**Author’s response to reviews**

**Title:** BUGSnet: An R Package to Facilitate the Conduct and Reporting of Bayesian Network Meta-Analyses

**Authors:**

Audrey Béliveau (audrey.beliveau@uwaterloo.ca)

Devon Boyne (devon.boyne1@ucalgary.ca)

Justin Slater (jslater@lighthouseoutcomes.com)

Darren Brenner (darren.brenner@ucalgary.ca)

Paul Arora (parora@lighthouseoutcomes.com)

**Version:** 1  **Date:** 31 Jul 2019

**Author’s response to reviews:**

Dear editor and reviewers,

Please find attached our revised manuscript, with tracked changes and a point-by-point response to the reviewers.

We thank the reviewers for their thoughtful feedback on our manuscript and for providing us with the opportunity to enhance our work. A response to each of the reviewer’s comments is provided below.

Please note that we have released a new version of BUGSnet (1.0.1) which has a few improvements (see file “WHATS NEW” for details). You may access those at the following link for the purpose of this review:

https://www.dropbox.com/sh/9oqqrcc4a2kv8ex/AABMidPtBPC_Fx4uMvgVyQpZa?dl=0

Hans-Peter Piepho (Reviewer 1)

Comment

(1) The models are denoted as generalized linear model. Given that two versions of NMA model are considered, i.e. fixed-effects and random-effects models, the term "generalized linear mixed models" (GLMM) may be more appropriate.
Response

We have chosen to use the term “generalized linear model” to be consistent with the NICE Technical Support Document 2 and the Network Meta-Analysis for Decision-Making book by Dias et al.

Comment

(2) The authors focus on contrast-based models. These are surely the most commonly used ones, but recently there has been some debate comparing arm-based and contrast-based models. Can the authors comment on options to also implement arm-based models? Such models may be easier to understand for researchers not yet familiar with NMA, but having some grasp of two-way analysis of variance (ANOVA) and two-way mixed models.

Response

At the moment, only contrast-based models have been implemented in BUGSnet. We decided to focus on contrast-based models because the use of arm-based models is controversial within the NMA community. To quote the Dias et al. book: “…[arm-based models] are not well suited to the current practice in decision making, because they would oblige investigators to use the same trial data to inform both the baseline model and the relative treatment effects.”

Arm-based models will be considered in future developments of BUGSnet. We have added the following sentence to the discussion: “At this point, arm-based models have not been implemented in BUGSnet; the R package pcnetmeta allows such analyses, although it does not readily provide a complete suite of outputs like BUGSnet.”

Comment

(3) On page 9 in line 21, the authors impose a non-negativity constraint on the linear predictor theta for binomial distributions. I do not think such a constraint is needed on the linear predictor with GLMMs. In fact, that's the whole idea of GLMMs: use a suitable link function so the possibly constrained range for the response is projected onto the whole real line using a suitable link function such that the linear predictor does not need to be constrained in any way. The authors use the log link for binomial data, which is unusual and has the disadvantage that binomial probability parameter is not constrained above as it should be. Other links are more suitable here, e.g. logit and probit. so I think the authors need to reconsider their general approach to usage of GLMMs for NMA.

Response

First, let us clarify that we have implemented the logit link for binomial outcome within BUGSnet (see Table 2). The non-negativity constraint was implemented only for the log link. This constraint is necessary when using the log link to “ensure the probabilities of an event in
each arm of each study remain between zero and one.“ (Network Meta-Analysis for Decision-Making book by Dias et al., p.53). The log link with binomial outcome is implemented in the gemtc package so we have implemented it in BUGSnet to offer the same capabilities.

(4) On page 10, line 32 the authors say they use vague priors on the parameters. What about variance parameters? It is not at all obvious what are vague priors for variances, so some comment seems in order here.

Our choice of automatic priors was based on the justifications made by van Valkenhoef et al. (2012). Those priors are described precisely in Table 3. Those are the same priors that are implemented in gemtc. We have added the following sentence to clarify:

“The user also has the option within the nma.model function to specify their own prior which is useful for conducting sensitivity analyses, namely for the comparison of prior distributions on the random effects standard deviation, σ, to insure that they do not have a significant effect on the posterior estimates.”

Comment

(5) The authors use the terms "consistency" and "transitivity". It is not clear from the text whether these terms can be used interchangeably and mean the same thing. Some explanation would be useful, and if the terms are indeed exchangeable it is perhaps better to stick to one term only.

