PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Why articles are retracted: a retrospective cross-sectional study of retraction notices at BioMed Central
AUTHORS	Moylan, Elizabeth; Kowalczuk, Maria

VERSION 1 - REVIEW

REVIEWER	David Resnik NIH USA
REVIEW RETURNED	05-Aug-2015

GENERAL COMMENTS	This is an interesting study of retractions in BioMed central journals. I have some minor comments. p. 3, 41 when you say "all retraction notices" I think you mean to include "in BioMed Central journals" p. 26-31 Did you analyze inter-rate agreement statistically? How good was the agreement? General: how do the result of this study compare to similar studies of retractions? Some have looked at all retractions in PubMed since 2000, for example.

REVIEWER	R. Grant Steen, PhD Bioventus LLC 4721 Emperor Blvd, Suite 100 Durham, NC 27703
REVIEW RETURNED	10-Aug-2015

GENERAL COMMENTS	Overview: This paper reads like an advertisement for BioMed Central (BMC). The manuscript touts BMC's adherence to COPE guidelines, its low overall rate of retraction, and its low rate of retraction specifically for dishonesty. Yet a low rate of retraction can arise in two ways; either editors are vigilant and weed out bad work pre-publication or editors are lax about oversight post-publication. It is impossible to know which process led to the low rate of retraction reported here. The fact that the paper was written by two people from the BMC Editorial office naturally led them to conclude that editorial oversight cannot be at fault. My experience with BMC journals has not been so positive that I want to help BMC conquer the world. Specific comments: Abstract: The proportions given in the sentence on, "The most common reason for retraction" do not appear to tally to 100%. This is a real problem because I am not certain how the authors have defined "Publishing misconduct". Is this meant to include data

fabrication or not? Why is no mention made of data falsification, which is likely to be the more important problem? Was no distinction made between data fabrication and data falsification?

Introduction: Here is where terms should be defined. What crimes are included under the heading of "unethical research"? Why isn't failing "to disclose a major competing interest which could influence the interpretation of the article" considered to be unethical research?

Methods: Is the "publisher's advance search function" available to all readers? If so, what does it add that the editors used that function? Surely not credibility!

It can be extraordinarily hard to distinguish between "honest error" and "research misconduct". The authors do not acknowledge this, but their category of "Unclear" is actually larger than their category of "Honest error" (Fig. 3). My hunch is that "Unclear" is actually "Publishing misconduct" in many cases, so these results are probably in line with reports that misconduct accounts for the majority of retractions.

Results: The sample size is 77, a tiny N. As noted, this tiny N could reflect lax editorial oversight, so it seems a mistake to be anything other than circumspect about conclusions. Yet we are told that "None of the retractions correlated with a particular journal, impact factor, article type or discipline…" Is this not simply a reflection of the small N?

Plagiarism is a vague term that can be used to hide a multitude of sins. Data plagiarism amounts to data fabrication, because a representation is made that old data is actually new. This is a considerable issue as there have been instances of the same data counting twice in meta-analyses. Word plagiarism, on the other hand, is a fairly trivial crime; it is an affront to the copyright holder and perhaps to the author, but it does not corrupt the scientific process in the same way that data plagiarism does. To lump data and word plagiarism together seems to me to be naive.

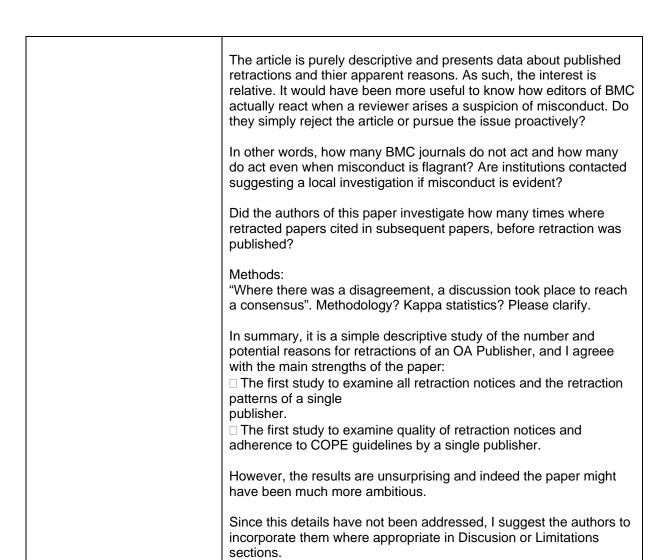Discussion: The main point, according to the first paragraph, is that BMC adheres to COPE guidelines. Why is that interesting?

The discussion is 4 pages long and it's about 3 pie charts and a bar graph. This is an excess of words….

This discussion just seems naïve. Retractions are not recommended for authorship disputes, but if an article is submitted without an author's knowledge, this is often to increase the odds of publication. The gloss of a high-profile name can mean that shoddy work gets published, reputations get tarnished, and flawed work stays in the literature.

Figures: I detest pie charts! These data would be so much easier to assess in a table. For example, I am not at all clear, from the Abstract and the figures, what specific crimes are considered to be "Research misconduct". The total given for "Research misconduct" (Fig. 3) is 14%, but I struggle to make any of the separate crimes in Fig. 4 tally to 14%. The greater clarity of a table would be deeply appreciated, but a table would make it perhaps too obvious how thin the data really are.

REVIEWER	Xavier Bosch
	Hospital Clinic, University of Barcelona, Internal Medicine Department
REVIEW RETURNED	02-Sep-2015

GENERAL COMMENTS	Interesting to know about how editors from a renowned open-access publisher deal with retractions.

