An investigation into the impact and implications of published papers from retracted research: systematic search of affected literature

Alison Avenell, Fiona Stewart, Andrew Grey, Greg Gamble

ABSTRACT

Objective Analyses of the impact of a body of clinical trial reports subject to research misconduct have been few. Our objective was to examine the impact on clinically relevant research of a group of researchers’ trial reports (‘affected trial reports’) affected by research misconduct, and whether identification of misconduct invoked a reappraisal.

Design In 2016, we used five databases and search engines to identify ‘citing publications’, that is, guidelines, systematic and other reviews, and clinical trials citing any of 12 affected trial reports, published 1998–2011, eventually retracted for research misconduct. The affected trial reports were assessed more likely to have had impact because they had hip fracture outcomes and were in journals with impact factor >4. Two authors assessed whether findings of the citing publications would change if the affected trial reports were removed. In 2018, we searched for evidence that the citing publications had undertaken a reassessment as a result of the potential influence of the affected trial reports.

Results By 2016 the affected trial reports were cited in 1158 publications, including 68 systematic reviews, meta-analyses, narrative reviews, guidelines and clinical trials. We judged that 13 guidelines, systematic or other reviews would likely change their findings if the affected trial reports were removed, and in another eight it was unclear if findings would change. By 2018, only one of the 68 citing publications, a systematic review, appeared to have undertaken a reassessment, which led to a correction.

Conclusions We found evidence that this group of affected trial reports distorted the evidence base. Correction of these distortions is slow, uncoordinated and inconsistent. Unless there is a rapid, systematic, coordinated approach by bibliographic databases, authors, journals and publishers to mitigate the impact of known cases of research misconduct, patients, other researchers and their funders may continue to be adversely affected.

BACKGROUND

We raised concerns about 33 randomised controlled trial (RCT) reports, ‘affected trial reports’, from one research group in Japan (see online supplementary appendix for list of 33 RCTs). Our systematic review published in November 2016 examined these affected trial reports published in the field of osteoporosis over 15 years. The affected trial reports ostensibly involved large numbers of older patients with significant comorbidities, such as stroke, Alzheimer’s disease and Parkinson’s disease. In September 2016, the editor of the journal that published our systematic review conveyed the results of its investigations to all the journals with affected trial reports. By May 2019, 27/33 of these affected trial reports had been retracted for reasons including fabrication, plagiarism, authorship misconduct and unresolved concerns about data integrity.

Retraction of a research paper may have important implications for clinical practice and present and future research initiatives. Patients and research participants may be put at risk if decisions are based on findings that are later retracted because they were incorrect or unreliable. It is therefore important to determine the extent of a retracted paper’s influence, for example, through citations in other influential publications, such as systematic reviews and guidelines, and its use in
initiating new research. There is evidence that authors of publications that cite retracted work remain unaware of the retraction, and this has potentially important consequences for their work, that of subsequent researchers, and for clinical practitioners and patients.

Analyses of the impact of a body of clinical trial reports subject to research misconduct have been few. Our objective was to examine the impact and influence of a selection of the published affected trial reports most likely to affect clinical guidance and practice and further research. We focused on affected trial reports with hip fracture outcome data in influential journals.

METHODS

Search criteria

We studied the impact of a subgroup of the 33 affected trial reports whose integrity was analysed in our systematic review. This subgroup of trial reports was used because these trials had hip fracture as an outcome, arguably the most important consequence of osteoporosis, and affected trial reports on this outcome are likely to have the greatest impact. We included all affected trial reports with hip fracture outcomes that had also been published in higher impact journals (ISI Web of Knowledge impact factor >4).

Evidence identification

In August 2016, we used Scopus and Web of Science to find citations of each affected trial report and the type of publication that cited each report ('citing publications'—guidelines, systematic and other reviews, and clinical trials). We also searched Google Scholar, PubMed and personal databases to identify systematic reviews, meta-analyses, narrative reviews and guidelines relating to hip fracture prevention, which potentially would include these affected trial reports. Finally, we sought other types of publications that cited the affected trial reports, through an iterative process, for example, using the following search command in Ovid MEDLINE: (sato.tw) and ((letter or comment$).pt). We excluded self-citing publications by authors of affected trial reports from our evaluations.

