Author's response to reviews

Title: Bridging health technology assessment (HTA) with multicriteria decision analyses (MCDA): field testing of the EVIDEM framework for coverage decisions by a public payer in Canada

Authors:

Michele Tony (michele_tony@biomedcom.org)
Monika Wagner (monika_wagner@biomedcom.org)
Hanane Khoury (hanane_khoury@biomedcom.org)
Donna Rindress (donna_rindress@biomedcom.org)
Tina Papastavros (Tina_Papastavros@wsib.on.ca)
Paul Oh (Oh.PaulDr@TorontoRehab.on.ca)
Mireille M Goetghebeur (mireille_goetghebeur@biomedcom.org)

Version: 2 Date: 30 August 2011

Author's response to reviews: see over
August 31st, 2011

Dr Hindrik Vondeling
Editor
BMC Health Services Research
e-mail: h.vondeling@utwente.nl

Dear Dr Vondeling,

Re: Bridging health technology assessment (HTA) with multicriteria decision analyses (MCDA): field testing of the EVIDEM framework for coverage decisions by a public payer in Canada

Attached please find the revised above-mentioned article to be considered for publication in BMC Health Services Research. We thank the reviewers for their very useful comments, giving us the opportunity to clarify and focus the scope of the work. Based on comments of the reviewers, we have extensively revised the manuscript. Detailed response to reviewers is attached with this letter. In addition, terminology was harmonized across the manuscript to facilitate understanding.

The manuscript was prepared according to the “Instructions for Authors”. All authors have given necessary attention to ensure the integrity of the work and seen and approved the final manuscript. No sources of funding were used to conduct this study.

We hope that the revised manuscript will be more suitable for publication in BMC Health Services Research, reporting on a field testing of a pragmatic multipurpose adaptable framework to support healthcare decisionmaking

I look forward to hearing from you.

Sincerely yours,

Mireille M Goetghebeur PhD

Vice-president Operations
BioMedCom Consultants inc., Montreal
mireille.goetghebeur@biomedcom.org
President, EVIDEM Collaboration
Reviewer’s report
Title: Bridging health technology assessment (HTA) with multicriteria decision analyses (MCDA): field testing of the EVIDEM framework for coverage decisions by a public payer in Canada
Version: 1 Date: 13 December 2010

Reviewer 1: Janine van Til

Reviewer’s report:
I would like to compliment the authors on a well written manuscript which discusses a relevant and emerging topic in decisionmaking and support systems. In general I would like to point out that the relevance of the current article is in the methodology, and not in the specific case. This means that I would like to see more attention to general properties of decision support framework and especially the importance of and outcome of this study with regard to the test-retest properties of the framework. Less attention can be given to the specific case, which would favor the readability of the manuscript, especially for less informed readers. Also, the length of the article can be reduced by focusing on the main objective.

Introduction
Page7. Line2. It’s not clear to me what the authors mean by “rooted in the natural thinking process”. It seems contra dictionary to the earlier mentioned criticism on MCDA being “not intuitive”, and much is still unclear about what the natural thinking process entails (Minor Essential Revision, MER).

Reply: The above statement was removed to avoid confusion.

Page7, last paragraph. Much relevant work related to the EVIDEM concept is mentioned. However, given these studies, the additional value of this particular study as stated in the objective is not clear/relevant. In my opinion, the test-retest properties of the framework are of paramount importance, and I would suggest emphasizing the concept and importance of stable preferences related to the goal of the framework in this paragraph and relating the objective to these outcomes of the study (Major Compulsory Revision, MCR).

Reply: The objective was restated and is now saying: “The objective of this study was to field-test a decision support framework (EVIDEM), explore its utility to a drug advisory committee and test the stability of estimates over time using tramadol for chronic non-cancer pain (CNCP) as a case study relevant to their context”. The other sections of the manuscript were adapted towards that goal including the abstract, the methods, results and discussion.

Methods
Although I find the information about data retrieval and presentation important to emphasize the scientific rigor of the decision process, I feel the amount of information can be cut back in favor of a more detailed explanation of the EVIDEM support structure and process.

Reply: This information was cut back as requested by the reviewer in both the methods and results sections, and the report is now an appendix.

F.i. what where the instructions given to the panel? Were the panel members allowed to deliberate, and was consensus sought? When did panel members read the information in Tikiwiki? Did panel members discuss the qualitative information among themselves, how much time was allowed for individual estimation and the whole process (MER)?

Reply: More details were added to explain the EVIDEM process and the structure of this case study. The methods section now states: “To explore individual values, during the workshop session (test), each member of the committee (n=9) were instructed to assign weight individually (on a scale of 1–low to 5–high) to each criterion of decision in the MCDA Core Model, from their perspective in the context of the health plan. For consistency across interventions, committee members were instructed to attribute these
weights independently of the intervention; these weights are expected to be defined once and then used throughout appraisals.

