Editors felt your paper covered an important and interesting topic that deserves greater attention and action. We would like to work with you towards publication pending your response to reviews. We think readers will be interested in the broader issues raised by this paper, such as more information about what paper mills are and how to detect a paper mill paper. We would like to consider commissioning related content, such as a linked editorial, to accompany your paper if it is accepted for publication.

We would like to thank the editors their time and thoughtful inputs on our paper besides considering it for its potential publication. The comments and suggestions from reviewers and Editors have undoubtfully increased the quality of the manuscript and the interest for the readership.

You will see that, following the Editors' and reviewers' suggestions, we have updated the data for our analysis, including paper-mill papers retracted until 26/06/2022, so the results section has been rewritten to include this update. To do this, Retraction Watch has provided us the available database.

You will find below a point-by-point answer to the reviewers' comments, using underlined text to indicate text added to the manuscript, and we also have attached a tracked-changes version of the manuscript in order you can see how and where we have changed the paper.

We ask that you first please revise your paper to respond to all of the comments by the reviewers. Their reports are available at the end of this letter.

Please also respond to these additional comments by the editorial committee below:

Statistical editor:

1. Some statements are about paper-mill papers in total (not just retracted). How are the paper-mill papers identified in general? This is not currently explained.

We appreciate the opportunity to provide a more detailed explanation of this point.
The main tool to flag potential paper-mill papers is based on image analysis, looking for manipulated and/or recycled images (stock or duplicates images) included in manuscripts. However, although there is software available for detecting manipulated images, at the moment there is no reliable software capable of analyzing recycled images in a manuscript. Although it is true that some journals are using artificial intelligence software, some experts indicated that automated checks might yield too many false positives. A further problem is that a low percentage of manuscripts include images that can be scanned for the potential identification of a paper-mill paper. We have included this information in the discussion section:

Page 19. "Although there are different softwares capable of detecting image manipulation, paper-mill papers often use duplicated images (or stock images) (5,19) , since they are more difficult to detect than manipulated ones. Currently there is no software capable of detecting image-duplication in a reliable way, thus leaving this task to editors and reviewers. That said, however, few papers contain images that allow for scrutiny."

Also, following this comment, we have included some information about strategies that allow editors and investigators to identify suspicious publications, such as image manipulation software, the “tortured phrases” detection software or other common indicators of paper-mill papers published recently by COPE (such as those for papers in the field of cellular and molecular biology, non-institutional email-addresses, substantial changes to the author list during revision or proof corrections, among others). However, it should be noted that these strategies do not identify unequivocally paper-mill products but can help in the screening of papers. A further problem is that these strategies are not applied systematically and not all of them can be approached using a software.
We have included this information in the discussion section:

Page 19. “Another strategy for screening questionable papers is the Problematic Paper Screener software. This software identifies “tortured phrases”, i.e. unusual phrases instead of established ones, which may be an indicator of suspected scientific misconduct (20). Also, COPE has published a list of common indicators for paper-mill papers which could serve as a screening tool for suspicious articles (18).”

Importantly, the identification and categorization of retractions as paper-mill papers is done by Retraction Watch, the organization that provided the data we used for this study, not by the research team. As Retraction Watch has explained, detection of paper mills is not an easy task, since they look like reliable articles. However, the organization employs several indicators and sources to assign “paper mill” as a cause for retraction. First, the notice of retraction. Some of them clearly indicate that the paper is from a paper mill or use a euphemism for this (“third-party editing service”). In other cases, journals/publishers issue an accompanying editorial indicating that the papers being retracted in the group were from paper mills. These issues usually use similar language: “resembles different papers from different authors”, i.e. paper-mill products. They also use PubPeer comments and the lists of articles of probable paper-mill products published by investigators such as Dr. Elisabeth Bik. Following this comment, we have added more information about how Retraction Watch assigns “paper mill” as a reason for retraction in the methods section:

Page 10. “In the specific case of paper-mill products, their identification is based on several indicators: notice of retraction, some clearly state that it is a paper-mill paper, others use a euphemism for paper mill such as “third-party editing service”. In other cases, journals/publishers retract a large number of articles accompanied by an editorial indicating that the retracted papers were from paper mills. These usually use a similar language, stating “resembles different papers from different authors”. Retraction Watch also uses PubPeer and the list of probable paper-mill papers published by Dr. Elisabeth Bik and other investigators (15).”