The article is purely descriptive and presents data about published retractions and thier apparent reasons. As such, the interest is relative. It would have been more useful to know how editors of BMC actually react when a reviewer arises a suspicion of misconduct. Do they simply reject the article or pursue the issue proactively?

In other words, how many BMC journals do not act and how many do act even when misconduct is flagrant? Are institutions contacted suggesting a local investigation if misconduct is evident?

Did the authors of this paper investigate how many times where retracted papers cited in subsequent papers, before retraction was published?

Methods:
"Where there was a disagreement, a discussion took place to reach a consensus". Methodology? Kappa statistics? Please clarify.

In summary, it is a simple descriptive study of the number and potential reasons for retractions of an OA Publisher, and I agreee with the main strengths of the paper:
☐ The first study to examine all retraction notices and the retraction patterns of a single publisher.
☐ The first study to examine quality of retraction notices and adherence to COPE guidelines by a single publisher.

However, the results are unsurprising and indeed the paper might have been much more ambitious.

Since this details have not been addressed, I suggest the authors to incorporate them where appropriate in Discusion or Limitations sections.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name : David Resnik
Institution and Country NIH
USA
Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below
This is an interesting study of retractions in BioMed central journals. I have some minor comments.
p. 3, 41 when you say "all retraction notices" I think you mean to include "in BioMed Central journals"

Response: Thank you for your feedback, we have now made it clear that the study is just about BioMed Central journals.

p. 26-31 Did you analyze inter-rate agreement statistically? How good was the agreement?

Response: The first author initially scored the retraction notices. The second author then checked the scoring. Where there was a disagreement, a discussion took place between the authors to reach an

agreement. We have now clarified this in the methods section (page 4).

General: how do the result of this study compare to similar studies of retractions? Some have looked at all retractions in PubMed since 2000, for example.

Response: We have included discussion of this in the discussion under 'Reasons for retraction' page 9, our findings are in agreement with these studies.

Reviewer: 2
Reviewer Name R. Grant Steen, PhD
Institution and Country Bioventus LLC
4721 Emperor Blvd, Suite 100
Durham, NC 27703
Please state any competing interests or state 'None declared': None declared.

Please leave your comments for the authors below
Overview:
This paper reads like an advertisement for BioMed Central (BMC). The manuscript touts BMC's adherence to COPE guidelines, its low overall rate of retraction, and its low rate of retraction specifically for dishonesty.

Response: This was not our intention. We wanted to audit all retractions at BioMed Central to examine the reasons why articles had been retracted and to put any improvements in place (e.g. plagiarism checks, policies supporting data deposition etc). Our frame of reference was to compare our findings against COPE guidelines. We also point out where we failed to follow COPE guidelines and what we have learnt.

Yet a low rate of retraction can arise in two ways; either editors are vigilant and weed out bad work pre-publication or editors are lax about oversight post-publication. It is impossible to know which process led to the low rate of retraction reported here.

Response: We cannot know to what extent our Editors spot issues pre-publication or suppress issues post-publication. We work on the basis that all parties operate with integrity and raise issues that are brought to their attention. If an Editor doesn't respond to an issue that has been raised, then the complainant could raise this directly with the Publisher or COPE.

The fact that the paper was written by two people from the BMC Editorial office naturally led them to conclude that editorial oversight cannot be at fault. My experience with BMC journals has not been so positive that I want to help BMC conquer the world.

Response: Our aim was to look at our retraction notices to find out: why we retract articles, whether the stated reason was clear and whether improvements could be made. Despite editorial oversight, notices were still published which failed to state who the retraction was coming from, and were ambiguous about underlying causes. We feel improvements could be made and additional checks are now in place. Our focus was on the reasons why BioMed Central has retracted articles and what we can learn from this.

Specific comments:
Abstract: The proportions given in the sentence on, "The most common reason for retraction" do not appear to tally to 100%. This is a real problem because I am not certain how the authors have defined "Publishing misconduct". Is this meant to include data fabrication or not?

Response: We are sorry this was not clear. In the abstract we simply wanted to highlight the most common reasons for retraction, not every single reason (as given in the results section) which is why they don't appear to tally to 100%. We have now revised the abstract to stress the most common reasons for retraction only.

The classification of retractions is given in the methods section. We felt that publishing misconduct does not include data fabrication, we feel this falls under the remit of research misconduct.

Why is no mention made of data falsification, which is likely to be the more important problem? Was no distinction made between data fabrication and data falsification?

Response: Thank you for pointing this out. We have now clarified that the 'data fabrication' category includes cases of data falsification and fabrication. We had 10 examples which fall into this category and while some appeared to be falsification (i.e. manipulation of material/processes/data to misrepresent results) others appear to be fabrication (i.e. reporting the results of experiments not actually performed). In some cases, even when an institutional investigation had taken place, it was not possible to distinguish data fabrication from data falsification and the retraction notices mentioned both. We have revised our manuscript to refer to this category as 'data falsification/fabrication'. We have clarified this approach in the methods.

Introduction: Here is where terms should be defined. What crimes are included under the heading of "unethical research"? Why isn't failing "to disclose a major competing interest which could influence the interpretation of the article" considered to be unethical research?

Response: In the introductory paragraph we are referring to COPE guidelines on retraction and the instances where COPE recommends that Editors should consider a retraction.
The rationale behind our classification of retractions is outlined in the methods. We feel that an 'undeclared competing interest' doesn't fall within 'research misconduct' but under 'publishing misconduct' because declarations are made at the stage of drafting the manuscript and submitting to a journal.