Time was allotted on an as-needed basis. This was followed by a period of discussion on each criterion, and committee members were allowed to modify their weights, on an individual basis.

To appraise the intervention, committee members were instructed to score (on a scale of 0 to 3) individually each criterion of the MCDA Core Model, using evidence synthesized for each of them (by-criterion HTA report). This was followed by a period of discussion on each criterion, and committee members were allowed to modify their scores, on an individual basis.

Committee members then explored the six contextual criteria and assigned the type of impact (negative, none or positive) each criterion would have on the appraisal of tramadol, using the colloquial and scientific evidence integrated into the Contextual Tool.

Feedback on the framework, criteria included and process was collected during discussion periods at the first workshop and at a follow-up workshop, and from a questionnaire administered during the follow-up workshop. To explore reliability, a retest was performed at least two weeks after the last session either on-line using the web-based prototype or a hardcopy document.”

Results
Page 13, line 1. Why independent of the specific case? Please state the relevance in the methods (Discretionary Revision, DR).

Reply: It is now stated in the methods section that: “For consistency across interventions, committee members were instructed to attribute these weights independently of the intervention; these weights are expected to be defined once and then used throughout appraisals.”

Are the weights significantly different from each other? In other words, what is the added value of criteria weights compared to equal weights (can also be addressed in the discussion) (Discretionary Revision, DR).

Reply: This is now addressed in the discussion which states: “Although at the individual level, relative weights varies largely between criteria, once weights are averaged at the committee level, extremes disappear and weights tend to be less distinguishable. Exploration of how the weight elicitation method (e.g., analytical hierarchy process [AHP], Simple Multi-Attribute Rating Technique [SMART], point allocation, ranking – with a new reference to the work of Jim Dolan (ref 71)), impacts weight attribution and the overall MCDA estimate is ongoing to further advance the approach and provide additional tools to adapt the framework to the preferences and needs of users.

Is the discussion on the relevance of disease severity as a relevant criterion in accordance with the high weight of 4 (low SD) in the evaluation (MER)?

Reply: Disease severity was considered an important criteria, but if always severe, then it was argued that there is no need to take it into consideration. The sentence in the text was modified and now states: “Although the criteria “Disease severity” was considered important by the committee (weight: 4.0±0), the necessity of this criterion was discussed; some committee members noted that because of the type of conditions covered by the health plan (i.e., work-related illness or injury), scores for criterion “Disease severity” may always be high in the committee context. However, it was also noted that some of the population covered by the health plan suffer from really severe diseases and that the framework would capture this aspect”

Page 14. Health Technology report is –in my opinion- not relevant in relation to the objective of the study. Maybe provide in attachment or online (MER)?

Reply: The “by criterion HTA report” is now provided as appendix 1 of the manuscript.

Page 21. How can it be that scores are more consistent than weights? Please explore possible explanations (DR).
Reply: Possible explanations are now provided in the discussion which states: “Higher consistency with scores, which are based on evidence, than with weights, which are based on individual perspective, were observed in this study and may reveal the difficulty to explicit one’s perspective. This might stem from the fact that perspectives are often implicit and that decisionmakers need to reflect on the implications of the criteria.”

Discussion
First paragraph: Please revise first paragraph of discussion according to higher relevance to test-retest properties as proposed as study objective (MCR).

Reply: Sentences describing the quantitative appraisal of tramadol have been removed to focus on the study objectives.

Second paragraph: HTA positioning

Third paragraph: Evidence

Fourth paragraph: Relevance of first two sentences is questioned (DR). Statement on source of data to third paragraph and omit rest of paragraph?

Reply: The paragraph was removed and the statement on the limitations of data was moved to the last paragraph of the Discussion dealing with limitations of the study.

Fifth paragraph: Weighting and scoring strategy. Please explore following issues:
Reasons for high SD in some of criteria. Are the criteria definitions clear to panel?

Reply: Possible explanation are now included in the discussion which states: “The weighting exercise and following discussion revealed the different perspectives of committee members as captured by the large SDs for some criterion. Variations may also be due to different understanding of the criteria, although detailed definition were provided and were further clarified during discussions.

What are the implications for total estimates (high, low, variation) for decisionmaking (MER)? The statement on the lack of reference points is the most important statement of this paragraph!

Reply: Implications of estimates are to be able to rank interventions based on a thorough evaluation to facilitate priority setting and justify decisionmaking. This is now clarified in discussion which states: “Indeed, the MCDA value estimates are meant to be used in a comparative manner for ranking healthcare interventions, which was beyond the scope of this case study. An MCDA model adapted from the EVIDEM framework by a district health board uses such approach to systematically evaluate and prioritize a wide range of healthcare interventions (Sharon Kletchko, MD, personal communication 2011).”