Assessment of impact

Where possible, meta-analyses which included data from affected trial reports were reanalysed to investigate whether the quantitative findings, such as summary risk ratios in forest plots, would change without the inclusion of those data. In the case of reviews in which data from affected trial reports were not included in quantitative synthesis, we used our judgement. One investigator (FS) initially assessed all citing publications for the influence of affected trial reports, which were then discussed in depth with a second investigator (AA). Agreement was reached between AA and FS on all affected publications, apart from two where AG and MB provided input leading to consensus. We categorised affected publications according to the likelihood of a change in findings if the affected trial reports were excluded:

1. Findings likely to change.
2. Uncertain if findings would change.
3. Findings unlikely to change.

In November 2018, we searched again Web of Science, Scopus or guideline websites to see if the affected systematic reviews, meta-analyses, narrative reviews and guidelines, identified in August 2016, had published any notice, update, correction or retraction on publishers’ websites resulting from recognition that the publication was potentially influenced by the affected trial reports. We searched Web of Science, or Scopus if not included in Web of Science, to identify the number of times the citing publications we had judged likely or possibly to have been influenced by affected trial reports had themselves ever been cited, and the date of the most recent citation.

In July 2019, we searched Web of Science for any publication that cited the affected trial reports after they had been retracted, to examine whether these publications mentioned that the affected trial reports had been retracted.

RESULTS

Twelve trial reports from the original 33 were identified by us for evaluation. These 12 affected trial reports all had hip fracture outcomes and were published between 1997 and 2011 in journals with impact factors>4, with 3182 reported participants (table 1).6–17 They were published in journals with a median impact factor of 5.8 (range from 4.5 to 30). All 12 affected trial reports were retracted between June 2016 and April 2019, but by July 2019 only 7 (58%) were marked as retractions on both Ovid Medline and PubMed, and two further affected trial reports were marked as retracted on PubMed but not on Medline.11,16

We examined 40 publications in July 2019 that cited any of the 12 affected trial reports after they were retracted. Thirty-four publications (85%) expressed no concern about the affected trial reports, and six (15%) cited the affected trial reports but discounted their findings as a result of misconduct.

Citations of affected trial reports

By August 2016, the 12 affected trial reports were cited a total of 1158 times in publications of any kind, identified by our literature searches. The median number of citations for affected trial reports was 84 (range 14 to 323).

Sixty-eight systematic reviews, meta-analyses, narrative reviews, guidelines and clinical trials cited at least one of the 12 affected trial reports. Each affected trial report was cited by a median of 11 of the 68 publications (range one to 25). Five citing publications, including Agency for Healthcare Research and Quality (AHRQ) comparative...
<table>
<thead>
<tr>
<th>Citation</th>
<th>Intervention</th>
<th>Journal impact factor</th>
<th>Control</th>
<th>Intervention</th>
<th>Times cited by any of the affected publications*</th>
<th>Google Scholar total citations</th>
<th>Hip fracture data</th>
</tr>
</thead>
</table>

Continued
### Table 1 Continued

<table>
<thead>
<tr>
<th>Citation</th>
<th>Intervention</th>
</tr>
</thead>
</table>

#Marked as retracted on both PubMed and Ovid Medline by July 2019. Retraction dates relate to online posting.

*Publications of interest: 68 systematic reviews, meta-analyses, narrative reviews, guidelines and clinical trials citing at least one of the 12 affected trial reports.

The 12 affected trial reports were included in 23 systematic reviews, meta-analyses and narrative reviews, 9 effectiveness reviews and guidelines, and 32 clinical trial reports.

### Reviews and meta-analyses

Fracture reviews and meta-analyses

Nine reviews and meta-analyses relating to hip fracture prevention were identified that cited at least one affected trial report. The findings of four were likely to change and fractures, treatment of psychiatric symptoms and the role of homocysteine in disease. Four further reviews and meta-analyses cited but did not include any data from affected trial reports in their analyses.