As noted in the methods we classified retractions according to the apparent underlying reason for the retraction i.e. honest error (mistakes on the part of the author or publisher) or research misconduct (data falsification/fabrication, failure to obtain ethical approval, failure to obtain permission for data) or publishing misconduct (plagiarism, duplicate publication, image duplication, authorship issues, compromised peer review) in line with previous studies [3] and using the definition of research misconduct given by [17].

Methods: Is the "publisher's advance search function" available to all readers? If so, what does it add that the editors used that function? Surely not credibility!

Response: Yes, the advanced search function is available to all readers, and a link is given in the reference. We have clarified this in the text. Anyone outside BioMed Central can repeat this study.

It can be extraordinarily hard to distinguish between "honest error" and "research misconduct". The authors do not acknowledge this, but their category of "Unclear" is actually larger than their category of "Honest error" (Fig. 3). My hunch is that "Unclear" is actually "Publishing misconduct" in many cases, so these results are probably in line with reports that misconduct accounts for the majority of retractions.

Response: Yes, we scored retraction notices based purely on their content and it is difficult to distinguish between honest error and research misconduct. Where an institutional investigation had

taken place and did not mention honest error had occurred, we inferred that misconduct had happened.

Yes, we take your point that 'unclear' probably does mean misconduct occurred but for whatever reason (e.g. fear of litigation) this wasn't mentioned.
In including another year's worth of data from 2015 to bring the study up to date, we also reach the conclusion that misconduct accounts for the majority of retractions now seen at BioMed Central.

Results: The sample size is 77, a tiny N. As noted, this tiny N could reflect lax editorial oversight, so it seems a mistake to be anything other than circumspect about conclusions.
Yet we are told that "None of the retractions correlated with a particular journal, impact factor, article type or discipline…" Is this not simply a reflection of the small N?

Response: The revised sample size with 2015 data is now 134, which is still small. We have clarified this as a limitation of the study and stated in the limitations section that we did not seek to explore other correlations.

Plagiarism is a vague term that can be used to hide a multitude of sins. Data plagiarism amounts to data fabrication, because a representation is made that old data is actually new. This is a considerable issue as there have been instances of the same data counting twice in meta-analyses. Word plagiarism, on the other hand, is a fairly trivial crime; it is an affront to the copyright holder and perhaps to the author, but it does not corrupt the scientific process in the same way that data plagiarism does. To lump data and word plagiarism together seems to me to be naïve.

Response: We did not come across any case where the term 'data plagiarism' was mentioned explicitly. In our study, a retraction notice which stated that the data were plagiarised would have been classified as 'data fabrication/falsification'.

Discussion: The main point, according to the first paragraph, is that BMC adheres to COPE guidelines. Why is that interesting?

Response: We disagree that the point of the first paragraph of the discussion is that BioMed Central adheres to COPE guidelines. We highlighted that COPE provides guidance for its members. The purpose of this study was to determine how far we adhere to the guidance on retraction notices. We have revised the Discussion to clarify this point.

The discussion is 4 pages long and it's about 3 pie charts and a bar graph. This is an excess of words….

Response: The figures and tables are part of the results. We have shortened the discussion where possible.

This discussion just seems naïve. Retractions are not recommended for authorship disputes, but if an article is submitted without an author's knowledge, this is often to increase the odds of publication. The gloss of a high-profile name can mean that shoddy work gets published, reputations get tarnished, and flawed work stays in the literature.

Response: Thank you for highlighting this. We have revised the discussion to address this point.

Figures: I detest pie charts! These data would be so much easier to assess in a table. For example, I am not at all clear, from the Abstract and the figures, what specific crimes are considered to be "Research misconduct". The total given for "Research misconduct" (Fig. 3) is 14%, but I struggle to

make any of the separate crimes in Fig. 4 tally to 14%. The greater clarity of a table would be deeply appreciated, but a table would make it perhaps too obvious how thin the data really are.

We are happy to present the data either way and have replaced the piecharts with Tables. We have included the following Table 3 in the manuscript to be explicit about our classification.

Reviewer: 3
Reviewer Name Xavier Bosch
Institution and Country Hospital Clinic, University of Barcelona, Internal Medicine Department
Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below

Interesting to know about how editors from a renowned open-access publisher deal with retractions.

Response: Thank you.

The article is purely descriptive and presents data about published retractions and thier apparent reasons. As such, the interest is relative. It would have been more useful to know how editors of BMC actually react when a reviewer arises a suspicion of misconduct. Do they simply reject the article or pursue the issue proactively?
In other words, how many BMC journals do not act and how many do act even when misconduct is flagrant? Are institutions contacted suggesting a local investigation if misconduct is evident?

Response: As our journals are members of COPE we expect our editors to follow COPE guidelines when research or publication ethics issues are raised. We acknowledge the limitation the reviewer has raised but this study was not designed to determine how far editors adhere to COPE guidelines.

Did the authors of this paper investigate how many times where retracted papers cited in subsequent papers, before retraction was published?

Response: Following your suggestion we looked at citations of the retracted articles before and after retraction, and citations to the corresponding retraction notice using Scopus. We found that articles continued to be cited after retraction (which has also been observed in larger studies). It is interesting that the rarely were retraction notices themselves cited. We have added a line to the methods, results and discussion.

Methods:
"Where there was a disagreement, a discussion took place to reach a consensus". Methodology? Kappa statistics? Please clarify.

Response: The first author initially scored the retraction notices. The second author then checked the scoring. Where there was a disagreement, a discussion took place between the authors to reach an agreement.

