# **Assessing the impact of retraction on the citation of randomized controlled trial reports: An interrupted time-series analysis**

**ABSTRACT**

**Objective**: To assess the impact of retraction on the citation of randomized controlled trials.

**Methods:** We used an interrupted time-series with matched controls. PubMed, CINHAL, Google and the Retraction Watch Database were searched. We identified retracted publications reporting the results of randomized controlled trials involving human participants with two years of available data before and after retraction. We obtained monthly citation counts across all articles for the 24 months before and after retraction, from Web of Science. We used a Poisson segmented regression to detect changes in the level and trend of citation following retraction. We also undertook a matched control analysis of unretracted randomized controlled trials and a sensitivity analysis to account for cases of large-scale, well-advertised fraud.

**Results**: We identified 387 retracted randomized controlled trial reports, of which 218 (56.3%) were included in the interrupted time-series analysis. A reduction of 22.9% (95% CI 4.0% to 38.2%, p=0.02) was observed in the number of citations in the month after retraction, and a further reduction of 1.9% (95% CI 0.4% to 3.5%, p=0.02) per month in the following 24 months, relative to the expected trend. There was no evidence of a statistically significant reduction among the matched controls. Authors with a large number of retractions saw a 48.2% reduction at the time of retraction (95% CI 17.7% to 67.3%, p=0.01). Other cases had a more gradual reduction with no change at the time of retraction and a 1.8% reduction per month in the following 24 months (95% CI 0.2% to 3.4%, p=0.03).

**Conclusions**: Retractions of randomized controlled trial reports can be effective in reducing citations. Other factors, such as the scale of the retractions and media attention, may play a role in the effectiveness of the reduction.

Keywords: Retraction, interrupted time-series, randomized controlled trials

**INTRODUCTION**

Retraction is the withdrawing of research from the literature when it is not valid. The Committee on Publication Ethics (COPE) retraction guidelines, which are endorsed by many journals, provide three reasons why an article should be retracted: i) the published results are unreliable due to misconduct or error; ii) the results have been previously published; and/or iii) the research reported is unethical.1 The number of retractions in the scientific literature has been rising,2 with the field of medicine having one of the highest rates of retraction amongst the scientific disciplines.3 This has been attributed to greater scrutiny and better understanding of retractions as a method of correction.4

Randomized controlled trials (RCTs) are the primary experimental method for establishing the effectiveness of interventions and informing systematic reviews and policy. It is therefore important that when articles reporting the results of RCTs require retraction, there is a process that it is effective in depressing the dissemination of the results. There are a number of systematic reviews that have been undermined by fraudulent work.5-7 In recent years, there have been a number incidences of large numbers of RCTs by the same author being retracted for fabrication (for example, Yoshitaka Fujii,8 Joachim Boldt,9 Scott Reuben,10 and Yoshihiro Sato11) as well as many individual retractions. Fujii and Boldt are of particular interest as between them they have over 200 retracted studies, the retractions of which received considerable media attention.12-16

One metric for measuring dissemination is the number of citations that an article receives. Previous studies have found a reduction in the number of citations an article receives following its retraction of 65-69% after one year.17, 18 However, these have used nearest-neighbour controls, where the nearest article by page number is matched with a retracted article, and have not considered the effect the study design may have on citation or wider practice. This is important as published meta-analyses and RCTs have been shown to be cited more than other study designs.19 Case studies have suggested that retracted RCTs continue to be cited, especially if they were the only available evidence for that particular intervention.20, 21

This study aimed to assess how retraction affected the citation of published RCTs for a period of two years after retraction.

**METHODS**

**Literature Search**

We searched PubMed, CINHAL, Google, and a beta version of the Retraction Watch Retractions Database (v.1.0.5.0) for all published and retracted RCTs up to August 2017. We used terms specific to each database and where relevant identified and used investigation documents to identify further retracted publications, such as the report published by the Japanese Society of Anesthesiologists Special Investigation Committee regarding Yoshitaka Fujii. Articles were eligible for inclusion if they: i) reported results from a randomized controlled trial involving human participants; ii) had been retracted; and iii) were available in English. Retraction was identified by the presence of a retraction notice or watermark. Those marked as withdrawn were also considered as retracted. For each retracted RCT, publication characteristics, including date of publication and retraction and citation data were collected.