2. How do retraction watch classify retractions? How many might they have missed?

Retraction Watch (RW) identifies retracted articles and retraction notices by conducting standardized manual searches of publishers’ and editors’ sites, which are the primary source for the identification of retractions. RW run these searches daily using some specific keywords (“retraction”, “withdrawn”, “retracted paper”). Moreover, they search reports of

research misconduct investigations and social media sites, as well as receive tips from their blog followers. Also, RW uses PubMed in order to double-check the retractions.
Taking this into account, RW has a greater coverage than PubMed and CrossRef and other tools, such as Web of Science, are now licensing their data.
Following the editor's comment, this information was added to the manuscript in order to clarify the methodology followed by the RW database:
Page 10. "The Retraction Watch database has higher coverage of retractions than PubMed and CrossRef since they use different sources to detect retracted articles and notices of retraction. The main sources for the identification of retractions are publishers' and editors' sites, but reports of scientific integrity investigations, social media sites and tips from their blog followers are also checked. In order to double-check the retractions, they use PubMed and Web of Science.
In order to identify retractions, Retraction Watch runs protocolized manual searches daily, using keywords such as "Retraction", "Withdrawn" or "Retracted Paper", among others. "
We have also added some information in the discussion section regarding the great coverage RW has, which we consider a strength of our study:
Page 21. "Retraction Watch database has three times the coverage of PubMed and five times the coverage of CrossRef. Taking this into account, we consider that he number of missing retractions should be minimal."
To categorize retractions by 'reason', RW uses not only the information included in notices of retraction, but they also check other sources for background and clarification of information. These sources include reports assessing institutional investigations and reports from the U.S. Office of Research Integrity (ORI). Occasionally, they review other media sites (The New York Times, Times Higher Education…) for background. We have added more information regarding this issue in the Methods section:

Page 10. "To classify retractions into different reasons, RW uses mainly the information included in the notice of retraction. Also, RW checks other sources for clarification of information such as institutional investigation reports and U.S. Office of Research Integrity (ORI) reports."

However, we are aware of the difficulty in assigning causes of retraction in some cases, so a certain risk of misclassification is possible, not only for paper-mill papers retractions but for all retractions. This was added as a limitation of the study:

Page 21. "Another limitation is the difficulty in assigning the cause of retraction in some cases, hence there may be some risk of misclassification."

3. Especially with covid papers, are retracted papers on the increase anyway (in general)? I this specific to paper-mill papers?

Several authors have analyzed the phenomenon of retractions in the recent decades. The absolute number of retractions is increasing so this is not specific to paper-mill papers. This increase could be in part because of the rising in the number of papers published but also of the availability of improved mechanisms to detect errors/misconduct.
Although some COVID-19 papers have been retracted, we think that so far this has not had an extraordinary impact on the number of retractions but on the publicity achieved by those

retractions. It should also be noted that the retraction of an article takes a long time, and it is possible that more retractions of COVID-19 papers will appear in the future. However, we consider that there has not been a significant increase in the number/proportion of retractions due to COVID-19 retractions.
Regarding COVID-19 papers and paper-mill papers, we identified no retracted paper-mill papers about COVID-19. However, we expect that as time passes, given the prominence of the COVID-19 pandemic, it is only a matter of time before retractions of COVID-19 papers are identified as paper-mill papers.

4. The last figure should say that the quartiles relate to journal impact factor.

Thank you for pointing this out. We have added the following title on the x-axis of the figure: “Journal quartile based on the Impact Factor”. Also, the title of this figure was modified for clarification. Please check the tracked changes version.

