

Why and how do journals retract articles? An analysis of Medline retractions 1988–2008

Elizabeth Wager,¹ Peter Williams²

¹Sideview, Princes Risborough, UK

²Department of Information Studies, University College, London, UK

Correspondence to

Dr Elizabeth Wager, Sideview, 19 Station Road, Princes Risborough, HP27 9DE, UK; liz@sideview.demon.co.uk

Received 23 October 2010

Revised 18 February 2011

Accepted 10 March 2011

Published Online First

12 April 2011

ABSTRACT

Background Journal editors are responsible for what they publish and therefore have a duty to correct the record if published work is found to be unreliable. One method for such correction is retraction of an article. Anecdotal evidence suggested a lack of consistency in journal policies and practices regarding retraction. In order to develop guidelines, we reviewed retractions in Medline to discover how and why articles were retracted.

Methods We retrieved all available Medline retractions from 2005 to 2008 and a one-in-three random selection of those from 1988 to 2004. This yielded 312 retractions (from a total of 870). Details of the retraction including the reason for retraction were recorded by two investigators.

Results Medline retractions have increased sharply since 1980 and currently represent 0.02% of included articles. Retractions were issued by authors (63%), editors (21%), journals (6%), publishers (2%) and institutions (1%). Reasons for retraction included honest error or non-replicable findings (40%), research misconduct (28%), redundant publication (17%) and unstated/unclear (5%). Some of the stated reasons might have been addressed by corrections.

Conclusions Journals' retraction practices are not uniform. Some retractions fail to state the reason, and therefore fail to distinguish error from misconduct. We have used our findings to inform guidelines on retractions.

INTRODUCTION

Since peer-reviewed journals provide a permanent record of research findings, editors have a duty to ensure the correctness of the material they publish.¹ When published material turns out to be seriously flawed, it should be retracted so that future readers are not misled by it. The Committee on Publication Ethics (COPE) *Code of Conduct* for editors (published in 2003) states that 'Editors should be responsible for everything published in their journals' and therefore should 'always be willing to publish corrections, clarifications, retractions and apologies when needed'.¹

However, retracting a published article may not be straightforward and, when we began this study, few guidelines were available for editors to determine when a retraction (rather than a correction) was warranted. Retraction is one of the most serious sanctions journals can take against authors in cases of misconduct, and can cause permanent damage to reputations and academic careers. Therefore, retractions should be handled carefully and journals should have processes for deciding when and how to retract articles. However, the literature shows that journals do not respond consistently to findings of serious research misconduct affecting their publications.^{2 3}

Discussion of cases at COPE indicates that editors and publishers are sometimes reluctant to retract articles. Reasons for such reluctance may include beliefs that retractions may be made only by the authors; author disputes in which some authors request retraction while others oppose it; or concerns about, or actual threats of, litigation from authors.

In order to inform the development of the COPE retraction guidelines,⁴ we examined retractions in Medline to understand trends in retraction behaviour over time and to record how and why articles are retracted. We also invited a sample of editors who had retracted articles recently to take part in qualitative research to learn more about their experience of retracting an article (this part of the project will be published separately).

Two studies of Medline retractions had been published before ours.^{5–7} Budd *et al* studied retractions published between 1966 and 1997, categorising them into misconduct or error⁵ and also examining citation patterns of retracted articles.⁶ Nath *et al* studied retractions from 1982 to 2002 and classified the reasons as mistakes or misconduct (defined as fabrication, falsification or plagiarism).⁷ To develop specific guidance for journal editors, we felt that an up-to-date survey using a more detailed classification of reasons for retraction (eg, distinguishing the type and extent of plagiarism) and who issued the retraction (eg, editor, author, publisher) was needed.

METHODS

Assessing the number of retractions over time

We searched Medline for all retractions since 1980 (when retractions were first listed as a specific publication category). We noted the number of retractions and the total number of articles for each year from 1980 to 2009.

Assessing characteristics of retractions

We searched Medline for retraction notices published between April 1988 and March 2008, where the retraction text was available in English. (Retractions are identified in Medline as a specific category and we used this tag for searching.) We retrieved the retraction notice for all retractions published between April 2005 and March 2008 which were accessible via the University College London library, and for one in three of the older (1988–2004) retractions (which were randomly selected by sorting the Excel spreadsheet containing the references in a random order then going down the list of those available until we had retrieved one-third of the references).

