PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Definitions of blinding in randomized controlled trials of interventions published in high-impact anesthesiology journals: a methodological study and survey of authors
AUTHORS	Penic, Antonija; Begic, Dinka; Balajic, Karolina; Kowalski, Martin; Marusic, Ana; Puljak, Livia

VERSION 1 – REVIEW

REVIEWER	Jon Henrik Laake National Hospital - Rikshospitalet, Dept. of Anaesthesiology
REVIEW RETURNED	07-Nov-2019

GENERAL COMMENTS	The authors provide empirical data demonstrating that the term "blinding" (in any version) is often misunderstood and misused in randomized trials published in seven anaesthesia journals. The methods are adequate, limitations discussed and my only reservations are 1) Language is mostly excellent, but there are a few errors (e.g. 'not' in conjunction with neither/nor vs either/or p 12, line 8). A thorough reading by an expert should sort this out quickly. 2) The discussion needs to be better balanced. The authors may find that the literature describes a differential effect of blinding on outcome assessment depending on severity of outcomes.

REVIEWER	Agnes Dechartres Sorbonne Université, France
REVIEW RETURNED	19-Nov-2019

GENERAL COMMENTS	Review BMJopen-2019-035168 This manuscript reports the results of a research on research study evaluating the reporting of blinding in 622 RCTs published between 2014 and 2017 in the 7 anesthesiology journals with the highest impact factor. The authors used various complementary approaches to study this question. First, they evaluated whether and how blinding was reported in title, abstract and methods section of the publication. Second, they evaluated whether key individuals (ie, participants, care providers, outcome assessors, data analysts) were blinded or not. Then, they searched for corresponding registration and evaluated the reporting of blinding in registries and whether there were discrepancies within the registry and between the registry and the publication. Finally, they conducted a survey of authors on the definitions of blinding. The topic is interesting and the manuscript is overall well-written. This represents an important amount of work and a significant contribution to the field. My main criticism is that the manuscript is

	a bit long and that some parts could be summarized with important points (in results and discussion) better highlighted. Please find below my detailed comments. - The authors should justify why they focus only on anesthesiology journals. - I think that the use of “protocol” is confusing to refer to registration as it is not really a protocol. In addition, the authors could have also searched whether protocols were available as a supplementary appendix or published - The authors analyzed whether there were discrepancies in the reporting of blinding between the abstract and the body of manuscript and between the registry and the publication. But it is unclear how they define a discrepancy. For example, if it is indicated in the abstract ‘double blind’ and if they describe that participants and care providers are blinded in the body of manuscript, is it considered as a discrepancy? It seems so later (cf page 13, line 56: “the most common differences were in the description of who was blinded”). This is questionable as it is normal to have less detail in the abstract and more details in the body of manuscript. - The first part of the results about descriptors of blinding in the title, abstract and body of manuscript is very repetitive. - Some parts of the results are also very long such as the results described in the paragraph ‘Other analyses related to blinding’ especially those concerning the reporting of blinding in the different parts of registration - The discussion is also very long. Some results from previous studies are too much detailed. I would focus on the recent literature on the field. In contrast, I would further detail why despite recommendations including the CONSORT Statement, description of who is blinded in a trial is so few reported. Is it because the editors of journals do not require submission of the CONSORT checklist? How could we improve that? - The following sentences or elements are not clear o Abstract: “In most of protocol-manuscript pairs without discrepancies within a protocol, we found discrepancy regarding reporting of protocol between the protocol and published manuscript” o Introduction: the reference to textbooks is unclear o Methods: in the part concerning the author survey, question 1: for authors who only indicated that study was blinded,… I don’t understand why there is “categorical answers” at the end of the sentence. There is no such mention for other questions. I also don’t understand the sentence “We emphasized to the survey participants that only aggregate data will be published” o In the results section, paragraph ‘Survey results’, it is unclear to mention the third question line 35, page 14 - I am not sure that “inadequate” is the good term when the authors say that the reporting of blinding is inadequate in conclusion and strengths and limitations. I would rather say that the reporting was poor or insufficient. - Additional questions or comments o Which filter was used to identify RCTs? This is not indicated. o Results section, paragraph “Blinding of individuals involved in trials”: the results reported there refer to the analysis of who was blinded for the 531 trials where they found at least some descriptors for blinding but later in the paragraph, the authors give a percentage calculated from the whole sample of 622 trials (the 14%). It should be 89/531=17%, no?

