Retractions in the scientific literature: is the incidence of research fraud increasing?

R Grant Steen

Correspondence to
R Grant Steen, Medical Communications Consultants, LLC 103 Van Doren Place, Chapel Hill, NC 27517, USA; g_steen_medicc@yahoo.com

Received 18 October 2010
Accepted 5 November 2010
Published Online First 24 December 2010

ABSTRACT

Background Scientific papers are retracted for many reasons including fraud (data fabrication or falsification) or error (plagiarism, scientific mistake, ethical problems). Growing attention to fraud in the lay press suggests that the incidence of fraud is increasing.

Methods The reasons for retracting 742 English language research papers retracted from the PubMed database between 2000 and 2010 were evaluated. Reasons for retraction were initially dichotomised as fraud or error and then analysed to determine specific reasons for retraction.

Results Error was more common than fraud (73.5% of papers were retracted for error (or an undisclosed reason) vs 26.6% retracted for fraud). Eight reasons for retraction were identified; the most common reason was scientific mistake in 234 papers (31.5%), but 134 papers (18.1%) were retracted for ambiguous reasons. Fabrication (including data plagiarism) was more common than text plagiarism. Total papers retracted per year have increased sharply over the decade ($r=0.96$; $p<0.001$), as have retractions specifically for fraud ($r=0.89$; $p<0.001$). Journals now reach farther back in time to retract, both for fraud ($r=0.87$; $p<0.001$) and for scientific mistakes ($r=0.95$; $p<0.001$). Journals often fail to alert the naïve reader; 31.8% of retracted papers were not noted as retracted in any way.

Conclusions Levels of misconduct appear to be higher than in the past. This may reflect either a real increase in the incidence of fraud or a greater effort on the part of journals to police the literature. However, research bias is rarely cited as a reason for retraction.

INTRODUCTION

Accusations that research is tainted by bias have become commonplace in the news media. The ClimateGate scandal arose when climate change critics hacked into a research database at the University of East Anglia, evaluated the data without authorisation and went public with accusations that data had been selectively published and perhaps even falsified.[1] More recently, a scientist at Harvard has been accused of biasing or falsifying data that show tamarin monkeys can learn algebraic rules.[2]

Yet it can be very hard to prove allegations of bias in the scientific literature. What has been called bias can potentially also be explained by unfavourable primary outcome findings that force a focus on secondary outcomes, problems in measurement of an outcome variable, changes mandated by reviewers, ambiguous or non-significant findings that may make authors reluctant to submit or journals reluctant to publish, errors that were corrected during data analysis, or simply lack of time and inclination to publish unexciting findings.[3]

For these reasons, it seems that probing allegations of bias may be an ineffective means to study the integrity of the scientific enterprise. If the goal is to characterise the veracity of science, it may be more useful to examine papers that have been retracted from the literature to determine the reasons for retraction.[4 5] This approach may be more objective than other approaches; bias is often in the eye of the beholder,[6] whereas retraction is akin to the death of a paper—an unambiguous end point.[7]

We postulate that the media focus on the integrity of science is a rational response to an actual increase in the rate of retraction. Our hypothesis is that the incidence of research fraud has indeed increased in recent years.

METHODS

Every research paper that was noted as retracted in the PubMed database from 2000 to 2010 was evaluated.[4] PubMed was searched on 22 January 2010 using the limits of 'items with abstracts, retracted publication, English.' A total of 788 retracted papers were identified, all of which were exported from PubMed and saved as a text file (available upon request). Formal retraction notices could not be obtained for 46 papers (5.8% of all retracted papers), so the reason these papers were retracted cannot be assessed.[5]

The 742 papers for which retraction notices could be obtained were evaluated to determine the reason(s) for retraction. Initially, reasons for retraction were dichotomised as fraud (manipulation of data) or error (all other causes).[5] Each retraction notice was then re-evaluated to determine detailed reasons for retraction. An effort was made to formalise a definition for every identified cause of retraction so that the reasons for retraction could be scored systematically. If a retraction notice noted several reasons for retraction, all reasons were tabulated. In cases where the retraction notice was vague, a 'best guess' was made as to the cause, although some retraction notices gave no information at all so an 'unstated' category was set up.