**Interrupted time-series**

An interrupted time-series analysis was used to assess how the rate of citation changed at the time of retraction and in the following 24 months. Studies were eligible for the interrupted time-series analysis if two years of citation data were available before and after retraction. Twenty-four pre-retraction time-points were used as it has been shown that studies with fewer time-points are unlikely to have sufficient power to detect even modest effect sizes.22

A full record of all citations was downloaded from Web of Science for each article, which was identified by PubMed identifier or by title, date of publication, journal of publication, and author names. The number of citations received in each calendar month was calculated and entered into a standardized Excel spreadsheet for the 24 months before and after retraction. Where details of the month of publication were unavailable in the Web of Science record for a citation, other details from the record were used to identify the original article and where possible the date of publication. Any citations of the retracted article included in the retraction notice were removed as these would artificially increase the number of citations in the month of retraction. The number of citations across all articles was totalled for each month relative to retraction at month 0.

A matched control article was identified for each of the articles eligible for the interrupted time-series analysis. This was the closest available non-retracted RCT by page number, in the same journal, published within three months of the retracted article. Where more than one retracted RCT was present in the same issue of a journal, a separate matched control was identified for each.

*Analysis*

Analyses were undertaken in Stata v14.2, using two-sided statistical tests at the 5% significance level. Poisson segmented regression analysis was used to test the null hypothesis that there would be no change in the level or trend of citations following retraction.23 The model included a time variable (month relative to retraction), an intervention variable (an indicator variable for retraction), and a time-by-intervention interaction term.

Due to an inability to confirm the exact day of retraction of the RCTs and day of publication of citations, a one-month time lag was allowed for by excluding the data in the month of retraction from the analysis to avoid misclassification of these citations.24 It was anticipated that some citing articles may remain within the publication process at the time of retraction and would be published without opportunity to acknowledge the retraction, however it was anticipated that publication of these articles would occur gradually and that this would be captured by a change in trend.

Where the citation data were over-dispersed, a scaling factor estimated by the Pearson Chi-squared statistic divided by the residual degrees of freedom was applied.25 As observations closer in time may have error terms which are more correlated than with distant observations, an assessment of autocorrelation was made using a plot of the residuals over time and autocorrelation and partial autocorrelation function plots.23 Where significant first-order autocorrelation was present, a model including an autoregressive term was applied.26 In the presence of any outlying data points an analysis was conducted omitting these to assess the impact on the model.24 This method was repeated for the control articles.

As the fraudulent cases of Fujii and Boldt were well-advertised, due to the scale of their fraud, these articles may have a larger reduction in citation after retraction due to the media attention rather than the retraction itself. This may result in an overestimation of the effect in the main analysis. Therefore, two sensitivity analyses were undertaken, considering Fujii and Boldt’s articles on their own, using the same method as the main analysis.

A flowchart of the articles included in each analysis is presented in Figure 1.

**RESULTS**

**General characteristics**

We identified 387 retracted RCTs, published between March 1983 and February 2017, of which 383 were indexed on PubMed, and a further three had the associated retraction notice indexed but not the original article. Of the 383 available on PubMed, 362 (94.5%) were marked as retracted on the PubMed Record with a link to the retraction notice. A date of retraction was available for 376 studies. The retractions were published between July 1988 and August 2017. The mean time to retraction was 82.0 months (standard deviation (SD) 68.5); however, the data were positively skewed with a median of 65.5 months. Citation data were available for 369 (95.3%) articles. The mean total number of citations was 28.4 (SD 52.9, range 0-658).

The articles were identified in 48 publishers and 156 author groups. Yoshitaka Fujii had 101 retracted RCTs and Joachim Boldt had 72 retracted RCTs, accounting for 26.1% and 18.6% of the total number of retractions, respectively. There were 139 author groups with only a single retraction, accounting for 35.4% of the total sample.

**Interrupted time-series**

There were 218 articles eligible for the main interrupted time-series analysis. A total of 2,178 citations were identified across all articles over the four years, a further 48 citations could not be attributed to a specific month and so were omitted. The citation data were over-dispersed with a mean of 44.1 and a variance of 260.6. The results of the analysis can be seen in Figure 2 and Table 1. There was no evidence of first order autocorrelation. There was a reduction of 22.9% (95% CI 4.0% to 38.2%, p=0.02) in citations in the first month after retraction compared to the expected level, followed by an increased predicted reduction of 1.9% per month (95% CI 0.4% to 3.5%, p=0.02) in the following 24 months.