Response

Although related, these two terms are not interchangeable. Transitivity is an unverifiable assumption that underlies network meta-analysis. One way to detect a violation of transitivity is by assessing consistency which refers to agreement between the indirect and direct evidence in the network. A lack of consistency indicates a violation of the transitivity assumption. However, the transitivity assumption can be violated despite the presence of consistency in the network. While we can not prove that the assumption of transitivity holds in a particular analysis, we can provide evidence that would suggest a violation of this assumption and one method of doing so is by examining consistency within the network. Another method of detecting a potential violation of transitivity is by detecting heterogeneity of effect modifiers across comparisons within the network. To further highlight the distinction between transitivity and the empirical methods used to assess it, the following sentence has been added to provide readers with a definition of transitivity that can be contrasted with the definition of consistency that follows:

“A fundamental assumption of an NMA is the assumption of transitivity. [2] Under this assumption, one assumes that one can estimate the difference in the effect of two treatments by subtracting the difference in the effects of the two treatments relative to a common comparator as follows: d_((t_i1,t_ik))=d_((1,t_ik))-d_((1,t_i1)). [2]”
Comment
(6) Page 12, line 4: Perhaps "Output" is a more suitable headline here.
Response
Changed.

Comment
(7) Figure 5: I would suggest to add text to the caption that explains what's shown in the cells of the table.
Response
Added.

Comment
(8) Figure 6: A sentence could be added to the caption to detail what a point in the graph represents.
Response
Added.

Satoshi Yokoyama (Reviewer 2)
Suggestions
1. The authors provided two of the significance of this paper: user friendly software and outputs for reporting guidelines. In particular, the former is not a point that can maximize the novelty of this paper, but the package is an attractive tool that facilitates good-quality reporting based on guidelines. To emphasize that again, the authors should consider how to report the output rather than just introducing package usage and analysis techniques.
Response
Additions: Detailed vignettes presenting how to use BUGSnet to conduct, interpret, and report the results from an NMA are available within the R package. We have included one of the vignettes as a supplemental file and have added the following sentence to highlight these vignettes:
“The BUGSnet code used to produce this analysis is available in the vignette titled survival in the BUGSnet documentation, and appended as a supplement to this article. Additional outputs are presented in the vignette as well as a more detailed description of how to conduct and report network meta-analysis, which is only presented here in brief.”

Comment

1. In the introduction section, differences with existing analysis packages should be emphasized. I could find that description in the discussion section, but this needs to be written in the introduction too.

Response

Added.

Comment

2. In data preparation, it is better to introduce the format of the data to be read. When end-users try to use a new analysis package, they often get confused about creating a data set. Although the author explains column specifications in dataset, the descriptions of the original data set and format can help the reader.

Response

To clarify the format of the data, we have added the following sentence to the manuscript:

“The data is organized in the long format (i.e. one row per treatment arm), with variables indicating the study ID, the treatment ID, the number of patients, the number of events, and the mean age (and standard deviation) of participants for each treatment arm (see Table 4). “

We have also added Table 4 to show the data.

Comment

3. Some reporting guidelines for network-meta-analysis such as the PRISM-NMA recommend reporting additional analysis including a subgroup analysis a stratified analysis, which is important for NMAs with complex structures in especially. Although authors set the section for meta-regression, wherever possible, a detailed description of how to report about the results rather than mathematical knowledge is also desired. I think this is an important point, as this paper introduces an analysis package with a report based on the guidelines in mind.

Response
When doing a meta-regression, outputs such as SUCRA plots and league tables can still be used to present the results. The key difference is that a value for the covariate must be specified. We have added the following sentences to clarify:

“When conducting a meta-regression analysis, the output plots and tables described in the Output section (league heat plot, league table, etc.) can also be produced but the user will need to specify a value for the covariate at which to produce treatment comparisons. Those treatment comparisons are calculated internally within BUGSnet by computing posterior quantities of interest at a specific covariate value \( x^0 \) as 

\[
\beta_{((t_{i1},t_{ik}))} = \beta_{((1,t_{ik}))} - \beta_{((1,t_{i1}))}.
\]

Comment

4. Many methods have been proposed to assess inconsistency as the authors noted. In the current NMA with various indicators and methods for inconsistency, the guidelines recommend that at least two approaches are conducted: global approach and local approach. It is necessary to be clearly described in which position authors’ approach corresponds. Comprehensively expressing in the term “inconsistency” may lead to false reports using this analysis package.

Response

We have added the following sentence to clarify the distinction: “Such methods are often categorized as being “global” or “local” depending upon whether they examine inconsistency within the entire network or within particular segments thereof. [2]”

The inconsistency model allows exploration of both global and local inconsistency. We have added the following to clarify:

“To examine inconsistency at the global level, the fit of the inconsistency model can be compared against a model in which consistency is assumed using the nma.fit() function and comparing the DICs. Local inconsistency can be explored on the leverage plots produced by nma.fit() and also using the nma.compare() function which produces a plot comparing the posterior mean deviance of each data point between the consistency and the inconsistency model.”

