Retractions in the scientific literature: do authors deliberately commit research fraud?

R Grant Steen

J Med Ethics published online November 15, 2010
doi: 10.1136/jme.2010.038125

JME Online First, published on November 15, 2010 as 10.1136/jme.2010.038125

Retractions in the scientific literature: do authors deliberately commit research fraud?

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 31 May 2010
Revised 29 July 2010
Accepted 13 August 2010

ABSTRACT

Background Papers retracted for fraud (data fabrication or data falsification) may represent a deliberate effort to deceive, a motivation fundamentally different from papers retracted for error. It is hypothesised that fraudulent authors target journals with a high impact factor (IF), have other fraudulent publications, diffuse responsibility across many co-authors, delay retracting fraudulent papers and publish from countries with a weak research infrastructure.

Methods All 788 English language research papers retracted from the PubMed database between 2000 and 2010 were evaluated. Data pertinent to each retracted paper were abstracted from the paper and the reasons for retraction were derived from the retraction notice and dichotomised as fraud or error. Data for each retracted article were entered in an Excel spreadsheet for analysis.

Results Journal IF was higher for fraudulent papers ($p<0.001$). Roughly 53% of fraudulent papers were written by a first author who had written other retracted papers ('repeat offender'), whereas only 18% of erroneous papers were written by a repeat offender ($\chi=88.40$; $p<0.0001$). Fraudulent papers had more authors ($p<0.001$) and were retracted more slowly than erroneous papers ($p<0.005$). Surprisingly, there was significantly more fraud than error among retracted papers from the USA ($\chi^2=8.71$; $p<0.05$) compared with the rest of the world.

Conclusions This study reports evidence consistent with the 'deliberate fraud' hypothesis. The results suggest that papers retracted because of data fabrication or falsification represent a calculated effort to deceive. It is inferred that such behaviour is neither naïve, feckless nor inadvertent.

INTRODUCTION

The scientific literature is a tangible record of the search for truth.[1] Fraudulent research, marred by an intention to deceive rather than by mere error or carelessness,[2] would subvert that truth and could potentially result in retraction of published papers.[3] Yet the motivation of authors who might try to commit research fraud is hard to fathom. Science is said to be self-correcting, so the literature purges itself of papers deemed fraudulent.[4] If indeed the literature self-corrects, then research fraud should be futile.

Because of the apparent futility of deliberate fraud, it was hypothesised that retracted papers arise from inadvertent error rather than from deliberate fraud.[3] If all retractions result from inadvertent error, one would expect retractions to be randomly distributed through the literature. Conversely, if all retractions result from deliberate fraud, one would expect retractions to be non-random and to focus on a few dishonest authors or a few poorly-edited journals or a few countries in which research infrastructure is weak. We tested the 'inadvertent error' hypothesis by determining whether retractions are randomly distributed throughout the literature and found evidence that retractions are, in fact, clustered. Retracted papers are more likely to appear in journals with a high impact factor (IF), are more likely to involve certain 'repeat offender' authors and are more likely to involve authors from the USA.[3]

Nevertheless, an apparently clustered distribution could result from a subset of deliberately fraudulent (non-random) authors in the context of a larger set of authors who made inadvertent (random) errors resulting in article retraction. This model suggests that fraudulent papers are fundamentally different from erroneous papers, perhaps in a way that will give insight into the motivation of fraudulent authors. A study was therefore undertaken to test the 'deliberate fraud' hypothesis that some authors deliberately commit research fraud. It is predicted that, compared with authors who produce erroneous papers, authors who produce fraudulent papers specifically target journals with a high IF, have other fraudulent publications, diffuse personal responsibility across a large number of co-authors, delay retracting the paper and collaborate with a cohort of co-authors who also have other retractions for fraud.

METHODS

Every research paper noted as retracted in the PubMed database from 2000 to 2010 was evaluated. PubMed was searched on 22 January 2010 with 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), as described elsewhere.[3]

Data pertinent to each retracted paper were abstracted from the paper itself or the retraction notice for that paper. Each retracted paper was categorised according to first author surname, number of authors, year of publication, year of retraction, time (in months) to retraction, journal of publication, journal IF and country of address of first author. Data for each retracted article were entered in an Excel spreadsheet for analysis. First authors with more than one retracted paper ('repeat offenders') were identified by sorting the final Excel spreadsheet.

