Author's response to reviews

Title: Quantitative summaries of treatment effect estimates obtained with network meta-analysis of survival curves to inform decision-making

Authors:

    Shannon Cope (scope@mapigroup.com)
    Jeroen P Jansen (jjansen@mapigroup.com)

Version: 2
Date: 21 October 2013

Author's response to reviews: see over
Oct 20, 2013

Dear Editor-in-Chief,

Enclosed is a revised manuscript entitled: “QUANTITATIVE SUMMARIES OF TREATMENT EFFECT ESTIMATES OBTAINED WITH NETWORK META-ANALYSIS OF SURVIVAL CURVES TO INFORM DECISION-MAKING”.

This manuscript presents alternative approaches to present rank probabilities in the context of a Bayesian network meta-analysis of parametric survival curves. The advantages and disadvantages of alternative treatment effects are explored in terms of an illustrative example in advanced melanoma using a fractional polynomial network meta-analysis of overall survival.

The models and methodology used in the research are not proprietary. This work has not been previously published, nor is it under consideration for publication elsewhere.

This research project was performed without specific funding.

The authors have revised the manuscript in response to the reviewer comments and have provided responses to each reviewer comment below.

This re-submission includes following files as Microsoft Word 97-2003 Documents (.doc):
1) Quantitative summaries of treatment effect estimates obtained with NMA survival curves_track_changes;
2) Quantitative summaries of treatment effect estimates obtained with NMA survival curves;
3) Figure 2_revised;
4) Appendix

Please address all correspondence to:

Shannon Cope, MSc
Associate Director
Mapi Group, HEOR and Strategic Market Access
33 Bloor Street East, Suite 1300
Toronto, ON
M4W 3H1 Canada
scope@mapigroup.com

I look forward to receiving your feedback.
Sincerely yours,

Shannon Cope
Author response to Referee 1 Major Compulsory Revisions:
The authors agree that the methods to synthesize the survival (or time to event) outcome applied in this study are not novel. Traditionally, hazard ratios are pooled (and indirectly compared) in order to assess the comparative efficacy of alternative treatments in a (network) meta-analysis, which assumes that hazard ratios are constant over time (i.e. ‘time independent measures’). Over the last few years, synthesis methods have been developed which do not rely on this assumption, which can be broadly classified as ‘time varying measures’ (Ouwens et al. 2010, Jansen 2011). This study applies one of these methods in order to illustrate the challenges of assessing the rank probabilities within the context of a time-varying result, and to discuss the advantages and disadvantages of alternative methods to illustrate these results. Therefore, what is novel about this study is the discussion regarding rank probabilities in the context of time-varying treatment effects. In other words, the alternative approaches to present rank information presented by Salanti et al. 2011 for time-independent treatment effects were extended to time-varying treatment effects in our study.

As with any meta-analysis the difference between the treatments is analyzed in order to preserve the randomization of patients within a trial. In our study, the treatment effect is indeed represented by the hazard ratio. However, this method is not based on a contrast based method but rather an arm based method. The data used for this analysis is derived from the survival percentages extracted from the Kaplan Meier curves for each treatment of each RCT. The underlying (ln) hazard functions are then derived for each treatment comparison as a function of the survival percentages (assuming a particular fractional polynomial model). Finally the ln(hazard ratios) are estimated based on the ratio of the ln(hazards) within each study.

Page 5 in Methods has been revised to clarify that an arm-based model was used:
A binomial likelihood distribution was used for the incident number of deaths for every two month interval, which was calculated based on the survival percentages from the Kaplan-Meier curves and the number of patients at risk at the beginning of the interval in each arm of each study, assuming a constant hazard rate within each interval.

The authors do not understand the Reviewer’s last question ‘Have the author considered about some direct measures regarding reporting the proper summaries?’. Clarification regarding this question would be appreciated.

Minor Essential Revisions

Abstract

1. The author mentioned rank probabilities associated with the alternative treatments can be obtained, but why do you use “however”?

Author response to Referee 1 abstract comment #1:
This sentence has been revised to exclude ‘however’.

In the Abstract, Background section the following sentence will be updated as follows:

In the context of survival analysis multiple treatment effect measures are available to inform the rank probabilities.