*Matched Control Articles*

A control RCT could be found within the same issue as the retracted article in 190 (87.2%) cases and within three months for 24 (11.0%) cases. In total, there were 2,034 citations that could be identified over the four years, a further 23 citations could not be attributed to a specific month and so were omitted.

The results of the analysis can be seen in Table 2. There was evidence of first-order autocorrelation in the control articles, after correction with a model including a significant autoregressive term there remained no significant changes in the level or trend of citations from the expected trend.

*Sensitivity analysis*

Of the 218 articles in the full analysis, 89 (40.8%) were published by Fujii and 65 (29.8%) by Boldt, leaving 64 articles published by other groups. The 154 Fujii and Boldt articles received 795 citations over the four-year period (mean 5.2 per article), whereas the other articles received 1,387 citations (mean 21.7 per article). The citation data were overdispersed in both cases, the Fujii and Boldt articles had a mean of 15.9 and a variance of 66.1 and the other articles had a mean of 28.1 and a variance of 95.3. The Fujii and Boldt data appeared to have an outlier 14 months after retraction. The results of the segmented regression are presented in Figures 3 and 4 and Table 3. There was no evidence of first order autocorrelation.

There was evidence that the Fujii and Boldt articles had a statistically significant reduction in monthly citations at the time of retraction of 48.1% (95% CI 17.7% to 67.3%, p=0.01) reduction. However, there was no significant change in the trend over time after retraction. The other articles saw no significant reduction in citations in the month after retraction, however the month-to-month trend showed a decrease of 1.8% per month (95% CI 0.2% to 3.4%, p=0.03) relative to the expected trend. Removal of the outlying point in the Fujii and Boldt data had a negligible impact on the results.

**DISCUSSION**

**Main Findings**

A systematic search of the literature identified a large number of retracted RCTs. Articles attributable to two authors (Fujii and Boldt) heavily influenced the sample, their articles accounted for 44.7% of the sample. We identified that following retraction there was an immediate reduction in the number of citations received by these retracted publications, and the rate of citations continued to decrease in the following two years. This effect was observed in the retracted RCTs and not in a paired set of control RCTs. Further investigations identified that articles that were part of large-scale retractions and received broad media attention had a larger and sustained reduction in the number of citations. Articles which did not receive such attention had no immediate reduction; however, the number of citations received declined, more than would be expected, over the two years following retraction.

In our study, the reduction in citations after one year was 45.7% from the expected trend. Previous studies have suggested reductions of 65-69% after one year.17, 18 Both Furman et al.17 and Azoulay et al.18 used difference-in-difference approaches to assess annual changes in the citation of retracted articles, they also utilized a broader sample of any retraction with different controls. The differences in design and sample make it difficult to draw direct comparisons between our results. However, our sensitivity analysis showed that the reduction in citations is smaller in articles where the case was not large-scale or well-advertised. Therefore, the estimates for the effect of retraction on citation in other studies may be overestimated as a result of external factors, such as media attention. We used monthly counts in this study, which allowed us to show the immediacy of an effect on citations. Furman et al.17 used a sample derived from articles retracted before 2007 and Azoulay et al.18 used articles retracted up to 2009. Therefore, only 12.5% of our sample could have been included in these studies.

**Limitations**

The interrupted time-series design is a strong quasi-experimental approach that has not previously been applied in this context. It cannot however prove causality and other factors may play a role in the number of citations received; we employed the use of a matched control analysis and sensitivity analysis to attempt to control for some of these factors. RCTs that had been retracted within the two years prior to data extraction or that were retracted within two years of publication were excluded as they did not have enough data available for analysis. These articles may have a different profile of citation as they have had less time to be disseminated. This could result in fewer citations following retraction as the lack of dissemination prior to retraction reduces erroneous citation, thus our analysis may underestimate the overall effect of retraction. Furman et al.17 also suggest that more recent retractions are more effective than older retractions, therefore those more recent retractions that were excluded may see a greater reduction in citations than those included in our analysis. As the data were aggregated in a standard time frame it was not possible to consider any seasonality to the data, however publishing of articles generally occurs throughout the year and is unlikely to vary with season.