Journal IF was determined using the ISI Web of Knowledge (Thomson Reuters) *Journal Citation Reports, Science Edition* for 2008 (last available year). Reasons for retraction were abstracted from the

retraction notice and dichotomised as either fraud or error. Research fraud includes data fabrication (manufacture of fictional data or data plagiarism) or data falsification (selective editing of data to confirm a hypothesis). It was noted that this definition classifies text plagiarism as a form of research error. Papers retracted for fraud were then contrasted with papers retracted for error using data in the Excel spreadsheet.

The prediction that fraudulent papers would come from countries with a weak research infrastructure was tested by collating the country of address of the first author of all retractions. Statistical significance was tested by a χ^2 analysis of a 2×2 test of homogeneity in which fraud and error by authors in each individual country was contrasted with fraud and error from all other countries pooled together.

Finally, the prediction that fraudulent authors collaborate with a cohort of co-authors who also have other retractions for fraud was tested. Testing this prediction on all 788 retracted papers was prohibitively complex since it entails tabulating every co-author and looking for multiple retractions among the co-authors. To limit the scope, the study focused on 63 clinical science papers (eg, research involving humans or human-derived material that does not involve competing treatments) retracted for fraud. All co-authors were tabulated in an Excel spreadsheet and co-authors with more than one retraction were noted.

RESULTS

Formal retraction notices could not be obtained for 46 papers (5.8% of all retracted papers, table 1); hence, the reason these papers were retracted is unknown. Papers with an unknown reason for retraction appeared in journals with a substantially lower IF than papers with a known reason for retraction ($p<0.001$). This is consistent with 'unknown' retraction notices appearing in journals that were not among the subscription holdings at a large public university. Papers retracted for an unknown reason were less often written by a 'repeat offender' author ($\chi^2=6.22$; $p<0.05$), and were retracted in significantly less time than other retracted papers ($p<0.05$).

Journal IF was higher among papers retracted for fraud than among those retracted for error ($p<0.001$, table 2). Roughly 53% of fraudulent papers were written by a first author who had written other retracted papers ('repeat offender') whereas only 18% of erroneous papers were written by a repeat offender ($\chi^2=88.40$; $p<0.0001$). The average number of authors was significantly higher for fraudulent papers ($p<0.001$), and fraudulent papers were retracted significantly more slowly than erroneous papers ($p<0.005$).

It has been postulated that retracted papers involve relatively few co-authors[5] since having many co-authors would seem to provide a backstop against both error and fraud. In contrast, we hypothesise that fraud is easier to perpetrate when personal responsibility is diffused. This study reports evidence that having many authors does not protect from fraud, since the number of authors per retracted paper ranged up to 26 (figure 1). The mean±SD number of authors per retracted paper was 5.1±3.3; the median number of authors per retracted paper was 4. Roughly 7% of retracted papers were written by a single author, but 18% of retracted papers had ≥8 authors and six retracted papers had >20 authors.

The country of affiliation for first authors of retracted papers was tabulated for countries with >10 retracted papers (table 3). Among all retractions, 33.0% were published by first authors in the USA and 32.3% of American retractions occurred because of fraud (fabrication or falsification). Asian nations together (China, Japan, India, South Korea) produced 30.1% of retractions and 26.6% of retractions were for fraud. China alone accounted for 11.3% of retractions and 22.5% of retractions were for fraud. There was significantly more fraud than error among retracted papers from the USA ($\chi^2=8.71$; $p<0.05$) and significantly less fraud than error among 34 other countries not tabulated ($\chi^2=6.47$; $p<0.05$). No other country differences achieved statistical significance, perhaps because sample sizes were generally small. When Asian nations were pooled to make a larger sample, 26.6% of retractions resulted from fraud whereas 24.3% of retractions in all other nations resulted from fraud ($\chi^2=0.87$; NS).

To test the prediction that fraudulent authors typically collaborate with an established cohort of co-authors, we evaluated the co-author list of all clinically-related papers retracted for fraud. Among 63 such papers there were a total of 404 authors (ie, authors were counted twice if they had two clinical papers retracted for fraud). Seven people first-authored 31 clinically-related fraudulent papers and co-authored 39 fraudulent papers in total (table 4), but there were 32 authors with just one clinically-related paper retracted for fraud. Among the first authors with multiple retractions for fraud (table 4), each first author had an average of 6.6 co-authors and each co-author had an average of 2.7 retractions. Thus, cohorts of authors collaborated to produce roughly half of all the retracted fraudulent clinical papers.