### Table 2 Number of reports citing affected trial reports, with assessment of impact of trial reports

<table>
<thead>
<tr>
<th>Topic</th>
<th>Number of affected publications</th>
<th>Findings likely to change</th>
<th>Unclear if findings would change</th>
<th>Findings unlikely to change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fracture reviews and meta-analyses</td>
<td>9</td>
<td>4</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Falls reviews and meta-analyses</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Other reviews and meta-analyses</td>
<td>12</td>
<td>3</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Effectiveness reviews and guidelines</td>
<td>9</td>
<td>5</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>32</td>
<td>13</td>
<td>8</td>
<td>11</td>
</tr>
</tbody>
</table>
following the removal of the affected trial reports (see table 2).18 21 22 24

Two systematic review authors did not express caution that their findings were derived from one group of investigators. The systematic review by Zhang et al25 (three citations, most recent December 2016) only included affected trial reports.21 However, the authors noted the lack of generalisability from Japanese-only populations. The systematic review by Zhao et al26 focused on hip fracture and bone mineral density outcomes in Alzheimer’s disease; affected trial reports were the only sources of bone mineral density data.

Cockayne et al28 undertook a meta-analysis of vitamin K for fracture prevention (217 citations, August 2018) which influenced Japanese osteoporosis guidelines. The reduction in hip fractures was statistically and clinically significant with an OR of 0.23 and narrow CI (95% CI 0.12 to 0.47). However, Cockayne et al28 also included a sensitivity analysis to investigate the effect of removing the three affected trial reports.8–10 This analysis changed the result to a statistically non-significant result with wide confidence intervals (OR 0.30, 95% CI 0.05 to 1.74). The reason given for conducting the sensitivity analysis was that the trial populations were from a single centre and included participants at much higher risk of fractures than other trials. The authors expressed some caution when interpreting the main findings of their review because of the uncertainty introduced by this sensitivity analysis and their conclusions—that vitamin K helps to prevent hip fractures—would be different if the affected trial reports were omitted. Importantly, this 2006 meta-analysis, without any caveat related to the sensitivity analysis, is the sole evidence cited for vitamin K preventing vertebral and non-vertebral fractures in the journal publication of the 2011 Japanese guidelines for osteoporosis23 (122 citations, October 2018). In 2018, in response to retractions, Cockayne’s group published a letter of explanation and corrected article,19 20 removing the three affected trial reports, with the revised OR for hip fracture of 0.30 (95% CI 0.05 to 1.74).

One affected trial report14 was judged to have influenced the strength of a review’s conclusions. This was a narrative review of B vitamins and bone health24 (eight citations, September 2018). The affected trial report14 showed that B vitamins significantly reduced hip fractures, contrary to the evidence cited that most studies did not demonstrate reduced fracture risk. The authors noted that the results of one affected trial report14 were unusual and speculated that improvements in neurological and cognitive function from B vitamins would prevent fall-related fractures. We judged that without the affected trial report the review’s conclusions of lack of efficacy of the intervention would be stronger.

Cases where we were unable to reanalyse meta-analyses after removal of affected trial reports would have been facilitated by authors providing open access to all their data. For four meta-analyses, it was unclear if omission of the affected trial reports would alter the findings (Richy et al25 79 citations, October 2018; Richy et al26 91 citations, February 2018; Murad et al27 26 citations, August 2018; Yang et al28 54 citations, October 2018). Clarification of the impact of the affected trial reports requires the reviews’ authors to repeat their meta-analyses with and without the affected trial reports. The citation of one affected trial report31 in the review by McCarus et al26 is little more than a passing reference and data from the trial report were not used.

Falls reviews and meta-analyses

Two affected reviews and meta-analyses related to the prevention of falls were identified, since the affected trial reports also provided data on falls.

The results from one affected trial report10 changed the findings for a combined treatment (calcium, vitamin D and vitamin K for falls prevention). One Cochrane review on the prevention of falls in the community30 (756 citations, November 2018) included an unpooled meta-analysis of data from one affected trial report10 and one other trial of calcium alone, relating to the number of fractures caused by falling. The analysis shows a large, statistically significant, reduction in fracture risk in the intervention group from the affected trial report (risk ratio 0.13, 95% CI 0.04 to 0.43), and a null effect in the other trial38 (risk ratio 0.90, 95% CI 0.69 to 1.16).

