Author's response to reviews

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

Authors:

Simon Innvær (simon.innvar@sam.hio.no)

Version: 2 Date: 3 April 2009

Author's response to reviews: see over
Author's response to reviews

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

Author:
Simon Innvær
Date: 03 April 2009

Dear editors and reviewers,

I thank all three reviewers for their useful and often quite similar comments. Their comments will certainly help to improve the article. In the following, I will respond to all of the major compulsory and discretionary calls for revision from the reviewers, beginning with reviewer two and three, since I believe that their comments and criticisms address several of the (more detailed) comments and criticisms forwarded by reviewer one. Sometimes, when comments from the reviewers refer to parallel concerns, I will refer to the previous response. The comments of reviewers are written in italic print. I have reviewed my references and minor changes now appear in four of the original references (5, 13, 15, and 31).

DETAILED RESPONSE TO REVIEWERS (“point by point description of changes”)

Response to reviewer two:

Reviewer two wants me to address two concerns:

1. “Can we assume then that EIDM [evidence-informed decision making, S.I.] was not happening at the policy level from 1994-98, given the NOUs rated so poorly on the ISQ. How might this be different now in 2009 in terms of what the Norway government is asking for? I think some attention to where most organizations were at with EIDM in the mid 90s, in contrast to where we have come by late 2000’s, is an important discussion that needs to be intertwined into this paper, given the data were collected such a long time ago. What hypotheses can be made about the value policy-makers place on evidence, given the mandates themselves did not specify that the use of high quality evidence had to be demonstrated. How does the language of the mandates need to change in order to facilitate use of high quality evidence? What other changes need to happen at the policy level in order to promote a valuing of high quality evidence in decision making.”

Response: The concerns identified above are important and very interesting. However, hypotheses about the value that policy-makers place on evidence seems to me to be an empirical question that is outside the scope of this article. The other concerns are discussed under the heading “Is the sample representative for current NOUs and relevant for comparison to governmental reports in other countries?”

The following new paragraphs have been inserted:
“A move to promote evidence-based policy making in Norway appears to have been made in January 2004, with the formation of an institution called The Norwegian Knowledge Centre for the Health Services (NKCHS). The Centre is organised and funded by the Norwegian Directorate of Health. NKCHS is meant to be scientifically and professionally independent. The Centre has two overarching tasks: The first is to promote evidence-based policy making, by producing health technology assessment reports, systematic reviews, overviews and reports that provide policy makers with early warnings. The second task is to support health services at all levels to incorporate evidence into their practices.

NOUs produced from 1994 through 1998 can be said to be representative for the 19 NOUs published since 1999. The formal framework for preparing a NOU has not been significantly changed, only minor revisions were made in 2005. By and large, it is up to the committee to determine its methods and procedures for producing the NOU report. When committees interpret their mandates and design their reports, they do not utilize a systematic approach that is similar to the ISQ. Only one of the NOUs published after 1998 reported that the committee preparing the report asked the NKCHS for a review of the relevant research. The NKCHS responded by stating that such a report would take the Centre from 6 to 24 months to produce. The committee deemed this to be beyond the framework provided to it to complete its work.

Two changes at the policy level may be able to promote greater reliance upon the evidence of effect in decision making: a) the language of the mandate can more clearly focus upon questions of effect, and b) the Ministry of Health and Care Services can give the NKCHS clear instructions and sufficient resources to systematically review questions of effect.”

2. “While the author discusses limitations of using the ISQ for this purpose, I did not see a discussion of limitations of this study itself. Perhaps some attention to limitations would be useful. For example, to what extent did pairs of researchers agree on extracted data. How much discussion had to occur to achieve consensus? The age of the data (94-98) and how this relates or not to current practice more than one decade later.”

Response: A discussion of “the limitation of the study itself” is now included in the article in the results section with the subheading ‘Sensibility of the ISQ: Degree of agreement among researchers’:

“Others have carried out a comprehensive sensibility analysis of the ISQ [31,32]. They found the ISQ acceptably reliable and credible, with one major problem, the need for judgement in making ratings. We found the same problems, but we were able to reach agreement by discussing our initial and minor disagreements.”