	o Results section, the title of the paragraph “other analyses related to blinding” could be replaced by a more informative title and split in two parts: “discrepancies between abstract and body of manuscript” and “Reporting of blinding in registries”, or something like that. The sentence concerning the CONSORT Statement should be better placed elsewhere for example in a paragraph about general characteristics of trials that is missing or in the paragraph ‘”Blinding of individuals involved in trials” as reporting who is blinded is a clear recommendation of the CONSORT Statement. o It would have been interesting to have a description of the included trials with the type of intervention (ie, pharmacological or not pharmacological), the year of publication, the sample size, the funding source o The authors included trials from 2014 to 2017. Is there any improvement over time? Is it possible to have a look at this? o Discussion section, for me the lack of comparison between journals is not a limit.

REVIEWER	Manoj Lalu Ottawa Hospital Research Institute, Ottawa, Canada
REVIEW RETURNED	28-Nov-2019

GENERAL COMMENTS	I have had the opportunity to review Penic et al’s study titled, “Definitions of blinding in randomized controlled trials of interventions published in high-impact anesthesiology journals: a methodological study and survey of authors”. A sample of anesthesiology related RCTs from 2014-2017 were analyzed. They demonstrated that only a small minority of trials report blinding adequately. A follow-up survey found that authors differed significantly in terms of how they defined blinding. As a disclosure: I am a practicing anesthesiologist and a clinical/preclinical meta-researcher. I have performed several reporting assessments. Overall, I found this study to be interesting. I had several questions, centered largely around the methods, that I have detailed below. MAJOR: 1. There are many other assessments of reporting and blinding in published RCTs. The authors should provide a synopsis in the introduction of the multiple similar studies to their own. The one by Devereaux et al is quite good, however there is a massive literature in this area that the authors must summarize for a naïve reader. Some of this is nicely summarized in the Discussion, however a very brief (e.g. short paragraph), high level overview of this literature should be provided in the introduction in order to provide context to the reader. 2. Was a protocol ever registered (e.g. Open Science Framework) or at least deposited in a time-stamped manner (e.g. institutional repository)? This is a basic component of best practices surrounding meta-research. 3. It is unclear why the authors used the search strategy they did. The preset filter in PubMed is not sensitive. Several groups have developed RCT filters for PubMed, e.g. Cochrane (https://work.cochrane.org/pubmed). Using this filter picks up ~10X many more hits than simply using the RCT “article type” filter in PubMed. I just compared the Cochrane filter in PubMed - ~4 million hits, versus simply selecting ‘RCT’ in the preset article type filter that only retrieved ~480 000 hits. In my mind, this is highly problematic, and raises critical questions with regards to the corpus of studies that were investigated.

	4. The definition around the primary outcome is unclear – how was 'clearly described' actually operationalized? What does 'clearly described' mean? In other words, what was the threshold to be 'clearly described'? This is critical, since the CONSORT checklist describes three parties participants, care providers, those assessing outcomes), whereas the longer form explanation describes four parties (participants, healthcare providers, data collectors, and outcome adjudicators). These details are critical and must be fully elaborated in the methods. 5. It is unclear to me how trials were assessed where it would be redundant to describe aspects of blinding. As a specific example, I would offer trials where ultrasound versus landmark technique for central line insertion were compared. Here it is obvious the operators could never be blinded; given the limited word counts in the journals, would the authors expect the study to report that the authors were not blinded? Although I am firm proponent of complete reporting, this particular instance would seem unnecessary to me. 6. What parts of the manuscript texts were assessed? Methods alone? The whole manuscript? What about supplemental material? (I see there is a brief mention of this in the discussion, but this needs to explained in detail with the Methods section. 7. Details of data extraction/verification are missing in the methods. How many people extracted data? Based on the CRediT taxonomy it appears five authors did? Was the data extracted in duplicate (i.e. like the study screening)? If so how were conflicts resolved? Was there any form of data extraction verification? These are critical methodological details that are absent from the manuscript. 8. Basic details around the development of the survey are missing. How were the questions developed? Was an expert in survey methodology engaged? Who piloted the survey? Were focus groups used to refine the questions? Was there a pretest with a smaller sample? Again, these are basic methodological details that I would anticipate be reported. 9. A copy of the actual survey questions, front matter etc should be provided as a supplemental file. 10. Within the questions, I was very unclear to me why #3 (i.e. estimating the prevalence of 'double-blind' trials in the literature) was asked. Moreover, I'm not sure I would initially interpret the question in the way the authors have presented the results of that question. 11. Within the results I would hope to see a flow diagram of the study selection process. How many came from each journal and each year? 12. Although the seven types of individuals that could be blinded (listed in Table 4) intuitively make sense to me, I'm unclear where these categories came from? As mentioned above, CONSORT mentions three individuals in the checklist, four in the longer form online (http://www.consort-statement.org/checklists/view/32--consort-2010/93-blinding). 13. A full listing of studies should be included in an online supplement. 14. The survey response rate is extremely low. This response rate needs to be highlighted in the abstract. Moreover, in the discussion this needs to be discussed further. The results of the survey are at very high risk of non-responder bias. 15. I'm not sure I agree with the use of the term 'retrospectively registered'. In my mind this would suggest the trial was complete and then registered. Likely the more precise term is 'after the trial was open for recruitment'. I say this because clinicaltrials.gov

