Why articles are retracted: a retrospective cross-sectional study of retraction notices at BioMed Central

Elizabeth C Moylan, Maria K Kowalczuk

ABSTRACT
Objectives: To assess why articles are retracted from BioMed Central journals, whether retraction notices adhered to the Committee on Publication Ethics (COPE) guidelines, and are becoming more frequent as a proportion of published articles.

Design/setting: Retrospective cross-sectional analysis of 134 retractions from January 2000 to December 2015.

Results: 134 retraction notices were published during this timeframe. Although they account for 0.07% of all articles published (190,514 excluding supplements, corrections, rejections and commissioned content), the rate of retraction was rising. COPE guidelines on retraction were adhered to in that an explicit reason for each retraction was given. However, some notices did not document who retracted the article (eight articles, 6%) and others were unclear whether the underlying cause was honest error or misconduct (15 articles, 11%). The largest proportion of notices was issued by the authors (47 articles, 35%). The majority of rejections were due to some form of misconduct (102 articles, 76%), that is, compromised peer review (44 articles, 33%), plagiarism (22 articles, 16%) and data falsification/fabrication (10 articles, 7%). Honest error accounted for 17 rejections (13%) of which 10 articles (7%) were published in error. The median number of days from publication to retraction was 337.5 days.

Conclusions: The most common reason to retract was compromised peer review. However, the majority of these cases date to March 2015 and appear to be the result of a systematic attempt to manipulate peer review across several publishers. Retractions due to plagiarism account for the second largest category and may be reduced by screening manuscripts before publication although this is not guaranteed. Retractions due to problems with the data may be reduced by appropriate data sharing and deposition before publication. Adopting a checklist (linked to COPE guidelines) and templates for various classes of retraction notices would increase transparency of retraction notices in future.

Strengths and limitations of this study
- The first study to examine all BioMed Central retraction notices published in 2000–2015 and the retraction patterns of a single publisher.
- The first study to examine transparency of retraction notices and adherence to COPE retraction guidelines by a single publisher.
- The study is limited by the number of retraction notices published to analyse and any correlations with a particular journal, article type, discipline or peer review model not been explored.

INTRODUCTION
Retractions are a permanent means of maintaining the integrity of the scientific literature and necessary to alert the reader to serious problems identified with a published article. The Committee on Publication Ethics (COPE) published retraction guidelines in 2009.1 These guidelines advise on retracting articles if the main findings are found to be unreliable (either as a result of misconduct or honest error), redundant (ie, previously published elsewhere in a citable format), plagiarised (text or figures) or if the authors have reported unethical research or failed to disclose a major competing interest, which could influence the interpretation of the article.

COPE recommends that retraction notices provide adequate information so that readers know who is retracting the article and why the findings are considered unreliable, while clearly distinguishing forms of misconduct from honest error. However, retraction notices often need to strike a balance between providing adequate information without being defamatory or libellous.2 In addition, retraction should be clearly identifiable; freely available; published promptly and be linked to the original article that is retracted (which should also be identified as a retraction).

Over the past few years, there have been reports that most cases of retraction are attributable to misconduct,3 with a notable rise in cases of fraud (data fabrication or
falsification). More recently, there have been retractions from several journals across different publishers, including BioMed Central, due to systematic manipulation of the peer review processes by the provision of fabricated contact details for peer reviewers.5–8 There have also been calls for journals to be more transparent regarding their retraction procedures and explicit in their retraction notices,9–12 especially as retraction notices have been found to vary between, and within, journals.13–15

Given this, we analysed all retraction notices published at BioMed Central between January 2000 and December 2015 to determine how transparent notices were in terms of reason for retraction and information provided, and if they complied with the COPE guidelines. We also wanted to determine if retractions were on the increase.

METHODS

All retraction notices published between January 2000 and December 2015 were identified using the publisher’s publicly available advanced search function16 using the search term ‘retraction’ within the article title. This time frame was selected because it spanned the largest window of retractions available to record—from the first retractions BioMed Central had ever published through to 2015, that is, 15 complete years to date. Articles which had ‘retraction’ in the title, but were not themselves retractions were excluded. Retractions were excluded if they were published by other publishers before the journal was transferred to BioMed Central.

Who issued the retraction notice and the reason for retraction were recorded. The time elapsed between publication of the original article and publication of the retraction notice was also recorded. After considering classifications of retractions in previous studies,3–12 retractions were broadly classified according to the apparent underlying reason for the retraction into the following broad categories:

► Honest error: defined as mistakes on the part of the author or publisher leading to publishing in error or unreliable data.
► Misconduct: defined as data falsification/fabrication, failure to obtain ethical approval or consent, failure to obtain permission for data, plagiarism, duplicate publication, image duplication, authorship issues, compromised peer review.
► Unclear: where it was not possible to distinguish ‘honest error’ from ‘misconduct’.