5. We are missing absolute numbers of retractions and total papers published per year (often just get told number per 100000 papers)

We have added the total number of publications and retractions during the study period (January 2004-June 2022) in the first paragraph of the Results section:

Page 12. “During the study period, a total of 58,278,163 papers were published and 33,741 were retracted for any reason, including being a paper-mill product, a rate of 57.9 retractions per 100,000 publications.“

Research editors:

1. Does RW cover all retractions? Is there a way to check this?

Below you will see our response to the Statistical Editor’s 2nd comment, in which we address this issue and provide more detail on how Retraction Watch database has been aggregated and updated to identify retractions:

Retraction Watch (RW) identifies retracted articles and retraction notices by conducting standardized manual searches of mainly publishers’ and editors’ sites, which are the primary source for the identification of retractions. RW run these searches daily using some specific keywords (“retraction”, “withdrawn”, “retracted paper”). Moreover, they search reports of research misconduct investigations and social media sites, as well as receive tips from their blog followers. Also, RW uses PubMed in order to double-check the retractions.
Taking this into account, RW has a greater coverage than PubMed and CrossRef and other tools, such as Web of Science are now licensing their data.
Following the editor’s comment, this information was added to the manuscript in order to clarify the methodology followed by the RW database:

Page 10. “The Retraction Watch database has higher coverage of retractions than PubMed and CrossRef since they use different sources to detect retracted articles and notices of retraction. The main sources for the identification of retractions are publishers’ and editors’

sites, but reports of scientific integrity investigations, social media sites and tips from their blog followers are also checked. In order to double-check the retractions, they use PubMed and Web of Science.
In order to identify retractions, Retraction Watch runs protocolized manual searches daily, using keywords such as “Retraction”, “Withdrawn” or “Retracted Paper”, among others. “
We have also added some information in the discussion section regarding the great coverage RW has, which we consider a strength of our study:
Page 21. “Retraction Watch database has three times the coverage of PubMed and five times the coverage of CrossRef. Taking this into account, we consider that the number of missing retractions should be minimal.”

2. How does RW assign ‘paper mill’ as the reason for retraction?

Below you will see our response to the Statistical Editor’s 1st comment, in which we provide a wider explanation about how Retraction Watch assign “paper mill” as a reason for retraction:

Importantly, the identification and categorization of retractions as paper-mill papers is done by Retraction Watch, the organization that provided the data we used for this study, not by the research team. As Retraction Watch has explained, detection of paper mills is not an easy task, since they look like reliable articles. However, the organization employs several indicators and sources to assign “paper mill” as a cause for retraction. First, the notice of retraction. Some of them clearly indicate that the paper is from a paper mill or use a euphemism for this (“third-party editing service”). In other cases, journals/publishers issue an accompanying editorial indicating that the papers being retracted in the group were from paper mills. These issues usually use similar language: “resembles different papers from different authors”, i.e. paper-mill products. They also use PubPeer comments and the lists of articles of probable paper-mill products published by investigators such as Dr. Elisabeth Bik.
Following this comment, we have added more information about how Retraction Watch assigns “paper mill” as a reason for retraction in the methods section:

Page 10. “In the specific case of paper-mill products, their identification is based on several indicators: notice of retraction, some clearly state that it is a paper-mill paper, others use a euphemism for paper mill such as “third-party editing service”. In other cases, journals/publishers retract a large number of articles accompanied by an editorial indicating that the retracted papers were from paper mills. These usually use a similar language, stating “resembles different papers from different authors”. Retraction Watch also uses PubPeer and the list of probable paper-mill papers published by Dr. Elisabeth Bik and other investigators (15).”

However, we are aware of the difficulty in assigning causes of retraction in some cases, so a certain risk of misclassification is possible, not only for paper-mill papers retractions but for all retractions. This was added as a limitation of the study:

Page 21. “Another limitation is the difficulty in assigning the cause of retraction in some cases, hence there may be some risk of misclassification.”

3. Fig 1 shows ”Percentage of paper-mill papers PUBLISHED per year with respect to total publications”, should it be “retracted”, or “published and then retracted”? Do we know the specific numbers of paper-mill papers published if they were not retracted and captured by Retraction Watch?

Thank you for pointing this out. Figure 1 shows the percentage of paper-mill papers published in one year and then retracted with respect to the total number of papers published in that same year. This is to assess the proportion of paper-mill papers published (and then retracted, since we do not know how many of these papers are still in the literature) per year.
We have clarified this issue in the methods section:

Page 11. “We then calculated the rate per 100,000, of paper-mill papers published in a given year which were then retracted, over the total number of papers published for the same year in order to assess the percentage that papers-mill publications regarding the total number of publications in each year of the study period.”