My response to your concern 1 answers “The age of the data (94-98) and how this relates or not to current practice more than one decade later.”

Response to reviewer three:
Reviewer three has called for nine major compulsory revisions:

1. “I had to read the text in the method section, page 5, the final 7 lines to page 6, line 1-3, several times and compare it with the text in Box 1, and the text under Results (The mandate, Strategy used to collect data, Scientific quality and Recommendations and Summary) to understand how these parts of the manuscript were linked. E.g. the first of the four components is labelled “The mandate’s description of the task” in Box 1, “the language in the mandate” in the Method section and “The mandate” in the Result section. Please use similar wording in the different parts of the paper which deal with these four components. It is also confusing that the four components are not presented in order in the Result section (component 4 is presented before component 3).”

Response: Similar wording is now used and the four components are presented in the same order as in box 1.

2. “I recommend a more comprehensive description of the ISQ, especially the modified version used by the author. Is this version of the ISQ tool validated? Box 2 is informative, but only regarding component 3 “The use of evidence in recommendations and in the summary”. A corresponding Box for component 4 “Evaluation of the quality of scientific evidence” is needed.”

Response: My expression “modified version of the ISQ” is misleading, since point four in box 1, “the evaluation of the quality of scientific evidence”, is identical with the ISQ. The modification here is that we added three points, the mandate’s description of the task, the strategy used by the committee to gather research information, and the use of evidence in recommendations and in the summary. A new box 3 for component 4 has been inserted and a more comprehensive description of the ISQ follows under the heading “The index of scientific quality”, and a new paragraph in the results section reads:

“Sensibility of the ISQ: Degree of agreement among researchers [subheading, S.I] Others have carried out a comprehensive sensibility analysis of the ISQ [31,32]. They found the ISQ acceptably reliable and credible, with one major problem, the need for judgement in making ratings. We found the same problems, but we were able to reach agreement by discussing our initial and minor disagreements.”

3. “Page 8 under Scientific quality. Since the 119 evaluations are not mentioned under the Method section I had to read this sentence several times to understand what was meant. A parallel to Box 2 would probably clear this up. The text under this heading repeat what is shown in Figure 1.”

Response: I hope the following sentences in the beginning of the paragraph will clear this up: “The 17 NOUs were evaluated by seven assessments of the quality of scientific evidence. This produced 119 evaluations distributed on a five-point scale. The scores for all the 119 evaluations appear in Figure 1.”

4. “Page 11, paragraph 2, line 4-5: remove the sentence “Dissemination is an
important matter to consider, as many evaluators have [37-39], it is not relevant.”

Response: The sentence has been removed.

5. “Explain the reference to Kingdon [42] or remove it.”

Response: The reference has been removed.

6. “The Method description in general has to be straightened up, e.g. the text under Four main components. The sentences on page 6, line 1-3, “The main focus in this paper is the fourth component, the way the committee USES scientific evidence. The results are contained in three figures that break down the evaluation into eight factors (each based on a specific question)” have to be clarified. This also applies to the text under the heading The selection of questions of effect, beginning with “A list of questions for all reports …”, page 7.”

Response: The first part of this list of concerns is partly answered under my response to this reviewer’s first criticism, but the sentence “The results are contained in three figures that break down the evaluation into eight factors (each based on a specific question)” is deleted since it is about results and the sentence is not needed under results.

The text “a list of questions for all reports …, page 7” has been changed and now reads: “The included NOUs circulated between four researchers until all agreed on the relevant questions of effect in the mandate (or on the correct interpretation of the mandate). Then these researchers worked in pairs to score all reports by the questions in box 1-3. After that, group discussions resolved opposing views.”

The heading “The selection of questions of effect in the mandate and the composition of the committees” has been moved and now appears in the beginning of the methods section, before the heading “Four main components”.

7. “The “predefined scheme developed from the questions in box 1 and box 2” has to be further described.”