Instances of data falsification or data fabrication were classified together as one category ‘data falsification/fabrication’. Where a retraction notice mentioned irregularities in the data and an institutional investigation, the notice was described as misconduct unless honest error was explicitly mentioned.

All notices were classified by one author (ECM) and checked for agreement by the other author (MKK) using the information given in the retraction notice alone (ie, no additional information was used). Where there was a disagreement in opinion, a discussion took place between the authors to reach a consensus. Where multiple reasons for the retraction were given, the main reason was described and the secondary reasons were noted. The descriptions of the retraction notices are given in online supplementary file 1 (Raw data for BMJ Open 2016) and the explanations (for the purposes of this study) are as follows:

► Plagiarism: duplication of text from previously published articles;
► Compromised peer review: compromises in the independent assessment of the manuscript by a peer;
► Data unreliable: data has errors;
► Data falsification/fabrication: data has been manipulated or made up;
► Published in error: article was accidently published twice as a result of publisher error;
► Duplicate publication: article was published twice (usually as a result of author misconduct);
► Image duplication: duplication of images from previously published articles;
► Authors unaware of manuscript submission: not all authors aware;
► No ethical approval: the study had no ethical approval;
► No consent: the study involved people who had not given consent;
► No permission for data: authors did not have permission to use the data reported;
► Undeclared conflict of interest: authors or reviewers did not declare a conflict of interest;
► Breach of editorial policy: the manuscript breached an editorial policy.

Citations for all retracted articles were counted before and after the date of retraction by searching for the article or authors in Scopus18 accessed on 26 February 2016. Citations to the retraction notice were also counted. Citation data are provided in online supplementary file 1. For further clarity, a checklist of the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) recommendations19 for the reporting of observational studies has been completed and is provided in online supplementary file 2.

RESULTS

Between January 2000 and December 2015, our search identified 134 retraction notices. This accounts for 0.07% of all articles published (a total of 190,514 articles excluding supplements, corrections, retractions and commissioned content).

All retraction notices were clearly labelled and linked to the retracted article except for cases where for legal reasons the original article could no longer be made available (eg, if there was sensitive information or if plagiarism infringed another journal’s copyright). Four retraction notices were not included because they were published by other publishers before the journal was transferred to BioMed Central.
Figure 1 shows the growth in rejections compared with growth in total articles published (excluding supplements, corrections, rejections and commissioned content). Proportionally, there was no increase in rejection rate until 2015 when the rejection rate increased due to 43 articles that were retracted because of compromises in peer review (see online supplementary file 1). The median time between publication of the article and its retraction was 337.5 days (with a minimum of 11 days and a maximum of 4147 days). Articles involving apparent misconduct took longer to retract (median of 386 days; minimum 17 days, maximum 4147 days) than honest error (median of 184 days; minimum 11 days, maximum 3361 days). The largest proportion of retraction notices were issued by the authors (47 articles, 35%), followed by the publisher and editor(s) jointly (43 articles, 32%) and editor alone (28 articles, 21%) as shown in table 1.

No cases were recorded where the authors’ institution issued a retraction. While the majority of retraction notices declared who was retracting the article, eight retraction notices (6%) did not explicitly state this information.

A detailed breakdown of the reasons for retraction is given in table 2.

The most common reason for retraction is compromises in peer review (44 articles, 33%), followed by plagiarism (22 articles, 16%) followed by problems with the data—that is, the data was found to be ‘unreliable’ (13 articles, 10%). Other reasons include lack of appropriate ethical approvals (5 articles, 4%) or permission to use data (5 articles, 4%), duplicate publication (11 articles, 8%), published in error (8 articles, 6%) where an article was accidently published twice,20–22 image duplication (6 articles, 4%) or because of a lack of awareness by some authors of the manuscript’s submission and publication (5 articles, 4%). A total of 10 rejections (7%) were due to data falsification/fabrication—reasons that were not seen in retraction notices before 2012. Three articles (2%) were retracted due to undeclared conflicts of interest either by the author (for example23) or by the reviewer (for example24). Not all rejections occurred for a single reason. In 13 cases of retraction there were two reasons (for example25,26) and in one case three reasons were given.27 If the individual reasons for retraction are categorised into the underlying reasons of honest error, misconduct or unclear (see table 2), then most rejections originated due to some form of misconduct.

Figure 2 shows the growth and variation in reasons for rejections year on year. Plagiarism occurred as a reason for retraction from 2010 onwards. Retractions due to compromised peer review process were not seen before 2014.

Analysis of citations to articles before and after retraction in Scopus revealed that of 128 articles listed (for which data were available), articles were cited in the range 0–830 times before retraction and 0–30 times after retraction. The distribution of values is highly skewed, but the median number of citations is higher after retraction (3) than before retraction (1).

