be that this information is hard to collect because some publishers do not adequately identify these notes, some articles are corrected or removed without any explanation, or the link between article and editorial notice is not correctly set up (da Silva & Dobránszki, 2017).

Literature review

Research into these editorial notices has been unequal, as studies on retraction notices have been far more prolific than on errata or expressions of concern. This could be because most retractions refer to misconduct and are used as an indicator of bad scientific practices. Many authors have undertaken exploratory studies to discover the incidence of this type of editorial literature. Wager and Williams (2011) used a Medline sample, observing a steady increase in retractions in the period 1999–2009. Similarly, Steen (2011) studied 742 articles from PubMed, perceiving that the levels of misconduct appeared to be higher than in the past. Fang et al. (2012) corroborated this perception, showing that the percentage of articles retracted had increased tenfold since 1975. Despite this growth, not all retractions are due to misconduct. Fang et al. (2012) showed that only 21% of retractions were attributable to mistakes. A slightly lower percentage (17%) was found by Moylan and Kowalczuk (2016) when they studied 173 articles from BioMed Central. Other studies have analysed the citation impact of retracted articles from a bibliometric perspective. Budd et al. (1999) were the first to develop a quantitative study about retractions, observing that many articles continue receiving citations after their retraction. Regardless of this, Furman et al. (2012) observed that biomedicine retractions cause an immediate, severe and long-lived decline in future citations to their authors. Shema et al. (2019) compared the altmetric impact of retracted papers, noticing a significant social media impact for these publications. However, retractions are not uniformed by disciplines or regions. Ribeiro and Vasconcelos (2018) analysed retractions from Retraction Watch at country level, but found no correlation between retractions and total publications, suggesting that local factors explain misconduct in certain countries. Stretton et al. (2012), detected a statistical association between retractions and first authors affiliated with lower-income countries. Huh et al. (2016) found that the main reason for retraction in Korean publications was duplication. Other studies have addressed different aspects of retractions. Grieneisen and Zhang (2012) performed one of the

most complete analyses with more than 4,000 retractions. They found that the percentages of retractions in medicine, life science and chemistry were higher than the number of publications in Web of Science of these same disciplines. Tripathi et al. (2019) detected that Oncology and Biochemistry & Molecular Biology were the research areas with the most retractions. Steen et al. (2013) studied retraction delay, observing that journals are retracting papers more quickly than in the past. More recently, Vuong (2020) studied a large sample of retracted papers and uncovered important imprecisions about the reasons and authors of the retraction notices.

However, other types of editorial notices have attracted less attention from the scientific community. The reasons could be that these actions imply fewer ethical questions (errata) or because they are not widespread in all research journals (expressions of concern). In any case, the first quantitative studies about errata are recent. Bhatt et al. (2014) analysed the errata of five top-tier medical journals, finding a mean of 1.3 errata per issue and a mean delay of 122 days for reporting an erratum. Hauptman et al. (2014), studying 20 English-language, general medicine and cardiovascular journals, observed an overall errata report occurrence of 4.2%. In a similar study, Farrah and Rabb (2019) appreciated an average time delay of 211 days and 20% of articles with errata. Pichardo-Corpus et al. (2019) carried out a disciplinary analysis when they found that biology and health sciences disciplines gather significantly more errata.

In terms of expressions of concern, the number of studies is even lower. Worthy of mention is the paper by Noonan and Parrish (2008), which examined the reasons and consequences of 16 expressions of concern; and the comprehensive article by Vaught et al. (2017) about more than 200 expressions of concern. Most were issued because of concerns with validity of data, methods or interpretation of the publication. However, no study has undertaken a joint analysis of these three editorial actions to explore their incidence and interactions.

Most of these studies are based in traditional academic databases (PubMed, Web of Science, Scopus) and specialized lists (Retraction Watch). However, no study has used data from a post-publication peer review site such as PubPeer. In fact, this platform has only been explored from an ethical (Blatt, 2015) and legal point of view (da Silva, 2018). The main advantage of this source is that it allows to track the relationship between critical comments and the later editorial notices.

Objectives

The aim of this study is to explore the relationships, evolution and time delay of errata, expressions of concern and retractions for a better understanding of how they originate and for what purpose. It also analyses the distribution of these document types according to journal and discipline in order to pinpoint where they appear more frequently as indicators of troublesome publications. In addition, the study analyses the PubPeer platform for the first time to discover how comprehensive it is and whether there is bias towards specific research areas and journals. These objectives arise from the need to know how editorial notices are related and where are more frequent because they inform us about errors and misconduct in specific research journals and topics.

Methods

This study takes a quantitative approach to respond the previous objectives. To know the distribution of editorial notices in the time, delays in their publication and their proportion

according to journals and subjects is necessary to adopt statistical tools that describe how the incidence of these documents is.

Source

PubPeer is a scientific forum or journal club where scientific publications are discussed after publication. A special feature of this post-publication peer review site, created in October 2012, is that comments can be anonymous. As a result, the forum has become a specialized site for reporting bad publishing practices. Considerable controversy has ensued because many authors feel defenceless in the face of unknown complainants. However, many users are grateful for this format because bad practices can be flagged with no reprisals. In addition to the comments, PubPeer announces when an erratum, an expression of concern or a retraction notice has been published. The source of these notices is mainly PubMed (65.9%), but notices can also come directly from academic publishers.

Data

Few sources manage this information properly. Citation indexes, bibliographic databases and academic search engines index this information but do not link the publication with the editorial note or they do not properly differentiate between notices. For example, when the Web of Science indexes an editorial notice, there is not any field that includes an identifier or link to the original paper. Instead, it supplies incomplete information (volume, issue, publication year) that it is not enough to identify the original article. Scopus does link the editorial notice with the original article, but the extraction of data from this paywall service is limited and it does not correctly identify different types of notices. PubMed is the most complete in this sense because it distinguishes between different editorial actions and links the notes with the original publication. However, this is a specialist tool for biomedical literature and does not cover other disciplines. For this reason, we selected PubPeer to obtain the data.