	(where the majority of your data was derived) allows for registration up to 21 days after the first patient is enrolled. This means a trial could open months before their first patient is enrolled, and then registered another 21 days after the first recruitment. Thus, I believe the term 'retrospectively' would be inaccurate to use. 16. Table 8 should be reformatted – for the questions regarding double and triple blind it would be most interesting within the table to clearly show the distribution of the dyads/triads that were listed by respondents. Minor: 1. I found it stunning that only 62% of RCTs recently published were registered – especially coming from these particular journals that are generally well regarded. This seems like a significant result that should be highlighted a bit more in the discussion. 2. The authors should briefly justify the use of JIF to select articles. I have used the same strategy for some of my group's reporting assessments as well, thus I am not opposed to this approach. There are inherent problems though, which the authors could very briefly address. 3. In a similar vein, the authors should very briefly justify the focus on anesthesia related literature. 4. Introduction reference #6 is used to refer to the CONSORT statement; the CONSORT 2010 statement should likely be referenced.

VERSION 1 – AUTHOR RESPONSE

Reviewer(s)' Comments to Author:

Reviewer: 1
The authors provide empirical data demonstrating that the term "blinding" (in any version) is often misunderstood and misused in randomized trials published in seven anaesthesia journals. The methods are adequate, limitations discussed and my only reservations are

1) Language is mostly excellent, but there are a few errors (e.g. 'not' in conjunction with neither/nor vs either/or p 12, line 8). A thorough reading by an expert should sort this out quickly.

Author response: We have now revised the language, as suggested.

2) The discussion needs to be better balanced. The authors may find that the literature describes a differential effect of blinding on outcome assessment depending on severity of outcomes.

Author response: We have now revised the Discussion section based on literature suggested by the reviewers.

Reviewer: 2
The topic is interesting and the manuscript is overall well-written. This represents an important amount of work and a significant contribution to the field. My main criticism is that the manuscript is a bit long and that some parts could be summarized with important points (in results and discussion) better highlighted.

Author response: We are grateful for the kind words. We have done our best to revise the manuscript in line with the suggestions, including its length.

1. The authors should justify why they focus only on anesthesiology journals.

Author response: Pain is a subjective outcome; there is no objective method for analyzing pain, and pain – by definition – is what the patient says it is. The analgesic placebo effect has been well documented. In the area of placebo effects, expectations play a major role, triggering a cascade of endogenous opioids and non-opioids, and altering the experience of pain. For this reason, the whole field of anesthesiology and pain is likely biased if blinding is not used adequately. This is now highlighted in the revised manuscript.

2. I think that the use of "protocol" is confusing to refer to registration as it is not really a protocol. In addition, the authors could have also searched whether protocols were available as a supplementary appendix or published

Author response: We have now revised the wording, and we used the term "protocol registration" instead of protocol to avoid any ambiguities. ClinicalTrials.gov also uses term "protocol registration" (link: https://clinicaltrials.gov/ct2/manage-recs/register)

We have searched whether protocols were available as a supplementary file or published, but we did not find such reports in the analyzed studies. We now report this in the revised manuscript.

3. The authors analyzed whether there were discrepancies in the reporting of blinding between the abstract and the body of manuscript and between the registry and the publication. But it is unclear how they define a discrepancy. For example, if it is indicated in the abstract 'double blind' and if they describe that participants and care providers are blinded in the body of manuscript, is it considered as a discrepancy? It seems so later (cf page 13, line 56: "the most common differences were in the description of who was blinded"). This is questionable as it is normal to have less detail in the abstract and more details in the body of manuscript.