4. How many of these papers were published in predatory journals?

Thank you for this relevant question. Following this comment, we have analyzed the 15 scientific journals which published the largest number of paper-mill paper retractions. We used ThinkCheckSubmit (https://thinkchecksubmit.org/) checklist to identify if a journal is reliable or not. This checklist is supported by publishers such as Springer, the Committee on Publication Ethics (COPE) and the Directory of Open Access Journals (DOAJ), among others. Although this checklist does not allow to identify predatory journals, a low score in it could detect suspicious journals. While it is true that several checklists to detect predatory journals have been published in the last years, none of them seem to be optimal and they lack assessment of reliability and validity.
Although it would be reasonable to think that paper mills would be published in predatory journals, according to the checklist used, all these 15 journals appear to be legitimate. Also, predatory journals might be less scrutinized.
This information you have suggested has been added in the methods and results sections of the manuscript:

Page 12. “We have determined if the journal of publication is reliable, i.e. not suspected of being a predatory journal, in the subsample of the fifteen scientific journals which published the most retracted paper-mill papers. To assess the reliability of a journal we used the ThinkCheckSubmit checklist (available in: https://thinkchecksubmit.org/journals/).”

Page 14. “Fifteen scientific journals published a total of 812 (68.8%) of all papers retracted for being paper-mill papers, and 166 (14.0%) were published in a single journal, the European Review for Medical and Pharmacological Sciences. Of these, all appear to be legitimate journals, i.e. non-predatory journals.”

Also, a sentence was included in the first paragraph of the discussion.

Page 17. “Despite what one might think, this phenomenon affects legitimate journals and does not seem to be exclusive to predatory journals.”

5. From the JCR category in table 1, it seems most of papers were out of the scope of interest for BMJ readers, namely, clinical medicine and public health? Could the authors comment on the kinds of topics covered by paper mill papers and whether the profile of topics has changed over time?

Thank you for this comment. The vast majority of the paper-mill papers identified until now cover basic research papers, especially cellular and molecular biology. The number of published papers in basic research areas is greater than that of clinical medicine. For example, according to Web of Science, 13.1% of the total publications in 2021 belonged to the field of Biochemistry and Molecular Biology and Chemistry publications accounted for the 11.6% of the total publications. In contrast, Oncology (the clinical medicine field with more publications in 2021) accounted for 6.6% of the total publications. Because of that, it is more likely that basic research fields have a greater number of retractions.
It should be noted that we have found retracted paper-mill papers published in clinical discipline journals, including Oncology, Endocrinology, Cardiovascular systems, or Psychiatry.
However, this pattern could change over time as COPE indicates in its last report. Also, in our results we have not observed important variations in the topics covered by paper-mill papers over time.
Following this comment, we have included the following paragraph in the discussion section:

Page 19. “The vast majority of identified paper-mill papers cover basic research topics, such as cellular and molecular biology or biochemistry. An explanation could be that publications pertaining to basic science fields, such as biology or chemistry, outnumber publications on clinical medicine and public health. Therefore, basic research fields are more likely to have a higher number of retractions. According to our results, no major variations over time have been observed in the topics covered by the paper-mill papers so far. However, the latest COPE report indicates that this may change over time (12). “

In your response please provide, point by point, your replies to the comments made by the reviewers and the editors, explaining how you have dealt with them in the paper.
Done

Comments from Reviewers

Reviewer: 1

Recommendation:

Comments:
Review of BMJ-2022-071517

This manuscript concerns an interesting and important analysis of a highly relevant and quite alarming topic: the rise of paper mills. It's well written and the study design is clear and adequately executed. A number of issues need to be repaired however and some minor points should be settled.

Thank you for the positive and thoughtful input on our paper. Please, check below a detailed reply to your comments and concerns.

Major issues

1. Please make very clear in the abstract and the main body of the text that retracted paper mill papers may differ substantially from (possibly the large majority of) paper mill papers not retracted.

Thank you for this comment. We agree with the reviewer regarding this point, and we do think that the characteristics of paper-mill papers retracted may be different of those of paper-mill papers non-retracted. On the one hand, retracted paper-mill papers may commit more flagrant misconduct than non-retracted ones. It is also highly possible that paper-mill companies are improving their technology/resources/ability to better jeopardize a real paper. On the other hand, since the relevant existence of these papers in the literature has been discovered, they have been associated with China and hospital-affiliated authors, as well as with basic research. Editors are using these characteristics to screen papers susceptible of being paper-mill products. It is possible that other paper-mill products do not match these characteristics and therefore have not been identified or thoroughly screened for the moment. It should be noted that it is likely that hundreds (if not thousands) of paper-mill products are circulating in scientific literature, and these may have different features than detected ones.
Following this comment, we have added a sentence regarding this issue in the discussion section:

Page 20. "It should be noted that the outward characteristics of retracted and non-retracted paper-mill papers may differ, which could explain why some were identified but not others, although all represent fraudulent science."