Two samples were extracted from PubPeer: in March 2019 and in January 2020. The second sample was used to update the first. After random queries (searching for the first letters of the alphabet: a, b, and c), all the entries published until December 2018 were retrieved: 39,985 posts about 39,449 publications, in total. Bibliographic data, authors and affiliations were extracted using web scraping techniques. According to Crossref, there are many dates for research articles (deposited, created, published, issued date, etc.) that would confuse the results, mainly about the delay of the editorial notices. Issued date, when a paper has been assigned to an issue, was used because is the most stable date.

We used the subject classification of the journals to perform disciplinary classification. This information was obtained from the Scimago Journal & Country Rank portal (scimagojr.com). This site uses All Science Journal Classification (ASJC) to categorize and rank journals. For documents published at other venues (repositories, conferences, books, etc.), we undertook a manual classification.

Editorial notices

This study focuses on three types of editorial actions about the integrity of a published research article (Allahbadia, 2014; NLM, 2018):

Erratum: a short note in which authors or editors correct errors in the article, e.g. a typographical error or a substantial omission of relevant information. It is also known as a

corrigendum or addendum. These errata are indexed when they are expressed separately, with issue, pagination and identifier.

Expression of concern: the most recent type. It alludes to a notice from the editor or editorial board drawing attention to potential problems in an article. The objective is to alert readers about possible irregularities before a final decision is made. This document marks the beginning of an investigation that may or may not result in a retraction. Like erratum, these documents are indexed separately.

Retraction: this document can be released by the authors, editors or institutions and refers to the withdrawal of an article due to serious irregularities in the publication or in the research. Retractions may be partial, when a section is erroneous, or complete, when all the research is not acceptable.

Results

Of the 39,449 publications extracted from PubPeer, 2,308 were the subject of an erratum (5.9%), 189 of an expression of concern (0.5%) and 1531 of a retraction (3.9%).

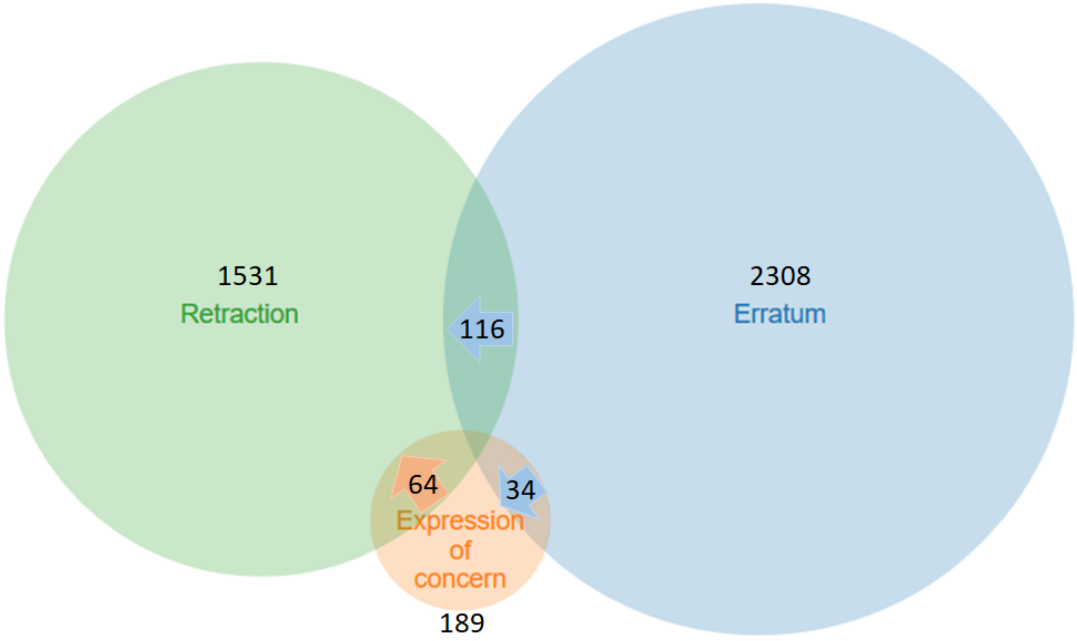

Figure 1. Venn diagram about the relationship and overlap between editorial notices.

Figure 1 shows the size of each notice and the relationships between them. Each overlap indicates the proportion of editorial notices that go from a type to other. The first finding reveals few connections between each type of document because the overlap is very low. This suggests that these notices rarely precede or follow the others. Thus, only 1.5% (N=34) of errata lead to an expression of concern and 29.4% (N=10) of these expressions of concern eventually become a retraction. Expression of concern interacts most with the other documents: 33.8% (N=64) become retractions, while 18% (N=34) stem from an erratum. Only 5% (N=116) of the errata involve a retraction. From a different stance, 6.9% (N=106) of all retracted articles were preceded by an erratum and 4.2% (N=64) by an expression of concern.

Articles publication date

This section analyses the distribution of editorial notices according to the publication date of the original publication. The aim is to observe the growing pattern of these documents and detect interactions among them.

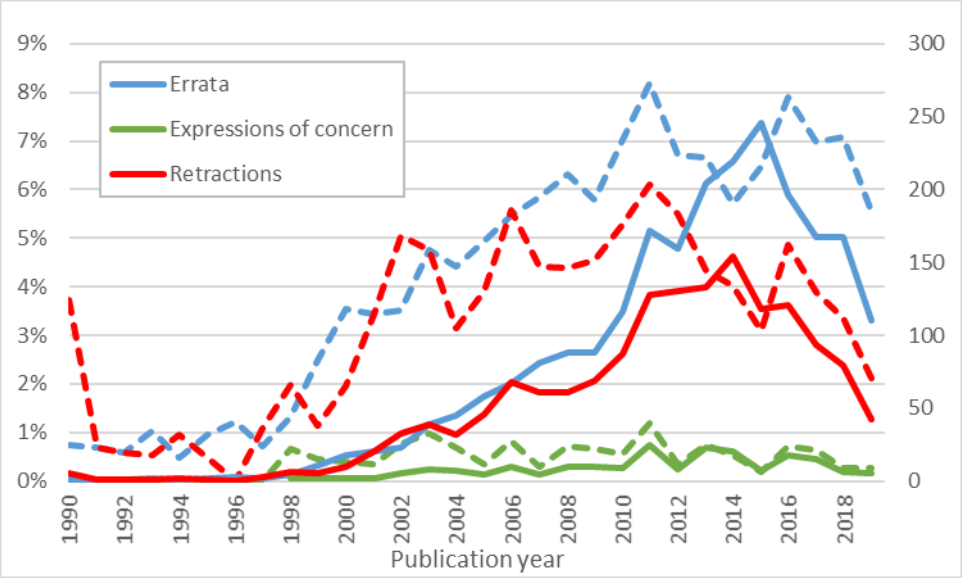

Figure 2. Distribution of publications with editorial notices in PubPeer by publication year (dotted line=percentage; solid line=total).