Author response: CONSORT for Abstracts extension stipulates that the following should be reported in an abstract about RCTs: "Blinding (masking): Whether or not participants, care givers, and these assessing the outcomes were blinded to group assignment". While we agree that it is common practice to give few details about blinding in abstracts, this should not be the case, according to the relevant reporting guideline for abstracts about RCTs. This is now highlighted in the revised manuscript. We also clarified in the methods that we considered it a discrepancy if the abstract only said "double-blind", and the body of the manuscript reported who was blinded.

4. The first part of the results about descriptors of blinding in the title, abstract and body of manuscript is very repetitive.

Author response: This section of text was reduced to two short sentences.

5. Some parts of the results are also very long such as the results described in the paragraph 'Other analyses related to blinding' especially those concerning the reporting of blinding in the different parts of registration

Author response: Thank you, we have now shortened this section in the revised manuscript.

6. The discussion is also very long. Some results from previous studies are too much detailed. I would focus on the recent literature on the field. In contrast, I would further detail why despite recommendations including the CONSORT Statement, description of who is blinded in a trial is so few reported. Is it because the editors of journals do not require submission of the CONSORT checklist? How could we improve that?

Author response: We have shortened the Discussion section in the revised manuscript, in order to reduce too many details on previous studies. We also added a comment regarding the lack of adherence to the CONSORT recommendations.

7. The following sentences or elements are not clear
Abstract: “In most of protocol-manuscript pairs without discrepancies within a protocol, we found discrepancy regarding reporting of protocol between the protocol and published manuscript”

Author response: We revised the sentence into: “Even when there were no discrepancies in the registration, we found discrepancy in the reporting of blinding between the majority of protocol registration and published manuscripts.“

8. Introduction: the reference to textbooks is unclear

Author response: This section was shortened, and hopefully now it is simpler and clearer.

9. Methods: in the part concerning the author survey, question 1: for authors who only indicated that study was blinded,… I don’t understand why there is “categorical answers” at the end of the sentence. There is no such mention for other questions. I also don’t understand the sentence “We emphasized to the survey participants that only aggregate data will be published”

Author response: We have now clarified that the 2nd question also used the same categorical answers as the 1st question. The categories used were chosen based on the manuscript by Haahr 2006. The verb “aggregate” was corrected into “anonymized”.

10. In the results section, paragraph ‘Survey results’, it is unclear to mention the third question line 35, page 14

Author response: We revised this to read “second” question.

11. I am not sure that “inadequate” is the good term when the authors say that the reporting of blinding is inadequate in conclusion and strengths and limitations. I would rather say that the reporting was poor or insufficient.

Author response: As suggested, we revised the wording, from inadequate into insufficient.

12. Additional questions or comments
Which filter was used to identify RCTs? This is not indicated.

Author response: We used the publication type filter; this is now clarified in the Methods.

13. Results section, paragraph “Blinding of individuals involved in trials”: the results reported there refer to the analysis of who was blinded for the 531 trials where they found at least some descriptors for blinding but later in the paragraph, the authors give a percentage calculated from the whole sample of 622 trials (the 14%). It should be 89/531=17%, no?

Author response: Yes, in this paragraph we aimed to provide a percentage for the whole sample, but we understand that it may be clearer to report the percentage as 89/531=17%. This is now revised.

14. Results section, the title of the paragraph “other analyses related to blinding” could be replaced by a more informative title and split in two parts: “discrepancies between abstract and body of manuscript” and “Reporting of blinding in registries”, or something like that. The sentence concerning the CONSORT Statement should be better placed elsewhere for example in a paragraph about general characteristics of trials that is missing or in the paragraph ‘”Blinding of individuals involved in trials” as reporting who is blinded is a clear recommendation of the CONSORT Statement.

Author response: Thank you for the useful comments. All these changes were made as suggested.

15. It would have been interesting to have a description of the included trials with the type of intervention (ie, pharmacological or not pharmacological), the year of publication, the sample size, the funding source

Author response: The study already has a lot of data, and a lot of text had to be reduced, so we could not further expand the study aim.

16. Is there any improvement over time? Is it possible to have a look at this?

Author response: As suggested, we analyzed for the primary outcome whether there was any change in the analyzed years, and we did find that the percent of trials that reported explicitly whether the three key groups of individuals were blinded improved over the analyzed time. This is now reported in the manuscript.