Additional information about each retracted article was tabulated including: first author surname; number of authors; country of address of first author; year of publication; year of retraction; journal of publication; and journal impact factor. Journal impact factor was determined using the ISI Web of Knowledge (Thomson Reuters) 'Journal Citation Reports, Science Edition' for 2008 (last available year).

To determine if the incidence of retraction for fraud (fabrication + falsification) has increased in recent years, the number of fraudulent papers was

tabulated by year. However, because the number of papers published per year is increasing, an increase in the number of papers retracted for fraud would not prove that fraud is becoming more common. To control for the total volume of papers published, it is assumed that scientific mistakes are an inevitable (and relatively constant) consequence of research, so mistakes are used to control for increases in research output. The relevant hypothesis is therefore that the number of papers retracted for fraud is increasing relative to the number of papers retracted for a scientific mistake. In addition, the number of published papers listed on PubMed per year was tabulated.

To determine if the time from publication to retraction has changed, the time (in months) from date of journal publication to date of formal retraction was calculated. This difference was analysed as a function of publication type, so time to retraction of fraudulent papers and of papers with a scientific mistake could be determined.

We also noted how papers are retracted from the literature. The relevant question is: how is retraction of a paper brought to the reader's attention if the reader is searching for the paper but doesn't already know that it has been retracted?

All data for each retracted paper were entered in an Excel spreadsheet for analysis.

RESULTS

The 742 retractions were distributed among 404 journals; 274 journals had one paper retracted while 130 journals had two or more retractions. Eight reasons for retraction were identified, split into the broad categories of fraud and error.

Fraud is defined as:

- Data fabrication: theft of data from an uncredited author (or generation of completely artificial data).
- Data falsification: editing or manipulation of authentic data to "prove" a hypothesis.

Error is defined as:

- Plagiarism: theft of text from an uncredited author.
- Duplication: self-plagiarism (using substantially the same data or words in different publications without citation).
- Scientific mistake: a mistake that invalidates research, typically explained carefully in a retraction notice.
- Ethical issues: violations of accepted publication practice (eg, publishing without co-author approval, patient-related IRB violations, etc).
- Journal error: accidental duplicative publication by a journal through no fault of the authors.
- Unstated: retraction notice is non-informative, simply stating that retraction has happened.

As a general rule, alteration of data in a table is considered fabrication whereas alteration of data in a figure is considered falsification.

Error is more common than fraud; 73.5% of papers were retracted for error (or an undisclosed reason) whereas 26.6% of papers were retracted for fraud (table 1). The single most common reason for retraction was a scientific mistake, identified in 234 papers (31.5%). Fabrication, which includes data plagiarism, was more common than text plagiarism. Multiple reasons for retraction were cited for 67 papers (9.0%), but 134 papers (18.1%) were retracted for ambiguous reasons (this total includes 61 papers retracted for unstated reasons (table 1) and 73 papers retracted with wording so vague that the reason for retraction could not be determined with certainty).

The total number of papers retracted per year increased sharply over the decade (figure 1). The trend to an increasing total number of retractions per year is significant (n=10; r=0.96; R=9.21; p<0.001), as is the trend to retractions happening after a longer period of time since publication (n=10; r=0.92; R=6.54; p<0.001). The increase in total retractions is correlated with the greater time to retraction (n=10; r=0.92; R=6.74; p<0.001). The total number of papers published per year also increased sharply over the decade; PubMed shows that 527 013 papers were published in 2000 while 852 325 papers were published in 2009, an increase of almost 62%. The rate of increase in retractions is greater than the rate of increase in publications, although the two are correlated (n=10; r=0.96; R=9.21; p<0.001).

The incidence of retraction specifically for research fraud has increased sharply in recent years (figure 2). However, the incidence of scientific mistakes has also increased over the same time period. The trend to an increasing number of retractions for fraud is significant (n=10; r=0.89; R=6.91; p<0.001), as is the trend to an increasing number of retractions for scientific mistake (n=10; r=0.94; R=7.84; p<0.001). There is a correlation between the number of papers retracted for fraud and the number retracted for a scientific mistake (n=10; r=0.80; R=3.75; p<0.01). The finding that more papers were retracted in 2009 than in 2000 does not prove that more fraudulent papers are being published, since journals could potentially be policing themselves more aggressively.