2. Please check the grammar in the “Methods” in Abstract.

Author response to Referee 1 abstract comment #2:
The Abstract, Methods section was revised as follows to improve the grammar:

A fractional polynomial NMA of overall survival in advanced melanoma was performed as an illustrative example. Rank probabilities were calculated and presented for the following effect measures: 1) median survival; 2) expected survival; 3) the mean survival at the follow-up time point of the trial with shortest follow-up; 4) hazard or hazard ratio over time; 5) the cumulative hazard or survival proportions over time; and 6) mean survival at subsequent time points. The advantages and disadvantages of the alternatives measures were discussed.

3. The author mentioned that “Since the NMA of survival curves can result in hazard and survival estimates that vary over time for the compared interventions, calculations of rank probabilities vary with the different effect measure”. If it is not NMA, there will not be hazard varying over time?

Author response to Referee 1 abstract comment #3:
This sentence in the abstract has been revised to clarify the meaning as follows:
Since hazard and survival estimates may vary over time for the compared interventions, calculations of rank probabilities for an NMA of survival curves may depend on the effect measure.

Background

4. In the second paragraph of the BACKGROUND, how could p-values reflect that there is no difference between treatments A and B? Please rewrite this sentence. And also “the risk of an unsupported positive interpretation” and “the risk of overlooking a true difference” are the same. Don’t need to use not only but also.

Author response to Referee 1 background comment #4:
In the Background section, the second paragraph the following sentence has been revised as follows to clarify the interpretation of p values in a frequentist framework:

 Therefore the p value reflects the probability of observing such a treatment difference assuming the null hypothesis is true.

Also, in the background section in the second paragraph, the following sentence was revised as suggested:

However, decision-makers are interested in minimizing the risk of an unsupported positive interpretation and the risk of overlooking a true difference.

5. In the fourth paragraph, check the grammar of the last sentence.
Graphical and numerical summaries of rank probabilities

Author response to Referee 1 background comment #5:
The authors did not identify a grammatical mistake in the title ‘Graphical and numerical summaries of rank probabilities.’ Please can you clarify the specific grammatical issue?

6. Check the grammar in the second paragraph.
Advantages and disadvantages of different effect measures in relation to ranking treatment ranking

Author response to Referee 1 background comment #6:
This title will be amended as follows:

ADVANTAGES AND DISADVANTAGES OF DIFFERENT EFFECT MEASURES IN RELATION TO TREATMENT RANKING

7. This part is overelaborated and some of them are not necessary. Please be concise.

Author response to Referee 1 background comment #7:
The section on the advantages and disadvantages of different effect measures in relation to treatment ranking has been revised to be more concise.
8. The difference between different measures among either the “time independent measures” or the “time varying measures” is also well known.

**Author response to Referee 1 background comment #8:**
This section has also been revised to be more concise.

9. But I like this part “probability of being best treatment, rankograms, or SUCRA”. And I think it is not restricted to NMA of survival curves. It can be NMA of others.

**Author response to Referee 1 background comment #9:**
The authors agree that these methods are not restricted to NMA of survival curves. These methods of presenting the rank probabilities have been presented previously for the NMA of time-independent outcomes. Therefore, this study focuses on the application of these methods to time-dependent outcomes.

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

**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

---

**REFEREE 2:**

**Reviewer's report**

**Title:** Quantitative summaries of treatment effect estimates obtained with network meta-analysis of survival curves to inform decision-making

**Version:** 1 **Date:** 30 August 2013

**Reviewer:** Jason Madan

**Reviewer's report:**
I found this paper to be generally well-written and interesting. I have some suggestions that I believe necessary to clarify the methods presented and strengthen the discussion and interpretation of results.

**Major compulsory revisions**

1) The authors use an approach to the network meta-analysis of survival data that they developed and presented in previous work. I believe the approach involves a Weibull survival model in which there are two jointly distributed treatment effects acting on the shape and scale of the Weibull distribution. However I cannot be sure, because the underlying model for the network meta-analysis is not presented in this paper, and it should be available without the need for readers to refer back to the authors' previous work.

**Author response to Referee 2 Major Compulsory Revision 1:**
The underlying model has been presented in an appendix of the revision.