Figure 2 plots the longitudinal distribution of publications with editorial notices covered by PubPeer by publication date. The solid line shows the total number of articles and the dotted line the percentage over the total number of publications in PubPeer. The graph depicts a continuous rise in publications with an editorial action up to 2014 for retractions and 2015 for Errata. The time delay between the publication date and the moment the errata and retractions are published would explain the decline in cases from 2014. As for expressions of concern, the low number of notices reveals no clear pattern. However, that the period 2011–2014 (38%, N=76) includes more expressions of concern suggests that this editorial notice suffers an important delay. However, as the percentage of the sample indicates, this increase in distribution is not so pronounced, and both errata and retractions show peaks and valleys from 2000.

Time delay

This section analyses the time delay between the publication of an article and the issue of the editorial notice. Differences between distributions disclose the speed with which these editorial decisions are made.

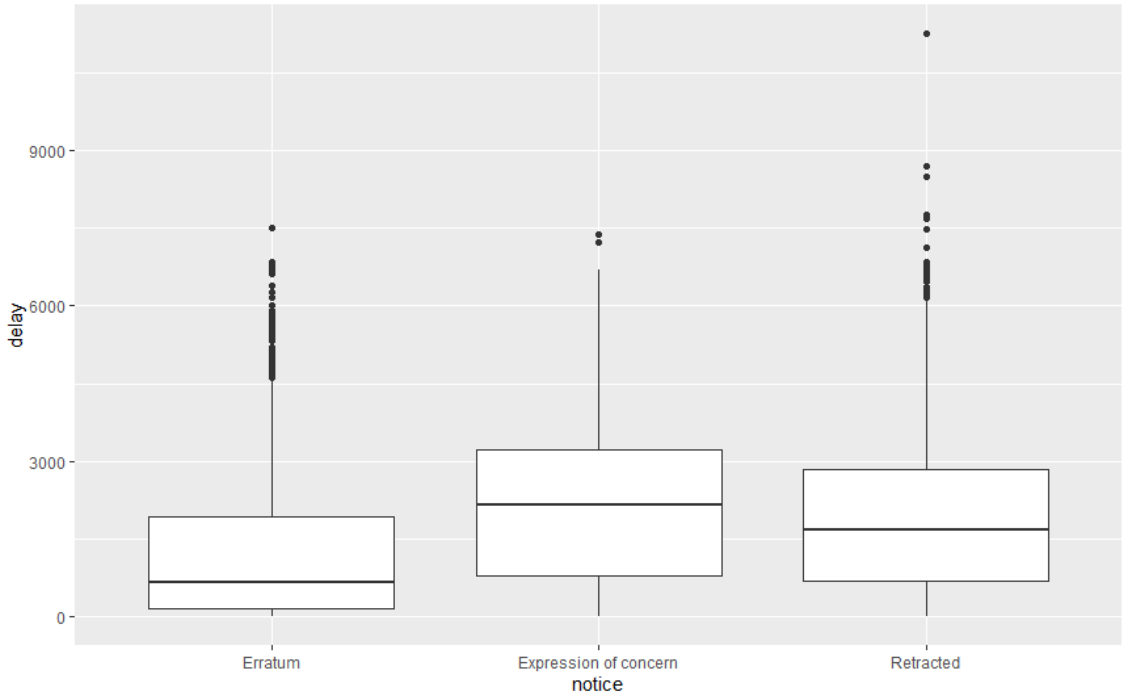

Figure 3. Distribution of the time delay (in days) of the three types of editorial notices.

Figure 3 depicts a Box and Whisker plot with the distribution of the time delay between the publication of a paper and the issue of the editorial notice according to the different types of notes. First of all, the graph shows that the distribution of the three types of notes is skewed, with some actions taking a long time (almost 28 years for an erratum and 30 years for a retraction) and with considerable dispersion (a standard deviation of roughly 4.3 years). This low homogeneity could be explained by multiple factors such as the type of problem, journal, discipline and so forth, which introduce a marked variation in the delay of these documents. On average, errata are published the most swiftly (taking 1,164 days, 3.2 years) because they usually allude to small corrections that can be rapidly offset. Next are retractions with a mean delay of 1,947 days (5.3 years). A retraction involves a drastic decision that, in some cases, is preceded by a long and careful investigation. Finally, expressions of concern are, on average, published 2,317 days (6.4 years) later. The considerable delay in this type of document could be explained by the fact these statements are issued after an in-depth enquiry with no clear conclusions.