---

Table 1 Number of rejections listed by who issued the retraction notice

<table>
<thead>
<tr>
<th>Who retracted?</th>
<th>Number of rejections</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authors</td>
<td>47</td>
</tr>
<tr>
<td>Authors and editor</td>
<td>1</td>
</tr>
<tr>
<td>Editor</td>
<td>28</td>
</tr>
<tr>
<td>Journal</td>
<td>1</td>
</tr>
<tr>
<td>Publisher</td>
<td>6</td>
</tr>
<tr>
<td>Publisher and editor</td>
<td>43</td>
</tr>
<tr>
<td>Information not provided</td>
<td>8</td>
</tr>
<tr>
<td>Grand Total</td>
<td>134</td>
</tr>
</tbody>
</table>

Figure 1 Growth in rejections compared with growth in total articles published (excluding supplements, corrections, rejections and commissioned content).
DISCUSSION

General observations

The median number of days from publication to retraction was 337.5. Articles involving apparent misconduct took longer to retract (median of 386 days) than honest error (median of 184 days) as has been previously reported by others. However, these times will all be overestimates of the actual time to retract because the issues leading to retraction are flagged after publication. For example, in one recent case, an article was retracted 11 years after publication due to a recent breach of editorial policy. The actual time taken to retract was not itself 11 years, but shortly after the issue was raised.

Analysis of citations to articles before and after retraction revealed that articles continue to be cited after retraction as noted previously by others. It is interesting that retracted articles continue to be cited much more than the retraction notices themselves, which are rarely if ever cited (even though clearly linked to the original article). The fact that retraction notices are so seldom cited suggests that readers are unaware of the article’s retraction.

Transparency of retraction notices

All BioMed Central journals have an overarching retraction policy to retract articles where necessary so as to

<table>
<thead>
<tr>
<th>Reason for retraction</th>
<th>Honest error</th>
<th>Misconduct</th>
<th>Unclear</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breach of editorial policy</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Co-authors unaware of manuscript submission</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Compromised peer review</td>
<td>44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data falsification/fabrication</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data unreliable</td>
<td>6</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Duplicate publication</td>
<td></td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>Image duplication</td>
<td>1</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>No consent</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>No ethical approval</td>
<td></td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>No permission for data</td>
<td></td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Plagiarism</td>
<td></td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>Published in error</td>
<td>10</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Undeclared conflict of interest</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sub-totals per broad category</td>
<td>17</td>
<td>102</td>
<td>15</td>
</tr>
</tbody>
</table>

Figure 2 Growth in retractions showing variation in reasons for retracting articles from 2000 to 2015.
maintain the integrity of the published literature. Retraction notices published during the time frame of this study were identifiable as retractions, linked to the retracted article, identified the retracted article in the heading and explained the reason for the retraction in accordance with COPE guidelines. While it was possible to classify retractions into discrete categories, in 15 notices (11%) it was not possible to distinguish the underlying issue, honest error or misconduct, which ultimately led to retraction. This may have been due to legal constraints or limited information available from institutions for editors to make the distinction between honest error and misconduct or perhaps simply due to oversight of the person writing the notice. In other cases, retraction notices were ambiguous. For example, the stated reason for one retraction was ‘published in error’ although the notice alludes to other problems with the data which likely also contributed to the retraction. Other articles were ‘published in error’ when a journal was transferred from another publisher and during this period an article was inadvertently published twice.

Eight notices (6%) did not state clearly who was retracting the article. In such cases, the retraction notices invariably simply stated ‘This article is retracted due to’ and it seems possibly the retraction could have come from the authors but was not explicit. These cases all occurred after the publication of the COPE guidelines on retraction and so the guidelines were not adhered to in this respect. In one case, the retraction notice came from the journal, this was likely an oversight and potentially the retraction could have come from the editor. While most retractions come from an individual, that is, the author or the editor, in some cases authors and editors or editors and publishers had shared responsibility for the retractions.

In five cases (4%), retractions occurred because not all authors had been aware of the manuscript submission. Retractions due to authorship disputes are not recommended by COPE because if the scientific integrity of the article is not affected it should be possible to resolve the issue by other means (eg, by publishing a correction). However, authorship disputes can sometimes be symptomatic of other more serious underlying issues such as data theft. Retractions solely due to a lack of awareness or agreement on behalf of all authors have not occurred at BioMed Central since 2009 when the COPE guidelines were published. It is possible that this is because straightforward authorship issues can be addressed by the publication of a correction and complicated disputes are eventually retracted for different reasons.

In summary, it is clear that COPE guidelines on retraction were not adhered to in all respects. Others have also raised this issue and called for the role of publishers in the process to be clarified. In order to further improve the transparency of retraction notices, publishers could enforce the use of an internal checklist capturing the main information required in a retraction notice along with template wording as previously proposed.