Table 1 Summary of reasons for retraction of 742 papers, 2000–10

Reason for retraction	Retracted papers, n (%)	IF, mean (SD)
Fraud		
Fabrication	111 (15.0)	9.79 (11.49)
Falsification	98 (13.2)	8.01 (8.33)
Error		
Scientific mistake	234 (31.5)	10.60 (10.66)
Duplicate publication	117 (15.8)	3.18 (2.63)
Plagiarism	107 (14.4)	2.31 (1.90)
Ethical violations	76 (10.2)	5.74 (8.51)
Unstated reasons	61 (8.2)	5.08 (8.82)
Journal error	27 (3.6)	2.66 (2.21)

IF, impact factor.

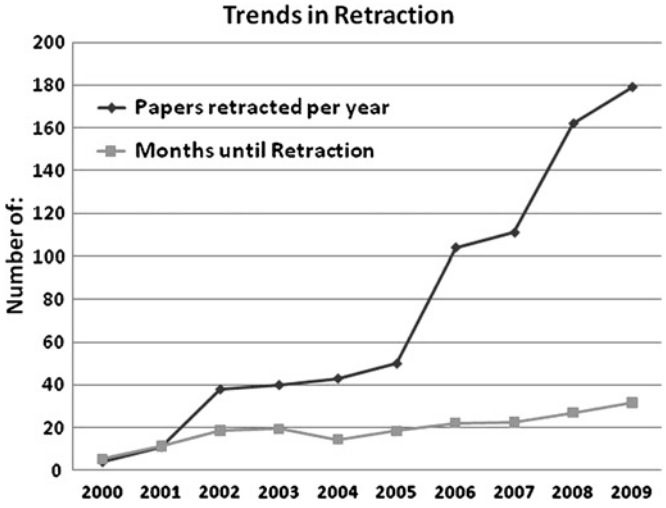

Figure 1 Trends in paper retraction showing the number of papers retracted by year, including retractions for all reasons. The average time until retraction is also shown.

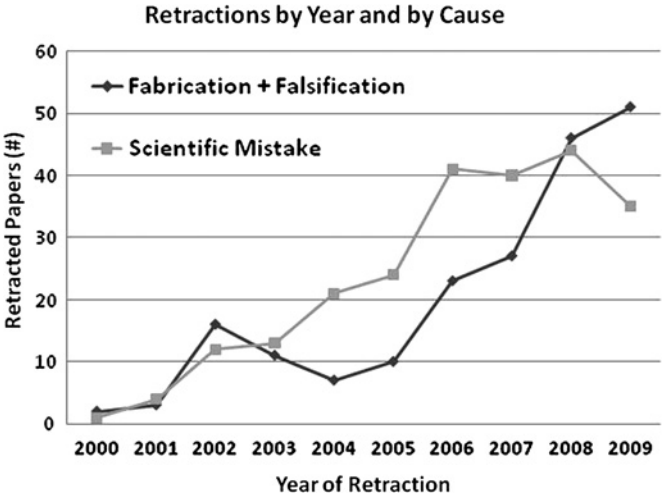

Figure 2 Number of papers retracted by year, including only 197 papers retracted for fraud (fabrication or falsification) and 234 papers retracted for scientific mistake. This figure therefore excludes all papers retracted for non-scientific error (eg, plagiarism, duplicate publication).

Journals are now reaching farther back in time to retract published papers (figure 3). The trend to increasing time to retraction for fraud is significant (n=10; r=0.87; R=5.05; p<0.001), as is the trend for increasing time to retraction for scientific mistakes (n=10; r=0.95; R=8.31; p<0.001). There is also a correlation between the time to retraction of papers retracted for fraud and for scientific mistake (n=10; r=0.94; R=7.80; p<0.001). In 2000, two papers were retracted for fraud and the longest time to retraction was 8 months; in 2005, 10 papers were retracted for fraud and the longest time to retraction was 44 months; in 2009, 51 papers were retracted for fraud and the longest time to retraction was 108 months. Similarly, in 2000, one paper was retracted for a scientific mistake after 3 months; in 2005, 24 papers were retracted for a mistake and the longest time to retraction was 67 months; and in 2009, 35 papers were retracted for a mistake and the longest time to retraction was 101 months.