DISCUSSION

This study reports evidence consistent with the 'deliberate fraud' hypothesis: authors of fraudulent retracted papers appear to target journals with a high IF (table 2), often have several other retracted papers (tables 2 and 4), tend to diffuse responsibility across ≥4 co-authors (table 2, figure 1), delay retracting fraudulent papers (table 2) and collaborate with co-authors who also have other retracted papers (table 4). Evaluated retractions appear to be representative of retractions in general, except that retractions for an unknown reason were generally published in journals with a low IF (table 1). Perhaps surprisingly, fraud occurs more often in the USA than the rest of the world (table

Table 1 Comparison of articles for which a retraction notice was available and the reason for retraction could be determined (known) versus articles for which a retraction notice was unavailable or the reason for retraction could not be determined (unknown)

	Known reasons		Unknown reasons			
	Mean	SD	Mean	SD	χ^2 or t value	p Value
Sample (n)	742	–	46	–	–	–
Journal impact factor	7.01	9.08	0.91	1.19	4.556	<0.001
Repeat offenders, n (%)	205 (27.6)	–	5 (10.9)	–	6.223	<0.05
No of authors per paper	5.11	3.29	4.37	2.51	1.501	NS
Months to retraction	24.23	21.53	16.73	19.33	2.317	<0.05

Table 2 Comparison of articles for which retraction was explained as the result of fabrication or falsification ('fraud') versus articles for which retraction was explained as the result of any other cause ('error')

	Fraudulent papers		Erroneous papers			
	Mean	SD	Mean	SD	χ^2 or t value	p Value
Sample (n)	197	–	545	–	–	–
Journal impact factor	8.99	10.24	6.29	8.52	3.595	<0.001
Repeat offenders, n (%)	105 (53.3)	–	100 (18.4)	–	88.403	<0.0001
No of authors per paper	5.82	3.63	4.86	3.13	3.529	<0.001
Months to retraction	28.41	22.87	22.72	20.84	3.193	<0.005

3), whereas one would expect the strong research infrastructure in the USA to limit fraud. Nevertheless, these results suggest that papers retracted because of data fabrication or falsification represent a deliberate effort to deceive, and that such behaviour is neither naïve, feckless nor inadvertent.

These results are substantially different from an earlier study which found that, relative to papers flawed by error, fraudulent papers are likely to have fewer authors and to be published in low-IF journals.[5] This earlier study evaluated 395 papers retracted between 1982 and 2002, so the period of overlap between the two studies is only 2 years and the earlier rate of retraction is about 25% of that reported in the later period. The study by Nath *et al*[5] also reported that 27.1% of papers were retracted for misconduct (falsification, fabrication or plagiarism) whereas we report that 26.6% of papers were retracted for fraud (only falsification or fabrication). Nath *et al*[5] concluded that single authorship, with its lack of accountability and oversight by colleagues, raises the risk of misconduct, but this study found no evidence to support that conclusion (figure 1). Nevertheless, the studies concur in finding that erroneous studies are withdrawn more rapidly than fraudulent studies, although we found that all retracted studies are now withdrawn more rapidly than they were in the past.[5]

The results of this study show unequivocally that scientists in the USA are responsible for more retracted papers than any other country (table 3). These results suggest that American scientists are significantly more prone to engage in data fabrication or falsification than scientists from other countries. There was no evidence to support a contention that papers submitted from China or other Asian nations and indexed in PubMed are more likely to be fraudulent.[6] China, despite several recent high-profile fraud cases,[7 8] did not commit a disproportionate share of research fraud. Chinese scientists published 11.5% of all scientific papers[6] and accounted for 11.3% of retractions in the world literature (table 3).