17. Discussion section, for me the lack of comparison between journals is not a limit.

Author response: This part of the Discussion was removed in the revised manuscript.

Reviewer: 3

MAJOR:
1. There are many other assessments of reporting and blinding in published RCTs. The authors should provide a synopsis in the introduction of the multiple similar studies to their own. The one by

Devereaux et al is quite good, however there is a massive literature in this area that the authors must summarize for a naïve reader. Some of this is nicely summarized in the Discussion, however a very brief (e.g. short paragraph), high level overview of this literature should be provided in the introduction in order to provide context to the reader.

Author response: We have now added a new paragraph to the Introduction, as suggested. This is a shortened version of the text that was previously in the Discussion section.

2. Was a protocol ever registered (e.g. Open Science Framework) or at least deposited in a time-stamped manner (e.g. institutional repository)? This is a basic component of best practices surrounding meta-research.

Author response: We did not register our protocol *a priori*. We started this study several years ago when it was not as common as today to publish a protocol online. While registering a protocol a priori is definitely a good practice, as of today there is still no mandatory requirement for prospective registration of all types of studies, such as ICMJE requirement for clinical trials.

3. It is unclear why the authors used the search strategy they did. The preset filter in PubMed is not sensitive. Several groups have developed RCT filters for PubMed, e.g. Cochrane (https://work.cochrane.org/pubmed). Using this filter picks up ~10X many more hits than simply using the RCT "article type" filter in PubMed. I just compared the Cochrane filter in PubMed - ~4 million hits, versus simply selecting 'RCT' in the preset article type filter that only retrieved ~480 000 hits. In my mind, this is highly problematic, and raises critical questions with regards to the corpus of studies that were investigated.

Author response: Our study was not a systematic review. For this reason, we used simpler search strategy, taking into account three components – journal name, publication type and chosen dates. We have now added in the Limitations that it is possible that we could have found different number of studies using a different search strategy.

4. The definition around the primary outcome is unclear – how was 'clearly described' actually operationalized? What does 'clearly described' mean? In other words, what was the threshold to be 'clearly described'? This is critical, since the CONSORT checklist describes three parties participants, care providers, those assessing outcomes), whereas the longer form explanation describes four parties (participants, healthcare providers, data collectors, and outcome adjudicators). These details are critical and must be fully elaborated in the methods.

Author response: We have now clarified that clearly described means that the authors of analyzed studies have explicitly reported who (which group of individuals) was blinded.

5. It is unclear to me how trials were assessed where it would be redundant to describe aspects of blinding. As a specific example, I would offer trials where ultrasound versus landmark technique for central line insertion were compared. Here it is obvious the operators could never be blinded; given the limited word counts in the journals, would the authors expect the study to report that the authors were not blinded? Although I am firm proponent of complete reporting, this particular instance would seem unnecessary to me.

Author response: We did not distinguish or excluded certain types of trials (such as the example mentioned by the reviewer) because blinding is a reporting item that is included in both the CONSORT guidelines and in CONSORT extension for Abstracts. If it was not possible, or needed, to

blind the participants, the trialists should explicitly report this in their manuscripts, in line with relevant reporting checklist.

6. What parts of the manuscript texts were assessed? Methods alone? The whole manuscript? What about supplemental material? (I see there is a brief mention of this in the discussion, but this needs to explained in detail with the Methods section.

Author response: We have now clarified in the Methods section that we assessed the full manuscript and supplementary material.

7. Details of data extraction/verification are missing in the methods. How many people extracted data? Based on the CRediT taxonomy it appears five authors did? Was the data extracted in duplicate (i.e. like the study screening)? If so how were conflicts resolved? Was there any form of data extraction verification? These are critical methodological details that are absent from the manuscript.

Author response: We have now clarified that data extraction for each study was done by two authors independently. Three authors (AP, KB, MK) participated in data extraction. We then compared extractions for any inconsistencies and resolved them via consensus. We also clarified who sent all the surveys, and who categorized survey data.

8. Basic details around the development of the survey are missing. How were the questions developed? Was an expert in survey methodology engaged? Who piloted the survey? Were focus groups used to refine the questions? Was there a pretest with a smaller sample? Again, these are basic methodological details that I would anticipate be reported.

Author response: We used the survey as reported by Haahr in 2006. We provided the reference before in the description of the survey, but we have now explicitly highlighted the provenance of the survey. This was now clarified in the revised manuscript.