Data from two affected trial reports were included in unpooled meta-analyses in the review by Batchelor et al31 (65 citations, September 2018), in which the affected trial report data were not outlying.

Other reviews and meta-analyses

Twelve other affected reviews and meta-analyses were identified. Removing affected trial reports from three would likely alter their conclusions. The conclusion of one systematic review on interventions for osteoporosis (Hermann et al,32 65 citations, 2018) that B-vitamins were likely to reduce the risk of osteoporosis was supported by data from an affected trial report,14 which showed a reduction in hip fractures in the intervention group. The review’s authors note several limitations in the affected trial report, but commented on its ‘very promising’ results.

In their review of vitamin D and Parkinson’s disease, Peterson et al33 (16 citations, 2017) base their conclusions almost entirely on data from four affected trial reports.14 16 16 16

Three affected trial reports8 15 16 were cited in the review by Binks and Dobson34 (one citation, 2017) as evidence for the benefit of vitamin D and bisphosphonates in people with Parkinson’s disease. Although Binks and Dobson were careful to draw attention to the limitations of the trial reports, nonetheless their conclusions would be substantially different without these data.

Affectd trial reports were included in three reviews (Alibhai et al35 118 citations, October 2018; Carda et al26 24 citations, September 2018; Simpson et al37 14 citations, April 2018) where it was unclear if findings would
be altered by the omission of the affected reports’ data. The conclusions of six systematic reviews were unlikely to change if data from affected trial reports were omitted.47–48

Systematic reviews excluding affected trial reports
A further four systematic reviews cited but did not include affected trial reports in their reviews as a result of existing concerns with data,44–46 or awaiting responses to enquiries about data.47 One was a Cochrane review by one of the authors of this paper, with concerns dating back to 2006.44 Another Cochrane review, whose authors corresponded with AA, excluded trials for not fitting study inclusion criteria.45 Latham et al.46 appeared to exclude one trial8 because of its poor quality from their review of vitamin D for falls prevention and other outcomes. Verheyden et al.47 categorised two affected trial reports as awaiting assessment11,12 in their Cochrane review of falls prevention after stroke.

Effectiveness reviews and guidelines
Affected trial reports were cited in nine effectiveness reviews and clinical guidelines (one published in Scotland, the others in the USA), for stroke,49 fracture prevention,50–55 and fall and injury prevention.56 Removing these affected trial reports would likely alter findings in five reviews and guidelines.31–34,56

The effectiveness review from the US AHRQ in 2007 on fracture prevention41 (no citation count available) included six affected trial reports in their Table 56,8,11–13,15 which are the only trials cited for bisphosphonates preventing fractures in high risk falls patients. In addition, three affected trial reports11–15 are the only evidence used to support the 2.5 mg dose of risedronate for preventing hip fracture. This dose of risedronate does not have marketing approval in the USA (https://www.fda.gov/downloads/drugs/developmentapprovalprocess/ucm071436.pdf), but does in Japan (https://www.ajinomoto.com/en/presscenter/press/detail/g200907_31.html).

The publication in the Annals of Internal Medicine52 from this AHRQ review has been cited 346 times, including September 2018, and it references four of the six above-mentioned affected trial reports, with these reports being the sole sources of data evidencing the reduction in fractures from bisphosphonates in patients with Parkinson’s disease, Alzheimer’s disease or stroke. The linked guideline from the American College of Physicians45 (114 citations, March 2018) references the same four affected trial reports as evidence for bisphosphonate use in populations at increased risk of falls.

When the AHRQ review was updated in 201244 (no citation count available), it included evidence from five affected trial reports,11–15,13,15,16 with no new trials from other authors providing data for risedronate 2.5 mg/day in the prevention of hip fracture. The effectiveness review also states that this dose is equivalent to higher doses of risedronate.