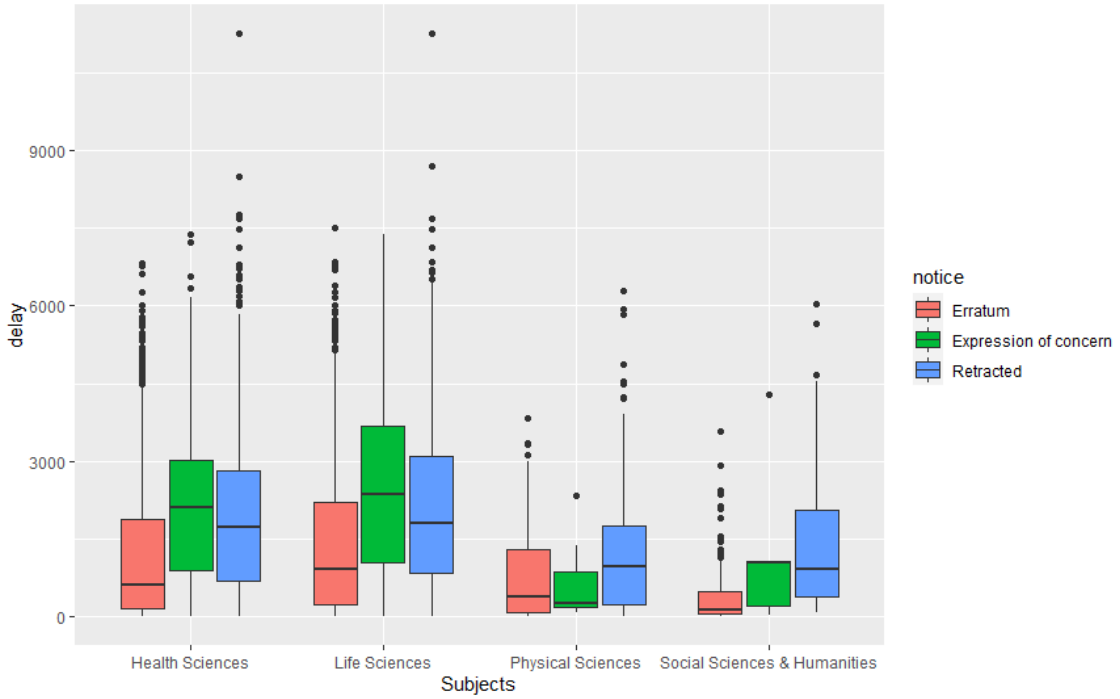

Figure 4. Distribution of the time delay (in days) of the three types of editorial notices by research area.

Editorial notice	Research Area	Count	Mean	Median	Standard Deviation	Confidence Level (95.0%)
Errata	Health Sciences	1399	1187.8	593.5	1365.5	76.9
	Life Sciences	2893	1402.4	912	1439.9	74.0
	Physical Sciences	651	749.2	365	877.2	113.2
	Social Sciences & Humanities	281	427.1	122	673.6	111.3
Expressions of concern	Health Sciences	148	2229.3	2108	1722.8	315.5
	Life Sciences	200	2510.5	2346	1758.8	345.5
	Physical Sciences	37	576.7	240.5	688.7	437.6
	Social Sciences & Humanities	5	1318.0	1030	1724.8	2141.6
Retractions	Health Sciences	969	1940.9	1711.5	1543.6	105.6
	Life Sciences	2160	2140.0	1798	1643.9	98.4
	Physical Sciences	546	1205.2	959	1267.8	180.0
	Social Sciences & Humanities	78	1479.6	898	1514.6	398.2

Table 1. Statistical parameters of the time delay (in days) in the publication of editorial notices by research area.

Figure 4 and Table 1 depict a more detailed view of the time delay in the publication of editorial notices, distinguishing between the four main research areas of ASJC. The results show that the three types of editorial notices have a longer mean delay in *Health Sciences* and *Life Sciences* than in *Physical Sciences* and *Social Sciences & Humanities*. In this sense, *Life Sciences* display the greatest mean delay in errata (1,402 days), expressions of concern (2,510 days) and retractions (2,140). *Physical Sciences* have the shortest delay in expressions of concern (576 days) and retractions (1,205 days), while errata are published most swiftly in *Social Sciences & Humanities* (427 days).

Journals

The distribution of the number of editorial notices by journal could be a fair indicator to detect the venues that attract more contentious publications. The projection of errata, expressions of concern and retractions by journal inform us about differences between venues.

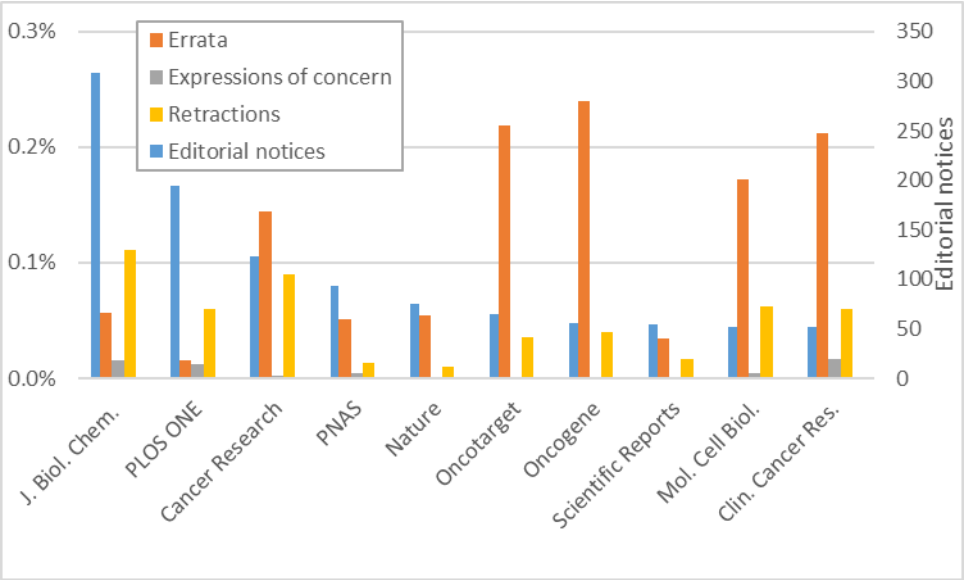

Figure 5. The ten journals with most editorial notices (secondary axis) and percentage of errata, expressions of concern and retractions per article (principal axis).

Journal	Editorial notices	articles	Errata	Errata ‰	Expressions of concern	Expressions of concern ‰	Retractions	Retractions ‰
J. Biol. Chem.	308	180,133	101	0.56	27	0.15	200	1.11
PLOS ONE	194	228,226	36	0.16	27	0.12	138	0.60
Cancer Research	123	52,026	75	1.44	1	0.02	47	0.90
PNAS	93	140,006	71	0.51	6	0.04	18	0.13
Nature	75	118,317	64	0.54	1	0.01	12	0.10
Oncotarget	65	25,560	56	2.19	0	0.00	9	0.35
Oncogene	56	19,999	48	2.40	0	0.00	8	0.40
Scientific Reports	55	110,434	38	0.34	0	0.00	18	0.16
Mol. Cell Biol.	52	22,659	39	1.72	1	0.04	14	0.62
Clin. Cancer Res.	52	18,353	39	2.12	3	0.16	11	0.60

Table 2. The ten journals with most editorial notices on PubPeer and proportion (one per mile) of these notices over the total number of publications (PubMed, 2020).