9. A copy of the actual survey questions, front matter etc should be provided as a supplemental file.

Author response: A copy of the survey sent to the participants via e-mail is now included as a Supplementary file 1.

10. Within the questions, I was very unclear to me why #3 (i.e. estimating the prevalence of ‘double-blind’ trials in the literature) was asked. Moreover, I’m not sure I would initially interpret the question in the way the authors have presented the results of that question.

Author response: We used this third question, as it was used in the Haahr 2006 article. There is always, of course, a possibility that survey respondents do not interpret the questions as we would like to. We did not have a single case where respondents asked for clarification of any question.

11. Within the results I would hope to see a flow diagram of the study selection process. How many came from each journal and each year?

Author response: As suggested, this information is now presented in a flow diagram, as Figure 1.

12. Although the seven types of individuals that could be blinded (listed in Table 4) intuitively make sense to me, I'm unclear where these categories came from? As mentioned above, CONSORT mentions three individuals in the checklist, four in the longer form online (http://www.consort-statement.org/checklists/view/32--consort-2010/93-blinding).

Author response: We used six categories of individuals as defined in the Haahr 2006 study. We highlighted this now in the revised manuscript.

13. A full listing of studies should be included in an online supplement.

Author response: We have now provided a full listing of included studies in the online supplement, as Supplementary file 2.

14. The survey response rate is extremely low. This response rate needs to be highlighted in the abstract. Moreover, in the discussion this needs to be discussed further. The results of the survey are at very high risk of non-responder bias.

Author response: As suggested, we highlighted the response rate in the abstract; we had to shorten a bit the other parts of the abstract to accommodate the limit of 300 words. We also highlighted the risk of non-responder bias in the Discussion section.

15. I'm not sure I agree with the use of the term 'retrospectively registered'. In my mind this would suggest the trial was complete and then registered. Likely the more precise term is 'after the trial was open for recruitment'. I say this because clinicaltrials.gov (where the majority of your data was derived) allows for registration up to 21 days after the first patient is enrolled. This means a trial could open months before their first patient is enrolled, and then registered another 21 days after the first recruitment. Thus, I believe the term 'retrospectively' would be inaccurate to use.

Author response: As suggested, the term 'retrospectively registered' was revised into 'registered after the trial was open for recruitment'.

16. Table 8 should be reformatted – for the questions regarding double and triple blind it would be most interesting within the table to clearly show the distribution of the dyads/triads that were listed by respondents.

Author response: The Table 8 would be very complex if we try to address the diversity of responses. For this reason, we left the Table 8 as it was, but we added Supplementary file 3, which shows all responses we received (de-identified), for the full transparency of responses.

Minor:
1. I found it stunning that only 62% of RCTs recently published were registered – especially coming from these particular journals that are generally well regarded. This seems like a significant result that should be highlighted a bit more in the discussion.

Author response: As suggested, this result is now highlighted in the Discussion section of the revised manuscript, with several examples from literature indicating that non-reporting of registration

and non-adherence with ICMJE policy on prospective trial registration has also been reported previously.

2. The authors should briefly justify the use of JIF to select articles. I have used the same strategy for some of my group's reporting assessments as well, thus I am not opposed to this approach. There are inherent problems though, which the authors could very briefly address.

Author response: We have addressed the issue of JIF in the Discussion section of the revised manuscript, as suggested.

3. In a similar vein, the authors should very briefly justify the focus on anesthesia related literature.

Author response: Pain is a subjective outcome; there is no objective method for analyzing pain, and pain – by definition – is what the patient says it is. The analgesic placebo effect has been well documented. In the area of placebo effects, expectancies play a major role, triggering a cascade of endogenous opioids and non-opioids, and altering the experience of pain. For this reason, the whole field of anesthesiology and pain is likely biased if blinding is not used adequately. This is now highlighted in the manuscript.

4. Introduction reference #6 is used to refer to the CONSORT statement; the CONSORT 2010 statement should likely be referenced.

Author response: As suggested, we revised the reference, and now we have referenced the CONSORT 2010 statement.

VERSION 2 – REVIEW

REVIEWER	Jon Henrik Laake Oslo University Hospital, Oslo, Norway
REVIEW RETURNED	03-Jan-2020

GENERAL COMMENTS	No further issues.