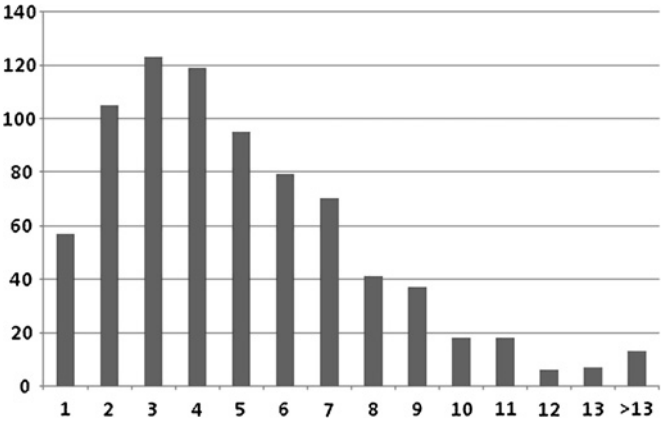

Figure 1 Number of authors per retracted paper. Among 788 papers, the mode was three authors per paper, but many papers had more than 10 authors. Only 57 papers had a single author.

The argument that there are deliberately fraudulent authors is compelling.[3] Jan Hendrik Schön, who fabricated results in at least 17 basic science papers in 2 years (9 in *Science*, 7 in *Nature*), at his peak produced a new paper every 8 days.[9] Yet Schön's extraordinary productivity and high visibility eventually led to his downfall; researchers noted that a figure in *Nature* was identical to a figure in *Science* although the two papers were about different subjects,[9] demonstrating that research fraud is risky. Because there is an expectation that fraud eventually will be noted by readers or revealed in the light of new experiments, it is hard to comprehend why an author would risk exposure and ruin by engaging in fraud.

A claim has been made that certain warning signs identify potentially fraudulent papers,[8] yet the work presented here does not identify any such signs. Clearly there are factors that may promote fraud, including a desire for personal fame or financial gain or competitive advantage; in addition, some scientists may exhibit the hubris of certainty before the results are fully known.[10] It is even possible that research fraud is aspirational; scientists may wish certain findings to be true to such an extent that eventually the truth seems proven by the belief.

It may be controversial that text plagiarism is a form of misconduct that is not as damaging as data plagiarism. Data plagiarism is theft of raw data, tables or figures from a published paper, often with modifications to obscure the theft, and data

Table 3 Country of affiliation for first authors of retracted papers 2000–10, tabulated for countries with more than 10 retracted papers

	Total retractions	Error (E)	Fraud (F)	Reasons unknown	Ratio E/F
USA*	260	169	84	7	2.0
China	89	60	20	9	3.0
Japan	60	41	18	1	2.3
India	50	27	17	6	1.6
UK	45	36	7	2	5.1
South Korea	38	27	8	3	3.4
Germany	25	22	3	0	7.3
Australia	17	13	3	1	4.3
Canada	17	15	2	0	7.5
Italy	17	11	6	0	1.8
Turkey	15	13	2	0	6.5
France	13	12	1	0	12.0
Greece	12	10	0	2	>10
Iran	11	9	1	1	9.0
Asian nations	237	155	63	19	2.5
34 other countries*	119	80	25	14	3.2
Overall total	788	545	197	46	2.8

Countries that differ significantly from the norm are noted.
*χ^2 analysis, 2×2 test of homogeneity, $p<0.05$.

Table 4 Summary of co-author cohort characteristics among first authors with more than one fraud retraction

First author	Retractions by first author	Co-authors with >1 retraction	Average number of retractions among co-authors
Ben-Gal Y	2	6	2.0
Lozano SC	4	3	4.0
Matsuyama W	9	10	5.4
Oka H	2	5	2.0
Paventi S	2	0	0.0
Reuben SS	13	8	2.6
Sudbø J	7	14	3.0
Mean	5.6	6.6	2.7

This table is limited specifically to clinically-related papers.

plagiarism is regarded as a major ethical breach.[11] Even the smallest degree of data plagiarism is deeply troubling, not only because it represents a theft of the time, effort and expense of another, but because it corrupts the essence of the scientific enterprise. Interpretations are built upon observations and hypotheses cannot be falsified if the link between interpretation and observation is broken; no one can vouch for plagiarised data, so the act of plagiarism must falsify the data.