Figure 5 and Table 2 include the ten journals with most editorial notices in PubPeer, along with their percentages of errata, expressions of concern and retractions over the total number of articles issued. The total number of publications of each journal was obtained from PubMed in April 2020. The journals that produce most editorial notices come mainly from biochemistry (*Journal of Biological Chemistry*, 7.8%, and *Molecular and Cellular Biology*, 1.3%), medicine (more precisely, oncology) (*Cancer Research*, 3.1%, *Oncotarget*, 1.6%, *Oncogene*, 1.4%, and *Clinical Cancer Research*, 1.3%) and multidisciplinarity (*PLOS ONE*, 4.9%, *Proceedings of the National Academy of Sciences*, 2.3%, *Nature*, 1.9%, and *Scientific Reports*, 1.4%). Interestingly, cancer journals are the venues with the highest proportion of errata. *Oncogene*, *Oncotarget* and

Clinical Cancer Research are thus the journals with more than two errata per thousand articles. A different distribution is appreciated in expressions of concern. In this case, the journals *Clinical Cancer Research* (.16‰), *Journal of Biological Chemistry* (.15‰) and *PLOS ONE* (.12‰) issue most expressions of concern. However, some of the journals that release most retractions are related to biochemistry such as the *Journal of Biological Chemistry* (1.11‰) and *Molecular and Cellular Biology* (.62‰). However, some oncology journals also have high proportions of retractions such as *Cancer Research* (.9‰) and *Clinical Cancer Research* (.6‰).

Subjects

This section examines the disciplinary distribution of editorial notices in relation to publications covered by PubPeer and Scopus. The aim is to observe which research areas have a greater proportion of errata, expressions of concern and retractions.

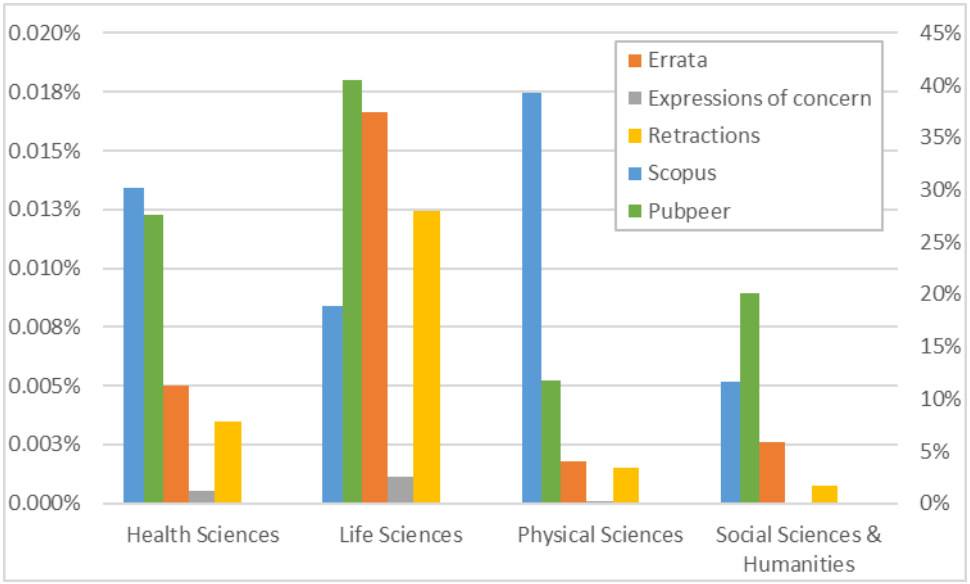

Figure 6. Distribution of publications in PubPeer and Scopus (secondary axis) and the percentage of errata, expressions of concern and retractions (primary axis) according to subject areas.

Figure 6 groups the editorial notices by ASJC subject areas. PubPeer and Scopus are used as a reference to compare the proportion of each type of notice against the total number of research publications both in PubPeer and Scopus. The comparison with Scopus allows to observe which research areas have more editorial notices in relation to their total production. Scopus was selected because it is disciplinarily more balanced than Web of Science (Martín-Martín et al., 2018). PubPeer is also included to indicate whether the differences according to Scopus is due to coverage biases in PubPeer. The percentage of editorial notices were calculated over Scopus publications. The graph reflects that PubPeer is significantly biased towards *Life Sciences* (PubPeer=41%, Scopus=19%) and *Social Sciences & Humanities* (PubPeer=20%, Scopus=12%) papers, while *Physical Sciences* (PubPeer=12%, Scopus=39%) publications are underrepresented in comparison to Scopus. These biases will be considered when the data are interpreted. The picture clearly shows that *Life Sciences* is the area that gathers the highest proportion of errata (0.17‰), expressions of concern (0.01‰) and retractions (.12‰), followed by *Health Sciences*, but with lower figures (errata=0.05‰, expressions of concern=0.01‰ and retractions=.03‰).