Number of citations is only one metric used to quantify the dissemination and reach of a research article, and its usefulness is a matter of debate. The nature and reason for the citation of an article can be positive, negative, or neutral. It has been estimated that around 2.4% of citations are negative within normal articles,27 however, one study found that amongst post-retraction citations 16% discussed the article favourably, 78% were neutral in their use of the article, and 94% made no reference to the retraction.28 The citation of an article is also not necessarily related to its impact on practice, there are many other metrics which could be used, such as readership or use in policy or guideline documents.29 It should also be noted that as citations had to be collected by hand there is a possibility of minor error in the data collection. The use of a language restriction was necessary to confirm the inclusion criteria for this study. However, there were some articles found in searching that may have met all the other criteria.

**Policy and Further Research**

The results of this study confirm the need for timely retraction, as reiterated by the COPE guidelines.1 Prolongation of the time to retraction, once a justification for retraction has been identified, can stem from journals and from institutions. Reasons that other studies have found for lack of retraction from journals have been the inability to validate the accusations, with few available methods of investigation, communication problems between involved parties, and a lack of appropriate standard guidelines that could be followed.30 However, it has also been reported that some journals are unwilling to even consider accusations of misconduct,31 which would allow citation to continue at its normal rate. Institutions also vary in their policy for investigations.32 Making retraction timely will ensure that citations do not unnecessarily continue at a normal rate. One solution to continued citation is systematic screening of citations by journals. Whilst doing this by hand would be time-consuming there is potential to link it to automated systems already in place within many journals.33

Further research is required to determine the context in which these articles are cited, before and after retraction, as this may explain some of the ongoing citations. Research should also consider what changes can be made to increase the effectiveness of a retraction notice in reducing citations, especially in the months immediately after retraction.

**CONCLUSIONS**

Our study offers evidence that the retraction of RCTs is effective in reducing citation of these articles. This reduction is smaller in articles that are not part of a large-scale retraction and do not receive broad media attention. Further research should focus on improving the effectiveness of retraction notices and on the effect these studies can have on practice.

**REFERENCES**

1. Wager E, Barbour V, Yentis S, et al. Retraction guidelines. Committee on Publication Ethics. 2009, https://publicationethics.org/files/retraction%20guidelines\_0.pdf.

2. Cokol M, Ozbay F and Rodriguez-Esteban R. Retraction rates are on the rise. *EMBO Rep*. 2008; 9: 2.

3. Grieneisen ML and Zhang M. A comprehensive survey of retracted articles from the scholarly literature. *PLoS One*. 2012; 7: e44118.

4. Fang FC, Steen RG, Casadevall A. Misconduct accounts for the majority of retracted scientific publications. *Proceedings of the National Academy of Sciences*. 2012; 109: 17028-33.

5. Carlisle JB. A meta-analysis of prevention of postoperative nausea and vomiting: randomised controlled trials by Fujii et al. compared with other authors. *Anaesthesia*. 2012; 67: 1076-90.

6. Marret E, Elia N, Dahl JB, et al. Susceptibility to fraud in systematic reviews: lessons from the Reuben case. *Anesthesiology*. 2009; 111: 1279-89.

7. Zarychanski R, Abou-Setta AM, Turgeon AF, et al. Association of hydroxyethyl starch administration with mortality and acute kidney injury in critically ill patients requiring volume resuscitation: a systematic review and meta-analysis. JAMA. 2013; 309:678-88.

8. Tramer MR. The Fujii story: a chronicle of naive disbelief. *Eur J Anaesthesiol*. 2013; 30: 195-8.

9. Wise J. Boldt: the great pretender. *BMJ*. 2013; 346: f1738.

10. Shafer SL. Tattered threads. *Anesth Analg*. 2009; 108: 1361-3.

11. Bolland MJ, Avenell A, Gamble GD, et al. Systematic review and statistical analysis of the integrity of 33 randomized controlled trials. *Neurology*. 2016; 87: 2391-402.