In summary, it is a simple descriptive study of the number and potential reasons for retractions of an OA Publisher, and I agreee with the main strengths of the paper:
• The first study to examine all retraction notices and the retraction patterns of a single publisher.
• The first study to examine quality of retraction notices and adherence to COPE guidelines by a single publisher.

However, the results are unsurprising and indeed the paper might have been much more ambitious.

Since this details have not been addressed, I suggest the authors to incorporate them where appropriate in Discusion or Limitations sections.

Response: Thank you, we have included these in the limitations section.

VERSION 2 – REVIEW

REVIEWER	David Resnik
	NIH, USA
REVIEW RETURNED	06-Apr-2016

GENERAL COMMENTS	Nice work

REVIEWER	Nadia ELIA
	Institute of Global Health, University of Geneva, Switzerland and Departement of Anaesthesiology, University hospitals of Geneva, Switezrland
REVIEW RETURNED	17-Jun-2016

GENERAL COMMENTS	Title: Study design is reported as being "a review of retraction notices" in the title and as "a retrospective analysis of retractions" in the abstract. I think the correct study design is a cross-sectional study (with a part as retrospective cohort analysis) of retraction notices published in Biomed central?
	The authors should follow the STROBE recommendations for the reporting of such observational studies. This will help complete the manuscript with important information.
	Objective: The objective of the study is not clearly defined "We also wanted to determine if retractions were on the increase. » Increase in absolute number ? in proportion of published articles ? In which case, which type of article enter the denominator ?
	Abstract:
	There is no "Method" section in the abstract. On should be added.
	Objectives
	A sentence summarizing what the COPE guidelines recommend regarding retraction notices should be added in the objective to make the result section understandable to the non-specialist reader.
	Results
	"Although they account for 0.07% of all articles published": It is unclear from the sentence which was the denominator: original research articles? All articles including commentaries letters and editorials?
	A sentence such as: the rate of retraction is rising should be accompanied by numbers, if not by statistics.
	Introduction:
	Page 3
	Lines 30 to 35: The COPE recommendations listed here are incomplete. COPE also recommends for example to identify the

name of the authors and the title of the retracted article in the retraction notice heading, and also that the notice should be clearly linked to the retracted article in all electronic sources.

Line 37: what is the difference between "misconduct" and "fraud"? These terms should be clearly defined since one of the major problem that research on misconduct is facing is the lack of clear definitions!

Methods:
Page 4, line 9: authors have searched for retraction notices using the search term "retraction" within the article title. The problem is that this is one of the COPE recommendations (that the notice be clearly identified as a retraction, not as a Correction or Comment). We have previously shown that this recommendation was not necessarily followed by all journals, and therefore, some retractions may have been missed (Elia et al, Fate of Articles That Warranted Retraction Due to Ethical Concerns: A Descriptive Cross-Sectional Study, PlosOne 2014) If there is a reason to believe that this cannot happen in BioMed Central journals, this should be described.
It is not clear who identified the retraction notices and finally decided which one to include or not. Following the STROBE recommendations closely should help the authors clarify these issues.

Line 27: Although some authors have previously made a difference between "research misconduct" and "publishing misconduct"… I would tend to call all this "fraud" since the difference is not quite clear… Research good practice includes the publication process; also why one should make a difference between plagiarism and image duplication?

Line 48: how was the "main" reason identified? Was it the first one cited in the text of the retraction notice?

Line 50: "The scoring of the retraction notice " should be changed to "the description of " since no score is being used.

There is no description of the plan of analyses, sample size justification. Even a short paragraph is needed.

Results:
Page 5, Line 5: "All retraction notices were clearly labeled and linked to the retracted article". I clicked on one retraction notice at random and figured that this was not the case: "Psychometric characteristics of the ankylosing spondylitis quality of life questionnaire, short form 36 health survey, and functional assessment of chronic illness therapy-fatigue subscale" Dennis a Revick et al health and Quality of Life Outocmes 2009 7:6. Although the link exists, the article is lacking. " The authors have withdrawn this article from the public domain because the conclusions drawn may not be supported by the data. In the light of this situation, BioMed Central regrets that this article is no longer available. The authors apologise to all parties for the inconvenience." I realize that this may be an exception but it would be interesting to describe.

Line 10: Figure 1: obviously something happened in 2015. Until then, it is not clear form the figure that the proportion of retracted articles has increased. A couple of data or satistics to quantify this "increase" would help. Also the « publications » should be

described. Are these only original articles? Or do they include Editorials? Letters ?

Line 12: measures of central tendency (median time to retraction) should be given with some description of the spread of the data (min and max, or inter-quartile range).

Table 1:
I understand who the authors, the Editors, and the publisher are… but who is the "journal"? Giving and exemple of a case where the journal was retracting the article would help understand.
I suggest that the "not stated" strata should be presented last.
Lines 38 and 39: I do not understand the sentence: "6% of retractions did not explicitly state who was retracting the paper "although it was implied that the notice was coming from the authors"..?" Maybe a link to an example could help the reader understand?

Table 2:
What does "published in error" stand for ?
What does "Data unreliable" mean ? Is this fabrication? Falsification? Wrong analyses? Although these terms may have been used in previous articles, they need to be defined this they are not generally recognised and accepted.
Table 2 reports on "image duplication" while the following text on page 6, line 17 reports on image "manipulation" which is not the same thing.
Table 2 and 4 can be deleted since all the information is reported in table 3.

Page 7 line 24: Obviously, the number of time an article is being cited after retraction is dependent on the time since retraction. Therefore authors should find a way to describe their data in an informative way.

