JAMA, The Journal of the American Medical Association, 1990;263(10):1420-1423

The continued use of retracted, invalid scientific literature. *Mark P. Pfeifer; Gwendolyn L. Snodgrass.*

Abstract: Research fraud, when a scientist intentionally reports data that have been fabricated or altered, has been discovered on numerous occasions in recent years. This phenomenon may result from the pressure on academic researchers to publish frequently; it is known as the "publish or perish" mentality. While the scientific community has debated the causes of research fraud and the appropriate disciplinary measures, little attention has been paid to determining how the invalid data and conclusions are used by other scientists after their publication. In some cases, formal retractions are published by the journal that originally carried the article, but many readers will find the original article without also encountering the retraction. Researchers frequently conduct searches of the literature to gain a background understanding of what other scientists have found on a particular topic. They then conduct their own studies and cite the work of others in their list of references. An analysis of the literature was performed to determine how often authors had cited 82 fraudulent articles that had already been formally retracted. The 82 invalid articles were cited 733 times by other authors who presumably did not know of the retraction. A control group of articles that had not been retracted was compared; this showed that citation of retracted articles was only 35 percent less frequent than citation of valid articles. The continued citation of retracted publications may result in part from the inconsistent procedures used by journals for stating and indexing retractions. Unfortunately the methods used by the scientific community for removing fraudulent literature from use are grossly inadequate at this time; this issue will require further attention. (Consumer Summary produced by Reliance Medical Information, Inc.)

The Continued Use of Retracted, Invalid Scientific Literature Little is known about the ultimate scientific fate of retracted, invalid literature. We identified 82 completely retracted articles by electronic and manual methods and measured their subsequent use in the scientific literature by performing citation analysis. After retraction, these studies were cited, for support of scientific concepts, 733 times. Comparison with a control group revealed that retraction reduces subsequent citation by approximately 35%. There was no evidence that small, obscure journals, non-US journals, or non-US authors were disproportionately responsbile for these citations. Although, after retraction, US authors accounted for a smaller percentage of citations, they continued to be the single greatest source. Several possible reasons why invalid information continues to be used were identified. These included a dearth of available information on retracted works> inconsistency in retraction format, terminology, and indexing> and an apparent lack of sufficient attention to manuscripts by some authors and editors.

OVER the last 20 years, the scientific community has become aware of numerous accounts of research fraud. These revelations have stirred debate about our "publish or perish" mentality, the structure and responsibilities of research institutions, the peer review process, the role of government intervention, and many issues in biomedical publishing. [1-7] There has been little evaluation, however, of the ultimate scientific fate of the invalid ideas and data generated by this fraud.

It may be impossible to measure accurately the use of retracted, invalid data by clinicians. Citation indexing, however, measures the use of such data in future scientific publications and would reflect the degree of removal of invalid work from at least the academic aspects of the scientific community.

We questioned whether current methods to disseminate knowledge of retracted publications were adequate to prevent their future use. Citation analysis was performed on a cohort of articles retracted because of fraud or error to determine whether they are commonly cited and, if so, to identify responsible factors.

METHODS

Scientific articles published since 1970 and eventually retracted were identified using Index Medicus and by manually reviewing editorials and reports on scientific fraud in both the scientific and lay literature. Articles known to be fraudulent but without published notices of retraction, those with ambiguous retractions, and those in which only part of the data or conclusions was retracted were excluded. While retraction notices were not required to be in the original journal, publication in a journal indexed by Index Medicus was necessary for inclusion. A total of 82 such retracted articles were identified.

Citation analysis, using the on-line database SCISEARCH (Institute for Scientific Information, Philadelphia, Pa), was performed on each article, giving information on every work subsequently citing the retracted articles. This included citations from editorials, letters to the editor, and original studies, both before and after retraction. Citations were considered "postretraction" beginning the calendar year after the retraction notice. Citing authors' and journals' geographic locations were grouped as the United States, Canada, Europe, Asia, and all others. In the

few articles with authors in more than one geographic category, the location of the first author was used. The geographic origins of the citing journals were determined using The Serials Directory. [8]

The title, source type, and subject headings obtained from the SCISEARCH electronic index were used to differentiate citations made from original works in reference to scientific concepts and those from articles and editorials on fraud in science. To validate this method, 25% of the retracted articles were randomly selected, and citations made to these articles after retraction were manually reviewed.