Journals do not do a careful job of alerting the naïve reader to a retraction notice (table 2). The most common way to alert readers about a retraction is with a watermark on the pdf (41.1% of retracted papers). A total of 149 papers were given such a watermark and noted as retracted at the website as well. Among 305 watermarked papers, 48.9% were also retracted at the website; of 248 papers retracted at the website, 60.1% were also watermarked. However, 31.8% of retracted papers were not noted as retracted in any way at all so the naïve reader would not be alerted to the fact that retraction had happened.

Table 2 Summary of how the naïve reader is alerted to paper retraction

Where retraction noted	Retracted papers, n (%)
Watermark on pdf	305 (41.1)
Journal website	248 (33.4)
Not noted anywhere	236 (31.8)
Note appended to pdf	128 (17.3)
pdf deleted from website	98 (13.2)

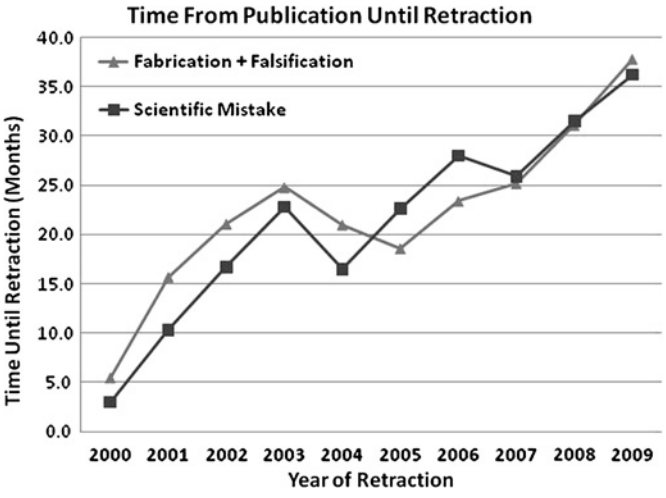

Figure 3 Average time from publication to retraction (in months) for 197 papers retracted for fraud (fabrication or falsification) and 234 papers retracted for scientific mistake.

DISCUSSION

This study has identified eight reasons for retraction, with the most common reason being a scientific mistake (table 1). In general, error is more common than fraud, but fabrication and data plagiarism are more common than text plagiarism (table 1). The total number of retractions has increased sharply in recent years (figure 1), although the time to retraction has also increased. Retractions specifically for fraud have also increased recently (figure 2), yet the apparent increase in the incidence of fraud is confounded by a similar and strongly correlated increase in the incidence of scientific mistakes (figure 2). It simply does not seem plausible that both fraud and mistakes are increasing rapidly and in tandem, even though the total number of publications is also increasing. Hence, the finding that journals are reaching farther back in time to retract papers (figure 1) and that both fraudulent papers and erroneous papers are retracted after a long period of time (figure 3) may be meaningful.

One possible interpretation of these results is that the incidence of research fraud truly is increasing, as it appears to be (figure 2). It is particularly striking that the number of papers retracted for fraud increased more than sevenfold in the 6 years between 2004 and 2009. During the same period, the number of papers retracted for a scientific mistake did not even double, and the total number of papers cited in PubMed increased by 35% (data not shown). This is consistent with the hypothesis that fraud is increasing more rapidly than scientific mistakes or publications overall. Yet it is hard to explain why fraud and mistakes are strongly correlated (figure 2). Very recently, a downturn in mistakes occurred at the same time as an upturn in fraud (figure 2), which may mean that some fraud in the past was excused as a scientific mistake whereas fraud is now more likely to be admitted as such. We note that some excuses given for a mistake seem implausible and no excuse at all is given for other retractions; it seems likely that at least some retractions for mistakes or for unstated reasons actually represent fraudulent papers.

Another possibility is that the incidence of fraud has not increased appreciably but journals are making a far more aggressive effort to self-police now than in the recent past. This could explain why the number of retractions for fraud and for scientific mistakes are both increasing (figure 2), and why the time to retraction for flawed papers is increasing in tandem for

both fraud and scientific mistakes (figure 3). If it is true that journals are self-policing more now than in the recent past, then the apparent increase in the incidence of research fraud may be an artifact of a change in behaviour by journal editors. This suggests that greater transparency by journal editors about the commission of fraud is perhaps seen as a tool to discourage misbehaviour.