Discussion

Page 7, line 36: the first sentence belongs to the result section. This is a (post-hoc) subgroup analysis. I did not find supplementary file 2
Page 7, line 40-43: result. I don't understand why there is a reference given (24).

Line 48: the authors highlight that the long delays are not due to a problem with the retraction process, but rather that the reasons for retraction have taken a long time to become known. Nevertheless, the cited article (with incorrect data) has remained 11 years in the published recognized as high quality published literature.

Line 57: true. This is why we (and others) had suggested that reader should be made aware that the article has been retracted through watermark (Elia, PlosOne 2014, Sox HC, Rennie D (2006) Research misconduct, retraction, and cleansing the medical literature: Lessons from the Poehlman case. Ann Intern Med 144: 609– 13.)
Page 8, line 15: another reason may be an error from the editor and publisher writing the retraction notice.
Page 9, line 3: some details regarding the process of "compromised peer-review" would add much to the interest of the manuscript.
Line 20: this last sentence deserves a reference.
Line 32-35: I am not sure that data sharing PRIOR to publication can reduce retraction due to errors in data analysis. On the peer-

	reviewers and/or editors could make this check, and I am not sure that they would actually do it (since it obviously will take a lot of time and energy).

REVIEWER	Daniele Fanelli Stanford University
REVIEW RETURNED	21-Jun-2016

GENERAL COMMENTS	The article is well written, but as it stands presents data of little originality that offer no significant new insights. Multiple previous studies have reported statistics on the prevalence of retractions and their causes, and showed retraction notices to be opaque. Their samples covered all available retractions including those examined here. It is therefore hard to see the point of reporting statistics about bmc per se, without any attempt to make comparisons or analyses that would produce usable new knowledge about them. The only real element of novelty of this paper, compared to older studies, is represented by having recent retractions from the last three years, most of which are shown to be due to peer-review abuse. But the fact that many recent retractions in bmc were due to peer-review abuses was known before the study was done, so, again, it is unclear what the analysis is meant to convey. The point of the work as it stands is made all the more doubtful by the fact that, as one of the references cited by the paper showed, retraction notices are known to be opaque. So one is not entirely sure what to do with the statistics reported. As the field stands by now, it makes little sense to just count retractions and report simple descriptive statistics whilst ignoring the element of journal policies. The rise in retractions is largely explained by a rise in the number of journals that retract papers, i.e. the rising number of journals that implement retraction policies. Fanelil 2013 (Plos Medicine, Why growing retractions are (mostly) a good sign) showed that retractions-per-retracting journals have not actually risen, showing misconduct not to be on the rise. Resnik 2015 (J Med Lib Assoc, Retraction policies of top scientific journals ranked by impact factor) showed that journals with retraction policies have tripled in ten years, but are still only about 2/3 of the total, and that's just amongst high-ranking journals. To make the analyses of this paper meaningful, I suggest that the authors try to move the field forward. Taking inspiration from the papers cited above (which should have been examined and discussed anyway, along with much other relevant literature), they should look at statistics at the journal-level and see if/how journal policies make a difference to results, and maybe also compare statistics in bmc with those of other publishers or journals. The authors should also pay attention to the phenomenon or prolific retractions. Grieneisen and Zhang 2012 (Plos ONE, A Comprehensive Survey of Retracted Articles from the Scholarly Literature) showed this to be an issue, and Fanelli et al 2015 (Plos ONE, Misconduct Policies, Academic Culture and Career Stage, Not Gender or Pressures to Publish, Affect Scientific Integrity) indeed showed how statistics look very different when this factor is accurately adjusted for.

	Finally, the authors could attempt to reproduce the work of Fang et al 2012 (cited) and follow up on all stories, to see how many are actually due to misconduct. Then again, perhaps compare bmc statistics with those of other journals.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name
David Resnik

Institution and Country
NIH, USA

Please state any competing interests or state 'None declared':
None declared

Please leave your comments for the authors below Nice work

Response: Thank you

Reviewer: 2

Reviewer Name
Nadia ELIA

Institution and Country
Institute of Global Health, University of Geneva, Switzerland and Departement of Anaesthesiology, University hospitals of Geneva, Switezrland

Please state any competing interests or state 'None declared':
none (regarding this mansucript)

Please leave your comments for the authors below
Title: Study design is reported as being "a review of retraction notices" in the title and as "a retrospective analysis of retractions" in the abstract. I think the correct study design is a cross-sectional study (with a part as retrospective cohort analysis) of retraction notices published in Biomed central?

Response: Thank you. We've revised the title to include 'a retrospective cross-sectional study'

The authors should follow the STROBE recommendations for the reporting of such observational studies. This will help complete the manuscript with important information.

Response: Thank you. We have completed a checklist following the STROBE recommendations and uploaded this as supplementary file 2.

Objective: The objective of the study is not clearly defined "We also wanted to determine if retractions were on the increase. » Increase in absolute number ? in proportion of published articles ? In which case, which type of article enter the denominator ?

Response: We meant in proportion of published articles. We have now clarified that the denominator is all articles published (190514) excluding supplements, corrections, retractions and commissioned content.

Abstract:
There is no "Method" section in the abstract. On should be added.

Response: The abstract follows BMJ Open guidelines and there is no method section (the study design is reported under the design heading).

Objectives
A sentence summarizing what the COPE guidelines recommend regarding retraction notices should be added in the objective to make the result section understandable to the non-specialist reader.

Response: The abstract has a word limit (300 words) We have provided this in the first line of the introduction on page 3, and expanded this as you suggest below.