To determine what impact, if any, retraction has on the future citation rate, we developed a control group consisting of all works published in the same years in the same journals in which the retracted articles originally appeared. Using the SCI Journal Citation Reports, [[9]sup.(pp1501-2399)] citations made to these "sister" articles were followed for up to 9 years after publication, and a cumulative citation curve over time for the control group was generated. This curve represented citations made to all works published in a total of 516 journal-years, and its absolute values were obviously much greater than those citations to the 82 retracted articles alone. However, the morphology of the control group's citation curve reflected the citation pattern expected for unretracted articles that appeared in the same years in the same journals in which the retracted works appeared. The curves for citations to the retracted and control articles were compared (Fig 1) by matching the number of citations received during the first year after publication (when few of the retracted articles were yet to be retracted), thereby placing the control curve on the same scale as the retracted curve. This allowed the morphology of the control curve to estimate the typical citation pattern for 82 unretracted articles published in the same years in the same journals as our 82 retracted articles.

Impact factors of journals are calculated by dividing the total number of citations received over the last 2 years by the number of source items published by the journal in the same period. Impact factors and impact factor ranks for the citing journals were obtained from the 1987 SCI Journal Citation Reports ranking list. [[9]sup.(pp81-120)]

RESULTS

We identified 82 original scientific articles published from 1973 to 1987 that were completely retracted between 1977 and 1988. These articles received 1835 citations through February 1989. Of these citations, 62 were from retractions themselves, 16 were from articles on fraud, and 1024 were made before the articles were retracted in reference to scientific concepts. The remaining 733 citations were used in reference to scientific concepts but were made after the studies were retracted. Because the average month of retraction was June and citations were considered postretraction beginning the next calendar year, our analysis allowed a 6-month average washout period for citing articles already in the publishing process at the time of the retraction. Original articles accounted for 637 (86.9%) of the 733 postretraction citations, while reviews cited 80 (10.9%). Editorials, abstracts, and letters to the editor were responsible for 10, 5, and 1 citations, respectively.

The 21 retracted articles randomly selected for manual review of their citations were cited 248 times after retraction, and 178 (72%) of these citations were in journals available in our medical center library. None of the 175 scientific citations or the 3 citations from articles on fraud (based on information from SCISEARCH) proved to be misclassified. Of the 175 scientific citations, 170 (97.1%) cited only the retracted study, while 5 (2.9%) cited both the retracted work and its retraction.

Comparison of the predicted (unretracted articles) cumulative citation curve and the actual (retracted articles) citation curve is shown in Fig 1. The ideal curve represents the number of cumulative citations for the retracted articles if no citations had been made following retraction. Retracting the articles reduced the number of postretraction citations by 35% from the number predicted by the controls. Citations continued to accumulate well after the majority of articles were retracted, with 552 (75.3%) occurring more than a year after the retraction. The total number of citations before and after retraction was reduced 18%.

The impact factor ranks for journals citing the retracted articles (Fig 2) revealed that the 733 postretraction scientific citations came from journals over the complete impact spectrum, from first to 3752nd of over 4100 journals ranked. The average impact factor represented a rank of 245th. A total of 475 (65%) citations came from journals ranking in the top quartile, with 9 citations made from journals in the exclusive top 10.

Because the majority (85%) of our retracted articles were from US journals, non-US journals and authors might have been unaware of their subsequent retraction. However, of the 301 citing journals, 136 (45.2%) were published in the United States, while 127 (42.2%) were European, 2 (0.7%) were Canadian, 7 (2.3%) were Asian, 12 (4.0%) were from other countries, and 17 (5.6%) were of indeterminate origin. The geographic location of citing authors is shown in Fig 3. While there was a shift after retraction away from US authors, they remained the dominant source of the citations, accounting for almost half> US, Canadian, and European authors together were responsible for 85% of the citations made after retraction.

COMMENT

We found that articles citing invalid, retracted work are abundant and ubiquitous in the scientific literature. The 82 retracted studies averaged nine scientific citations each after retraction. While retraction does have some impact on the future citation rate, about two of every three citations expected for these studies, if valid, are still made despite their retraction. Invalid work is not being effectively purged from the scientific literature.