Text plagiarism is less problematic in that it is a theft of something less tangible; it is hard to establish the provenance of an idea at the best of times. There are times when the words of one author so clearly express an idea that those words may be used, deliberately or inadvertently, by other authors.[11] Text plagiarism can also happen by accident, or it can be naively done by an author who believes it to be a form of flattery, or non-native English writers may be unable to alter a sentence sufficiently to avoid a charge of plagiarism.[11] Finally, some text—such as a methods description—may be difficult to reword without creating confusion, especially if the method is widely used and has been described many times previously.[11]

We believe that data plagiarism and text plagiarism should not be regarded as equally serious ethical breaches. In fact, it is unclear how even to define text plagiarism. For example, self-plagiarism is clearly different from non-self-plagiarism. The fraction of material that can be reused by an author in a subsequent publication without self-plagiarisation has arbitrarily been set at 30% by some journal editors,[12] yet this limit would be well outside the range of what is acceptable to steal from another author. Retraction notices are rarely explicit as to the extent of theft involved in text plagiarism; clearly, theft of a few unattributed sentences is more likely to be inadvertent than is the appropriation of whole chunks of text, yet retraction notices virtually never differentiate between these two behaviours.

Journal retraction notices are often quite cryptic as to why a particular paper was retracted, although greater clarity could potentially benefit everyone. If a paper with original data is retracted for text plagiarism, then the published data could still guide future research and conserve research funding that might otherwise be wasted in repeating the study. If a paper in which data are corrupted by an error is subjected to the same fate as a paper in which an experimental dataset is entirely fabricated, then punishment is too harsh for those authors who made an error. Embarrassment on the part of a journal editor is not sufficient reason to publish a coded retraction notice that does not explicitly describe what was wrong with a retracted paper.

While it is compelling that some authors are deliberately fraudulent, it is also apparent that some forms of transgression that result in retraction can result from naivety or fecklessness. For example, there can be no possible benefit from duplicative publication unless an author is transparent about the several published articles, and transparency must inevitably lead to exposure.

This research is limited by the fact that only papers that were formally retracted were evaluated and there is little doubt that scientific misconduct is under-reported.[13] A survey of 3247 active scientists in the USA, all of whom were funded by the National Institutes of Health, revealed that 0.3% of scientists admitted falsifying data in the past while 1.4% admitted to plagiarism.[14] If the average scientist who was surveyed published 10 papers in their career, then 552 of these papers (or 1.7% of total output) might require retraction. Another survey of 2212 scientists revealed 201 instances of likely research misconduct over a 3-year period, for an incidence rate of 3 per 100 scientists per year.[15] Among 163 professional biostatisticians, more than half personally knew of a fraudulent project in the 10 years prior to the survey and nearly one-third had been engaged in a project in which fraud had actually occurred.[16] Careful screening of 660 publications in surgery concluded that 3% were double publications (ie, self-plagiarism) and another 8% were substantially similar to other previously published work.[17] A meta-analysis including 11 647 scientists assessed in 21 separate studies (including three that were described above) concluded that 2% of scientists had committed research fraud at least once in their career.[18] If these numbers are credible—and it is likely they are conservative because people are loath to admit scientific misconduct—then there may be many fraudulent papers that have not been retracted.

Another limitation of this study is that we relied upon PubMed exclusively to study retraction of scientific papers. PubMed is administered by the National Library of Medicine in the USA and contains nearly 20 million records (http://www.nlm.nih.gov/). The total number of journals included in PubMed in 2008 was 6000; in contrast, Scopus (administered in the Netherlands) included 12 850 journals in 2008, although the focus was less exclusively on biological or medical science.[19] This may mean that PubMed is biased in its coverage towards English language journals or perhaps even towards journals from the USA. This could potentially affect the results in a systematic way; if more papers by Chinese authors are indexed in Scopus than in PubMed, then Scopus will be more likely than PubMed to contain retracted papers from China.

A third limitation of this study is that retraction of a paper for fraud probably makes it more likely that other papers by the same author will also be retracted, whether or not they are truly fraudulent. While this might inflate the apparent rate of fraud, it might also strike some papers from the literature that are either not fraudulent or not misconduct at all. At present we cannot distinguish between two possibilities: that fraudulent authors are more likely than other authors to produce multiple fraudulent papers; or that fraudulent authors are vigorously expunged from the literature, whether they warrant such harsh treatment or not.