Results
"Although they account for 0.07% of all articles published": It is unclear from the sentence which was the denominator: original research articles? All articles including commentaries letters and editorials? A sentence such as: the rate of retraction is rising should be accompanied by numbers, if not by statistics.

Response: Thank you, this was an oversight. Between January 2000 and December 2015, our search identified 134 retraction notices. This accounts for 0.07% of all total articles published (190514 excluding supplements, corrections, retractions and commissioned content). We have added this to the abstract and the results section.

Introduction:
Page 3
Lines 30 to 35: The COPE recommendations listed here are incomplete. COPE also recommends for example to identify the name of the authors and the title of the retracted article in the retraction notice heading, and also that the notice should be clearly linked to the retracted article in all electronic sources.

Response: Yes. We have now added the following sentence: "In addition, retractions should be clearly identifiable; freely available; published promptly and be linked to the original article that is retracted (which should also be identified as a retraction)." In the discussion and conclusions we also clarify that in these respects BioMed Central adhered to COPE guidelines and also discussed issues where relevant information was unfortunately not provided.

Line 37: what is the difference between "misconduct" and "fraud"? These terms should be clearly defined since one of the major problem that research on misconduct is facing is the lack of clear definitions!

Response: Thank you, the ORI definition of misconduct is fabrication, falsification or plagiarism. We are using fraud to refer to fabrication or falsification. We have clarified the terms in the methods. As you point out, because of ambiguity in the literature about what constitutes misconduct – and whether plagiarism is research misconduct or publishing misconduct we now simply describe the individual reasons for retraction (plagiarism, duplicate publication, image duplication etc) and only broadly classify retractions into three general categories of misconduct, honest error or unclear as defined in

the methods. We raise this in the discussion under 'reasons for retraction'.

Methods:
Page 4, line 9: authors have searched for retraction notices using the search term "retraction" within the article title. The problem is that this is one of the COPE recommendations (that the notice be clearly identified as a retraction, not as a Correction or Comment). We have previously shown that this recommendation was not necessarily followed by all journals, and therefore, some retractions may have been missed (Elia et al, Fate of Articles That Warranted Retraction Due to Ethical Concerns: A Descriptive Cross-Sectional Study, PlosOne 2014) If there is a reason to believe that this cannot happen in BioMed Central journals, this should be described.

Response: It has been BioMed Central policy to publish all retraction notices as a separate retraction article type, so they all start with "Retraction: xxxx" so we don't believe we missed any.
Thank you for drawing our attention to your article, we have also cited it in the discussion.

It is not clear who identified the retraction notices and finally decided which one to include or not. Following the STROBE recommendations closely should help the authors clarify these issues.

Response: We have clarified this in the methods of the paper (page 4) and the STROBE checklist.

Line 27: Although some authors have previously made a difference between "research misconduct" and "publishing misconduct"… I would tend to call all this "fraud" since the difference is not quite clear… Research good practice includes the publication process; also why one should make a difference between plagiarism and image duplication?

Response: Thank you, yes people define these terms in different ways. As we commented above, we now simply describe the individual reasons for retraction (plagiarism, image duplication etc) and only broadly classify them into general categories of 'misconduct' (which would include research misconduct, publishing misconduct, fraud), 'honest error' or 'unclear' as defined in the methods. With regard to image duplication, we did feel this was a separate category rather than text duplication, especially when looking to explore possible ways in which retractions at BioMed Central could be reduced in future.

Line 48: how was the "main" reason identified? Was it the first one cited in the text of the retraction notice?

Response: It was the main reason why an article was ultimately retracted, in most cases it was also the first.

Line 50: "The scoring of the retraction notice " should be changed to "the description of " since no score is being used.

Response: Thank you. We have changed this throughout the entire manuscript to be precise.

There is no description of the plan of analyses, sample size justification. Even a short paragraph is needed.

Response: The analysis was based on all retractions that BioMed Central has ever published in 15 years of complete publishing, from the time we first started publishing retraction notices after 2000 up until 2015. So we felt it best to have the most complete data set we could in whole years (the study was submitted in March 2016). We have clarified this in the methods.

Results:
Page 5, Line 5: "All retraction notices were clearly labeled and linked to the retracted article". I clicked on one retraction notice at random and figured that this was not the case: "Psychometric characteristics of the ankylosing spondylitis quality of life questionnaire, short form 36 health survey, and functional assessment of chronic illness therapy-fatigue subscale" Dennis a Revick et al health and Quality of Life Outocmes 2009 7:6. Although the link exists, the article is lacking. " The authors have withdrawn this article from the public domain because the conclusions drawn may not be supported by the data. In the light of this situation, BioMed Central regrets that this article is no longer available. The authors apologise to all parties for the inconvenience." I realize that this may be an exception but it would be interesting to describe.

Response: Thank you. We had overlooked in certain cases where the retraction article had been removed for legal reasons (e.g. if there was sensitive information or if plagiarism infringed another journal's copyright). We have now rephrased the text to say:

"All retraction notices were clearly labelled and linked to the retracted article except for cases where for legal reasons the original article could no longer be made available (for example, if there was sensitive information or if plagiarism infringed another journal's copyright)."

Line 10: Figure 1: obviously something happened in 2015. Until then, it is not clear form the figure that the proportion of retracted articles has increased. A couple of data or satistics to quantify this "increase" would help. Also the « publications » should be described. Are these only original articles? Or do they include Editorials? Letters ?