Response: The sentence (on page 7); “Then these researchers worked in pairs to score all reports in a predefined scheme developed from the questions in box 1 and box 2”, has been changed to read; “Then these researchers worked in pairs to score all reports by the questions in box 1-3.”

8. “The method limitations, e.g. the modified version of the ISQ, ought to be discussed.”

Response: See my response to the criticism raised in point 2.

9. “Under Conclusions, the text beginning with “Reports that focus upon the ISQ criteria …” should fit better under Discussion than under Conclusion. The “Other frameworks exist [14-29] should either be remover or developed under the
Discussion section.”

Response: The sentence beginning with “Reports that focus …” has been moved to the discussion section under the heading “Narrow focus and broad mandates”, first paragraph. I have also deleted “Other frameworks exist[14-29]. They can and probably will express different perspectives and generate quite different results. An interesting future research agenda would be to use the same governmental documents as data to compare the findings, usefulness and evidence base in other frameworks with the ISQ. Applying the ISQ criteria is only one way to evaluate health policy-making.”

Reviewer three also wants me to make nine minor essential revisions:

1. “The aim of the study is quite clear to me, but the text is not consistent between the abstract and the background section.”

Response: The text is now consistent.

2. “Page 5, line 6-7: The final sentence, including references 14-29, in the background section is not necessary (outside the scope as the author states).”

Response: I find the background section to be the right place to address both the focus of the paper and references to the broader picture. Readers with interest in other ways of evaluating evidence will find valuable references. I really hope this sentence will be accepted.

3. “Under Results, page 7, after the first sentence, add something like: “Three NOUs were excluded due to ….., resulting in a total of 17 analysed NOUs”.

Response: This has been done. I’ve written: “Three NOUs were excluded due to lack of questions of effect in their mandate, resulting in a total of 17 analysed NOUs.”

4. “Page 8, under Strategy used to collect data. Expert participation is not mentioned in the Method section but appears under Results.”

Response: The following sentence was added to the Method section under the heading “The selection of questions of effect in the mandate and the composition of the committees” and reads: “To find out whether the composition of the committees influenced the strategy used by committees for information collection, we mapped out the relative part that was played by researchers, policy makers and representatives from interest organisations.”

5. “Consistent use of numbers, page 10 Figure three or Figure 3 (Figure 1 and Figure 2 on previous page).”

Response: Thank you, this now reads “Figure 3”. Another inconsistency has been corrected in regards to box 3.

6. “A short description of the main results either at the bottom of the result section or at the beginning of the Discussion section provides a good service to potential
Response: The sentence on page 11, line 8-10, has been moved to the beginning of the Discussion section.

7. "Box 1. Consistent use of (Yes/No), either after the component (as in 1) or after the questions (as in 2). Bold or not bold?"

Response: Thank you, this has been done (as in 1) and bold has not been used.

8. "Box 3. Explain also in the box why the excluded NOUs are excluded."

Response: The sentence now reads: “Three NOUs were excluded due to the lack of questions of effect in their mandate, resulting in a total of 17 analysed NOUs.”

9. "Figure 1. Refer to Box 1, component 4, question A-G. Judgements wrong spelt. Why is overall quality not included?"

Response: Figure 1 now refers to the new box 3. Spelling of the word ‘Judgements’ has now been corrected. Overall quality is not included because its inclusion would mean that the scores would be counted twice.

**Response to reviewer one:**

Reviewer one points out two main problems which are discussed in detail by referring to points 1 and 2 below:

1. You need to be clearer about the different levels of analysis, as I read it now it is often confusing. There are several levels to keep track of in this study; your own meta-analysis (review and assessment) of the NOUs; the meta-analysis of the NOUS; your review of the ISQ instrument (for use by the NOUS to evaluate scientific reports or for evaluation of NOUs?). This is a complex picture which needs be very orderly and clearly explained for the reader to follow your arguments.

Response: I hope that the confusion has been resolved by my response to the two other reviewers and by my response to your detailed comments (see for example my response to your points 18-22). I evaluate how NOU committees appraise evidence of the questions of effect according to an index of scientific quality (ISQ). Then I discuss shortcomings of the ISQ. That’s it.

