Reviewer's report

Title: The use of evidence in public governmental reports on health policy: An analysis of 20 Norwegian official reports (NOU)

Version: 2 Date: 16 April 2009

Reviewer: Benedicte Carlsen

Reviewer's report:

Reviewer report on revised version of the paper The use of evidence in public governmental reports on health policy: An analysis of 20 Norwegian official reports (NOU).

Generally I find that the paper is clearer and better structured then before, but I still miss satisfactory responses to some of my questions from the first review and I find that some new problems have appeared in the new version. My main objection now is that the discussion is messy, parts of it are not directly related to the results, the results are not scrutinized and the discussion of methodological limitations is weak. I explain details below.

Firstly, I comment on the former comments where I find your responses incomplete.

2. I find the following argumentation problematic: You start by deciding (without telling the reader why) to use the ISQ to evaluate the scientific quality of the NOUs and (on p 4) that if the reviews of effects in the NOUs are transparent and explicit, the policy makers may weigh the evidence of effects against other factors that they would like to consider, but then you go on to argue that ISQ may not be a relevant instrument to inform policy makers of the NOUs qualities because it only assesses scientific evidence. This makes me wonder why you chose that instrument in the first place and whether you intended to evaluate the usefulness of the ISQ instruments for wider purposes then simply assessing the scientific quality? Maybe the paper would have been clearer without involving an evaluation of the ISQ instrument?

Response: See my response to your points 5, 8, 24, 25, 30-32. I did not write “that ISQ may not be a relevant instrument to inform policy makers”.

New comment (see also my comment to point no 24 below): Firstly; I was not quoting you here, but still you end the manuscript with the following: “the information provided by applying ISQ criteria is narrow and other kinds of information may be more relevant for policy-makers.”

The motivation for using the ISQ is now clearer, but I am still not comfortable with the argument in the discussion. You state elsewhere that the evaluation of the
ISQ instrument has been played down, but it is still part of the aim stated in the abstract, and your discussion is still much about the usefulness and relevance of the ISQ instrument in evaluating NOUs. I still find this problematic as your initial argument is: “Instead of one more discussion paper on evidence-based policy-making, this is an empirical study which examines whether the scientific evidence of effect used in health policy documents in Norway were transparent and explicit, and whether the examination of the evidence of effect was clearly identified and scrutinised.”

The above also shows that the aim stated in the abstract is still not consistent with the aim stated in the background section.

If you disagree with me on this point, I suggest that we leave it to the editor to decide whether this is of importance or just a hang up of mine.

3. The aim or parts of the aim of the study are repeated six times during the manuscript and the meaning is not consistent. I would advice you to reduce this to no more than three; describe the aim in the abstract, in the introduction and perhaps once in the discussion or conclusion.

Response: The meaning is now consistent. I have deleted two of my references to the aim (page 12 and 14, see my response to your points 25 and 40).

New comment: The meaning is not consistent. The aim in the abstract is double and the aim on p 4 (quoted above) is single. See also my previous comment.

5. It is unclear in the abstract whether the ISQ is developed for evaluating scientific reports or government reports, reviews, meta-analyses or single studies?

Response: Thank you, the methods section in the abstract now includes the following sentence:” The ISQ is based on common scientific criteria for all research based on rigorous methods that aim at answering questions of effect.”

New comment: This information should also be included in the methods section where the ISQ is described.

7. In the end of the section, it would be relevant to state what expectations or hypothesis you had regarding your study of the Norwegian NOUs.

Response: Unfortunately, we did not state any hypothesis about what we expected to find.

New comment: But you do in the following on p 14 in the discussion section: “I expected some issues to be more challenging for the committees to assess, for example, how organisational changes affect health outcomes. In these cases, I expected lower quality scores.”

Please follow up on my earlier comment.
19. Also the information that only five of the reports discuss more than one study should be placed in the beginning of the results, as this is basic information about the sample of NOUs.

Response: This is still about consistency.

New comment: Still this is key information that should have been given in the overview of the included NOUs. In addition; if only five reports could be assessed according to consistency, I find the statistical information about the average score of only five reports meaningless. Here the variation in scores or all scores should have been reported.

20. Then you go on to say that “The other reports provided no information about consistency between studies.” Does this refer to the other 12 studies and does this mean that these reports are based on one single study or are they based on several but do not discuss consistency between studies? Please clarify.

Response: I study how 17 NOUs describe consistency between studies. “The other reports” do not refer to 12 studies, but to 12 reports (NOUs).

New comment: I am sorry; I meant 12 reports, not 12 studies; but nonetheless it is unclear to me whether the other reports are based on a single study or only discuss one study. If the former is true; how could these 12 studies consider consistency between studies? The question is irrelevant to them. Please clarify this in the manuscript.

24. The discussion: Based on the results I think the first thing to discuss should be the scientific quality of the NOUs. I fail to see that the results of this study call for conclusions about the usefulness of the ISQ-instrument and the discussion regarding this appears speculative (see my main point about argumentation above). Nor do I find this question relevant as long as the policy makers in Norway have not explicitly asked the researchers to evaluate the NOUs according to the ISQ. If the evaluation of the ISQ shall be part of this paper at all, we need more information about former evaluations of the ISQ, about which policy makers, in which countries argue for evidence based government reports, and when these statements are given.