Response: Yes. Figure 1 shows that proportionally there has been no increase in retraction rate when growth in the total number of articles published (excluding supplements, corrections, retractions and commissioned content) is accounted for. However, in 2015 we published 43 retractions in March related to repeated and inappropriate attempts to manipulate the peer review process of several journals. With the addition of 2015 data on retractions the retraction rate increases. We have clarified this in the results section on page 5.

Line 12: measures of central tendency (median time to retraction) should be given with some description of the spread of the data (min and max, or inter-quartile range).

Response: We have now added the minimum and maximum times to give some indication of spread of data. In addition all data is provided in supplementary file 2.

Table 1:
I understand who the authors, the Editors, and the publisher are… but who is the "journal"? Giving and exemple of a case where the journal was retracting the article would help understand.

Response: In accordance with COPE guidelines the retraction notice should clarify who is retracting the article. In one case the retraction notice was worded as coming from 'the journal' that the retraction was published in. The authors of this paper did not oversee the wording of all the retraction notices published by BioMed Central during the time frame of this study so we do not know why 'journal' was used in this instance it was probably an oversight for Editor.

I suggest that the "not stated" strata should be presented last.

Response: Agreed. We have moved the category 'not stated' to the bottom of Table 1 and called it 'information not provided' for further clarity.

Lines 38 and 39: I do not understand the sentence: "6% of retractions did not explicitly state who was retracting the paper "although it was implied that the notice was coming from the authors"..?" Maybe a link to an example could help the reader understand?

Response: This was a textual explanation of Table 1. While it is usual for retraction notices to state who is retracting the article for example, the Editor, the Publisher, the authors. In 6% of retraction notices this information wasn't explicitly spelt out, so it was ambiguous who actually retracted the article, although one could infer it was potentially the authors. In such cases the retraction notices invariably simply stated 'This article is retracted due to' and it seems it should have come from the authors in all likelihood but it's not explicit. We've added an explanation to the text.
Table 2:
What does "published in error" stand for ?

Response: We discuss this in the section under transparency of retraction notices where we explain that articles were 'published in error' when a journal was transferred from another publisher and during this period an article was inadvertently published twice. However, we've made this clear earlier on in the paper by adding the explanation to the results section too. In addition we have defined the 'individual' reasons for retraction in the methods section.

What does "Data unreliable" mean ? Is this fabrication? Falsification? Wrong analyses? Although these terms may have been used in previous articles, they need to be defined this they are not generally recognised and accepted.

Response: We have now defined all the individual descriptions of retraction in the methods section. Data unreliable is where data is not reliable due to mistakes or errors in data. In recording this category in some cases this was due to honest error, in some cases it was unclear if it was honest error or misconduct so some were classified as 'unclear' (see Table 2).

Table 2 reports on "image duplication" while the following text on page 6, line 17 reports on image "manipulation" which is not the same thing.

Response: We double checked all retraction notices where there was a particular issue with images. They all referred to image duplication, the text was amended to be consistent and refer to image duplication. This was a mistake.

Table 2 and 4 can be deleted since all the information is reported in table 3.

Response: Done. Table 3 now renamed Table 2 for continuity.

Page 7 line 24: Obviously, the number of time an article is being cited after retraction is dependent on the time since retraction. Therefore authors should find a way to describe their data in an informative way.

Thank you. After consideration we think that given the low numbers of citations (apart from a few outliers) this would not be informative. We have provided all the information in the additional supplementary file 1 so that readers can see the year of publication and number of citations. We just wanted to make the point that similar to other studies we also found that articles do continue to be cited after retraction.
Discussion

Page 7, line 36: the first sentence belongs to the result section. This is a (post-hoc) subgroup analysis. I did not find supplementary file 2 Page 7, line 40-43: result. I don't understand why there is

a reference given (24).

Response: We have moved the sentence. We give a reference to clarify that this wasn't a new finding of this study, but that others have reported this also.

Line 48: the authors highlight that the long delays are not due to a problem with the retraction process, but rather that the reasons for retraction have taken a long time to become known. Nevertheless, the cited article (with incorrect data) has remained 11 years in the published recognized as high quality published literature.

Response: In this particular case the author only recently changed their position on making their software available (within the last year).

Line 57: true. This is why we (and others) had suggested that reader should be made aware that the article has been retracted through watermark (Elia, PlosOne 2014, Sox HC, Rennie D (2006) Research misconduct, retraction, and cleansing the medical literature: Lessons from the Poehlman case. Ann Intern Med 144: 609– 13.)

Response: Thank you for flagging. We have added this reference to the discussion and conclusion.

Page 8, line 15: another reason may be an error from the editor and publisher writing the retraction notice.

Response: Agreed. We have added a line here.

Page 9, line 3: some details regarding the process of "compromised peer-review" would add much to the interest of the manuscript.

Response: We have added a line but the details have been previously outlined in a blog (reference 7).

Line 20: this last sentence deserves a reference.

Response: We think we did have a reference here [41] but we have moved it to the end of the sentence to make this really clear.

Line 32-35: I am not sure that data sharing PRIOR to publication can reduce retraction due to errors in data analysis. On the peer-reviewers and/or editors could make this check, and I am not sure that they would actually do it (since it obviously will take a lot of time and energy).

Response: We mean that by encouraging data sharing and data deposition prior to publication authors collate their data to make it 'publication ready' and this exercise in itself can help resolve honest errors. We have re-worded this sentence. In addition we hope that some peer reviewers and editors would be able to check the data and thus increase the chances of catching any errors or possible misconduct prior to publication. There are data journals in existence where peer review of data is the norm.

Reviewer: 3

Reviewer Name
Daniele Fanelli

Institution and Country

Stanford University

Please state any competing interests or state 'None declared':
None declared

Please leave your comments for the authors below The article is well written, but as it stands presents data of little originality that offer no significant new insights.