For each retraction, we recorded the article topic (eg, clinical medicine, chemistry, biology), article

type (eg, original research, review article, letter), number of authors, who issued the retraction (eg, authors, editor, publisher) and the reason for the retraction (eg, plagiarism, redundant publication, honest error). We used a detailed system for classifying reasons for retraction to identify situations where editors might have concerns about issuing a retraction (see table 1 for details of the categories used). For example, we considered that editors might feel more confident about retracting a completely plagiarised copy of another paper than an article containing only some plagiarised fragments. Similarly, we wanted to distinguish between complete duplicate publication of an entire article by the original authors from overlapping publications that contain largely the same data but with some new elements (since we knew from cases presented to COPE that these can be controversial). We therefore developed our own classification system for the reasons for retraction using an iterative process.

At the start of the project, the two investigators worked together and jointly agreed the data extraction and classification system to ensure inter-investigator agreement. Retraction statements were reviewed by both investigators and the reasons for retraction were independently coded. Coding discrepancies were resolved by discussion and, where necessary, further analysis of the retraction text. Other publication details (such as number of authors) were extracted by one or other of the two investigators.

RESULTS

Number of retractions over time

The number of retractions listed in Medline has increased considerably since 1980. During the 1980s, only seven retractions were listed each year on average (range 2–12). This rose to 25/year in the 1990s (range 15–47). The annual number of retractions in Medline increased steadily between 2000 and 2005

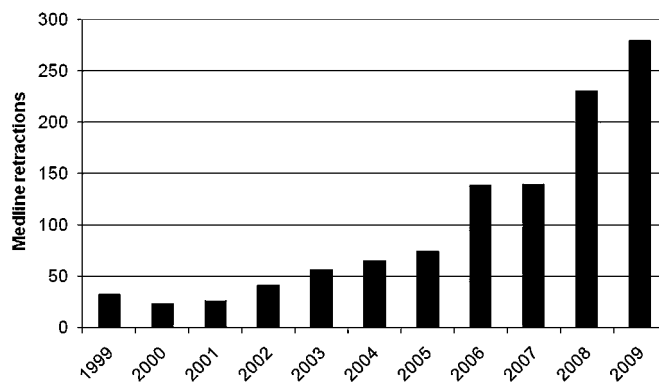


Figure 1 Retractions listed in Medline 1999–2009.

from 23 to 74, then from 138 in 2006 to 279 in 2009 (figure 1). While the number of articles listed in Medline each year has also increased during this time (from around 300 000/year in the 1980s to over 800 000/year in 2009) the proportion of retractions has increased tenfold from 0.002% in the early 1980s to 0.02% in 2005–2009 (figure 2).

Analysis of retractions

Medline lists 870 retraction notices, published in English, in the 20 years from 1988 to 2008. We analysed 312 of these (176 from 2005 to 2008 and 136 from 1988 to 2004). The retractions came from 187 different journals. Most journals only had one retraction in the sample but the following journals were notable for having retracted at least five articles by different authors (*Contraception*, 6; *Journal of Biological Chemistry*, 5; *Nature*, 12; *NEJM*, 6; *PNAS*, 11; *Science*, 18).

The most common reasons for retraction were honest research errors (28%), redundant publication (17%) and plagiarism (16%) (table 1). Overall, the retractions could be categorised as honest error or non-replicable findings (40%), research misconduct (28%), redundant publication (17%) and other reasons (15%).

In our sample, 63% (197) of the retractions were issued by one or all of the authors. In most of these cases (152 or 77% of the sub-group), all authors signed the retraction, but 25 (13% of the sub-group) were signed by only one author (of several) and 20 (10% of the sub-group) by some (but not all) of the authors. Several of the retraction statements specifically noted that at least one author had refused to sign the retraction. In two cases, a retraction was issued by the authors' institution following a finding of misconduct.

When retractions were not issued by the authors, the wording used to identify the source of the retraction was variable. We found 65 retractions by editors (21% of the total analysed), 20 by journals (6%), six by publishers (2%) and some by combinations of these (eg, editor and publisher). In 22 cases (7% of the total) it was unclear who had retracted the article.