And also in the abstract:

Page 6. "Nevertheless, detected paper-mill papers might be substantially different from those non detected"

2. Please make very clear in the abstract and the main body of the text whether citations were counted until retraction or until the end of the observation period. Ideally this information should be presented stratified (before versus after retraction).

The citations were collected from publication until 28 June 2022, since we have updated the database, including paper-mill paper retractions up to June 26th 2022. These citations were received during both pre-retraction and post-retraction periods and, in this analysis, we have not differentiated between pre and post-retraction citations. We are really

aware of this aspect and in fact published a paper recently where we have shown that retractions hardly had any impact on subsequent citations (Candal-Pedreira C, Ruano-Ravina A, Fernández E, Ramos J, Campos-Varela I, Pérez-Ríos M. Does retraction after misconduct have an impact on citations? A pre-post study. BMJ Glob Health. 2020;5(11)). We agree with the reviewer and this issue should be taken into account in future studies. Following this comment, this issue was clarified in the methods section and was added as a limitation in the discussion section.

Page 10. “For every paper included, the total number of citations received, both pre-retraction and post-retraction, from date of publication until 28 June 2022 were collected.”

Page 21. “A limitation regarding the citation analysis is that pre-retraction and post-retraction citations have not been differentiated in this study and this issue should be considered in future research.”

3. You declare that the Retraction Watch database is the gold standard but did you check whether Web of Science or Medline feature retractions not contained in the RW database?

Thank you for this question. As explained in our prior response to the Statistical Editor’s 2nd comment, Retraction Watch database uses different sources for the identification of retractions.
Retraction Watch (RW) identifies retracted articles and retraction notices by conducting standardized manual searches of mainly publishers’ and editors’ sites, which are the primary source for the identification of retractions. RW run these searches daily using some specific keywords (“retraction”, “withdrawn”, “retracted paper”). Moreover, they search reports of research misconduct investigations and social media sites, as well as receive tips from their blog followers. Also, RW uses PubMed in order to double-check the retractions.
Taking this into account, RW has a greater coverage than PubMed and CrossRef and other tools, such as Web of Science are now licensing their data.
Following the editor’s comment, this information was added to the manuscript in order to clarify the methodology followed by the RW database:
Page 10. “The Retraction Watch database has higher coverage of retractions than PubMed and CrossRef since they use different sources to detect retracted articles and notices of retraction. The main sources for the identification of retractions are publishers’ and editors’ sites, but reports of scientific integrity investigations, social media sites and tips from their blog followers are also checked. In order to double-check the retractions, they use PubMed and Web of Science.
In order to identify retractions, Retraction Watch runs protocolized manual searches daily, using keywords such as “Retraction”, “Withdrawn” or “Retracted Paper”, among others. “
We have also added some information in the discussion section regarding the great coverage RW has, which we consider a strength of our study:
Page 21. “Retraction Watch database has three times the coverage of PubMed and five times the coverage of CrossRef. Taking this into account, we consider that the number of missing retractions should be minimal. “

4. You seemed to have missed the work by Guillaume Cabanac on ‘Tortured Phrases’ and his wonderful software package the Problematic Paper Screener (https://dbrech.irit.fr/pls/apex/f?p=9999:1::::::). Please add this.

Thank you, we have added this information in the discussion section. These “totured phrases” very well defines one of the distinctive aspects of paper-mill papers. Please, check the tracked changes version.

5. You seem to have missed that software for detecting image manipulation (e.g. FigCheck, Forensically, DARPA, ImageTwin, Proofig and MediaFor) is already available, although not yet functioning impressively. Please add this.

Thank you for pointing this out, we have added this in the introduction section. Please see the tracked changes version of the manuscript.

6. You present IF quartiles in the main text, tables and figures but fail to explain whether Q1 consists of the lowest or the highest IFs. Please add this information.