2) I would have found it informative to see mean and credible interval estimates for all three of the univariate outcome measures (median survival, mean survival up to 22
months, extrapolated mean survival). The authors only present mean estimates (and no credible intervals) for median survival. I had a general view that the presentation and interpretation of results on pages 7-13 could be improved and organised better. This is, I accept, a vague criticism, and I give a few thoughts on how this might be achieved below.

**Author response to Referee 2 Major Compulsory Revision 2:**
An additional table (Table 1) has been added to summarize the mean survival at follow-up and expected survival (including the 95% credible intervals).

3) There should be a clear separation between a section on results, in which findings are presented objectively, and a separate section with discussion/interpretation of those results.

4) I would like to see a more nuanced discussion of the merits of the different approaches presented in the paper. To me, it seems that the authors have a clear opinion on how ranking information should be presented, and their discussion is a little one-sided, not presenting counter-arguments in sufficient detail. For example, they have a clear preference against univariate measures, which they argue 'do not provide the wealth of information captured by rank probabilities over time'. It isn't clear to me what figures 5 and 7 add that I can't get more easily from figures 3 and 4. This may depend on what the reader is trying to get out of the figures, which brings me to a final point:

5) Present a clearer discussion of what the purpose of all these graphs is, and how they help achieve that purpose. The 'best' way to present the information produced in these analysis depends on the user, and it would be useful to discuss who might find the various figures useful, how they should interpret them, and what they could use them for.

**Author response to Referee 2 Major Compulsory Revision 3, 4, and 5:**
The sections have been revised to present: 1) Rankograms for the illustrative example (explaining Figures 4 and 5); 2) Probability of being best treatment and SUCRA for the illustrative example (explaining Figures 6 and 7); and 3) Advantages and disadvantages of different effect measures in relation to treatment ranking.

The section evaluating the advantages and disadvantages of different effect measures in relation to treatment ranking has been revised to be more concise and more balanced. Within this section, the subsection ‘Probability of being the best treatment, rankograms, or SUCRA?’ has been updated to provide guidance regarding the alternative graphs.

**Discretionary Revisions**

For the univariate outcome measures, it should be possible to fit a network meta-analysis directly, without making any assumptions about survival time distributions. It
would be interesting to compare rankings derived using this method to the more complex approach taken by the authors, although it is not an essential addition.

**Author response to Referee 2 Discretionary Revisions:**
The authors agree it is possible to assess alternative network meta-analysis models. For example, if the proportional hazards assumption is valid, it may be possible to pool and indirectly compare the treatments of interest based on a reported hazard ratio for example. However, for the purpose of this study, a comparison between alternative network meta-analysis models may make it difficult to understand what is driving the differences in the treatment ranking probabilities (i.e. differences due to models versus differences due to other factors). Therefore, the authors prefer to focus on the current objective to avoid overcomplicating the study.

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

**Quality of written English:** Acceptable

**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.

**REFEREE 3:**

**Reviewer's report**

- **Title:** Quantitative summaries of treatment effect estimates obtained with network meta-analysis of survival curves to inform decision-making
- **Version:** 1  **Date:** 4 September 2013
- **Reviewer:** Cinzia Del Giovane

**Reviewer's report:**
The paper provides a summary of different alternative approaches to summarize and present results of a network meta-analysis (NMA) (in a Bayesian context) of survival data based on parametric curves with rank probabilities. The different effect measures in relation to ranking the treatment ranking can be divided into time independent measures and time-varying measures.

The manuscript is innovative, interesting and well-structured. The network meta-analysis methodology is a recent statistical method used to synthesize evidence from RCTs for multiple intervention and it increasingly used in the medical literature. One of the potentiality of the NMA, when it is performed in a Bayesian framework, is that a ranking of the treatments included in the network can be calculated according to their effectiveness. Such possibility facilitates decision-making. However many and more methodological aspects still have to be investigated. I guess that in this paper the authors explored one of these aspects when survival data are analyzed using NMA with appropriated methods and conclusions well supported by the data. Therefore for me only 1 minor essential revision should be done:

1. report that in the legend of the figure 2B and in the document text confidence intervals are showed.

**Author response to Referee 3 Minor Essential Revision:**
The legend for Figure 2B has been revised to clarify that the dotted lines reflect the 95% credible intervals and the text has been revised on page 6 for the corresponding description.

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

Quality of written English: Acceptable

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.