12. Akst J. Anesthesiologist fabricates 172 papers. The Scientist, 3 Jul 2012, https://www.the-scientist.com/the-nutshell/anesthesiologist-fabricates-172-papers-40780.

13. Royal College of Anaesthetists. Dr Joachim Boldt. News and Statements: Royal College of Anaesthetists, 2011., https://www.rcoa.ac.uk/news-and-bulletin/rcoa-news-and-statements/dr-joachim-boldt (accessed [add date])

14. Torjesen I. Journals question integrity of almost 200 papers by Japanese anaesthetist. *BMJ*. 2012; 344: e2490.

15. Blake H, Watt H, Winnett R. Millions of surgery patients at risk in drug research fraud scandal. *The Telegraph*, 3 March 2011 , https://www.telegraph.co.uk/news/health/8360667/Millions-of-surgery-patients-at-risk-in-drug-research-fraud-scandal.html.

16. Cyranoski D. Retraction record rocks community. *Nature*. 2012; 489: 346-7.

17. Furman JL, Jensen K, Murray F. Governing knowledge in the scientific community: Exploring the role of retractions in biomedicine. *Research Policy*. 2012; 41: 276-90.

18. Azoulay P, Furman JL, Krieger JL et al. Retractions. *Review of Economics and Statistics*. 2015; 97: 1118-36.

19. Patsopoulos NA, Analatos AA, Ioannidis JP. Relative citation impact of various study designs in the health sciences. *JAMA*. 2005; 293: 2362-6.

20. Fulton A, Coates A, Williams M, et al. Persistent citation of the only published randomised controlled trial of omega-3 supplementation in chronic obstructive pulmonary disease six years after its retraction. *Publications*. 2015; 3: 17-26.

21. Bornemann-Cimenti H, Szilagyi IS, Sandner-Kiesling A. Perpetuation of retracted publications using the example of the Scott S. Reuben case: Incidences, reasons and possible improvements. *Sci Eng Ethics*. 2016; 22: 1063-72.

22. Zhang F, Wagner AK, Ross-Degnan D. Simulation-based power calculation for designing interrupted time series analyses of health policy interventions. *Journal of Clinical Epidemiology*. 2011; 64: 1252-61.

23. Bernal JL, Cummins S, Gasparrini A. Interrupted time series regression for the evaluation of public health interventions: a tutorial. *Int J Epidemiol*. 2017; 46: 348-55.

24. Wagner AK, Soumerai SB, Zhang F, et al. Segmented regression analysis of interrupted time series studies in medication use research. *Journal of Clinical Pharmacy and Therapeutics*. 2002; 27: 299-309.

25. Bhaskaran K, Gasparrini A, Hajat S, et al. Time series regression studies in environmental epidemiology. *Int J Epidemiol*. 2013; 42: 1187-95.

26. Tobias A, Campbell MJ. Time series regression for counts allowing for autocorrelation. Stata Technical Bulletin. 1998; 46: 33-37 (https://www.stata.com/products/stb/journals/stb46.pdf)

27. Catalini C, Lacetera N, Oettl A. The incidence and role of negative citations in science. *Proc Natl Acad Sci U S A*. 2015; 112: 13823-6.

28. Budd JM, Coble ZC, Anderson KM. Retracted publications in biomedicine: Cause for concern. *Declaration of Interdependence: The Proceedings of the ACRL 2011 Conference, March 30-April 2, 2011, Philadelphia, PA*. 2011: 390-5.

29. Guthrie S, Krapels J, Lichten CA, et al. 100 Metrics to assess and communicate the value of biomedical research. *Research Reports*. Association of American Medical Colleges, 2016, p. 71.

30. Williams P, Wager E. Exploring why and how journal editors retract articles: findings from a qualitative study. *Sci Eng Ethics*. 2013; 19: 1-11.

31. Koziol M. A shadow was cast on a bone researcher's work. What are journals doing about his papers? *Retraction Watch* https://retractionwatch.com/2017/04/25/shadow-cast-bone-researchers-work-journals-papers/ (2017, accessed September 2017).

32. Resnik DB, Rasmussen LM, Kissling GE. An international study of research misconduct policies. *Account Res*. 2015; 22: 249-66.

33. Cosentino AM, Verissimo D. Ending the citation of retracted papers. *Conserv Biol*. 2016; 30: 676-8.