2. I find the following argumentation problematic: Your 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”.

The abstract:
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).

4. Description of aim: Please be clear about whether both the NOUs’ processes of systematic searches and their synthesis of evidence are to be evaluated. It is also unclear to me if only NOUs that assess questions of effect are to be included in this study. This should also be clarified in the description of inclusion criteria. In addition I would like to know if you actually are evaluating the synthesis of recommendations as you say here. I get the impression elsewhere that you are evaluating the synthesis of evidence, i.e. are the NOUs to be evaluated as reviews or as meta-analyses?

Response: This is hopefully clear in box 1 (2b) and in the new box 3. Inclusion criteria in the abstract and in the beginning of the Method section state what you ask for (abstract summary of criteria: “health reports at the NOU level in 1994-1998, where questions of effect are relevant in the mandate or in the committees’ interpretation of the mandate.” Method section: “All NOUs commissioned by the Norwegian Department of Health and Care Services from 1994 through 1998 were collected and those that reported evidence of the effect on health were included in this study.”)

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.”

6. The background: Page 4, second paragraph, last sentence: Please explain which kind of documents you are referring to. UK documents? If similar studies have not previously been conducted on Norwegian government reports this information would be relevant here.
Response: The UK, WHO and Mali are now included in the text.

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.

8. The methods: I find the description of methods messy and difficult to follow: I think you need to explain some more about the ISQ-instrument. For what purposes was it designed and why did you choose this instrument? What are the questions taken from the ISQ? Are they similar to point 4 in box 1? Or is perhaps box 2 taken form the ISQ-score? How does box 1 relate to box2? Box 2 seems to be an elaboration of component 3 in box 1, but this is not very explicit as the reference to box 2 is placed after the description of the ISQ-instrument in the text. Also you state that component four is the main focus of the study, so why is component 3 highlighted in a new box?

Response: Thank you, this problem is hopefully solved in my response to reviewer three, point 1 and 2.

9. The reader needs to know about the selection process and your evaluation. Were the NOUs not evaluating effectiveness excluded from the study? And if so, from which component of the study are they not included? Why did you choose NOUs to represent government reports and why was the time span 1994-8 chosen?

Response: The article already states that effect is an inclusion criterion. When questions of effect have been excluded from the mandate of the NOU or from the interpretation of the mandate by the committee, the NOU was excluded from this study. The NOUs are reports that study “complex policy questions of national importance” and this justifies why we chose to study NOUs. In order to justify the choice of NOUs further, the following sentence has been added to the background section, second paragraph: “NOUs are clearly the governmental policy documents that to the greatest extent refer to research evidence.”

10. In the description of the first component, arguments as for example for why the data search should be transparent, are repeated. These should be moved to the introduction or discussion sections.

Response: I have deleted the sentence: “A comprehensive and transparent search for data provides a good starting point for summing up all that is known on the subject.”

11. Page 6, first paragraph, last sentence: Are you referring to the results of only the last component here? Also it seems that only figure 2 illustrates the eight factors.

Response: The sentence “The results are contained in three figures that break down the
evaluation into eight factors (each based on a specific question)” has been deleted since it is about results and the sentence is not needed under results.

12. Page 7, first paragraph: The first half of this paragraph belongs with the results. It is also confusing that this paragraph apparently describes the first component of the study, but is presented last.

Response: Thank you, the first half of this paragraph is moved to the beginning of the Results section, second sentence. The rest of the paragraph is moved to the beginning of the Method section.

13. The results: It is difficult to understand the logic in the structure of the results. After some searching, it appears that the subtitles are partly consistent with the component 1,2 and 4 in box 1, but why is component 3 left out?

Response: Thank you, component three was presented last, but it is now presented in its proper order.

14. Also I suggest that the eight questions that you claim are the main results (figure 2) should be highlighted in the structure of the results in the text.

Response: The text is now highlighted (in italics).