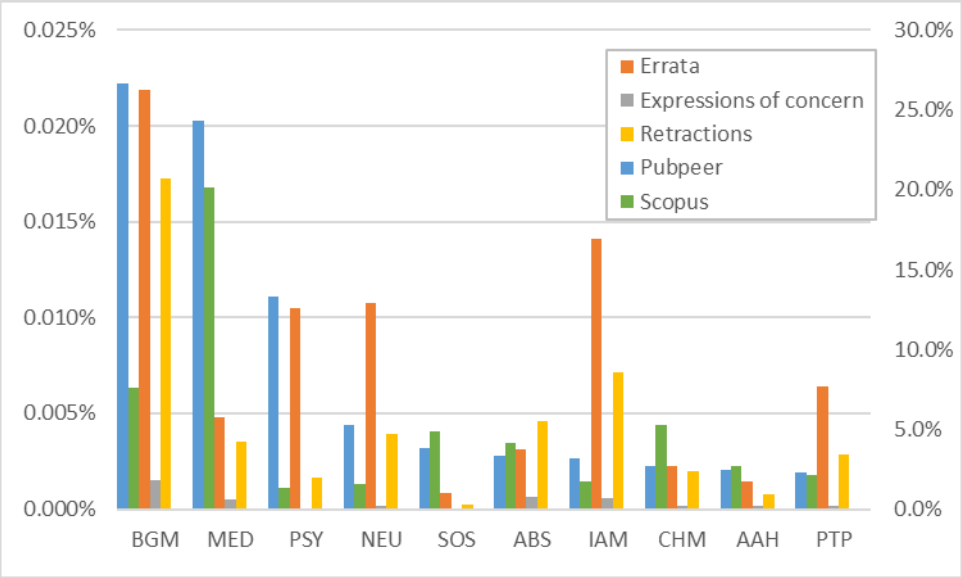

Figure 7. Distribution of publications in PubPeer and Scopus (secondary axis), and percentage of errata, expressions of concern and retractions (primary axis) in relation to Scopus articles by research fields.

Figure 7 shows a more detailed view of which disciplines are more discussed (the first ten) on PubPeer and the proportion of editorial notices in each field. *Biochemistry, Genetics and Molecular Biology* (BGM) (26.7%, N=21,962), *Medicine* (MED) (24.3%, N=20,017) and *Psychology* (PSY) (13.3%, N=10,947) are the most commented field in PubPeer. In particular, *Biochemistry, Genetics and Molecular Biology* (Scopus=7.6%) and *Psychology* (Scopus=1.3%) are significantly overrepresented in PubPeer in relation to Scopus. However, *Engineering* (PubPeer=1.6%, Scopus=11.6%) and *Physics and Astronomy* (PubPeer=1.5%, Scopus=7.1%) are the least commented research fields in PubPeer. This result suggests that PubPeer could be a specialist network that focuses on life and health sciences disciplines.

According to the proportion of errata, expressions of concern and retractions, the figure again shows that *Biochemistry, Genetics and Molecular Biology* (BGM) (0.22‰, N=2059) and *Immunology and Microbiology* (IAM) (0.14‰, N=305) are the fields with the highest percentages of errata, as well as the greatest proportion of retractions (BGM=0.17‰, N=1624; IAM=0.07‰, N=154). For expressions of concern, *Biochemistry, Genetics and Molecular Biology* (BGM) (0.015‰, N=145) and *Agricultural and Biological Sciences* (ABS) (0.007‰, N=34) are the fields with most publications. These results emphasize that biology and biochemistry are the disciplines that produce more troublesome publications. Conversely, *Social Sciences* (SOS) and *Arts and Humanities* (AAH) have very few editorial notices.

Discussion

This pioneering study about the joint analysis of the three most important editorial notices has enabled us to discover how these document types are interrelated and the incidence of these notices by journal and discipline. The results show a limited connection between these editorial notices. 0nly 6% of errata become a retraction (5%) or an expression of concern (1.5%), making it clear that most errata overcome the integrity problems of the publications. More interesting are the expressions of concern, since they are a transitory statement about whether a publication will ultimately be retracted or not. In this sense, only 18% of expressions of concern stem from an erratum and 34% of these end in a retraction. Lastly, most retractions are not

alerted by previous statements because only 6.9% are preceded by an erratum and 4.2% by an expression of concern. These results suggest that these editorial types are highly independent of each other and that only in a few cases is an editorial action based on previous reports.

The longitudinal distribution of these notices shows a constant increase similar to other studies (Fang et al., 2012; Steen et al., 2013; Wang et al., 2019). Figure 1 shows how errata and retractions significantly increase, both in number and proportion. However, percentage-wise, these editorial notices show a discontinued growth, marked by several fluctuations that describe an irregular growth. These increases would be caused by the small number of notices and the selection of data from PubPeer, which could have favoured the inclusion of recent publications.

According to the delay in the publication of these notices, the results have shown that on average these documents take more than three years to be released. This finding contrasts with previous studies that observed dissimilar delays, from 122 (Bhatt et al, 2014) and 211 days (Farrah and Rabb, 2019), in the case of errata, to more than three years for retractions (Nath et al., 2006; Rosenkrantz, 2016; Park et al., 2018; Bhatt, 2019). These differences could be due to the great variability in the delay between article publication and the editorial notice because they depend on the type of error, discipline, journal and so forth. In our case, an important reason is that this is a multidisciplinary study, with significant differences between research areas, which increases variability. Despite these factors, the considerable average delay in editorial notices, more in expressions of concern and retractions, can be an example of how slowly the publishing system takes decisions in the face of scientific errors.

The findings deriving from the incidence of editorial notices in disciplines and journals clearly indicate that these documents are more common in biology and the health sciences. Biochemistry journals such as the *Journal of Biological Chemistry* and *Molecular and Cellular Biology* have more retractions. These results coincide with Bhatt (2019), who identifies the *Journal of Biological Chemistry* as the first in retractions, whereas oncology journals, such as *Oncotarget* and *Oncogene*, gather a higher proportion of errata. A possible explanation could be that these journals show high rates of false discoveries (Hall and Hendricks, 2019) and the proliferation of fraudulent papers about cancer biomarkers (Byrne et al., 2019).

By research field, *Biochemistry, Genetics and Molecular Biology* and *Immunology and Microbiology* are the disciplines with the highest rates of errata, expressions of concern and retractions (Pichardo-Corpus et al., 2019; Tripathi et al., 2019). These results may be interpreted in two ways. The first would suggest that these research disciplines produce more misconduct, perhaps because they are extremely competitive (Kennedy, 2000; Anderson et al., 2007). However, a second interpretation could be that journals from these disciplines are especially involved in the detection of errors and fraud in publications and that they therefore release more editorial notices than other disciplines due to their experience and rigour. By contrast, *Psychology*, which is the third discipline in commented articles, shows very low percentages of editorial notices. This could mean that the proportion of misconducts is low in this discipline or that these journals could be less committed to reporting errors. Undoubtedly, more studies are needed to clarify these questions.