REVIEWER	Agnes Dechartres Sorbonne Universite
REVIEW RETURNED	03-Jan-2020

GENERAL COMMENTS	Review Manuscript BMJ Open-2019-035168-R1 I would like to thank the authors for answering my comments. Some remaining minor comments: Although the manuscript is overall well-written, I think it would be good to have a check for English. There were also some typos for example in the conclusion, page 22, line 31 "which is not sufficiently informative for readers". The abstract format is not very adapted to such type of studies in particular the "setting" subsection.

	I was surprised that discrepancies within protocol registrations and between protocol registrations and publications were not listed as secondary outcomes. I was confused with some numbers because the results were sometimes difficult to follow. For example, in the first page of the results, it is indicated that 357 trials were explicitly described as blinded, 29 were explicitly described as not blinded and in 236 the information about blinding was not clearly reported. However, the analysis of who was blinded was done for 531 trials where they found at least some descriptors of blinding. Why not 357? This is a bit confusing. Later, page 14, it is indicated “among 126 registered trials that did not have discrepancies in the protocol registration,…” where does this number of 126 come from? Page 13, line 8: I would rather say : This percentage seemed to improve over analyzed time The discussion is still very long in particular the Discussion of the study by Haar and Hjobjartsson and the limitations. The warning of Montori about the avoidance of the term double blind is mentioned twice in the Discussion. Perhaps discussion about that once is enough.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 2
Reviewer Name: Agnes Dechartres
Institution and Country: Sorbonne Universite

1. Although the manuscript is overall well-written, I think it would be good to have a check for English. There were also some typos for example in the conclusion, page 22, line 31 “which is not sufficiently informative for readers”.

Author response: We have now asked native English speaker, Prof. Shelly Pranic, PhD, to conduct language editing. We have acknowledged her assistance in the Acknowledgements section.

2. The abstract format is not very adapted to such type of studies in particular the “setting” subsection.

Author response: We agree that the recommended BMJ Open abstract format, with pre-specified subtitles in the abstract, may not be perfectly suitable for our study format. Based on this recommendation, we have now slightly revised the abstract format, and we hope that the Editors will endorse this revision.

3. I was surprised that discrepancies within protocol registrations and between protocol registrations and publications were not listed as secondary outcomes.

Author response: The reviewer is right; we should have listed these as secondary outcomes. We have revised this now, both in the abstract and in the text of the methods in the full.

4. I was confused with some numbers because the results were sometimes difficult to follow. For example, in the first page of the results, it is indicated that 357 trials were explicitly described as blinded, 29 were explicitly described as not blinded and in 236 the information about blinding was not clearly reported. However, the analysis of who was blinded was done for 531 trials where they found at least some descriptors of blinding. Why not 357? This is a bit confusing.

Author response: In this first analysis, we were looking for explicit descriptions of study design, i.e. whether blinding was mentioned in the description of the study design. In the first paragraph we

described whether study design was explicitly described as study that was open-label, blinded, double-blinded, etc. We have tried to clarify this in the revised manuscript.

5. Later, page 14, it is indicated “among 126 registered trials that did not have discrepancies in the protocol registration,…” where does this number of 126 come from?

Author response: The number 126 comes from this calculation:

We analysed 174 protocols, of which 157 were registered in ClinicalTrials.gov and 17 in other registries. Among 157 protocols in ClinicalTrials.gov, 92 had information about blinding only in the field Masking (so those **92** protocols had no discrepancies within the protocol), while in the remaining 65 protocols with information about blinding in multiple fields, there were 47 protocols with discrepancies found, and **18** protocols without discrepancies. Among 17 protocols registered in databases other than ClinicalTrials.gov, only one had discrepancies, while the remaining 16 did not.

So, when we summarize 92+18 protocols from ClinicalTrials.gov without discrepancies, with 16 from other registries, we get 126 protocols without discrepancies. We clarified this now in the manuscript with an additional short text inserted in several places.

6. Page 13, line 8: I would rather say : This percentage seemed to improve over analyzed time

Author response: This text was revised, as suggested.

7. The discussion is still very long in particular the Discussion of the study by Haar and Hjobjartsson and the limitations. The warning of Montori about the avoidance of the term double blind is mentioned twice in the Discussion. Perhaps discussion about that once is enough.

Author response: We revised the Discussion as suggested. We reduced the part of the discussion about the Haar and Hjobjartsson study, and the reference to the Montori study. We also shortened the part on study limitations, but preserved the mention of all limitations identified in the study and suggested in the previous rounds of review.