A final limitation of this study is that articles for which the reason for retraction is unknown (tables 1 and 3) may actually have been retracted for fraud. Systematic bias could potentially be introduced if retraction notices from certain countries are less likely to disclose fraud (table 3). This could alter the estimate of which countries have the highest rate of retraction, although the USA leads the pack in retractions by any estimate. We note that the low proportion of repeat offenders among the papers retracted for an unknown reason (table 1) argues that these papers probably represent inadvertent errors. The proportion of

repeat offender authors among fraudulent papers (table 2) is nearly fivefold higher than among papers with an unknown reason for retraction (table 1).

These findings are potentially important for several reasons. First, the idea that certain authors may be deliberately trying to deceive should make journal editors and general readers profoundly cautious. Although most editors of scientific journals are apparently not concerned about the frequency and severity of ethical issues in their own journal,[20] the duplicity of some authors is cause for concern. Second, although the evidence has not been described in depth, there is clearly variation in the severity of the offence that leads to retraction. Retraction is the most serious sanction that can be applied to a published paper, and it should be reserved for papers that are so seriously flawed that their findings or conclusions cannot be relied upon.[21] Some papers were retracted because pervasive and systematic fraud had tainted the research output of an entire career (table 4), yet 13 papers were retracted because of an error at a journal office, an offence for which authors are apparently blameless (data available upon request). At present, retraction is a very blunt instrument used for offences both gravely serious and trivial.

Competing interests None.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. **Sox HC,** Rennie D. Research misconduct, retraction, and cleansing the medical literature: lessons from the Poehlmann case. *Ann Intern Med* 2006;**144**:609—13.
2. **Franzen M,** Rödder S, Weingart P. Fraud: causes and culprits as perceived by science and the media. *EMBO Rep* 2007;**8**:3—7.
3. **Steen RG.** Retractions in the scientific literature: who is responsible? *AMWA J* In press.
4. **Korpela KM.** How long does it take for scientific literature to purge itself of fraudulent material? The Breuning case revisited. *Curr Med Res Opin* 2010;**26**:843—7.
5. **Nath SB,** Marcus SC, Druss BG. Retractions in the research literature: misconduct or mistakes? *Med J Aust* 2006;**185**:152—4.
6. **Anon.** Scientific fraud: action needed in China. *Lancet* 2010;**375**:94.
7. **Saunders R,** Savulescu J. Research ethics and lessons from Hwanggate: what can we learn from the Korean cloning fraud? *J Med Ethics* 2008;**34**:214—21.
8. **Harrison WTA,** Simpson J, Weil M. Editorial. *Acta Crystallogr Sect E* 2010;**66**: e1—2.
9. **Service RF.** Bell Labs fires star physicist found guilty of forging data. *Science* 2002;**298**:30—1.
10. **Jaffer U,** Cameron AE. Deceit and fraud in medical research. *Int J Surg* 2006;**4**:122—6.
11. **Mason PR.** Plagiarism in scientific publications. *J Infect Dev Ctries* 2009;**3**:1—4.
12. **Giles J.** Taking on the cheats. *Nature* 2005;**435**:258—9.
13. **Kakuk P.** The legacy of the Hwang case: research misconduct in biosciences. *Sci Eng Ethics* 2009;**15**:545—62.
14. **Martinson BC,** Anderson MS, de Vries R. Scientists behaving badly. *Nature* 2005;**435**:737—8.
15. **Titus SL,** Wells JA, Rhoades LJ. Repairing research integrity. *Nature* 2008;**453**:980—2.
16. **Ranstam J,** Buyse M, George SL, *et al*. Fraud in medical research: an international survey of biostatisticians. ISCB Subcommittee on Fraud. *Control Clin Trials* 2000;**21**:415—27.
17. **Schein M,** Paladugu R. Redundant surgical publications: tip of the iceberg? *Surgery* 2001;**129**:655—61.
18. **Fanelli D.** How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. *PLoS One* 2009;**4**:e5738.
19. **Falagas ME,** Pitsouni EI, Malietzis GA, *et al*. Comparison of PubMed, Scopus, Web of Science, and Google Schoolar: strengths and weaknesses. *FASEB J* 2008;**22**:338—42.
20. **Wager E,** Fiack S, Graf C, *et al*. Science journal editors' views on publication ethics: results of an international survey. *J Med Ethics* 2009;**35**:348—53.
21. **Wager E,** Barbour V, Yentis S, *et al*. Retractions: guidance from the Committee on Publication Ethics (COPE). *Maturitas* 2009;**64**:201—3.