A 2008 evidence-based handbook for nurses56 (no citation count available) contains the statement that risedronate is effective in preventing fractures in older women, older men who have had a stroke and older women with Alzheimer’s disease, based entirely on two affected trial reports.12,13

It was unclear whether exclusion of the affected trial reports would alter findings in one report. American stroke guidelines49 (1230 citations, November 2018) used evidence from one report14 of vitamin B12 and folate supplementation as the only evidence when discussing fracture prevention among patients with a recent ischaemic stroke. However, ‘routine’ supplementation of vitamins was not recommended, so we judged that it was unclear if findings, related to higher risk patients, would change without this one report.

Findings of three reviews were unlikely to change following exclusion of affected trial reports. The updated 2017 American College of Physicians’ guidelines55 (74 citations, October 2018) includes two of the affected trial reports on 2.5 mg daily risedronate13,16 in its overview of the evidence for the use of risedronate from the AHRQ review,54 but does not discuss the specific issue of the lower dose of risedronate. Guidelines from Scotland relating to osteoporosis and fractures50 express caution about using the affected trial report on vitamin B12 and folate supplementation14 in recommendations: ‘As this was a Japanese population that had suffered a stroke, it is not certain how relevant the findings are to a Scottish population.’ A guideline from AHRQ47 excluded one trial report6 from its review of interventions to prevent falls in older people. The reason for exclusion was that the report did not focus on the outcome of interest, that is, the rate of falls or number of fallers, despite what appeared to be relevant falls data in the affected trial report.

Trials
We identified 32 clinical trial reports (including 27 RCTs) which cited affected trial reports. In eight cases,58–65 affected trial reports contributed to the rationale for undertaking further RCTs. These RCTs are listed in table 3. Seven trials discussed one or more of the affected trial reports in their introduction sections, and five trials in their discussion sections. The strongest suggestion of influence in study design or rationale comes from the RCT by van Wijngaarden et al.,65 published in 2014, which discusses two RCTs in people at risk of cardiovascular disease or with cerebrovascular disease which had been unable to demonstrate B vitamins preventing fractures. These RCTs were contrasted with the affected trial report,13 which reported a reduction in hip fractures in stroke survivors. van Wijngaarden et al. then state that ‘Given the conflicting results and low generalizability to the general older population, further investigation is needed.’ van Wijngaarden et al’s trial randomised 2919 participants to B vitamins or placebo for 2 years, and found no treatment effect on osteoporotic fractures.63
In another eight RCTs (not shown in table 3), the authors cited affected trial reports to draw attention to the disparities between their own findings and those reported. It appeared unlikely that the affected trial reports contributed to the rationale for these trials.

**DISCUSSION**

Our analysis suggests that affected trial reports are likely to have had an adverse impact on clinical care and other research. By 2016, affected trial reports were widely cited in the published literature of particular relevance to older people with Parkinson’s disease, stroke or Alzheimer’s disease, where, despite their generally small sample size and number of events, they dominated the literature for fracture prevention. Despite recommendations for caution in deriving conclusions from data from a very limited number of authors and centres, authors of reviews that included affected trials rarely expressed caution. We were unable to identify published or registered (ClinicalTrials.gov) RCTs of bisphosphonates in these patient groups by other research groups. Thus, other researchers (and funders) may have been dissuaded from undertaking further trials by evidence from these affected trial reports. It was apparent that some systematic reviews and guidelines, particularly for the earlier three patient groups, would be different for vitamin K and risedronate 2.5 mg/day if the affected trial reports were removed, and that some affected systematic reviews and guidelines have themselves been widely disseminated. However, we do not know which parts of these systematic reviews and guidelines have been influential. With one exception, authors and/or journals of citing publications have either not identified that their publications have been compromised, or decided no action is required, although the latter seems unlikely. To our knowledge, bibliographic database/journal/publisher/guideline developer structures are not established that permit systematic identification and correction of publications that are affected by the inclusion of research with compromised integrity. Even if removing the affected trial reports did not influence their conclusions, citing authors should publish an update. This should give details of their examination of the impact of the correction or retraction on their own work, and confirm that changes are not required or have been made. This would remove uncertainty in the interpretation of their work.