However, these results are limited to PubPeer and they need to be contrasted with other sources. The distinctive feature of this platform is that it is not a bibliographic database but a journal club. This fact could distort the coverage of articles and editorial notices. For example, this web platform is a specialist forum for discussing research articles, mainly for reporting bad publishing practices. This fact could bias the sample including more troublesome articles, and

therefore editorial notices, than other sources. The comparison between Scopus and PubPeer has led us to observe that PubPeer is significantly biased towards biomedicine disciplines, with biochemistry, medicine and psychology being the most commented fields (Fang et al., 2020). *Physical Sciences*, on the other hand, are seriously underrepresented. These differences may have influenced the incidence of editorial notices, resulting in the undervaluing of the *Physical Sciences*.

Conclusions

This descriptive analysis on the characteristics and distribution of errata, expressions of concern and retractions in PubPeer has allowed us to draw several conclusions. The first is that the relationship between these editorial notices is scant. Only in a few cases does an expression of concern cause a final retraction.

The growth of these notices is ongoing, but they have increased in proportion to the scientific literature since 2000. The distribution of the time delay between the publication of the article and the release of the editorial notice is heavily skewed, with a very wide publication window and differences between disciplines. On average, errata take more than three years to be released, while expressions of concern and retractions take more than five. By discipline, the longest periods are in the *Life Sciences* and *Health Sciences*.

According to the incidence of editorial notices the *Journal of Biological Chemistry* and *PLOS ONE* gather more editorial notices, *Oncotarget* and *Oncogene* have more errata, while the *Journal of Biological Chemistry* and *Cancer Research* include a greater proportion of retractions. By discipline, editorial notices are more frequent in *Life Sciences* and *Health Sciences*. Concretely, *Biochemistry, Genetics and Molecular Biology* and *Immunology and Microbiology* are the disciplines with the highest rates of expressions of concern and retractions.

Finally, with regard to the coverage of PubPeer, the data show that this site tends to discuss publications from biomedicine areas, especially, biochemistry, medicine and psychology, whereas publications from *Physical Sciences* are underrepresented. The results suggest that Pubpeer could be an interesting source to study editorial notices, although they should be interpreted in the scope of a post peer-review site.

References

Allahbadia, G. N. (2014). Why Correcting the Literature with Errata and Retractions is Good Medical Practice? Journal of Obstetrics and Gynecology of India, 64(6), 377–380. https://doi.org/10.1007/s13224-014-0643-z

Anderson, M. S., Ronning, E. A., De Vries, R., & Martinson, B. C. (2007). The perverse effects of competition on scientists' work and relationships. Science and engineering ethics, 13(4), 437-461. https://doi.org/10.1007/s11948-007-9042-5

Bhatt, B. (2019, 24 March). On Retractions in Biomedical Literature. Toward Data Science. [Web log post]. Retrieved from: https://towardsdatascience.com/on-retractions-in-biomedical-literature-1565e773559e

Bhatt, V. R., Aryal, M. R., Panta, S., Mosalpuria, K., & Armitage, J. O. (2014). A retrospective analysis of reported errata in five leading medical journals in 2012. Journal of community hospital internal medicine perspectives, 4(5), 25738. https://doi.org/10.3402/jchimp.v4.25738

Blatt, M. R. (2015). Vigilante science. Plant Physiology, 169(2), 907-909. https://doi.org/10.1104/pp.15.01443

Brainard, J., & You, J. (2018). What a massive database of retracted papers reveals about science publishing's 'death penalty'. Science, 25(1), 1-5.

Budd J. M., Sievert, M. E., Schultz, T. R., & Scoville, C. (1999). Effects of article retraction on citation and practice in medicine. Bulletin of the Medical Library Association, 87(4), 437–43.

Budd, J. M., Coble, Z. C., Anderson, K. M. (2011). Retracted publications in biomedicine: cause for concern. ACRL Conference, Philadelphia, PA.

Byrne, J. A., Grima, N., Capes-Davis, A., & Labbé, C. (2019). The possibility of systematic research fraud targeting under-studied human genes: causes, consequences, and potential solutions. *Biomarker insights*, *14*, 1177271919829162. https://doi.org/10.1177/1177271919829162

Fang, F. C., Steen, R. G., & Casadevall, A. (2012). Misconduct accounts for the majority of retracted scientific publications. Proceedings of the National Academy of Sciences, 109(42), 17028-17033. https://doi.org/10.1073/pnas.1212247109

Fang, Z., Costas, R., Tian, W., Wang, X., & Wouters, P. (2020). An extensive analysis of the presence of altmetric data for Web of Science publications across subject fields and research topics. 124(2), 2519–2549. https://doi.org/10.1007/s11192-020-03564-9.

Farrah, K., & Rabb, D. (2019). Errata for trial publications are not uncommon, are frequently not trivial, and can be challenging to access: a retrospective review. Journal of the Medical Library Association: JMLA, 107(2), 187. https://doi.org/10.5195/jmla.2019.629

Furman, J. L., Jensen, K., & Murray, F. (2012). Governing knowledge in the scientific community: Exploring the role of retractions in biomedicine. Research Policy, 41(2), 276-290. https://doi.org/10.1016/j.respol.2011.11.001

Gingras, Y. (2016). *Bibliometrics and research evaluation: Uses and abuses*. MIT Press.

Grieneisen, M. L., & Zhang, M. (2012). A comprehensive survey of retracted articles from the scholarly literature. *PloS one*, *7*(10). https://doi.org/10.1371/journal.pone.0044118

Hall, L. M., & Hendricks, A. E. (2019). Analysis of> 30,000 abstracts suggests higher false discovery rates for oncology journals, especially those with low impact factors. BioRxiv, 500660.