Sixth paragraph: Limitations. Importance of panel vs. online testing (DR)?

Reply: The framework is designed to support the deliberation process that takes place during committee meetings while the online prototype facilitates access to information at different levels of details. The importance of online versus panel testing was not evaluated in this study.

Level of interest: An article of importance in its field
Quality of written English: Acceptable
Statistical review: No, the manuscript does not need to be seen by a statistician.
Reviewer 2

Reviewer's report

Title: Bridging health technology assessment (HTA) with multicriteria decision analyses (MCDA): field testing of the EVIDEM framework for coverage decisions by a public payer in Canada

Version: 1 Date: 9 May 2011

Reviewer: Paule Poulin

Reviewer's report:

This objective of this study was to field-test a decision support framework (EVIDEM) and explore its utility for a drug advisory committee. Essentially there appear to be three Field test components to the paper’s argument:

1. Weight (weighting the importance of each decision criterion)
2. Score to appraise Tramadol on each criterion followed by the calculation of a “Quantitative Appraisal Score” and a systematic consideration of qualitative impacts of contextual decision criteria on the appraisal
3. Utility and validity of the framework was explored using discussion, survey and test-retest

Arguments are presented to show that the committee found the framework useful in supporting the systematic consideration of a broad range of issues when appraising technologies, and to show that the framework is good to lead group discussions and ensure transparent and consistent consideration of important elements that may affect decision making. This is very important and well demonstrated.

The paper presented an HTA report tailored to each decision criterion regarding Tramadol – Synthesized info for each criterion is important and the lack of available data for many of them is also well documented. This is good because it clearly documents the lack of data and this may be helpful in describing gaps in our knowledge. This is excellent and well demonstrated.

The paper makes these arguments in an articulate manner, drawing on the literature and an illustrative example. The paper makes a useful contribution to the need for the development of practical framework to help bring evidence along with others factors that needs to be considered when evaluating technologies (drugs in this case) for funding consideration. Overall paper well written and the information provided useful.

I have a number of comments and suggestions which translate into a series of recommendations, some optional, others considered compulsory. These are:

1. The main difficulty I have with this paper is the “Quantitative Appraisal Score” which has limited appeal. It attempts to squish many dimensions into a single number. People have a lot of trouble with costs per quality. I struggle with these, even though they are simple ratios. I'm always apprehensive of attempts to compress complicated information into a single number. Quantitative appraisal of Tramadol has an estimate of 0.44 (normalized Weight x Score) –what does that mean for decision makers? Are they able to interpret such number? The Authors mentioned that there is no reference point so interpretation is difficult, but even if there was some reference points, are such numbers useful? Would there be other analytical approach that may be more helpful? For this issue I recommend that a statistician be consulted to investigate alternative quantitative analytical approach.

Reply: The implications of estimates are to be able to rank interventions based on a thorough evaluation to facilitate priority setting and justify decision making. This is now clarified in discussion which states: “Indeed, the MCDA value estimates are meant to be used in a comparative manner for ranking healthcare interventions, which was beyond the scope of this case study. An MCDA model adapted from the EVIDEM framework by a district health board uses such approach to systematically evaluate and prioritize a wide range of healthcare interventions (Sharon Kletchko, MD, personal communication 2011).”

The manuscript was evaluated by a statistician who mentioned that several methods could be used to generate weights and value estimates with this framework, which is currently explored by a research group, but was beyond the scope of this case study. This is now included in the Discussion which states: “Exploration of how the weight elicitation method (e.g., analytical hierarchy process [AHP], Simple Multi-Attribute Rating Technique [SMART], point allocation, ranking), impacts weight attribution and the overall MCDA value estimate is ongoing to further advance the approach and provide additional tools to adapt the framework to the preferences and needs of users.”
2. The second difficulty I have with this paper is the claim that the framework is useful in making decision. While the Field-Test of EVIDEM was well carried out, and clearly demonstrated that the framework promoted the explicit consideration of a wide range of criteria by decision makers, because there was no “decision” reported from the case study, the authors cannot claim that the EVIDEM framework supported decision making in practice. It would be more accurate if the authors stated that the EVIDEM framework supported the evaluation of a technology in practice. Alternatively, the authors could present the decision made with regards to the Tramadol case study and how has the framework helped in making that decision.

*Reply:* The text was revised throughout the manuscript to shift the emphasis on the evaluative process rather than ultimate decision.

Minor essential amendments

1. A more critical treatment of the “Quantitative Appraisal Score” including a discussion of potentially more appropriate analytical approach (with the help of a statistician) or the removal of the quantitative analysis.

