

PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Has the reporting quality of published randomised controlled trial protocols improved since the SPIRIT statement? A methodological study
AUTHORS	Tan, Zet; Tan, Aidan; Li, Tom; Harris, Ian; Naylor, Justine; Siebelt, Michiel; van Tiel, Jasper; Pinheiro, Marina; Harris, Laura; Chamberlain, Kira; Adie, Sam

VERSION 1 - REVIEW

REVIEWER	Livia Puljak Center for Evidence-Based Medicine and Health Care, Catholic University of Croatia, Zagreb, Croatia
REVIEW RETURNED	15-Mar-2020

GENERAL COMMENTS	<p>Dear Editor,</p> <p>I have reviewed the manuscript bmjopen-2020-038283. I find it interesting and relevant. It is yet another proof that merely having reporting checklists does not guarantee better (or much better) reporting. Despite significant findings of 8.8% improvement, this is not something that we should be happy about. I would suggest the following revisions that could further improve this fine manuscript:</p> <p>Study design labelling The authors have labelled this as a meta-epidemiological study, both in the title and throughout the manuscript. However, there is much ambiguity about what actually means "meta-epidemiological". The research community has recently taken up the use of this descriptor as a synonym for methodological research, but we have recently conducted a study (unpublished data; manuscript accepted) showing that there is a lot of heterogeneity in how we use this term. With the lack of consensus in what this term actually means, it would be better to use (throughout the manuscript) a more neutral expression such as "methodological study/research".</p> <p>Abstract -It would be good to write in the abstract when the SPIRIT was published; for readers not aware of this information. -"journal policy of compliance" – I would change this into "publicly reported journal policy of compliance" – the authors have reported in the methods that they have only analyzed information that was publicly available in 2019</p> <p>Reporting The authors declared that they have reported their study in line with</p>
-------------------------	--

PRISMA statement. However, PRISMA is for systematic reviews, and this was not a systematic review. I would suggest to delete this sentence.

Included protocols

-The authors reported that they excluded protocols registered on a clinical trial registry but not published in a peer-reviewed journal. This is confusing, because the authors searched MEDLINE, EMBASE and CENTRAL. I would not expect that any such cases would be retrieved via searching of those three databases.
-The authors write “We retrieved two equal, arbitrary samples of 150 RCT protocols...”. But, while reading the manuscript it is obvious that the authors did not retrieve 300 RCT protocols. Instead, the authors retrieved much larger number of records, and then used screening to get to their pre-defined sample of 150 RCT protocols in each study period.

Variables

-The authors used affiliation with epidemiology/statistics department as a proxy of qualification in epidemiology or statistics. I am not sure this is appropriate; I personally know many people employed in such departments, but who do not have either of those formal qualifications. Therefore, I am not sure that this proxy measure was adequate. At least the authors could do another analysis where they would repeat this analysis, but without using this proxy, i.e. by including in the analysis as “having qualification” only those authors that have actually reported such qualification in the manuscript.
-The authors analyzed “journal policy of compliance”. However, those policies can vary in the way they are formulated, ranging from encouragement (the authors “should” use the checklist) to requirement (the authors have to/must use the checklist). It would be interesting to see whether difference in the wording/demanding adherence to the checklist makes any difference to the reporting results.

Results, Included protocols

-In the first section of results, the authors frequently use the term “studies”; I would suggest using the term “articles” instead. We have no idea whether all those excluded articles actually described “studies”.
-In the results, I would appreciate seeing information related to source of funding in Table 1, and also in the analysis described in the Table 2.
-The finding that none of the analyzed protocols was fully adherent to all checklist items is very interesting, and if the word count allows, it would be good to report this in the Abstract.

Limitations

-The authors should mention as a limitation the fact that the journal web sites were analyzed for endorsement of SPIRIT checklist in 2019. Since these manuscripts were published over a long period of time, we cannot know whether those journal web sites endorsed SPIRIT checklist at the time when the manuscript was being prepared and submitted. We cannot be sure about this unless we have information available about the content of journal instructions for authors over a long period of time.

Interpretation

-The authors adequately discussed that a study of industry-initiated RCTs reported that the majority of RCTs protocols used ghost

	<p>authors. Honestly, I would fully expect this from industry, but not from academia, i.e. studies funded via non-profit sources. Personally, I would never have funds to engage professional writers, and I do not know of any grant funding agency that would approve funds for professional writers (which does not mean that they do not exist, of course). It would be interesting to discuss whether there were any differences in reporting depending on the affiliation and source of funding of the RCT protocols analyzed in this study.</p> <p>-The findings that merely self-reporting compliance with SPIRIT checklist cannot be relied on is very important. I could be mentioned that the authors probably consider they are expected (or they are asked explicitly by reviewers/editors) to mention the checklist, without actually taking time to report everything in line with the checklist. So, they are superficially fulfilling an obligation/requirement, without actually taking time to do what the checklist requires.</p> <p>-Regarding the word length, there could be a difference between print and electronic journals. Electronic journals do not have to worry about printing too many pages, so theoretically there could be a difference in wording allowance with electronic journals. It would be interesting to see how many print vs electronic journals were in the sample.</p> <p>-Technical check of manuscripts in journals should include checklist adherence check. Instead, what they do at technical check, is harassing authors if they did not submit the manuscript with proper referencing style (which can always be fixed later if the journal intends to publish the manuscript) or did not upload appendices in pdfs. Journals should instead focus during initial technical check on what really matters, and that would include mandatory checklist compliance check.</p> <p>-One solution that I would expect to improve reporting radically is to change the format of the manuscript submission for RCT protocols in a way to require authors that instead of the "classic" introduction-methods-discussion format the authors must use subtitles corresponding to each SPIRIT item. With this change of format, it should be easier to check whether the authors have actually use all the mandatory checklist items.</p> <p>-Use of software such as PENELOPE should help in checking reporting adherence at submission.</p> <p>Sincere regards, Livia Puljak</p>
--	--

REVIEWER	soumya tiwari Lady Hardinge Medical College, New delhi, India
REVIEW RETURNED	21-Mar-2020

GENERAL COMMENTS	<p>The authors have compared the reporting quality of RCT protocols before and after publication of SPIRIT guideline. Following are the comments:</p> <ol style="list-style-type: none"> 1) The Kappa values are negative and very low (<0.2) at multiple places indicating flaws in data collection and/or analysis. 2) The low kappa values and a simultaneous fair percentage agreement indicates a possibility of guessing/assumptions in marking by raters. It is overestimating the true agreement between raters. 3) Reference 11 has not been cited correctly (journal name not
-------------------------	---

	<p>mentioned)</p> <p>4) Reference 14 has been cited wrongly.</p> <p>5) The authors have accepted that the improvement in reporting quality may just be due to secular trend and the associations may not be causal. To overcome this limitation the authors may have used interrupted time series design.</p> <p>6) The conclusion is contradictory, "Although there is a significant improvement in reporting quality, a substantial proportion of items remain poorly reported". The authors should not have highlighted a statistically significant improvement as significant impact of SPIRIT guideline.</p> <p>7) PRISMA is not the standard reporting guideline for this type of meta-epidemiological studies.</p> <p>8) In exploratory regression analysis (table 2) the contribution by various predictors studied was only about 20%. Authors have missed the important factors which could explain 80% change in reporting quality. Also, the details about goodness of fit of regression model is not reported.</p> <p>9) Figure 1 does not provide any information about significance of difference between pre and post SPIRIT era. This is contrary to the legend of figure 1.</p>
--	---

REVIEWER	Odile Sauzet Bielefeld University, Germany
REVIEW RETURNED	25-Mar-2020

GENERAL COMMENTS	<p>This review assess the adherence to SPIRIT Statement in the year 2019 compared to published Protocols just before the publication of the statement.</p> <p>The main findings are that the journal guidelines are a predictor of adherence as well as the presence among the authors of someone working in a Public Health/Epidemiology/Biostatistics department. In short, better bigger studies have a better adherence or if the journal request it.</p> <p>There are some methodological point that need addressing:</p> <p>Data collection: Only 100 studies were double reviewed and the authors point out that the inter-rater agreement could be quite poor for some items, indicating a rather subjective assessment. Systematic reviews are usually based on consensus but there are no indication that disagreement were resolved and 200 studies were only red by one person. While this is not a systematic review as such, the data collection should be of the same quality.</p> <p>Selection Bias: It is not mentioned in the manuscript which proportion of all protocols present in protocol databases are published. Given that some databases require specific items to be fill in I would imagine that the quality of reporting in databases is quite high. A discussion about how relevant are published protocols compared to data bases would be welcome.</p>
-------------------------	---

VERSION 1 – AUTHOR RESPONSE

Reviewer 1	
Comments	Responses
<p>I have reviewed the manuscript bmjopen-2020-038283. I find it interesting and relevant. It is yet another proof that merely having reporting checklists does not guarantee better (or much better) reporting. Despite significant findings of 8.8% improvement, this is not something that we should be happy about. I would suggest the following revisions that could further improve this fine manuscript:</p>	<p>Thank you for your review.</p>
<p>Study design labelling The authors have labelled this as a meta-epidemiological study, both in the title and throughout the manuscript. However, there is much ambiguity about what actually means “meta-epidemiological”. The research community has recently taken up the use of this descriptor as a synonym for methodological research, but we have recently conducted a study (unpublished data; manuscript accepted) showing that there is a lot of heterogeneity in how we use this term. With the lack of consensus in what this term actually means, it would be better to use (throughout the manuscript) a more neutral expression such as “methodological study/research”.</p>	<p>Thank you for this advice. We have revised the manuscript to replace all references to “meta-epidemiological study” with “methodological study”.</p>
<p>Abstract - It would be good to write in the abstract when the SPIRIT was published; for readers not aware of this information.</p>	<p>Thank you for this suggestion. We have revised the manuscript to include the year of publication of the SPIRIT statement (2013).</p>
<p>- “journal policy of compliance” – I would change this into “publicly reported journal policy of compliance” – the authors have reported in the methods that they have only analyzed information that was publicly available in 2019</p>	<p>Thank you for this suggestion. We have revised the manuscript to replace all references to “journal policy of compliance” with “publicly reported journal policy of compliance”.</p>
<p>Reporting The authors declared that they have reported their study in line with PRISMA statement. However, PRISMA is for systematic reviews, and this was not a systematic review. I would suggest to delete this sentence.</p>	<p>Thank you for this suggestion. We have revised the manuscript to remove this sentence.</p>
<p>Included protocols - The authors reported that they excluded protocols registered on a clinical trial registry but not published in a peer-reviewed journal. This is confusing, because the authors searched MEDLINE, EMBASE and CENTRAL. I would not expect that any such cases would be retrieved via searching of those three databases.</p>	<p>Thank you for this comment. We initially included this information as both CENTRAL and MEDLINE “In Process” databases index articles that may not be published, and only contained in online registries. We agree however that this detail may be confusing and have revised the manuscript to remove this exclusion criteria.</p>
<p>- The authors write “We retrieved two equal, arbitrary samples of 150 RCT protocols...”. But, while reading the manuscript it is obvious that the authors did not retrieve 300 RCT protocols. Instead, the authors retrieved much larger</p>	<p>Thank you for this comment. We have revised the manuscript to clarify the screening process.</p>

number of records, and then used screening to get to their pre-defined sample of 150 RCT protocols in each study period.

Variables

- The authors used affiliation with epidemiology/statistics department as a proxy of qualification in epidemiology or statistics. I am not sure this is appropriate; I personally know many people employed in such departments, but who do not have either of those formal qualifications. Therefore, I am not sure that this proxy measure was adequate. At least the authors could do another analysis where they would repeat this analysis, but without using this proxy, i.e. by including in the analysis as “having qualification” only those authors that have actually reported such qualification in the manuscript.

Thank you for this comment, which allowed us to realise that our use of the term “proxy” was in error. We have removed this from the manuscript. We did not measure affiliation with an epidemiology/statistics department as a “proxy” for those qualifications. This variable was satisfied if an author either had an epidemiology/statistics degree OR were affiliated with an epidemiology/statistics department. It was our experience from our previous studies (Adie S, Harris IA, Naylor JM, Mittal R. CONSORT compliance in surgical randomized trials: are we there yet? A systematic review. *Ann Surg.* 2013;258(6):872-8. and Adie S, Ma D, Harris IA, Naylor JM, Craig JC. Quality of conduct and reporting of meta-analyses of surgical interventions. *Ann Surg.* 2015;261(4):685-94) that formal qualifications were seldom listed in publications, usually due to journal restrictions

We agree that affiliation to a clinical epidemiology, public health or biostatistics department is an imperfect proxy for authors' qualifications, but this is irrelevant as both have been captured by our definition above. We would argue that even if authors do not have these formal qualifications, affiliation with an epidemiology/statistics department provides access to knowledge and expertise that would otherwise not be available in other departments. However, to explore the reviewer's suggestion further, we did review all our included studies – and we found that none of the included publications included information on author qualifications. It was also impractical to individually search

	<p>each author online, as it would be impossible to determine the timeline of when their qualifications were achieved in relation to their publication.</p>
<p>- The authors analyzed “journal policy of compliance”. However, those policies can vary in the way they are formulated, ranging from encouragement (the authors “should” use the checklist) to requirement (the authors have to/must use the checklist). It would be interesting to see whether difference in the wording/demanding adherence to the checklist makes any difference to the reporting results.</p>	<p>Thank you for this comment. We agree that there is variation in the formulation of publicly reported journal policies of compliance with the SPIRIT statement. This variation in formulation is spread across a continuous spectrum from suggestions, to recommendations, to requirements. This variation in formulation is further confounded by a variation in enforcement. For example, some journals do not enforce all requirements and other journals expect all suggestions to be strictly followed. We attempted to review the journals which published our included articles, and found it impossible to differentiate the language of “suggestion” from the language of “enforcement”. We conclude that the only method to verify this would be to actually submit articles for publication to each included journal. Rather than risk the subjectivity and irreproducibility of arbitrarily categorising such a complex and continuous spectrum, we decided to use an objective and reproducible dichotomous measure which was independent of investigator interpretation. We feel that this strengthens the association which we have found, and allows it to be verifiable.</p>
<p>Results, Included protocols - In the first section of results, the authors frequently use the term “studies”; I would suggest using the term “articles” instead. We have no idea whether all those excluded articles actually described “studies”.</p>	<p>Thank you for this suggestion. We have revised the manuscript to replace the references to “studies” or “study” in the included protocols section with “articles” or “article”.</p>
<p>- In the results, I would appreciate seeing information related to source of funding in Table 1, and also in the analysis described in the Table 2.</p>	<p>Thank you for this suggestion. It is important to note that for the purposes of compliance with SPIRIT, we only measured whether the funding source was reported, not the actual source of funding. However, to address this</p>

	<p>important suggestion, we reviewed the included protocols and collected data on the source of funding – categorised as either industry (for-profit) or non-industry funding. We have revised the manuscript to include the funding source in trial characteristics in Table 1.</p> <p>8/131 (6%) of RCT protocols received industry funding before the SPIRIT statement was published, while 10/145 (7%) received industry funding after the SPIRIT statement was published. Industry funding was not associated with compliance with the SPIRIT statement, with only a 0.3% difference in mean SPIRIT scores (95% CI -4.9% - 5.6%, p=0.9) between industry and non-industry funded trials. As such, industry funding was not included in the regression analysis as our pre-planned regression modelling limited the inclusion of variables to only those with potential statistical influence.</p>
<p>- The finding that none of the analyzed protocols was fully adherent to all checklist items is very interested, and if the word count allows, it would be good to report this in the Abstract.</p>	<p>Thank you for this suggestion. We have revised the manuscript to report this finding in the abstract.</p>
<p>Limitations</p> <p>- The authors should mention as a limitation the fact that the journal web sites were analyzed for endorsement of SPIRIT checklist in 2019. Since these manuscripts were published over a long period of time, we cannot know whether those journal web sites endorsed SPIRIT checklist at the time when the manuscript was being prepared and submitted. We cannot be sure about this unless we have information available about the content of journal instructions for authors over a long period of time.</p>	<p>Thank you for this comment. The publicly reported journal policy of compliance with the SPIRIT statement was identified from the instructions to authors on the journal website on 29 March 2019. The 150 RCT protocols from after the SPIRIT statement were published between 25 January 2019 and 20 March 2019. Whilst it is possible that the journals changed their policy of compliance with the SPIRIT statement in this narrow time frame, we feel that it is unlikely.</p>
<p>Interpretation</p> <p>- The authors adequately discussed that a study of industry-initiated RCTs reported that the majority of RCTs protocols used ghost authors. Honestly, I would fully expect this from industry, but not from academia, i.e. studies funded via non-</p>	<p>We agree with this perspective. Please see our response to the comment above. We have revised the manuscript to include the funding source in trial</p>

profit sources. Personally, I would never have funds to engage professional writers, and I do not know of any grant funding agency that would approve funds for professional writers (which does not mean that they do not exist, of course). It would be interesting to discuss whether there were any differences in reporting depending on the affiliation and source of funding of the RCT protocols analyzed in this study.

characteristics in Table 1. We found no association between industry funding, and publication before or after the SPIRIT statement, or with compliance with the SPIRIT statement. It is important to note that the SPIRIT statement is a reporting tool, and was not designed as a measure of overall protocol "quality".

- The findings that merely self-reporting compliance with SPIRIT checklist cannot be relied on is very important. I could be mentioned that the authors probably consider they are expected (or they are asked explicitly by reviewers/editors) to mention the checklist, without actually taking time to report everything in line with the checklist. So, they are superficially fulfilling an obligation/requirement, without actually taking time to do what the checklist requires.

We agree with this opinion. We have revised the manuscript to state that a possible explanation for this finding is that some authors who are aware of either the SPIRIT statement or a journal's policy of compliance with the SPIRIT statement may decide to report compliance with the SPIRIT statement without applying the checklist. In short, awareness of the SPIRIT statement does not translate into application of the checklist.

- Regarding the word length, there could be a difference between print and electronic journals. Electronic journals do not have to worry about printing too many pages, so theoretically there could be a difference in wording allowance with electronic journals. It would be interesting to see how many print vs electronic journals were in the sample.

Thank you for this suggestion. We have reviewed our included protocols and collected data on which RCT protocols were published in journals which published in either print or exclusively online. Only 17/300 (6%) of RCT protocols were published in journals which published in print, while the remainder (94%) were published in journals which published exclusively online. We found that the mean word count of RCT protocols published in online journals and print journals was 4387 words and 3581 words, respectively, with a 806 word difference in mean word count (95% CI 26 -1586 words, $p=0.04$). Publication type (either print or exclusively online) was not associated with compliance with the SPIRIT statement ($p=0.29$). As such, publication type was not included in the regression analysis as our pre-planned regression modelling limited the inclusion of

	variables to only those with potential statistical influence. This important information has been added to the manuscript. We would argue that the more important measure (which was included in the regression modelling) is word count.
- Technical check of manuscripts in journals should include checklist adherence check. Instead, what they do at technical check, is harassing authors if they did not submit the manuscript with proper referencing style (which can always be fixed later if the journal intends to publish the manuscript) or did not upload appendices in pdfs. Journals should instead focus during initial technical check on what really matters, and that would include mandatory checklist compliance check.	We agree with this opinion. We have revised the manuscript to state that the association between publicly reported journal policy of compliance with the SPIRIT statement and higher reporting quality supports the role of journals and editors in checking adherence to the SPIRIT statement and this may be aided by mandated author completed pre-submission checklists.
- One solution that I would expect to improve reporting radically is to change the format of the manuscript submission for RCT protocols in a way to require authors that instead of the "classic" introduction-methods-discussion format the authors must use subtitles corresponding to each SPIRIT item. With this change of format, it should be easier to check whether the authors have actually use all the mandatory checklist items.	Thank you for this suggestion We have revised the manuscript to state that checking adherence to the SPIRIT statement may be aided by structured online manuscript submission systems.
- Use of software such as PENELOPE should help in checking reporting adherence at submission.	Thank you for this suggestion We have revised the manuscript to state that checking adherence to the SPIRIT statement may be aided by automated manuscript reporting quality checks.
Reviewer 2	
Comments	Responses
The authors have compared the reporting quality of RCT protocols before and after publication of SPIRIT guideline. Following are the comments: 1) The Kappa values are negative and very low (<0.2) at multiple places indicating flaws in data collection and/or analysis. 2) The low kappa values and a simultaneous fair percentage agreement indicates a possibility of guessing/assumptions in marking by raters. It is overestimating the true agreement between raters.	Thank you for this comment. We agree that some of the kappa values are negative and low, however the majority (55%) are ≥ 0.4 , indicating at least moderate agreement. The final datapoints used for the analysis were the results of the duplicate data collection and discussion of disagreements. We calculated the kappa values based on the initial datapoints extracted by the duplicate data collection before discussion of disagreements. We had a pilot phase to independently appraise the data extraction form

	<p>and assessment criteria, several round table discussions to harmonise the definitions of adequate and inadequate/unclear for each checklist item, and regular fortnightly meetings to discuss discrepancies. The low kappa values reflect the difficulty interpreting these items and highlights problems with adaptability and interpretation of the SPIRIT items for authors and journals. Most previously published similar studies do not include any formal assessment of agreement, and we feel that this is indeed a strength, rather than a weakness, of our study.</p>
<p>3) Reference 11 has not been cited correctly (journal name not mentioned)</p>	<p>Thank you for this comment. We have revised Reference 11 to include the journal name (PLoS Med).</p>
<p>4) Reference 14 has been cited wrongly.</p>	<p>Thank you for this comment. Unfortunately, we are unsure as to how Reference 14 has been cited wrongly and would appreciate any clarification.</p>
<p>5) The authors have accepted that the improvement in reporting quality may just be due to secular trend and the associations may not be causal. To overcome this limitation the authors may have used interrupted time series design.</p>	<p>Thank you for this comment. We agree that an interrupted time series design would contribute to overcoming this limitation and will consider it for future studies. However, there would also be issues with this design - as the exact time of publication is often unclear, with many articles being made available online prior to "print"</p>
<p>6) The conclusion is contradictory, "Although there is a significant improvement in reporting quality, a substantial proportion of items remain poorly reported". The authors should not have highlighted a statistically significant improvement as significant impact of SPIRIT guideline.</p>	<p>We respectfully disagree with the reviewer. Our statement is accurate in that there was an observed increase in the overall quality as assessed by a significant increase in the overall proportion of adequately reported SPIRIT items over the time period. However, we also noted that many individual SPIRIT items remain poorly reported. In order to clarify this, we have edited the conclusion statement to: "There has been significant improvement in the</p>

	<p>overall reporting quality of RCT protocols since the SPIRIT statement, although a substantial proportion of individual checklist items remain poorly reported.” Further, we have shown that a journal policy of compliance with SPIRIT is associated with SPIRIT compliance in our modelling, which supports the conclusion that the SPIRIT statement has improved the reporting of protocols. We have also edited most references to reporting quality and checklist items to clarify the delineation between the overall findings and individual findings.</p>
<p>7) PRISMA is not the standard reporting guideline for this type of meta-epidemiological studies.</p>	<p>Thank you for this suggestion. We have revised the manuscript to remove the reference to PRISMA, and have removed the term “meta-epidemiological study” as there is ambiguity related to this, as per Reviewer 1’s comments above.</p>
<p>8) In exploratory regression analysis (table 2) the contribution by various predictors studied was only about 20%. Authors have missed the important factors which could explain 80% change in reporting quality. Also, the details about goodness of fit of regression model is not reported.</p>	<p>Thank you for this comment. We are unsure where the reviewer concluded that the “contribution by various predictors was only about 20%”. Table 2 clearly describes the data as the increase in proportion of SPIRIT items with each covariate in our modelling. However, we have realised that we did not report our goodness of fit data according to our data analysis plan and have now added this to the manuscript. The final model had an adjusted R^2 value of 0.37, indicating 37% of the variability in SPIRIT score was explained in our model. We would argue this is relatively high given our study design and the huge amount of potential factors that could influence a trial protocol.</p>
<p>9) Figure 1 does not provide any information about significance of difference between pre and post SPIRIT era. This is contrary to the legend of figure 1.</p>	<p>Figure 1 visualises the checklist items with a significant increase in adequate reporting after the SPIRIT statement. The checklist items included are only those which demonstrated a significant increase in adequate reporting</p>

	from the pre-SPIRIT era to the post-SPIRIT era.
Reviewer 3	
Comments	Responses
<p>This review assess the adherence to SPIRIT Statement in the year 2019 compared to published Protocols just before the publication of the statement.</p> <p>The main findings are that the journal guidelines are a predictor of adherence as well as the presence among the authors of someone working in a Public Health/Epidemiology/Biostatistics department. In short, better bigger studies have a better adherence or if the journal request it.</p> <p>There are some methodological point that need addressing:</p> <p>Data collection:</p> <p>Only 100 studies were double reviewed and the authors point out that the inter-rater agreement could be quite poor for some items, indicating a rather subjective assessment. Systematic reviews are usually based on consensus but there are no indication that disagreement were resolved and 200 studies were only red by one person. While this is not a systematic review as such, the data collection should be of the same quality.</p> <p>Selection Bias:</p> <p>It is not mentioned in the manuscript which proportion of all protocols present in protocol databases are published. Given that some databases require specific items to be fill in I would imagine that the quality of reporting in databases is quite high. A discussion about how relevant are published protocols compared to data bases would be welcome.</p>	<p>Thank you for your review.</p> <p>Thank you for this comment. Please see our response to Reviewer 2 above. We had a pilot phase to independently appraise the data extraction form and assessment criteria and several round table discussions to harmonise the definitions of adequate and inadequate/unclear for each checklist item. Although data was collected individually for some articles, it was checked at regular fortnightly meetings where any doubts or discrepancies were discussed. Any disagreements were resolved by discussion between the two investigators and, if required, arbitration by a third investigator. The final datapoints used for the analysis were the results of the duplicate data collection and discussion of disagreements.</p> <p>All included RCT protocols were published in full-text in a peer-reviewed journal. RCT protocols published only in protocol databases or online registries were excluded. We did not search online registries and did not measure the number of protocols published in these registries - so this was beyond</p>

the scope of this paper.

However, we have added a section to the discussion regarding the role of the SPIRIT statement in clinical trial registries (e.g. ClinicalTrials.gov, ANZCTR and ISRCTN).

VERSION 2 – REVIEW

REVIEWER	Livia Puljak Center for Evidence-Based Medicine and Health Care, Catholic University of Croatia, Zagreb, Croatia
REVIEW RETURNED	27-Jun-2020

GENERAL COMMENTS	The authors have revised the manuscript adequately.
-------------------------	---

REVIEWER	Odile Sauzet Bielefeld University Germany
REVIEW RETURNED	14-Jul-2020

GENERAL COMMENTS	<p>The complement of information about agreements with data extraction are satisfactory. Unpublished protocol may well not be in the scope of this article, they are nonetheless relevant to any study discussing the quality of reported protocols and therefore at least some information about the part of all protocols which are published would be useful to understand how much this article covers relative to all protocols. Whether published protocols cover 90% of all protocols does not have the same impact than if it covered only 10%.</p> <p>Another minor remark: the authors added the following sentence: "None of the 300 RCT protocols adequately reported all individual checklist items from the SPIRIT statement". I am wondering whether this could not be a trivial statement because not all items are relevant to all studies and thus might not be an indication of lack of quality in reporting.</p>
-------------------------	--

VERSION 2 – AUTHOR RESPONSE

Reviewer 1

Comments

Responses

The authors have revised the manuscript adequately.

Thank you for your review.

Reviewer 2

Comments

Responses

The complement of information about agreements

Thank you for your review.

with data extraction are satisfactory.

Unpublished protocol may well not be in the scope of this article, they are nonetheless relevant to any study discussing the quality of reported protocols and therefore at least some information about the part of all protocols which are published would be useful to understand how much this article covers relative to all protocols. Whether published protocols cover 90% of all protocols does not have the same impact than if it covered only 10%.

Thank you for this comment. To the best of our knowledge, there is no data on published randomised controlled trial protocols as a proportion of all randomised controlled trial protocols. As per the scope of the study, we report the results from published protocols; the extent these results reflect all protocols, including unpublished, is not possible to determine, though we assume the latter may be less compliant as they have not been subjected to peer-review. We have already stressed the importance of adequate reporting in trial registries, and have added the additional statement to the discussion: "This is particularly relevant given many trials may be registered but do not have published protocols."

Another minor remark: the authors added the following sentence: "None of the 300 RCT protocols adequately reported all individual checklist items from the SPIRIT statement". I am wondering whether this could not be a trivial statement because not all items are relevant to all studies and thus might not be an indication of lack of quality in reporting.

Thank you for this comment. We agree that some SPIRIT checklist items may not be relevant to all randomised controlled trial protocols, and have added the additional statement to the discussion: "Additionally, whilst none of the 300 RCT protocols adequately reported all individual checklist items from the SPIRIT statement, some checklist items may not be relevant to all RCT protocols and thus the level of under-reporting observed here may be a slight overestimate."