The 733 scientific citations made after retraction were not the result of articles being too far along in the publishing process to be corrected when retractions were announced. By defining "postretraction" as beginning the calendar year after retraction, a 6-month washout period was allowed on average for studies about to be published. In addition, the postretraction citations were not the result of simple delays before authors became aware the works were retracted, with most citations occurring years after the retractions. The number of citations an article receives in a given year appears to be more dependent on when it was published than whether it was retracted.

Is the typical scientist vulnerable to reading articles citing invalid work? Studies citing such work were not found only in small, lesser known journals. Instead, they were from journals distributed along the full spectrum of impact factors, from well-known, influential journals to obscure, rarely cited journals. While the specific journals that cite a work are largely determined by their shared subject matter, more than half of the citations to our 82 retracted studies were from journals ranking in the top quartile of impact factors. Likewise, non-US journals or non-US-based authors, perhaps less aware of the news of retraction than their US counterparts, were not found to be largely responsible for the problem. Not only were US journals and authors the single greatest source of studies citing retracted works, but it follows that US scientists are perhaps the most likely to read and use such studies.

We have identified several possible reasons retracted information continues to be used. No reliable source exists to locate fraudulent or erroneous work, and the majority of our retractions were not indexed by their own journals. In addition, journals lacked a consistent approach to the format of retractions, which further impaired their indexing by bibliographic databases. We found retractions as letters to the editor, dramatic full-page statements, and in small notices at the back of issues. Some did not include the word "retraction," using phrases such as "withdrawal" [10,11] or "questionable" [11,12] or "not reliable." [13] For 12 journals we could not locate any type of retraction and identified their retracted works only through notices in other journals.

Bioscience libraries, because of a lack of both information and available methodologies, have not historically addressed the issue of information removal. The National Library of Medicine added a "retraction of publication" subject heading to Index Medicus only in 1984, and since that time indexing has proved erratic, most likely because of the above-mentioned variances in retraction terminology and format. We found only 46 (56%) of our 82 retractions listed in Index Medicus. In 1988 the MEDLINE database began retroactively tagging articles that have been retracted, using the subject heading "retracted publication," but because the National Library of Medicine requires a printed statement of retraction signed by the article's author or the journal editor, some full retractions remain untagged. [14] It is not surprising that, if those attempting to identify retracted works are finding it challenging, the typical author, perhaps having never considered this issue, could unknowingly forward retracted data and ideas as valid fact.

However, the citing of retracted papers does not seem entirely the result of a lack of information. One year after Gullis and Rowe [15] published a study on neurotransmitters, Robert Gullis [13] admitted in a letter that the "values published are mere figments of my imagination." While these events would no doubt trouble an editorial staff, the original article was cited by the same journal a year after the letter was published. [16] The John Darsee affair, in which numerous reports of human and animal studies were found to contain fraudulent data, caused an uproar in academic cardiology and drew wide coverage in both scientific and lay media. [2,17] Yet, despite this, Darsee's 14 completely retracted articles were subsequently cited 123 times after their retraction. These scenarios seem more likely to result from a lack of attention to manuscripts than from a lack of information regarding retractions.

Because authors, editors, reviewers, and librarians are all involved, a multi-faceted approach will be required to address the problem of continued use of invalid data and ideas. Several steps have already been taken. In 1987 the International Committee of Medical Journal Editors produced and published guidelines for the format and terminology of retraction notices. [18] While it is too early to judge their success, none of the retraction notices for the four articles we found retracted since their inception have met those criteria. As mentioned, the National Library of Medicine has also recently implemented changes that should improve indexing of retracted works.

Driven by the desire for professional achievement, researchers are producing a tremendous number of studies, which may impair adequate attention to their detail [1,3] and makes it likely that journals will continue to receive manuscripts citing invalid information. Editors, then, appear well positioned to play a key role in limiting its use.

In addition to improving the consistency of retraction notices, journals should consider listing retractions in their own cumulative indexes. Editors could also offer more explicit expectations of their authors regarding the accuracy and integrity of their references and reexamine their in-house and external review processes.