A final possibility is that, because the overall number of retractions is rather small, our results are strongly affected by random events. It is clear that the recent history of retractions for fraud has been dominated by several prolific 'repeat offender' authors.[5] For example, Jan Hendrik Schön[8] fabricated results in 15 retracted papers in our database and all of his papers were retracted in 2002 and 2003, so he alone was responsible for 56% of papers retracted for fraud in 2002–3. Similarly, Scott Reuben[9] fabricated results in 13 retracted papers and all of his papers were retracted in 2009; thus he alone was responsible for 25% of retracted fraudulent papers in 2009. Overall, these two 'repeat offender' authors were responsible for 14% of all articles retracted for fraud over the last decade. When their papers were retracted has therefore had a large impact on the year-by-year incidence of retraction.

The levels of misconduct reported are generally higher than in previous studies. One study reported that 235 papers had been retracted from 1966 to 1997,[10] a rate of retraction roughly 10% of that reported here. A study of 395 retracted articles by Nath *et al*[11] published from 1982 to 2002 found a rate of retraction roughly 27% of that reported in the current study. Furthermore, Nath *et al* reported that 27.1% of retractions were due to fabrication, falsification or plagiarism,[11] whereas the current study found that 41.0% of retracted papers were withdrawn for such causes. Still, we concur with the conclusion that scientific mistakes—including the inability to reproduce results—are the most commonly cited reasons for retraction[12 13] and agree that the reasons for retraction published in the retraction notice are not always reliable.[13]

Retraction is the most serious sanction that can be applied to a published paper, and it is imperative that this sanction be applied fairly and is reserved for cases of scientific misconduct.[7] Yet the results of the present study show that 27 papers were retracted because of an administrative error at a journal office (table 1), an offence for which authors are apparently blameless. Thus, retraction is a very blunt instrument used for offences both gravely serious and trivial.[5] Economic theory suggests that societies develop norms of behaviour to sustain fairness in ephemeral exchanges.[14] In small societies—such as the society of scientists in 18th century England—reputational systems were largely sufficient to maintain integrity. However, as societies grow in size, fairness of exchange must increasingly be maintained by punishment and it is expected that large societies—such as the society of scientists in the modern world—will punish more vigorously.[14] Given that retraction of a paper is the harshest possible punishment for a scientist, we must work to assure that it is applied fairly.

This research is limited by several weaknesses. It may be contentious that plagiarism is regarded as error rather than fraud. However, we previously defined fraud as manipulation of data and error as all other infractions[5]; this definition gives primacy to the originality of data rather than to the originality of words. Many retractions are cryptic as to the actual reason(s) for retraction; 8% of retractions were for unstated reasons (table 1) and up to 18% of retractions were for ambiguous reasons. Another weakness is that a retraction notice may not specify every reason for retraction. Multiple reasons for retraction were cited for 67 papers (9%), yet research fraud (fabrication or falsification) might tend to trump everything else so that lesser offences might not be reported in the context of research fraud. Arguing against this interpretation, we found that two-thirds of the papers retracted for multiple reasons had fraud noted as a problem (data not shown).

A final weakness of the study is that it is unclear what role bias may play in retraction. This gap in our knowledge is important because bias, which has been defined as distorted presentation or 'spin', has been a central concern in much of the angst about pharmaceutical research.[15] There is a fear that clinical trial results may be interpreted in a distorted manner by certain authors—perhaps authors in the pay of a pharmaceutical company—and one study concluded that 40% of randomised clinical trials show evidence of 'spin' in the published text.[15] These findings feed a fear that medical journals have become an extension of the marketing department of pharmaceutical companies.[16] Yet bias is in the eye of the beholder, and the paper entirely free of bias has not yet been written. We note that one study concluded that only 4% of retractions for misconduct had declared pharmaceutical sponsorship,[17] while another study reported that pharmaceutical sponsorship of obesity-related randomised clinical trials was actually associated with higher quality reporting of results.[18]

A key issue for studies of bias is that published data may not accurately reflect the underlying database that relates to a specific medication.[19] A review of 164 new drug applications (NDAs) submitted to the US Food and Drug Administration found that only 78% of NDAs were subsequently published in the open literature and that the information available in that literature did not necessarily reflect all of the information available in the original NDA.[3] This was taken as evidence of bias, which was defined as not publishing data at all, selectively reporting data or framing data to make it look more favourable.[3] However, our work suggests that selectively reporting only clinical data that confirms a hypothesis might actually be a form of data falsification, which is research fraud.