Most retractions (270, 87%) related to primary papers reporting research findings but we also found retracted literature reviews (17), case reports (7), letters (7), abstracts (6) and other types of publication (5) such as news items relating to retracted publications and editorials. Most of the retracted reviews, case reports and abstracts were retracted because of plagiarism (12/30=40%) or redundancy (6/30=20%). The retracted publications covered basic biomedical research (181, 58%), clinical medicine (71, 23%) and other subjects (60, 19%) reflecting the composition of the Medline database.

The retracted articles were published by a single author (9%), two (14%), three (17%), four (15%), five (11%) or more than five

Table 1 Classification of retractions (N=312)

Code	Reason	Sub-category	N	%
1	Honest error	(a) based on incorrect/fraudulent* data	10	28
		(b) research error (eg, wrong cell line)	44	
		(c) calculation error/inaccurate data	33	
2	Could not replicate findings	(a) same authors	31	11
		(b) different authors	4	
3	Journal (administrative) error		3	1
4	Redundant publication	(a) complete (ie, duplicate/triplicate publication)	34	17
		(b) partial (ie, overlapping publications)	20	
5	Plagiarism	Amount unspecified	6	16
		(a) total (ie, whole paper copied)	4	
		(b) partial—data only	7	
		(c) partial—text only	18	
		(d) partial—data and text	14	
6	Data used without permission/author dispute		17	5
7	Fabricated data		16	5
8	Falsified data		13	4
9	Ethical problems with research		4	1
10	Misconduct (unspecified)		10	2
11	Inaccurate/misleading reporting		12	4
12	No reason/unclear		17	5

*For example, calculations based on published work by another group that later turned out to be fraudulent, therefore the inclusion of the fraudulent data was an honest error on the part of the second group of researchers.

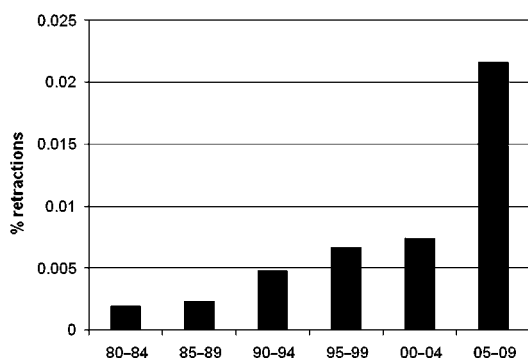


Figure 2 Retractions expressed as a percentage of the total number of publications listed in Medline 1980–2009.

authors (34%). Articles with single authors included a higher proportion retracted because of disputed authorship (5/29=17%) and fewer due to research errors (four, 14%) than the total sample (which had 5% and 28% respectively), but rates of plagiarism (17%) and redundancy (24%) were similar. However, as this subgroup was small (N=29) these findings should not be over-interpreted. We did not identify any clear ‘warning signs’ of articles more likely to be retracted, such as single-authored clinical trials.

In 17 cases (5%) no reason for retraction was stated, or the language was so unclear that the reason could not be determined. Some retraction statements appeared to use deliberately ambiguous wording which made it difficult to distinguish honest errors from suspected (or proven) misconduct. Examples are given in box 1. We included a category of ‘inaccurate/misleading reporting’ since these terms were sometimes used, but in some cases we suspected that such phrases were used as euphemisms for data falsification. For example, one retraction notice stated ‘Following further work at [X] University it has been discovered that the paper contains several examples of

Box 1 Examples of ambiguous wording used in retraction statements

‘important irregularities’

‘the authors ... withdraw the paper because of some inaccurate data and other irregularities’

the authors ‘no longer stand by their results’

‘incorrect data were found to have been included on the study Case Report Forms’

This figure is withdrawn ‘due to lack of supporting data’

‘Retraction... is being done for legal reasons based on the advice of counsel’

‘the Review contained numerous errors in the text and references that were not discovered until after publication, although neither novel ideas nor data were misappropriated’

‘This article has been removed at the request of the authors in order to eliminate incorrect information’

‘unknown to the authors, at the time of publication some of the data on which the article was based were not in the public domain’

misrepresentation of scientific data’. Another stated that one author ‘has acknowledged that she misreported data supporting an exploratory... analysis’ while another stated that there were ‘errors in the presentation of data in some of the figures’. Given the small numbers, we cannot draw firm conclusions, but our sample showed no trend in the clarity of retractions over time. Six of the retractions that did not clearly state the reason were published in 2006 (in five different journals).