15. In the paragraph, “Strategy used to collect data”, the effect of types of committee participants is described and analysed. I find no reference that this was supposed to be part of your study in the introduction, aim or methods section, nor is this finding mentioned in the discussion.

Response: A new section: “The selection of questions of effect in the mandate and the composition of the committees” is presented in the beginning of the Method section. It reads: “The included NOUs circulated between four researchers (see the acknowledgements at the end of the paper) until all agreed on the main questions of effect in the mandate (or on the correct interpretation of the mandate). Then these researchers worked in pairs to score all reports by the questions in box 1-3. After that, group discussions resolved opposing views. To find out whether the composition of the committees influenced the strategy used by committees for information collection, we mapped out the relative part that was played by researchers, policy makers and representatives from interest organisations.”

16. How are the 119 evaluations extracted? How does this number appear? How are they distributed across the 17 studies?

Response: The following sentence now appears under the heading Evaluation of the quality of scientific evidence:
“The 17 NOUs were evaluated by seven assessments of the quality of scientific evidence. This produced 119 evaluations distributed on a five-point scale.” (17x7=119).
17. In the discussion I miss a comparison of these results with the results of other evaluations of government reports that have used the same instrument. Are these NOUs scoring higher or lower than other government reports? (And what about the mean quality scores of scientific publications?)

Response: I’ve added two sentences after the two first sentences under the heading “The Index for Scientific Quality”, which reads: “Most studies used the index to evaluate health reporting in newspapers. We know of no other studies using the ISQ to evaluate governmental reports.”

18. Page 10, first paragraph: I do not understand what is meant by “how the evidence related to other findings in the field”. Do you refer to how the evidence found in the NOU (based on a review or meta-analysis, or review of meta-analyses) relates to other types of evidence or how evidence from different individual studies used in the NOU relates to each other? I suspect my confusion here is due to the general confusion of levels that the manuscript conveys.

Response: See new box 3: “Question F. Consistency: Is the consistency of the evidence (between studies) considered?” The sentence is now changed on page 10 and now reads: “Only one report specifically discussed how the evidence related to other findings in the field (question F on consistency between studies).

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.

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).

21. I would also like to be informed about the level of consistency; whether “consistency” here refers to consistency between studies in the reports or consistency in the reports’ meta-analyses of different studies.

Response: It is about consistency between studies that answer the same question of effect that is posed by the NOU being investigated.

22. Also regarding the question about consequences: Please specify whether this point assesses A) the reports’ assessment of consequences or B) whether the individual studies used in the reports report consequences.
Response: We always evaluate the assessments that are made in the report, but those assessments can be based on individual studies. We do not assess individual studies. I hope my answers to your previous responses clarify this question.

23. The subtitle “Recommendations and summary” is confusing because you might think it is about recommendations and summary of this study. It is also confusing that this section, which relates to component 3, is placed after the section describing component 4. Also I miss any reference to question 3B.

Response: The ‘Recommendations and summary’ references are now properly placed. Thank you for discovering the lack of reference to question 3b. The following sentences have been added: “The committees wrote very clear and structured summaries, with an average value of 3.4. 10 NOUs scored 4, only three scored lower than 3.”

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. Also, it must be clear whether component 3B is part of the ISQ-instrument and also the results from point 3B must be presented.

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 policy-making in other ways than those that are asked for by policy-makers. When and why policy-makers 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. Your comment about component 3B is answered in my response to your point 23.

25. Page 12, first paragraph, second phrase: Once again the aim of the study is repeated inconsistently: “The main concern in this article is the difficulty in determining how recommendations relate to evidence.” To me this is unclear. To whom is it difficult? Is the main aim the problems for the researcher to evaluate the NOUs according to the ISQ? The same point could be made about the first sentence of the second paragraph. Here you also need to be clear about the level of analysis. Some issues were difficult to assess to you or the authors of the
NOUs?

Response: The sentence about the aim of the study has been deleted. It is unnecessary (you may want to review my responses to your points 3, 4 and 40 here). The first sentence of the second paragraph now reads:
“I expected some issues to be more challenging for the committees to assess …”

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.”