Key unanswered questions remain. How many cases of perceived bias are actually cases of research fraud? Conversely, if accusations of bias cannot be supported with evidence of fraud, is bias a serious problem or is it really just evidence either that a researcher is naïve or that an accuser has a different viewpoint? Is it possible that bias is not a crime at all, but more nearly the normal state of affairs in every paper? It may be noteworthy that, among 742 retracted articles, only one retraction notice[20] described biased viewpoint as the reason for retraction ("authors have confounded the elegant analysis …with a political viewpoint representing only one side"). We believe that no researcher is entirely free of bias, no matter the funding source, and careful scientists already know that.

Competing interests None.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. **Broder JM.** Scientists taking steps to defend work on climate. *New York Times*, 2 March 2010.
2. **Wade N.** Inquiry on Harvard lab threatens ripple effect. *New York Times*, 12 August 2010.
3. **Rising K,** Bacchetti P, Bero L. Reporting bias in drug trials submitted to the food and drug administration: review of publications and presentation. *PLoS Med* 2008;**5**:1561–9.
4. **Steen RG.** Retractions in the scientific literature: who is responsible for scientific integrity? *AMWA J* 2010.
5. **Steen RG.** Retractions in the scientific literature: do authors deliberately commit research fraud? *J Med Ethics* 2011;**37**:113–17.

6. **McKee M,** Stuckler D. How cognitive biases affect our interpretation of political messages. *BMJ* 2010;**340**:936—7.
7. **Wager E,** Barbour V, Yentis S, *et al*. Retractions: guidance from the Committee on Publication Ethics (COPE). *Maturitas* 2009;**64**:201—3.
8. **Service RF.** Bell labs fires star physicist found guilty of forging data. *Science* 2002;**298**:30—1.
9. **Marcus A.** Fraud case rocks anesthesiology community. *Anesth News* 2009;**35**:3.
10. **Budd JM,** Sievert M, Schultz TR. Phenomenon of retraction: reasons for retraction and citations to the publications. *JAMA* 1998;**280**:296—7.
11. **Nath SB,** Marcus SC, Druss BG. Retractions in the research literature: misconduct or mistakes? *Med J Aust* 2006;**185**:152—4.
12. **Budd JM,** Sievert M, Schultz TR, *et al*. Effects of article retraction on citation and practice in medicine. *Bull Med Libr Assoc* 1999;**87**:437—43.
13. **Redman BK,** Yarandi HN, Merz JF. Empirical developments in retraction. *J Med Ethics* 2008;**34**:807—9.
14. **Henrich J,** Ensminger J, McElreath R, *et al*. Markets, religion, community size, and the evolution of fairness and punishment. *Science* 2010;**327**:1480—4.
15. **Boutron I,** Dutton S, Ravaud P, *et al*. Reporting and interpretation of randomized controlled trials with statistically nonsignificant results for primary outcomes. *JAMA* 2010;**303**:2058—64.
16. **Smith R.** Medical journals are an extension of the marketing arm of pharmaceutical companies. *PLoS Med* 2005;**2**:e138.
17. **Jones N.** Analysis of retractions puts spotlight on academia. *Nat Med* 2009;**15**:1101.
18. **Thomas O,** Thabane L, Douketis J, *et al*. Industry funding and the reporting quality of large long-term weight loss trials. *Int J Obes* 2008;**32**:1531—6.
19. **Spielmans GI,** Parry PI. From evidence-based medicine to marketing-based medicine: evidence form internal industry documents. *J Bioeth Inq* 2010. Published Online First: 21 January 2010. doi:10.1007/s11673-010-9208-8.
20. **Suciu-Foca N,** Lewis R. Retraction notice. *Hum Immunol* 2001;**62**:1063.

Retractions in the scientific literature: is the incidence of research fraud increasing?

R Grant Steen

J Med Ethics 2011 37: 249-253 originally published online December 24, 2010
doi: 10.1136/jme.2010.040923