Author's response to reviews

Title: Concern noted: A descriptive study of editorial expressions of concern in PubMed and PubMed Central.

Authors:

Melissa Vaught (melissa.vaught@nih.gov)
Diana Jordan (jordandc@ncbi.nlm.nih.gov)
Hilda Bastian (hilda.bastian@nih.gov)

Version: 1 Date: 24 Apr 2017

Author’s response to reviews:

Dear Dr. Boughton,

We sincerely thank you and the reviewers for the timely and thoughtful review of our manuscript "Concern noted: A descriptive study of editorial expressions of concern in PubMed and PubMed Central" (RIPR-D-17-00005). Changes have been made to the manuscript using “Track changes.” Below we provide a point-by-point response to the reviewers’ comments. Where applicable, we have noted the page and line numbers where changes have been made in the revised manuscript.

Sincerely,

Melissa Vaught

Reviewer #1: “Background: Although this is at the core of their study and made clear later in the Method section, it would be helpful if the authors could explain early how they define ‘issued’ Expressions of Concern (i.e. that they do not restrict their analysis to formally published and indexed EEoCs).”

We have added a brief description to the Background (p. 7, line 132-5) to introduce the different types of EEoCs considered in this study.

Reviewer #1: “Method: (Additional data file 1) I couldn't see a rationale for selecting the 9 publisher websites that were included in the search...”

We have added the rationale for publisher selection to the Methods in the main text (p. 10, lines 197-201) and in Additional File 1. Publishers were selected either because publish a significant
share of literature in relevant subjects; or because they published at least 1 EEoC that was not captured by title search in PubMed/PMC, and their website facilitated single search of multiple journal titles indexed in PubMed/PMC. One reference (#19 in revised manuscript) was added for support.

Reviewer #1: “I couldn't see a rationale… for focusing on the 2014-2016 time frame for further analysis.”

We have also added rationale for selecting Aug. 2014 to Aug. 2016 for further analysis (pp. 12-13, lines 255-58). We sought to provide an analysis of recent trends for EEoCs, but also wanted to select a period that would not include an EEoC in early 2014 affecting 41 publication, which would skew the analysis.

Reviewer #1: “The authors analysed the time for retractions to occur after the EEoCs were issued. Can they discuss how this delay affects the 2014-2016 sample? Are there any limitations because many retractions would not have come through for more recent EEoCs?”

We have expanded discussion of this point in the Discussion (p. 22, lines 435-41). The mean and median time to retraction after an EEoC is issued are less than 1 year. As of 8 December 2016, the proportion of publications retracted in the 2-year subset was similar to that of the larger set. However, we estimated that about a third of EEoCs in the 2014-2016 subset remain unresolved. Publisher practices could also change. So additional retractions may be observed for this cohort.

Reviewer #1: “Figure 3: It would be helpful to have an expanded legend here… (“Unretracted” should be “Publication unretracted”).”

Upon further consideration, we decided the way the term “Unretracted” was used here was confusing. We have modified Figure 3 to reverse the order and change the label from “Unretracted” to “Other or no action.” We have also revised the title of Figure 3 (p. 35, lines 719-21). We think these revisions, along with accompanying text in the Results, provide more clarity.

Reviewer #1: “The discussion states that 62 publications represent 82% of the 76 retracted publications shown; 2 retracted papers couldn't be analysed fully - what are the remaining 12?”

We have revised the Results to explicitly account for all retractions (p. 17, lines 345-46). Figure 3 illustrates observed outcomes for all 300 affected publications that we identified. Survival analyses were restricted initially to 260 affected publications that had unique PubMed records for the publication and the associated EEoC, resulting in exclusion of 12 retracted publications. Another 2 retracted publications were dropped from the time-to-retraction analysis because unique PubMed records were not provided for the retraction.
Reviewer #1: “Discussion: The authors stated ‘Although we had no language restrictions, our search strategies would not have identified an EEoC entirely in a language other than English’? I understand what they mean, but as written the statement seems to a contradiction.”

We have edited this sentence for clarity (p. 23, lines 459-462).

Reviewer #1: “It would be interesting to have some references to the total number of retractions in the investigated time frames (e.g. how many retractions are issued without being preceded by an EEoC; what is the delay between publication and retraction when there is no EEoC).”

We have added a paragraph to the Discussion to place our findings in the context of others’ published work on the time to retraction for PubMed-indexed publications (pp. 22-3, lines 443-53); two references (#35 and 36 in revised manuscript) were added. In the Results, we provided the time between EEoC and retraction, but to support this section of the Discussion, we have also added summary statistics for the time from publication to retraction (p. 17, lines 354-56).

Reviewer #1: “This is clearly beyond the scope of this study, which was motivated to focus on PubMed, but if the authors know how common EEoCs are in other disciplines and how they are handled, they might want to include a brief discussion about this.”

As the reviewer notes, our study focused on PubMed-indexed publications. Our searches found 40 EEoCs affecting publications outside PubMed, which were excluded from our analysis. We have added a paragraph to the Discussion to note this limitation, particularly as it concerns the broader scholarly literature (p. 23, lines 466-71).

Reviewer #1: “The Additional data files are all very clearly labelled and structured. It's not immediately obvious by looking at the sheet what the numbers for "Excluded - not EEoCs" refer to (in Additional data file 3) - are these also PubMed IDs?”

The description for Sheet 4 in Additional File 3 has been updated to indicate that these may be PMIDs, PMCIDs, or links to source pages.

Reviewer #2: “This paper offers a well-written and clear report about descriptive data on the prevalence and nature of editorial expressions of concern, a type of publication that is as controversial as it is poorly studied. Findings and limitations are adequately discussed in relation to the relevant literature, and the manuscript is concise and informative to an extent that is fully appropriate for the preliminary nature of the data presented.”

We thank the reviewer for the comments.