27. The subtitle “From sample to population” is about strengths and limitations of the study. It would be clearer to reformulate the title accordingly and perhaps remove the other subheadings that relate to the NOUs and not to this study.

Response: The text under this subheading is changed substantially, see reviewer 2, point 1.

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.

29. On page 12 a suggestion for future research is given. This would fit better in the conclusion were other implications are drawn. In addition the suggestion is a bit unclear: For what do you suggest “mapping out the pros and cons from different practices”?

Response: The sentence about future research has been removed (see my comments to reviewer three, point 9, major revisions). I have not suggested any “mapping out the pros and cons from different practices”. I have written: “If the report refers to evidence in an explicit and transparent way, readers will, to a greater extent be able to make reliable judgements by comparing the pros and cons of the evidence. Explicit and transparent references to the evidence will help to identify bias in the recommendations in governmental reports”.

The conclusion:
30. Both in the conclusion and in the discussion I miss a thorough discussion of the results. The discussion of the usefulness of the ISQ-instrument is not supported in the data and should be played down. Also, as the reader has no reason to expect that future NOUs will focus upon the ISQ criteria, the focus on the usefulness of the ISQ seem exaggerated.

Response: I have played down the focus on the ISQ criteria in the conclusion.
31./32. I have problems following the main argument in the conclusion: I read the argument as follows: A systematic, explicit, and transparent approach in NOUs generate evidence based decision making, and is necessary to use the ISQ. Policy makers ask for scientific evidence, but maybe scientific evidence is not the most relevant information for policy makers? This argument is not convincing and not according to the main aim or findings in the study. In addition it appears a bit arrogant: If the politicians want evidence, then who are we as researcher to tell them to remember to also consider their political aims and other pragmatic concerns?

Response: Nowhere in the article do I “tell them to remember”. In the conclusion, I write:

“A systematic, explicit and transparent approach – such as that required by the ISQ – may help to generate the evidence-based decision-making that Norway, the UK, the EU and WHO desire and seek. On the other hand, the information provided by applying ISQ criteria is narrow and other kinds of information may be more relevant for policy-makers.”

In the conclusion to the abstract, I write: “The ISQ information leaves out broader perspectives, political decisions and aims, budgetary procedures, cost-effectiveness, prudent practices and commonsense understandings that are obvious but which are not supported by research evidence.”

In fact, the reason why I chose to go beyond merely evaluating the NOUs by the ISQ and to include a discussion of the shortcomings of the ISQ is the following: I believe that we can learn something by evaluating our presuppositions. We may become wiser, and, hopefully, less arrogant.

Minor essential revisions:
33. The abstract: I would like to know if Norway is one of the countries you mention in the first sentence.

Response: This is the broad picture and I give references to Norway in the Background section. See also my response to your point 24 (second paragraph).

34. The Background: The sentence: Previous studies of health policy documents investigated...
Please add “have” after “documents”.

Response: Done.

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.

36. The results: Please mention why three NOUs are excluded. Are the seventeen included
NOUs all about effects?

Response: Yes, questions of effect are necessary in the mandate or in the interpretation of the mandate for inclusion in our study. See my response to your points 4, 5, 9, and 15.

37. “Recommendations and summary”, first sentence: Please change to: “Part 3A in box 1 asks whether the committee stated how they valued and weighted the different aspects of costs, …”

Response: This has been rewritten because of comments from reviewer 3 and now reads: “The third component is an overall evaluation of the use of evidence in recommendations and in the summary”.

38. The discussion: Page 13, first sentence: The sentence does not make sense. I believe either it should say “Even if the NOUs had included” or “When the NOUs included”. Please correct.

Response: The sentence has been corrected and now reads: “Even if the NOUs did include a comprehensive ISQ review of questions of effect, there would still be problems.”

39. Page 14, last paragraph of the discussion: Another version of the aim of the study. This contains an important clarification that would have been informative for the reader to see in the introduction of the study.

Response: This is now included in the abstract and removed from page 14.