At least two of the stated reasons for retraction were issues that might have been addressed by corrections. For example, one paper was apparently withdrawn because of the improper use of company trademarks. In another case, a partial retraction was issued to correct a statement about the reason for discontinuation of a company product.

DISCUSSION

Analysis of the retractions on Medline suggests that journals’ retraction practices are not uniform. Detailed analysis of the retraction texts also revealed that reasons for retraction are not always clearly stated. This is unfortunate, since retractions may be triggered both by genuine mistakes and by misconduct. It is therefore important to indicate the reason for the retraction so that authors who have acted responsibly in alerting the journal to problems with their work are not stigmatised along with those who have committed misconduct. This is emphasised in the COPE retraction guidelines⁴ (published in 2009 after this study was completed) and has been recommended by the International Committee of Medical Journal Editors since at least 1997. However, we found several, quite recent, examples of retraction notices that failed to state the reason for retraction, including six published in 2006 alone.

Cases presented to COPE indicate that some editors believe retractions can only be made by the original authors. This can cause problems if some or all of the authors do not agree or cooperate with a retraction. Our survey clearly shows that retractions can be issued by journal editors, with or without the cooperation of the authors. In two instances, we found retractions that had been issued by the authors’ institution following inquiries into research misconduct. These provide an excellent example of institutions taking their responsibility for research integrity seriously and we suggest this practice should be encouraged.

Two previous studies on this topic have been published. Budd *et al* analysed 235 Medline retractions issued between 1966 and 1997.⁵ They found 81% of retractions were issued by authors (compared with 63% in our sample). They also found that 9% of retraction notices did not state the reason for retraction clearly (while we found 5%). They also examined whether retracted articles were cited after retraction and found 2034 citations to the 235 articles in their sample. Further analysis of these citations found that most were positive (ie, not criticising the retracted paper or mentioning the fact that it had been retracted).⁶

Nath *et al* examined retractions listed in Medline between 1982 and 2002 and categorised the reasons for retraction as either misconduct or unintentional errors.⁷ They found 27% were retracted because of misconduct and 62% because of errors. Given the similarity of sample base for this study and ours, the similarity of findings is reassuring (we found 28% of retractions were due to misconduct) but our categorisation provided more detail than previous studies.

Anecdotal evidence from COPE Forum meetings, and analysis of some high-profile cases, suggests that journals are not consistent in their retraction practices.^{2,3} Liu has noted that the

highest impact journals are also 'world leaders in publishing retractions' calculating that *Science*, *Nature*, *PNAS* and *Cell* contributed 9% of all Medline retractions in the early 2000s.⁸ Liu suggests that this may be because these journal publish controversial research, but it may also be because these well-staffed journals are more diligent in pursuing or responding to misconduct allegations and errors than less well-resourced journals. The qualitative arm of our study (to be published separately) aimed to address reasons that might prevent editors from issuing retractions.

Inconsistency in retracting articles may be due to lack of external guidance (the COPE guidelines were only published in late 2009) or lack of a clear internal (ie, journal or publisher) policies. Atlas reviewed the instructions to authors of 122 'major biomedical journals' in 2003 and found that only 18% reported having a policy on retractions.⁹ However, editors' uncertainty about how and when to retract articles may also reflect the fact that the research community does not have agreed norms about responses to research misconduct. Korenman *et al* surveyed 606 scientists receiving US National Science Foundation grants and 91 representatives of research institutions with responsibility for the ethical conduct of research (listed by the US Office of Research Integrity) about whether certain acts were unethical and, if so, the appropriate response and punishment for the act.¹⁰ They found that the scientists were more likely to suggest retraction as an appropriate response than the institutional representatives. They suggest that this may be because the institutional representatives believe 'their main role is to protect, preserve and enhance the institution' so they favour 'sanctions aimed at punishment and deterrence' while 'damage to the social fabric of science may be a lesser consideration'. Whereas, for scientists, Korenman *et al* suggest, the 'integrity of the scientific community is essential'.