This issue was clarified in the methods section for a better understanding, adding the following:

Page 10. “…JCR category, relative position (JCR quartile, with the highest IF journals belonging to first quartile or Q1 and the lowest IF journals belonging to fourth quartile or Q4.).”

7. The first author of retracted paper mill papers has a hospital affiliation in 91.5% (82.8+8.7) the included cases. Please comment in the Discussion section whether this may be an artefact due to more intense scrutiny of biomedical publications.

As the reviewer rightly pointed out, we have observed that the vast majority of authors were affiliated with hospitals and this result is consistent with previous works. There might be different reasons for this. On the one hand, there are a greater number of publications in the biomedical field in comparison with other disciplines or areas of knowledge and consequently, the number of paper-mill papers will be also greater in the biomedical literature. On the other hand, the specific phenomenon of paper mills may have a greater impact in hospital-affiliated authors since the pressure to publish is greater in biomedical sciences and publications are needed in order to get a university degree or a promotion. This, combined with the fact that physicians do not always have time (or resources) to do research may explain this finding.
Following this comment, we have added some information about this issue in the discussion section:

Page 19. “Also, it was observed that the great majority of authors of identified paper-mill papers were hospital-affiliated, which is consistent with previous works (15). The main reason for this may be that Chinese doctors are not affiliated with medical schools, but with hospitals. Also, it should be noted that pressure to publish is greater in biomedical sciences and publications are usually needed to get a university degree or a promotion (15).”

Minor issues
8. Why did the time window stop in (December?) 2019? Consider updating for at least 2020 and 2021.

Thank you for this comment. We have updated the database, including paper-mill papers retracted until 26/06/2022. Please check the tracked changes version.

9. Line 222 and 212 are not clear. Do you really say here that 51.0% of all retracted paper mill papers you found were published in the second quartile of the JCR category of Biochemistry and Molecular Biology?

Thank you for pointing out this mistake. It was corrected in the manuscript.

Reviewer: 2

Recommendation:

Comments:
Candal-Pedreira and colleagues have presented the results of a comprehensive cross-sectional analysis on paper mill articles published and retracted between 1st Jan 2013 to 27th Sept 2021 (as identified in the Retraction Watch database). They identified 622 retracted articles, all of which included authors from China (majority of which are affiliated with hospitals), and note that 51.9% of these articles are published in journals in Q2 journals within the Biochemistry and Molecular Biology JCR category. The authors have listed the number of retracted paper mill articles by journal title (including OA status) and publishing house, and note that retractions are issued more quickly in Q1 and Q2 journals. In addition, the authors have noted that paper mill articles published in Q4 journals received more citations compared to paper mill articles published in Q1-Q3 journals.

Detection of paper mills is a significant and evolving issue, and affects many different parts of the publishing process – e.g. dishonest authorship practices, fake data (including stock images), plagiarism, unethical citation practices, peer review manipulation, special issue manipulation, fake guest editors, etc. There have been many editorials, news articles and perspectives published on the topic, but this is the first cross-sectional analysis covering retraction of paper mill articles between 2013 – 2021 that I'm aware of.

This is an issue affecting scholarly articles across the industry, and is therefore of relevance to researchers from various disciplines who of course rely on the integrity of published content to inform their own research/education/policy work etc. This cross-sectional analysis addresses retracted paper mill articles within biomedical journals, most of which (to my knowledge) are non-clinical research articles (although can often involve the use of human tissue samples).

As someone who has been actively involved in dealing with paper mills, I'd like to thank the authors for producing this research, as this comprehensive analysis will be of great interest to many in the field. I have the following specific comments for the authors to consider:

Thank you very much for the positive comments on our work, which continues a research line started some years ago and is starting to produce very relevant results.

1. The Retraction Watch database was used to identify paper mill articles which have been retracted due to being paper mills. Can the authors comment on how to Retraction Watch identify these as being paper mill articles? Do they rely on the reasons stated within the retraction note (which varies in level of detail across publishers) or do they verify these as being paper mill articles by other means?