The scientific community has yet to agree on the level of inaccuracy required to mandate a retraction vs an erratum or on who may retract an article. In 1985 an article evaluating the results of a trial of tretinoin in 22 patients stated, "The present study represents the first demonstration of the efficacy of topical retinoids in treating dry-eye disorders," and noted, "All patients demonstrated clinical improvements." [19] A recent letter from the authors, [20] labeled as a correction, requested the deletion of 5 patients for whom records could not be bound> noted a sixth was treated with etretinate, not tretinoin> and observed that 3 patients who were treated without improvement were not mentioned. Some of our 82 retracted papers contained lesser inaccuracies, and there is little agreement on what requires a retraction. In another controversy, the work of Stephen Breuning has highlighted issues regarding who controls a study's destiny after publication and whether coauthors, funding agencies, or editors may retract. [4]

Several methodological aspects of our study warrant further discussion. While all located complete retractions were included, the cohort of 82 articles is likely incomplete and may be especially biased in omitting non-English language publications. However, the two electronic databases we used, MEDLINE and SCISEARCH, are international in scope, indexing over 3300 and 3800 journals, respectively. It is also impossible to predict accurately the number of citations our retracted articles would have received if they were not retracted. Thus, the control group's citations and the 35% noted reduction after retraction can be considered only rough estimates. In addition, about 3% of the works citing retracted studies cited the subsequent retraction as well and would be considered appropriate use of the literature. Extrapolated to our complete results, this still leaves over 700 inappropriate citations after retraction. Finally, we cannot be sure that the modest shift in the citing author's location away from the United States after retraction does not occur routinely. It is possible that non-US authors would cite these predominantly US studies later after publication than US authors, regardless of their retraction status.

It has been said that every study will eventually be published if the author is persistent. If so, the road that beckons so freely is clearly one way. Methods currently in place to remove invalid literature from use appear to be grossly inadequate. Regardless of strides made in controlling fraud, error is generally considered an inherent and inevitable aspect of research, and efficient removal of invalid information from the literature would serve science well.

This work was supported by a grant from the University of Louisville (Ky) School of Medicine Research Committee.

References

[1.] Stossel TP. Volume: papers and academic promotion. Ann Intern Med. 1987>106:146-149.

[2.] Knox RA. Deeper problems for Darsee: Emory probe. JAMA. 1983>249:2867-2876.

[3.] Engler RL, Covell JW, Friedman PJ, Kitcher PS, Peters RM. Misrepresentation and responsibility in medical research. N Engl J Med. 1987>317:1381-1389.

[4.] Freedman DX, ed. Request for retraction. Arch Gen Psychiatry. 1988>45:685-686.

[5.] Holden C. NIMH review of fraud charge moves slowly. Science. 1986>234:1488-1489.

[6.] Broad WJ. Coping with fraud. Science. 1982>215:479.

[7.] Petersdorf RG. A matter of integrity. Academic Med. 1989>64:119-123.

[8.] The Serials Directory: An International Reference Book. 3rd ed. Birmingham, Ala: Ebsco> 1988-1989.

[9.] Garfield E, ed. SCI Journal Citation Reports. Philadelphia, Pa: Institute for Scientific Information> 1987:5a-9a, 81-120, 1501-2399.

[10.] Schooler NR, ed. Notification to readers. Psychopharmacol Bull. 1987>23:587-588.

[11.] Siegelman SS, ed. Notice to the readership. Radiology. 1986>158:279.

[12.] Felig P. Data on insulin binding in growth hormone deficiency are questionable. N Engl J Med. 1980>303:1120.

[13.] Gullis RJ. Letter to the editor. Nature. 1977>265:764.

[14.] Lindberg DA. Retraction of research findings. Science. 1987>235:1308.

[15.] Gullis RJ, Rowe CE. Pharmacological studies on the stimulation of the phospholipase A2-acylation system of synaptic membranes of brain, by neurotransmitters and other agonists. J Neurochem. 1976>26:1217-1230.

[16.] Webb JG, Saelens DA, Halushka PV. Biosynthesis of prostaglandin E by rat superior cervical ganglia. J Neurochem. 1978>31:13-19.

[17.] Culliton BJ. Coping with fraud: the Darsee case. Science. 1983>220:31-35.

[18.] International Committee of Medical Journal Editors. Retraction of research findings. Ann Intern Med. 1988>108:304.

[19.] Tseng SCG, Maumenee AE, Stark WJ, et al. Topical retinoid treatment for various dry-eye disorders. Ophthalmology. 1985>92:717-727.

[20.] Maumenee AE, Stark WJ, Maumenee IH, Jensen AD, Green WR. Correction. Ophthalmology. 1989>96:730.