Response: Three new paragraphs are now present in the beginning of the Discussion section, where results are more directly and closely discussed.

I will reserve for myself and others (as researchers and citizens) the right to evaluate policymaking in other ways than those that are asked for by policy-makers. When and why policymakers ask for an evaluation is not the most important matter to scrutinize, in my opinion. It is more important to study the way in which policy is made. It is possible that governments that have not called for evidence-based policy making are more evidence-based in their own policy-making than governments that do make such a call. That would certainly be an interesting empirical finding, one that would require honed interpretive
skills. For other responses to your comment about the usefulness of the ISQ, see my responses to your points 2, 5, 8, 17, 25, and 30-32.

New comment: I appreciate the three new paragraphs discussing the results in the beginning of the discussion, but still a substantial part of the discussion is about whether it is a good idea to use ISQ and similar measures of scientific quality to assess policy documents. Hence, I uphold my comment; I just think this part of your focus knocks the bottom out of your own evaluation.

26. In this chapter the possibility of bias in the report is mentioned. Here I would expect a reference to the information in the results about different types of NOU authors.

Response: I find it unnecessary to discuss bias related to different types of NOU authors as long as the results are: “Some committees consisted mainly of experts in the field, while other groups consulted external experts and included their recommendations as appendices. Expert participation did not seem to influence the strategy used for information collection.”

New comment: It seems to me as a contradiction when you find that type of author do not affect the NOUs in this study but on the other hand states that: “Politicians, scientists, bureaucrats, experts and lay people often have different (and legitimate) interests. […] Explicit and transparent references to the evidence will help to identify bias in the recommendations in governmental reports.” In my opinion this may confuse the reader.

28. Also the discussion of methods should be more elaborate.

Response: I hope the new version is appropriate. If it is not deemed to be appropriate, I will need more information about what you want me to discuss.

New comment: While some methodological strengths are mentioned, there seem to be no methodological limitations discussed either in the section called “Is the sample representative for current NOUs and relevant for comparison to governmental reports in other countries?” or elsewhere. I do not see it as my job to suggest limitations to you, but relevant themes could be: Disagreement about how to evaluate the reports, the size and generalisability of sample of NOUs, the great variation in scientific quality, problems with treating such a low number of observations statistically etc.

I also suggest you rename this paragraph to something more traditional such as: “Strengths and limitations of the study”.

35. Page 4-5, last paragraph. Some of the sentences are difficult to understand. Please improve the language.
Response: Please inform me about the sentences that you find to be difficult to understand and in need of improvement.

New comment: This is one of the sentences I was referring to:
"Instead of one more discussion paper on evidence based policy-making, this is an empirical study which examines whether the scientific evidence of effect used in health policy documents in Norway were transparent and explicit, and whether the examination of the evidence of effect was clearly identified and scrutinised."

Was it the examination (or rather the evidence) that should be clearly identified and scrutinised?

Comments to the new version of the manuscript:

P 6: The ISQ is referred to here, but the description of the ISQ is not given until p 7.

P 8, third paragraph: Belong in the discussion section, under Strengths and limitations of the study.

P 12: In the following sentence “not interested” should be replaced by “uninterested”:
"This may be due to a belief that policy makers are not interested in the science behind recommendations."

There are several other sentences in the new text that I suspect represent errors in the English language, but as I am no English speaker, I believe you should have the new text revised by an expert in the English language.

P12, second paragraph: This is unclear and seems contradictory. Here you claim that there is a lack of clarity in the reports (a score of 3 is deemed low here, but not elsewhere?), but on p 9 you say: "The committees wrote very clear and structured summaries, with an average value of 3.4." Please clarify in the text.

P14: first and last paragraphs: Here you are suggesting implications. These should be grouped together and mentioned in the end (e.g. under Conclusions and implications).

Also in the last paragraph: The claim that the selected NOUs are relevant also in other European countries need to be documented or argued for in a more convincing way.

P15: The second paragraph about the Norwegian Knowledge Centre seems out of place in the manuscript (I know that you have added it as a response to one of the other referees), and certainly out of place in this section discussing strengths and limitations of the study.

Also, most of the second paragraph is not about strengths and limitations of this study.
When discussing shortcomings of the ISQ, why do you not suggest adding the elements that you yourself have added in your study? (Points 1-3 in Box 1). I find this to be the main practical value of this study and would like to see a discussion of this and how the findings in 1-3 relates to the findings regarding scientific qualities. I suggest you concentrate on discussing your own findings in stead of discussing the relevance of assessing NOUs according to scientific criteria.

P17: I find the second paragraph repetitive and therefore unnecessary, it also adds to the confusion of the aim of the study. I advise you to remove it.

**Level of interest:** An article of importance in its field

**Quality of written English:** Needs some language corrections before being published

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**

I declare that I have no competing interests.