To classify retractions into different reasons, RW uses not only the information included in notices of retraction, but also checks other sources for background and clarification of information. These sources include reports assessing institutional investigations and reports from the U.S. Office of Research Integrity (ORI). We have added more information regarding this issue in the Methods section:

Page 10. "To classify retractions into different reasons, RW uses mainly the information included in the notice of retraction. Also, they check other sources for clarification of information such as institutional investigation reports and U.S. Office of Research Integrity (ORI) reports. "
In the specific case of paper-mill papers retractions, due to the difficulty in their detection, RW uses different sources and indicators: notice of retraction, PubPeer and the list of probable paper-mill papers published by Dr. Elisabeth Bik and other investigators. Regarding this, we have included a paragraph in the methods section:
Page 10. "In the specific case of paper-mill products, their identification is based on several indicators: notice of retraction, some clearly state that it is a paper-mill paper, others use a euphemism for paper mill such as "third-party editing service", in other cases, journals/publishers retract a large number of articles stating "resembles different papers from different authors" and issue an editorial indicating that papers being retracted in a group were from paper mills. Retraction Watch also uses PubPeer and the lists of probable paper-mill papers published by Dr. Elisabeth Bik and other investigators (14)."

2. Can the authors specify the date(s) the Retraction Watch database was accessed to source the data? Was this all done on the same day or over several days (and if so, would this have changed the results)?

The database was first accessed on September 27th 2021 and data were extracted during the following week. However, following the suggestion of the reviewers, we have updated the database. The last access was June 26th 2022. This was added to the methods section. Results were not different following this update. Please check the tracked changes version of the manuscript.
Regarding the citations, we have collected all citations from publication until a specific date (June 28th 2022), as stated in the manuscript.

3. I'm surprised that the authors have not cited Elisabeth Bik's post from 2020 titled 'The Tadpole Paper mill' on Science Integrity Digest https://scienceintegritydigest.com/2020/02/21/the-tadpole-paper-mill/ , as this post contained a list of over 400 potential paper mill articles. This had a big impact on raising awareness about the widespread issue of papermills, noted several red flags which could help identify paper mills, and therefore led to large scale investigations being conducted by several publishers from 2020 onwards. Are the authors familiar with this report, and if so have they considered cross-referencing the articles listed in Bik's post with the data they sourced from RW? This could give an indication of differences/similarities in dealing with paper mills between journals/publishers.

Thank you for pointing out this work, we have cited it now in our manuscript. We are familiar with this post, and we think it is of great importance as the reviewer has mentioned. In fact, RW database uses information from the list in order to classify a retraction of an article for being a paper-mill product. However, the articles included in that database are potential paper-mill products and many of them have not been retracted. The inclusion of papers from this list may imply the inclusion of articles that are not paper-mill products, and this could be a limitation of the study. Because of that we have only included formally retracted papers. Nevertheless, we recognize that journals are slow and sometimes reluctant to retract papers and many of the Bik's identified papers will be finally retracted. We have included the following sentences highlighting that this is a limitation.

Page 21. "In this study, we have included formally retracted paper-mill papers, not taking into account suspicious papers (i.e. those from the list elaborated by Dr. Elisabeth Bik and others) and this may be a limitation of the present research. The inclusion of papers not formally retracted might incur in a risk of misclassification of those papers if they are not finally retracted as paper-mill products. "

4. Regarding the observation about paper mill articles being cited by other articles, can the authors comment on who is citing these- i.e. are these are paper mill articles which have also been identified, or are those other articles which don't seem to be paper mills?

Thank you for this question. Unfortunately, we have not analyzed who cites paper-mill products. Due to the unusual pattern of citations observed in our work, we believe that this might suppose a fraudulent pattern of citations. However, at this time, we have not analyzed this aspect. We are really interested in citation patterns of retracted papers and have some work published on this (Candal-Pedreira C, Ruano-Ravina A, Fernández E, Ramos J, Campos-Varela I, Pérez-Ríos M. Does retraction after misconduct have an impact on citations? A pre-post study. BMJ Glob Health. 2020;5(11)). The next step will be to identify who is citing this, and specifically regarding paper mills paper we are convinced that the same paper-mills are citing again and again the previous retracted papers to use this as marketing for those buying their services.

5. As we're now in May 2022, have the authors considered updating their search and results? This would be a worthwhile exercise given how quickly this field has been moving and how many new retractions are likely to have been issued since Sept 2021.

As we mentioned above, we have updated our database until June 26th 2022 following the comments of both reviewers. Please check the tracked changes version.