*Reply:* As mentioned above, the implications of estimates are to be able to rank interventions based on a thorough evaluation to facilitate priority setting and justify decisionmaking. This is now clarified in discussion which states: “Indeed, the MCDA value estimates are meant to be used in a comparative manner for ranking healthcare interventions, which was beyond the scope of this case study. An MCDA model adapted from the EVIDEM framework by a district health board uses such approach to systematically evaluate and prioritize a wide range of healthcare interventions (Sharon Kletchko, MD, personal communication 2011).” The use of other weighting methods and exploration of scoring scales and ranking is currently ongoing with a research group including a statistician to further advance the approach and provide additional tools to adapt the framework to the preferences and needs of users.

2. Removal of the claim that the framework was useful in supporting decision making unless the decision is presented in the context of the framework.

*Reply:* The text was revised throughout the manuscript to shift the emphasis on the evaluative process rather than ultimate decision.

3. Tramadol was selected by the drug advisory committee as a good case study–Why?

*Reply:* Tramadol is relevant to the population covered by the WSIB. This is now clarified in the text which states: “Tramadol for CNCP was selected by the committee as a relevant case study to the context of the covered population by the WSIB”

4. Background 3rd para: The authors stated “if kept simple, it facilitates….”, My question is: Did the Committee find the process simple? How long did the Committee take to go through the weighting and scoring process? The authors reported that there was a concern about developing HTA report but was there an issue with the time involved in using the framework to make a decision? Was there a decision made? It would be useful to have a discussion on this point.

*Reply:* MCDA can be very complex and we believe it is important to keep it as simple as possible for new users. This is developed further in the discussion which states: the framework was built to use a simple approach. More information about the process is now provided in the methods section. Although we did not measure the appraisal time in this study, we did so in another study in which 13 panelists evaluated 10 drugs and it took on average 31 minutes [SD 15] for scoring each drug (Goetghebeur 2011 (ref 23).
5. Also, it would be useful for the audience to know if the authors have done work to describe whether or not the criteria overlap? If they overlap, this would indicate one limitation of the framework.

*Reply:* The framework was designed taking into consideration all the requirements of MCDA, including non-redundancy which was extensively explored at the design stage, which is stated in the methodology paper that is referred to in the manuscript (ref 21)

The order and the headings of the items described in the methods and results and discussion section is not always consistent. Revision for consistency of presentation would be useful.

*Reply:* The order and the headings of the results section is now fully aligned with that of the methods section. The discussion section was also reorganized to follow more closely the methods and results sections.

Discussion 5th para: While the authors describe that (…MCDA value estimates obtained by applying this framework are committee-specific….) it would also be informative how much similarity or differences there may be between different committees that have field-tested the framework.

*Reply:* Unfortunately, we do not have comparative data for tramadol across several committees that we could report in the manuscript.

Health Technology Assessment Report: I am not sure if describing how the reports were done is relevant to this manuscript unless the decision made is also reported.

*Reply:* The HTA report methodology is an integral part of the EVIDEM framework and we felt that it should be described to emphasize the scientific rigor of the process but we have cut back on the extent of the description to address the reviewer comment, in both the methods and results sections, and the report is now an appendix.

The authors should be careful in their use of the term “Value”. It has different meaning in different line of the manuscript – sometimes it represent the “Weights” (the importance of a criterion), sometimes it represents the calculated MCDA Matrix Estimate.

*Reply:* To limit confusion, we are now referring to “individual perspective” when committee members are assigning weights and to “value estimate” when reporting on the MCDA estimation.

Result, 2nd paragraph: MCDA value estimate paragraph may be best as part of the discussion.

*Reply:* The feedback from the committee was reported as results and the discussion section was expanded to discuss the use of the MCDA estimates and the limitations of this case study.

Results: I was not able to evaluate the validity of the HTA report and appraisal of tramadol for chronic non-cancer pain section. I believe this section should be reviewed by an expert in the field.

*Reply:* The report was reviewed by one of the author who is a clinician practicing in the field of rehabilitation and pain management.

**Level of interest:** An article whose findings are important to those with closely related research interests.

**Quality of written English:** Good

**Statistical review:** Yes, the manuscript needs to be seen by a statistician

*Reply:* The document was revised by a statistician. Following her recommendations:

* we are now reporting only ICC (3,1) which is the most appropriate for this type of analysis;
we developed the discussion section on testing other methods to establish the MCDA estimate and added a reference to the work of Jim Dolan (ref 71)
we added a reference to MCDA methods and linear models. (ref 30)

Quality of written English: Acceptable
Statistical review: Yes, but I do not feel adequately qualified to assess the statistics.
Declaration of competing interests: 'I declare that I have no competing interests'