Our assessment in August 2016 relates to publications up to that time. New, affected publications have continued to accumulate. We only assessed the impact of affected trial reports, because published data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports. Data in affected publications were insufficient to allow us to replicate analyses after excluding affected trial reports.

### Table 3: RCT reports in which affected publications by Sato and colleagues are included in the justification for the trial

<table>
<thead>
<tr>
<th>RCT</th>
<th>Affected trial report cited</th>
<th>Intervention, patient group and outcome</th>
<th>Sample size</th>
<th>Follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bauman 2005</td>
<td>6</td>
<td>1 alpha-hydroxyvitamin D&lt;sub&gt;2&lt;/sub&gt; for reducing bone loss in spinal cord injury patients</td>
<td>40</td>
<td>24 months</td>
</tr>
<tr>
<td>Berendsen 2013</td>
<td>14</td>
<td>Vitamins D, B&lt;sub&gt;12&lt;/sub&gt;, and folate for slowing functional decline in people over 65 years</td>
<td>1250</td>
<td>12 months</td>
</tr>
<tr>
<td>Binkley 2009</td>
<td>7 10</td>
<td>Vitamin K for bone density and biochemical markers in postmenopausal women</td>
<td>381</td>
<td>12 months</td>
</tr>
<tr>
<td>Emaus 2013</td>
<td>7 9 10</td>
<td>Vitamin K for bone density and biochemical markers in postmenopausal women</td>
<td>334</td>
<td>12 months</td>
</tr>
<tr>
<td>Grieger 2009</td>
<td>14</td>
<td>Multivitamins for improving bone quality, falls and nutritional status in care home residents</td>
<td>92</td>
<td>6 months</td>
</tr>
<tr>
<td>Hermann 2007</td>
<td>14</td>
<td>B-vitamins for bone density and biochemical markers in people with osteoporosis</td>
<td>47</td>
<td>12 months</td>
</tr>
<tr>
<td>Rucklidge 2012</td>
<td>14</td>
<td>Multivitamins and minerals for stress in adults</td>
<td>91</td>
<td>2 months</td>
</tr>
<tr>
<td>Van Wijngaarden 2014</td>
<td>14</td>
<td>Vitamins B&lt;sub&gt;12&lt;/sub&gt; and folate for preventing fractures in people ≥65 years with elevated homocysteine status</td>
<td>2919</td>
<td>24 months</td>
</tr>
</tbody>
</table>

RCT, randomised controlled trial.
and assessing impact in narrative reviews was often more challenging, as others have found.74

We only investigated affected trial reports’ impact on published research. They could also have influenced grant applications, educational events, media coverage and social media, evaluation of which require a very broad range of information sources. Most importantly, we could not directly establish the effect on patients from clinical practices informed by the unreliable research. We did not examine the impact of reviews and systematic reviews authored by the group of researchers who published the affected trial reports, which includes more than 30 reviews and meta-analyses. Such active dissemination by self-citation in cases of prolific misconduct also occurred in the Reuben and Fujii cases.74,79

We have probably missed guidelines in our evaluation of citing publications, since these are poorly covered by indexing databases. AHRQ full guidelines51 54 were identified through linked journal articles, and Scottish Inter-collegiate Guidelines Network (SIGN) guidelines from personal databases.56 Thus, we have probably underestimated the impact of these 12 trial reports.

Our findings are consistent with those of others who have investigated the impact of publications affected by research misconduct on subsequent publications and systematic reviews.74 76 80–83 In the Scott Reuben case, almost half of Reuben’s articles on perioperative analgesia were still being cited more than 5 years after their retraction,81 and his reports widely infiltrated literature in this area.84

Retractions of affected trial reports examined here started only in 2016, but concerns about research by this Japanese research group had been expressed as early as 2004–2007 by other groups, so that delays in investigation also increased the impact of this misconduct.85–89

Mott and colleagues found a 46% reduction in citations of randomised clinical trial reports in the first year after retraction,83 and retractions also reduce subsequent publication by authors associated with misconduct.90