Another reason why editors may be uncertain what to do is that they may be faced with inconclusive evidence.¹¹ Parrish noted that 'substantial delays are endemic in notifying the journal... about allegations and findings of scientific misconduct'.³ In a later study, she examined the use of 'expressions of concern' and concluded that these may be useful 'for ensuring the integrity of the research record during the sometimes lengthy period that research misconduct investigations can take'.¹² The COPE retraction guidelines recommend that editors should consider issuing an expression of concern if they receive inconclusive evidence of misconduct, the authors' institution refuses to investigate the case, they believe an investigation will not be impartial, or a judgement will not be available for some time.⁴ Medline does not tag expressions of concern so we were unable to include them in this study.

This study was used to inform the development of the COPE retraction guidelines, in addition to consideration of cases presented at COPE Forum meetings.⁴ In our study, for example, although we did not record the proportion of retraction notices that were freely available to all readers (ie, not only to journal subscribers), we were surprised how hard some retraction

notices were to find. The COPE guidelines therefore recommend that retraction notices should 'be freely available to all readers (ie, not behind access barriers or available only to subscribers)'. Cases presented at COPE Forum meetings suggested that some editors were not clear about the primary function of retractions, so the guidelines state that 'The main purpose of retractions is to correct the literature and ensure its integrity rather than to punish authors who misbehave'. The COPE guidelines also emphasise the fact that retractions may be needed in cases of misconduct or of genuine error and therefore stress the importance of clearly stating the reason for each retraction.

We conclude that journals indexed in Medline display inconsistency in their practices regarding retractions but there has been a marked increase in the number of articles being retracted. Despite published guidance, some journals continue to retract articles without clearly stating the reason for the retraction: this is regrettable. While we recognise that variations in journals' practices may reflect genuine differences in belief about the correct approach, we suggest that a more uniform approach is desirable to safeguard the reliability of the published research record and to ensure that authors are not unfairly stigmatised by unwarranted retractions. We therefore recommend that editors follow the COPE retraction guidelines which were based on the experience of COPE members and the findings of this study.⁴

Funding This study was funded by a research grant from the Committee On Publication Ethics (COPE).

Competing interests EW developed the COPE guidelines on retractions and is currently Chair of COPE; PW has no competing interests.

Contributors EW designed the study, collected and analysed data, and wrote the first draft of the paper. PW contributed to the classification system, collected data, carried out the qualitative arm of the study (reported separately), and contributed to the paper.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. **Committee on Publication Ethics**. COPE code of conduct. <http://publicationethics.org/code-conduct> (accessed 22 Sep 2010).
2. **Sox HC**, Rennie D. Research misconduct, retraction, and cleansing the medical literature. Lessons from the Poehlman case. *Ann Intern Med* 2006;**144**:609–13.
3. **Parrish DM**. Scientific misconduct and correcting the scientific literature. *Acad Med* 1990;**74**:221–30.
4. **Wager E**, Barbour V, Yentis S, *et al*. Retractions: guidance from the Committee on Publication Ethics (COPE), 2009. http://publicationethics.org/files/u661/Retractions_COPE_gline_final_3_Sept_09_2_.pdf (accessed 22 Sep 2010).
5. **Budd JM**, Sievert ME, Schultz TR. Phenomena of retraction. Reasons for retraction and citations to the publications. *JAMA* 1998;**280**:296–7.
6. **Budd JM**, Sievert ME, Schultz TR, *et al*. Effects of article retraction on citation and practice in medicine. *Bull Med Libr Assoc* 1999;**87**:437–43.
7. **Nath SB**, Marcus SC, Druss BG. Retractions in the research literature: misconduct or mistakes? *Med J Aust* 2006;**185**:152–4.
8. **Liu SV**. Top journals' top retraction rates. *Sci Ethics* 2006;**1**:91–3.
9. **Atlas MC**. Retraction policies of high-impact biomedical journals. *J Med Libr Assoc* 2004;**92**:242–50.
10. **Korenman SG**, Berk R, Wenger NS, *et al*. Evaluation of the research norms of scientists and administrators responsible for academic research integrity. *JAMA* 1998;**279**:41–7.
11. **Horton R**. Revising the research record. *Lancet* 1995;**346**:1610–11.
12. **Noonan BM**, Parrish D. Expressions of concern and their uses. *Learn Publ* 2008;**21**:209–13.