Response: Thank you. Yes, we appreciate this is a retrospective and descriptive study but nevertheless we felt our findings would be of interest to other editors and publishers and would fit the scope and remit of BMJ Open which considers "small and specialist studies and negative studies".

We also feel other editors and publishers may have something to learn from our findings and it may inspire them to conduct their own audits and put processes in place to reduce retractions (e.g. adopting plagiarism checks or data deposition policies may help to reduce the number of retractions which arise due to honest errors with the data by making data 'publication ready') and adopt a checklist for the appropriate wording of the retraction notices

Our objectives were to

• examine all BioMed Central retraction notices since retractions started being published at BioMed Central from 2000 onwards and the retraction patterns of a single publisher.
• examine transparency of retraction notices and adherence to COPE retraction guidelines by a single publisher.

Multiple previous studies have reported statistics on the prevalence of retractions and their causes, and showed retraction notices to be opaque. Their samples covered all available retractions including those examined here.

Response: We don't think there has been a single study looking at consistency and transparency of retractions notices at this scale within a single publisher. We don't think there has been a single publisher study of specific adherence (or otherwise) to COPE guidelines. We also include more recent retraction data from the last three years.

It is therefore hard to see the point of reporting statistics about bmc per se, without any attempt to make comparisons or analyses that would produce usable new knowledge about them. The only real element of novelty of this paper, compared to older studies, is represented by having recent retractions from the last three years, most of which are shown to be due to peer-review abuse. But the fact that many recent retractions in bmc were due to peer-review abuses was known before the study was done, so, again, it is unclear what the analysis is meant to convey.

Response: While isolated cases of peer review manipulation had been previously documented from individual authors or peer review rings, the systematic manipulation of peer review across unrelated papers possibly by a third party is new and first documented in 2015.

The point of the work as it stands is made all the more doubtful by the fact that, as one of the references cited by the paper showed, retraction notices are known to be opaque. So one is not entirely sure what to do with the statistics reported.

Response: The paper also attempted to quantify the 'impact' of the COPE guidelines at a single publisher which is a member of COPE. Post-2009 retractions are more consistent but there are still occasional errors. We suggest there is a need for Publishers to take responsibility for ensuring

retraction notices are consistent, transparent and appropriately worded.

As the field stands by now, it makes little sense to just count retractions and report simple descriptive statistics whilst ignoring the element of journal policies.
The rise in retractions is largely explained by a rise in the number of journals that retract papers, i.e. the rising number of journals that implement retraction policies. Fanelil 2013 (Plos Medicine, Why growing retractions are (mostly) a good sign) showed that retractions-per-retracting journals have not actually risen, showing misconduct not to be on the rise. Resnik 2015 (J Med Lib Assoc, Retraction policies of top scientific journals ranked by impact factor) showed that journals with retraction policies have tripled in ten years, but are still only about 2/3 of the total, and that's just amongst high-ranking journals.

Response: BioMed Central has always implemented retraction policies for its journals so this has been a constant since it first started publishing in 1999/2000. The same umbrella policies apply across all our journals, we have clarified this in the text under the discussion 'transparency of retraction notices'.
Resnik (reviewer 1) reviewed an earlier version of this paper and re-reviewed our revisions and is satisfied with our revisions.
Unfortunately our study is limited by the available retraction notices that have been published for us to describe, so we have not looked at correlations with a particular journal, impact factor, article type, discipline or peer review model because the individual numbers would be too low.
However, we have responded to earlier reviewer comments and expanded the study by including data on citations of retractions and time taken to retract.

To make the analyses of this paper meaningful, I suggest that the authors try to move the field forward. Taking inspiration from the papers cited above (which should have been examined and discussed anyway, along with much other relevant literature), they should look at statistics at the journal-level and see if/how journal policies make a difference to results, and maybe also compare statistics in bmc with those of other publishers or journals.

Response: We have included relevant literature in the discussion, we have also discussed the general problem of defining terms as raised by Reviewer 2.

The original objectives are set out in the paper and it was not one of our objectives to compare statistics from BioMed Central with those of other publishers or journals. But by publishing our findings we hope others would find our study useful and then be able to make comparisons and draw more conclusions. Indeed, this could be scope for a future study.

The authors should also pay attention to the phenomenon or prolific retractions. Grieneisen and Zhang 2012 (Plos ONE, A Comprehensive Survey of Retracted Articles from the Scholarly Literature) showed this to be an issue, and Fanelli et al 2015 (Plos ONE, Misconduct Policies, Academic Culture and Career Stage, Not Gender or Pressures to Publish, Affect Scientific Integrity) indeed showed how statistics look very different when this factor is accurately adjusted for.

Response: Thank you, we were conscious of this phenomenon but we don't feel it impinged on this study too much, we had 3 Mori retractions.

Finally, the authors could attempt to reproduce the work of Fang et al 2012 (cited) and follow up on all stories, to see how many are actually due to misconduct. Then again, perhaps compare bmc statistics with those of other journals.

Response: While this would certainly be interesting to do, this was not an objective of the original study.

<div align="center">**VERSION 3 – REVIEW**</div>

REVIEWER	Daniele Fanelli Stanford University, USA
REVIEW RETURNED	22-Sep-2016

GENERAL COMMENTS	The authors made an appreciable effort to accommodate the numerous issues raised and have otherwise responded fairly to the objections raised. I am happy to recommend the paper for publication, and hope that it will inspire them and other researchers to conduct more work on the subject.