Hauptman, P. J., Armbrecht, E. S., Chibnall, J. T., Guild, C., Timm, J. P., & Rich, M. W. (2014). Errata in medical publications. The American journal of medicine, 127(8), 779-785. https://doi.org/10.1016/j.amjmed.2014.03.012

Huh, S., Kim, S. Y., & Cho, H. M. (2016). Characteristics of retractions from Korean medical journals in the KoreaMed database: A bibliometric analysis. PloS one, 11(10). https://doi.org/10.1371/journal.pone.0163588

Kennedy, D. (2000). Reflections on a Retraction. Science, 289(5482), 1137. https://doi.org/10.1126/science.289.5482.1137

Khadilkar, S. S. (2018). Scientific Misconduct: A Global Concern. *Journal of Obstetrics and Gynaecology of India*, *68*(5), 331-335. https://doi.org/10.1007/s13224-018-1175-8

Knoepfler, P. (2015). Reviewing post-publication peer review. Trends in Genetics, 31(5), 221-223. https://dx.doi.org/10.1016%2Fj.tig.2015.03.006

Martín-Martín, A., Orduna-Malea, E., Thelwall, M., & López-Cózar, E. D. (2018). Google Scholar, Web of Science, and Scopus: A systematic comparison of citations in 252 subject categories. *Journal of Informetrics*, *12*(4), 1160-1177. https://doi.org/10.1016/j.joi.2018.09.002

Meadows, A. J. (1979). The Scientific Journal. London : Aslib,

Moylan, E. C., & Kowalczuk, M. K. (2016). Why articles are retracted: a retrospective cross-sectional study of retraction notices at BioMed Central. BMJ open, 6(11), e012047. http://dx.doi.org/10.1136/bmjopen-2016-012047

Nath, S. B., Marcus, S. C., & Druss, B. G. (2006). Retractions in the research literature: misconduct or mistakes?. Medical Journal of Australia, 185(3), 152-154. http://dx.doi.org/10.5694/j.1326-5377.2006.tb00504.x

National Library of Medicine (2018). Errata, Retractions and, and Other Linked Citations in PubMed. Retrieved 8 May 2020, from: https://www.nlm.nih.gov/bsd/policy/errata.html

Noonan, B. M., & Parrish, D. (2008). Expressions of concern and their uses. Learned publishing, 21(3), 209-213. http://dx.doi.org/10.1087/095315108X288857

Park, J., Lee, J. Y., & Kwon, O. J. (2018). How the retracted publications are managed and used? A South Korean case. In 23rd International Conference on Science and Technology Indicators (STI 2018), September 12-14, 2018, Leiden, The Netherlands.

Pichardo-Corpus, J. A., Contreras-Nuño, G., & de la Peña, J. A. (2019). Proporción y distribución de erratas en publicaciones científicas. Investigación Bibliotecológica: archivonomía, bibliotecología e información, 33(80), 97-116. http://dx.doi.org/10.22201/iibi.24488321xe.2019.80.58000

Ribeiro, M. D., & Vasconcelos, S. M. R. (2018). Retractions covered by Retraction Watch in the 2013–2015 period: prevalence for the most productive countries. Scientometrics, 114(2), 719-https://doi.org/10.1007/s11192-017-2621-6

Rosenkrantz, A. B. (2016). Retracted publications within radiology journals. American Journal of Roentgenology, 206(2), 231-235. https://doi.org/10.2214/ajr.15.15163

Shema, H., Hahn, O., Mazarakis, A., & Peters, I. (2019). Retractions from altmetric and bibliometric perspectives, *Information - Wissenschaft & Praxis*, *70*(2-3), 98-110. doi: https://doi.org/10.1515/iwp-2019-2006

Scimago (2020). Scimago Journal & Country Rank. Retrieved 8 May 2020, from: https://www.scimagojr.com/countryrank.php

da Silva, J. A. T., & Dobránszki, J. (2017). Notices and policies for retractions, expressions of concern, errata and corrigenda: their importance, content, and context. *Science and engineering ethics*, *23*(2), 521-554. https://doi.org/10.1007/s11948-016-9769-y

da Silva, J. A. T. (2018). The Issue of Comment Ownership and Copyright at PubPeer. Journal of Educational Media & Library Sciences, 55(2), 181-191. http://dx.doi.org/10.6120/JoEMLS.201807_55(2).e001.BC.BE

Smith, R. (1997). Misconduct in research: editors respond. *BMJ: British Medical Journal*, *315*(7102), 201. https://dx.doi.org/10.1136%2Fbmj.315.7102.201

Steen, R. G. (2011). Retractions in the scientific literature: is the incidence of research fraud increasing?. Journal of medical ethics, 37(4), 249-253. http://dx.doi.org/10.1136/jme.2010.040923

Steen, R. G., Casadevall, A., & Fang, F. C. (2013). Why has the number of scientific retractions increased?. PloS one, 8(7), e68397. https://dx.doi.org/10.1371%2Fjournal.pone.0068397

Stretton, S., Bramich, N. J., Keys, J. R., Monk, J. A., Ely, J. A., Haley, C., & Woolley, K. L. (2012). Publication misconduct and plagiarism retractions: a systematic, retrospective study. Current medical research and opinion, 28(10), 1575-1583. https://doi.org/10.1185/03007995.2012.728131

Tijdink, J. K., Verbeke, R., & Smulders, Y. M. (2014). Publication pressure and scientific misconduct in medical scientists. *Journal of Empirical Research on Human Research Ethics*, *9*(5), 64-71. https://doi.org/10.1177/1556264614552421

Tripathi, M., Sonkar, S., & Kumar, S. (2019). A cross sectional study of retraction notices of scholarly journals of science. *DESIDOC Journal of Library & Information Technology*, *39*(2), 74-81. https://doi.org/10.14429/djlit.39.2.14000

Vaught, M., Jordan, D. C., & Bastian, H. (2017). Concern noted: a descriptive study of editorial expressions of concern in PubMed and PubMed Central. Research integrity and peer review, 2(1), 10. https://doi.org/10.1186/s41073-017-0030-2

Vuong, Q. H. (2020). The limitations of retraction notices and the heroic acts of authors who correct the scholarly record: An analysis of retractions of papers published from 1975 to 2019. Learned Publishing, 33(2), 119-130. https://doi.org/10.1002/leap.1282

Wager, E., & Williams, P. (2011). Why and how do journals retract articles? An analysis of Medline retractions 1988–2008. Journal of medical ethics, 37(9), 567-570. https://doi.org/10.1136/jme.2010.040964

Wang, T., Xing, Q. R., Wang, H., & Chen, W. (2019). Retracted publications in the biomedical literature from open access journals. Science and engineering ethics, 25(3), 855-868. https://doi.org/10.1007/s11948-018-0040-6