It seems systems have not changed to mitigate the impact of misconduct, once it is identified, more than 10 years since these issues were highlighted by Sox and Rennie.76 van der Vet and Nijveen91 argued on the basis of a single, preclinical case study that indirect citations did not contribute to the propagation of research misconduct. However, for randomised trials in clinical areas affecting systematic reviews and guidelines further propagation is likely, as we show in the case of the systematic review by Cockayne et al.88 and its influence on Japanese osteoporosis guidelines.92 In the case of Fujii’s extensive publications, the effect of his misconduct on the management of postoperative nausea and vomiting appears to have only been minimised by the large volume of publications from other authors.79 In a recent paper, analyses by Fanelli and Moher93 suggested that meta-analyses may overestimate their summary effect sizes when they include studies later retracted for issues with data, methods or results.

Delays in the processes of investigating, correcting or retracting research misconduct add to the impact on patients, funders and other researchers. Delays in retraction by journals, even in response to official notification by investigating authorities, continue to be problematic and contribute to the impact of retracted work.93 Once a retraction is posted by a journal all bibliographic data, or requiring authors to check their reference lists, or ensure their guidelines are listed on bibliographic databases.

Box 1 Some possible solutions for minimising the impact of retracted research reports

- Journals and publishers should ensure that expressions of concern, retractions or corrections are appropriately flagged so that they are immediately available to be listed as such on bibliographic databases, including that of Retraction Watch, and search engines.
- Publishers should sign up to The CrossMark (https://www.crossref.org/services/crossmark/), an initiative to take readers to the current version of the paper, which should include expressions of concern, retractions or corrections.
- After institutional investigations have found that misconduct has taken place, institutions could notify corresponding, first authors and senior authors of citing publications.
- Listing an expression of concern, retraction or correction on bibliographic databases should generate automatic alerts to corresponding, first authors and senior authors of citing publications.
- Retraction Watch’s database of retractions, linked to reference management software, should be used to regularly scan researcher’s personal reference libraries.
- Journals and their publishers could help to prevent inappropriate citations by themselves checking or requiring authors to check their reference list for expressions of concern, retractions or corrections.
- Organisations responsible for publications, which are not usually listed on bibliographic databases, e.g. clinical guideline groups, should regularly check Retraction Watch’s database against their reference lists, or ensure their guidelines are listed on bibliographic databases.
- Authors of citing publications should publish an amendment, or a reassurance that the publication is unaffected, with a link to the affected publication.

Research misconduct can have widespread detrimental effects on subsequent research initiatives and clinical practice. Some possible solutions to minimise the impact of retracted publications are given in box 1, but there remains no overarching body with the commitment

Open access
to coordinate managing the consequences of proven research misconduct.

Contributors AA, FS, AG, GG and MB conceived and designed the study. FS and AA collected data. FS, AA, AG and MB analysed data. FS and AA wrote the draft manuscript. AA, FS, AG, GG and MB revised and approved the final submission. AA is the guarantor of the paper.

Funding The Health Services Research Unit at the University of Aberdeen is funded by the Chief Scientist Office of the Scottish Government Health and Social Care Directorates.

Competing interests All authors have completed the ICMJE unified disclosure form competing interest at www.icmje.org (available on request from the corresponding author). AG is a shareholder in Auckland Bone Density, a company that provides bone mineral density measurements. MB and AG report grants from the Health Research Council of New Zealand during the conduct of this study. AA reports grants from NIH during the conduct of this study. FS and GDG have no financial conflict of interest to declare.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iDs
Andrew Grey http://orcid.org/0000-0002-7803-0096
Mark Bolland http://orcid.org/0000-0003-0465-2674

REFERENCES
in elderly individuals with an elevated plasma homocysteine concentration: B-PROOF, a randomized controlled trial. Am J Clin Nutr 2014;100:1578–86.


Correction: An investigation into the impact and implications of published papers from retracted research: systematic search of affected literature


This article has been corrected since it was first published online. Details of retractions and correspondence relating to some references were inadvertently omitted in the first version; the following references have since been corrected to:


