

## 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

<b>TITLE (PROVISIONAL)</b>	The STROBE extensions: A protocol for a qualitative assessment of content and a survey of endorsement
<b>AUTHORS</b>	Sharp, Melissa K; Utrobičić, Ana; Gómez, Guadalupe; Cobo, Erik; Wager, Elizabeth; Hren, Darko

### VERSION 1 – REVIEW

<b>REVIEWER</b>	Koen Pouwels Public Health England, United Kingdom
<b>REVIEW RETURNED</b>	14-Aug-2017

<b>GENERAL COMMENTS</b>	<p>This manuscript describes a study protocol focusing on STROBE extensions. The study aimed to identify changes made compared to the original STROBE checklist and will also assess endorsement and uptake by journals and authors. In general, the manuscript is well-written, however I do have some comments:</p> <ol style="list-style-type: none"><li>1. Page 6, line 117-120. '... identify the strengths and weaknesses of the original STROBE checklist; this will identify problem areas or deficiencies conveyed in extension additions.' It is unclear to me how this will identify problem areas and strengths and weaknesses. It would be helpful if it would be explained what the authors mean and preferably give an example. I presume that STROBE extensions are generally focussed on a specific field/type of analyses, while the original STROBE was probably more general. Inevitably this will lead to important differences between the checklist, but does not necessarily highlight any strengths or weaknesses.</li><li>2. Page 12, line 243-244. 'By identifying the content areas that authors have difficulties with, the groundwork will be laid for an assessment into how authors currently use and understand STROBE and what difficulties they encounter with its implementation.' Could the authors give an example to further clarify what they mean with this? And how do the researchers of this study know whether authors of other studies tried to adhere to guidelines and whether they had difficulties with parts of the guidelines or that they just did not bother to adhere to the guidelines in the absence of sufficient mechanisms to guarantee adherence?</li></ol> <p>Minor:</p> <ol style="list-style-type: none"><li>1. Abstract, line 44. 'Upon disagreement, consensus will be reached. Shouldn't reached be replaced by sougheed, as one cannot be sure whether consensus will be reached.'</li></ol>
-------------------------	--

	<p>2. Page 5, line 94-95; page 6, line 129-130; page 7, line 144-145. There is some inconsistency in the manuscript regarding the number of STROBE extensions that do exist. Sometimes sentences in the manuscript suggest there are only 13 extensions in total (page 5, line 94-95; page 7, line 144-145; Table 1 and Table 2) while other sentences suggest that the study will try to find additional extensions.</p> <p>3. How many papers do the authors expect to identify. Will the search strategy result in a feasible number of publication and will it be enough for the planned statistical tests?</p> <p>4. Page 10, line 207-208. Will the text describing guideline documents be manually searched or by text mining, by 1 or by 2 independent researchers?</p> <p>5. Page 20, line 82-92. Why did the authors only include certain antibiotics and completely ignored other antibiotics in the search terms for STROBE-AMS?</p>
--	--

<b>REVIEWER</b>	<p>Carl Lachat Ghent University, Belgium</p> <ul style="list-style-type: none"> <li>- I led the development of a STROBE extension for nutritional epidemiology</li> <li>- I know the lead author and have discussed a study protocol on the use of a writing aid to enhance completeness of reporting (not related to the present study)</li> <li>-no other competing interests are declared</li> </ul>
<b>REVIEW RETURNED</b>	21-Aug-2017

<b>GENERAL COMMENTS</b>	<p>The review will assess the use of STROBE-extensions and endorsement by journals. The study has the potential to address quality of reporting of observational studies in epidemiology. Some elements require clarification or attention</p> <ul style="list-style-type: none"> <li>- The exclusion of almost half of the STROBE-extensions is unfortunate. By the time the review will start, most of these guidance documents will have been published at least for one year. A sensitivity analysis might be more appropriate here instead of simple exclusion</li> <li>- Contacting the journal staff and authors of the STROBE-extensions would yield interesting extra information on the reasons and experiences for endorsing and/or developing STROBE extensions</li> <li>- It remains unclear how the endorsement of non-STROBE guidelines will be assessed and how comparisons will be made</li> <li>- clarify how the findings of this study will be communicated to STROBE and inform development or future efforts</li> <li>- I am not sure how informative the coding of field and non field specific STROBE guidelines will be. I assume all STROBE extensions are developed to provide additional guidance for particular aspects / field in epidemiology.</li> <li>- Some clarification is needed on what is understood with "endorsement". Does this also comprise journals that recommend the use of the guideline or refer to EQUATOR and reporting guidelines in general?</li> </ul>
-------------------------	--

## VERSION 1 – AUTHOR RESPONSE

### Reviewer: 1

Reviewer Name: Koen Pouwels

Institution and Country: Public Health England, United Kingdom

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

Please leave your comments for the authors below

This manuscript describes a study protocol focusing on STROBE extensions. The study aimed to identify changes made compared to the original STROBE checklist and will also assess endorsement and uptake by journals and authors. In general, the manuscript is well-written, however I do have some comments:

1. Page 6, line 117-120. '... identify the strengths and weaknesses of the original STROBE checklist; this will identify problem areas or deficiencies conveyed in extension additions.' It is unclear to me how this will identify problem areas and strengths and weaknesses. It would be helpful if it would be explained what the authors mean and preferably give an example. I presume that STROBE extensions are generally focussed on a specific field/type of analyses, while the original STROBE was probably more general. Inevitable this will lead to important differences between the checklist, but does not necessarily highlight any strengths or weaknesses.

Response: Thank you for this question. We have reworded and clarified the objective for the qualitative assessment of the extensions (pages 5-6). We recognize that many of the items on the STROBE extensions will be field-specific, as they should be, however we reason that they should \*all\* be specific to their field as STROBE should cover all non-specific items. Our interest lies largely in identifying the non-specific suggestions to see if they are already covered in STROBE (and/or its explanation and elaboration document) or highlight information that could be helpful for STROBE to add.

2. Page 12, line 243-244. 'By identifying the content areas that authors have difficulties with, the groundwork will be laid for an assessment into how authors currently use and understand STROBE and what difficulties they encounter with its implementation.' Could the authors give an example to further clarify what they mean with this? And how do the researchers of this study know whether authors of other studies tried to adhere to guidelines and whether they had difficulties with parts of the guidelines or that they just did not bother to adhere to the guidelines in the absence of sufficient mechanisms to guarantee adherence?

Response: Thank you for this comment. We addressed this point on page 13. We are currently planning a survey for authors which will be informed by results of this study. It will ask authors about their adherence to and use of STROBE and the difficulties they have encountered. Our qualitative assessment will allow us to identify key areas (i.e., subsets of the methods, results, conclusions sections) that we can probe authors further about. We can then compare between results of our qualitative assessment and authors opinions to see if the non-specific additions are in fact the areas that authors generally have the most difficulties with.

Minor Comments:

1. Abstract, line 44. 'Upon disagreement, consensus will be reached. Shouldn't reached be replaced by sought, as one cannot be sure whether consensus will be reached.

Response: Thank you for this suggestion. We have corrected the text accordingly to state that consensus will be sought as this semantic change better represents the process of deliberative consensus-decision making.

2. Page 5, line 94-95; page 6, line 129-130; page 7, line 144-145. There is some inconsistency in the manuscript regarding the number of STROBE extensions that do exist. Sometimes sentences in the manuscript suggest there are only 13 extensions in total (page 5, line 94-95; page 7, line 144-145; Table 1 and Table 2) while other sentences suggest that the study will try to find additional extensions.

Response: Thank you for this point. We agree that the wording throughout reflected a degree of uncertainty regarding the number of extensions. We have changed this wording.

3. How many papers do the authors expect to identify. Will the search strategy result in a feasible number of publication and will it be enough for the planned statistical tests?

Response: We expect that the article pool will be quite large as previous research has shown that for certain clinical specialties, 9 out of 10 studies could be observational. [1] However, the time period for the searches is quite restricted so this should help to narrow down the article pool. As this study does not include the analysis of the articles themselves, but rather utilizes articles to screen journals for inclusion, the number of articles is more a resource and time concern for screening, rather than a concern for statistical tests. In order to ensure that our study was feasible, we looked at the number of journals listed within the Broad Subject Terms that we defined. This includes 2,500 unique journals which is divided amongst 7 extensions.

1. Funai EF, Rosenbush EJ, Lee MJ, Del Priore G null. Distribution of study designs in four major US journals of obstetrics and gynecology. *Gynecol. Obstet. Invest.* 2001;51:8–11.

4. Page 10, line 207-208. Will the text describing guideline documents be manually searched or by text mining, by 1 or by 2 independent researchers?

Response: Thank you for this question. We have added the clarifying information in the text under the Data Extraction section.

5. Page 20, line 82-92. Why did the authors only include certain antibiotics and completely ignored other antibiotics in the search terms for STROBE-AMS?

Response: The search strategy was based on a systematic review conducted by the United States Veterans Administration (reference 34). The use of the search item 17 [(antibiot\$ or antimicrob\$).tw.] should pick up antibiotics not explicitly mentioned in the other items. Of note, the search strategy is a combination of OR terms, not AND so the search is quite broad and should not ignore other antibiotics. The inclusion of broad MeSH terms like Anti-Infective Agents, Infection, and Drug Resistance, Microbial also ensure a wide scope. In particular, "Drug Resistance, Microbial" includes beta-lactam, chloramphenicol, multiple drug, kanamycin, tetracycline, trimethoprim, and vancomycin resistance.

**Reviewer: 2**

Reviewer Name: Carl Lachat

Institution and Country: Ghent University, Belgium

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

- I led the development of a STROBE extension for nutritional epidemiology
- I know the lead author and have discussed a study protocol on the use of a writing aid to enhance completeness of reporting (not related to the present study)
- no other competing interests are declared

Please leave your comments for the authors below

The review will assess the use of STROBE-extensions and endorsement by journals. The study has the potential to address quality of reporting of observational studies in epidemiology. Some elements require clarification or attention.

Comment: The exclusion of almost half of the STROBE-extensions is unfortunate. By the time the review will start, most of these guidance documents will have been published at least for one year. A sensitivity analysis might be more appropriate here instead of simple exclusion

Response: We agree with that the exclusion is unfortunate. However, the work contained within this manuscript is the result of lengthy developments for search strategies resulting in a large volume of journals and accompanying articles. Current screening efforts have proven to be time-consuming, encompassing nearly 2500 unique journals. Sensitivity analyses would still require the creation of search strategies and some screening efforts. At this time, the resources are not available to include all extensions. However, this work will provide data and methodological grounds for subsequent efforts in further investigating this issue and/or replicating our results.

Comment: Contacting the journal staff and authors of the STROBE-extensions would yield interesting extra information on the reasons and experiences for endorsing and/or developing STROBE extensions

Response: Thank you for this suggestion. We agree that a survey involving journal staff and authors of the STROBE extensions would provide valuable insight regarding the process of the STROBE extensions. However, we believe that this should be a part of another study as it is outside the scope of the proposed study. We are approaching the problem from the perspective of a typical author who is not aware of the "behind the scenes" motives and experiences of the extension authors. Generally, the extensions include rationale for their creation and details of the methods used. Taking extension information at face value, including the elaboration and explanation documents, is the typical experience of authors of observational studies. We have added information to the manuscript which attempts to clarify the qualitative evaluation of the content of the extensions. Also, of note, this study will collect all the publicly available information for included journals, making follow-up studies with journal staff easier.

Comment: It remains unclear how the endorsement of non-STROBE guidelines will be assessed and how comparisons will be made.

Response: A classification scheme will be created for endorsement which is based on, and will most likely reflect, prior research. Generally speaking, endorsement can fall into a few categories ranging from strong to weak. Examples include a requirement of a completed checklist with manuscript submission, a suggestion that authors “should” reference or follow a specific guideline, a general suggestion that authors should adhere to reporting guidelines, a vague suggestion that authors should adhere to certain standards which include reference to reporting guidelines, or no explicit mention at all.

Comment: clarify how the findings of this study will be communicated to STROBE and inform development or future efforts.

Response: As a part of the first author’s doctoral studies, she is working directly with the EQUATOR Network which is deeply involved in guideline development and promotion.

I am not sure how informative the coding of field and not field specific STROBE guidelines will be. I assume all STROBE extensions are developed to provide additional guidance for particular aspects / field in epidemiology.

Response: We agree with this point and emphasize that the coding is informative for an educational intervention and results may not be extrapolated to other domains. This point was also raised by the other reviewer and we have added information on pages 5 and 6 that further clarifies the goals and expected outcomes of the coding.

Comment: Some clarification is needed on what is understood with "endorsement". Does this also comprise journals that recommend the use of the guideline or refer to EQUATOR and reporting guidelines in general?

Response: Thank you for this question. We have provided examples of different types of “endorsement” with further detail (p.12 of the manuscript). Based on prior research (references 6, 38, 39), we anticipate that there will be several types of “endorsement” found in author guidelines.

## VERSION 2 – REVIEW

<b>REVIEWER</b>	Koen Pouwels Public Health England, United Kingdom
<b>REVIEW RETURNED</b>	13-Sep-2017
<b>GENERAL COMMENTS</b>	<p>The comments have been addressed sufficiently. I do have one remaining minor suggestion:</p> <p>Wouldn't it be better to already predefine how many and which categories of endorsement will be compared, instead of giving examples of types of endorsement? It is not clear yet whether the authors will only make the suggested yes/no comparison regarding endorsement or whether they will also formally assess associations for different types of endorsements (e.g. strong (i.e. checklist), intermediate, weak). If such a comparison is not planned; wouldn't it be interesting to see e.g. whether different types of endorsement correlate with impact factors as well?</p>

<b>REVIEWER</b>	Carl Lachat Ghent University Belgium -I led the development of a STROBE extension for nutritional epidemiology -vl know the lead author and have discussed a study protocol on the use of a writing aid to enhance completeness of reporting (not related to the present study) - no other competing interests are declared
<b>REVIEW RETURNED</b>	16-Sep-2017
<b>GENERAL COMMENTS</b>	Thank you for addressing my comments

### VERSION 2 – AUTHOR RESPONSE

Thank you for the time spent reviewing our response and comments. As we have not reviewed all possible semantics of endorsement from journals, we would like to keep our approach flexible. However, we have added some more specifics regarding our initial coding approach (page 12, lines 234-242).

In reference to the question regarding comparisons regarding endorsement, we note in the statistical analysis section (page 13, lines 261-264) that we plan on using operationalizing endorsement as binary (yes/no) for analyses investigating associations between endorsement of other guidelines (e.g., STROBE, yes/no) and impact factor. We are not yet certain that the size of the subgroups of endorsement (i.e., endorsement can be split into several types) will allow for us to conduct such analyses, however are open to the idea if sample sizes allow for it.