Reasons for retraction

The majority of retractions were a result of misconduct, as found in other larger studies (3, 13, 28). However, definitions of misconduct vary and while many agree that fabrication, falsification and plagiarism are serious forms of misconduct, we ultimately adopted a broader definition here and included other forms of misconduct (eg, manipulations to the peer review system which resulted in compromised peer review)—see table 2. In the absence of clear definitions for what constitutes misconduct, others acknowledge there is a spectrum of lessor forms of misconduct, which comprise ‘questionable research practices’.

For clarity in this study, we focus on the individual reasons for retraction.

Compromised peer review did not occur as a reason to retract at BioMed Central prior to 2014 (figure 2). However, the majority of cases reported here date to a set of retractions in March 2015 related to attempts to positively influence the outcome of peer review process of several journals by authors or third-party agencies suggesting fabricated reviewers.

Plagiarism was found to be the second main reason for retraction (table 2) and has also been a predominant reason for retraction highlighted in other studies. The rise in software to detect plagiarism (alongside development of sophisticated approaches to check figure manipulation) has gone hand-in-hand with a rise in retractions due to plagiarism in recent years. While the use of plagiarism detection software before publication may prevent the occurrence of retractions due to plagiarism in future, we have seen cases where authors disguise the plagiarism, for example, by substituting different key words to evade detection. Often it is the order of identical references within a section of text, rather than the exact words used that reveals that plagiarism has occurred. Also peer reviewers frequently detect ‘disguised plagiarism’ more accurately than software programmes given their familiarity with previously published literature.

The third main reason for retraction was that the published data has subsequently been found to be unreliable in some way. A total of 15 articles (10%) of retractions were due to problems with the data. Often these issues occurred through honest error in how the data were handled, for example, although in some cases it is difficult to determine whether honest error or misconduct was the cause. Recent initiatives towards increased transparency and reproducible research through encouraging sharing and deposition of data prior to publication could have an impact on reducing instances of retraction due to errors with the data in future. In preparing data to be ‘publication ready’, many issues may be caught and fixed before publication.
In some cases, notices related to misconduct were transparent,\(^37\) in other cases less so. Several retractions were due to lack of appropriate ethical approvals (five articles, 4%) or permission to use data (five articles, 4%). It is difficult to pinpoint measures that editors or reviewers can take to detect fraudulent and unethical practices before publication or even prevent them happening at all.\(^30\) However, having policies in place to encourage explicit author contributions, declaration of conflicts of interests (for authors and reviewers), data sharing, adherence to reporting guidelines and ensuring the correct ethical approvals and permissions to publish data were obtained are vital. Most recently, the BMJ introduced a ‘transparency declaration’, requiring the lead author to confirm that the manuscript is an honest, accurate and transparent account of the study being reported.\(^39\) It will be interesting to see if this has any effect on reducing retractions in future. While the decision to act unethically rests with the researcher,\(^40\) the tremendous pressures that continue to be placed on researchers to ‘publish or perish’\(^41\) \(^42\) may unintentionally fuel acts of misconduct.\(^9\) \(^43\) Clearly, there is a real need for integrity and transparency at all levels, from those in research (researchers and their institutions) to those making editorial recommendations (peer reviewers and editors) as previously suggested.\(^40\)

**CONCLUSIONS**

We found that COPE guidelines on retraction were adhered to in that an explicit reason for retraction was given in all cases of retraction evaluated from 2000–2015. Retractions were also readily identifiable, linked to the retracted article and identified the retracted article in the heading. However, in some cases notices did not document who issued the notice and there were ambiguities as to the underlying cause (honest error or misconduct). In future, we agree with others that adopting a checklist (linking to COPE guidelines) and a standard template formats for various classes of retraction notices would facilitate increased transparency and consistency of retraction notices. There may also be a need for making the retraction notice more obvious on the original article\(^31\) given that the retracted articles are always more highly cited than the retraction notice, post retraction.

In general, across the publishing industry, editors are adopting procedures and policies which may help to reduce certain classes of retraction in future. For example, many journals now screen for plagiarism and image manipulation and so we would predict a fall in retraction due to these issues in coming years. By encouraging data sharing and data deposition prior to publication, authors collate their data to make it ‘publication ready’ and this exercise in itself can help resolve honest errors. However, robust publication ethics does not fall solely to editors. Publication ethics is inclusive—authors, peer reviewers, editors, publishers and institutions all have their part to play to foster a culture of trust and transparency and maintain the integrity of the published literature.

**REFERENCES**


26. Ngemu EK, Khayeka


12. Bilbrey E, O’Moylan EC, Kowalczuk MK.