We currently plan to implement additional methods including “local” methods such as the Bucher method. We chose to implement the inconsistency model method because it can account for dependencies within the network arising from trials with more than three arms and because it has a straightforward interpretation in the presence of a consistency violation. As noted in the NICE-DSU TSD 4, it is difficult to interpret results from several repeated “local” tests for inconsistency because one must adjust for multiple comparisons and there is currently no definite method for doing so within this context. The following has been added to our manuscript:
“More options for assessing inconsistency at both the global and local levels will be considered in further BUGSnet releases.”

Comment

5. For each function, materialize at least the specification of the arguments (e.g., chose the fixed/random model, burn-in, and link function) specified in this demonstration. If the authors intend to explain them on the web, please declare them along with the planned web links. Specifying arguments can be a frequent operation by users.

Response

See response to suggestion 1 (arguments specifications are given in the vignette, provided as Supplement to the manuscript).

Comment

6. Unlike two existing packages (NetMetaXL and GeMTC), the authors said that this analysis package can be apply to handle various outcomes. Is there any difference other than being a unified analysis package? For example, if data has a dichotomous outcome, will the BUGSnet and NetMetaXL output the same results? Please show these comparisons to ensure the credibility of the analysis results.

Response

BUGSnet, NetMetaXL and gemtc should output the same results. To check this, we have analysed the diabetes dataset using all three packages and obtained the same league tables. Those results are in line with those from the NICE-DSU TSD 2. The R code used for the BUGSnet and gemtc analyses as well as the outputs are provided as supplementary material.

The following has been added to the manuscript: “The results of our package are concordant with those reported in the TSD as well results obtained with GeMTC (code provided as supplement to this article) and NetMetaXL.”

Comment

7. In the discussion section, it is necessary to describe from the comparison with the existing analysis packages, whether the outputs needed to satisfy from this demonstration has been obtained, which also relates to the title in this paper. Otherwise, it can not be judged that the link between the result and the consideration in the paper is appropriate.

Response
We have heavily revised the introduction and discussion sections to better compare the capabilities of BUGSnet with NetMetaXL and GeMTC as follows:

Introduction: “BUGSnet is intended to be used by researchers when assessing the clinical efficacy of multiple treatments within the context of a submission to a journal or a health technology assessment agency. For conducting a contrast-based Bayesian NMA, the two main competing software packages that one may consider are GeMTC [22] and NetMetaXL. [17] While NetMetaXL does produce much of the output necessary to satisfy reporting guidelines, it is limited in the types of analyses it can carry out. Specifically, one cannot use NetMetaXL to analyze outcomes that are not dichotomous, to conduct meta-regression, or to analyzing evidence bases with more than 15 treatments. [17] While GeMTC provides an enhanced suite of functions for conducting NMA relative to NetMetaXL, its reporting capabilities are limited. For example, GeMTC does not readily produce key reporting items for an NMA such as tabular overview of the evidence base or a SUCRA plot and league table of the NMA results on the original scale.”

Discussion: “With BUGSnet, we aimed to create a single tool that would compete with the reporting capabilities of NetMetaXL and the analytic capabilities of GeMTC. We have also aimed to provide users with enhanced reporting options not included in existing software such as a function to produce bar graphs that show the distribution of effect modifiers by trial or by treatment arm and an option to print study names and highlight certain treatment comparisons within the network plot. To help facilitate the use of BUGSnet among new users, we have provided three vignettes (with more vignettes forthcoming) in the R help files that walk users through conducting an NMA using BUGSnet by providing detailed R code and interpretations of the statistical output. Despite these benefits, there are limitations of BUGSnet. BUGSnet is currently limited to exclusively analyzing arm-level data. In contrast, GeMTC can be used to conduct an NMA using entirely arm-level or entirely contrast-level data. [21] Relative to GeMTC, another limitation of BUGSnet is that GeMTC currently provides a broader range of methods of assessing inconsistency such as the node-splitting method and a broader range of meta-regression analyses such as subgroup meta-analysis. Since it is implemented within the R environment, some users may find BUGSnet more difficult to use relative to NetMetaXL, which is implemented within Microsoft Excel. We plan to address these shortcomings in future iterations of BUGSnet and interested users should check the previously mentioned URL for updates.”

Comment

8. Many readers due to make some effort in order to reach the usefulness of Figure 2. Please more annotations for figure 2. What I point out is not the addition of the interpretation of this figure. It is an explanation of individual bar captions and summarized elements.

Response

To better highlight the usefulness of the figure, we have added a detailed explanation to the corresponding paragraph in which it is introduced as follows:
“After the data was prepared using the data.prep() function, the net.plot() and the net.tab() functions were used to describe the network of studies in a graphical (Figure 1) and tabular format respectively (Table 5). As previously discussed, the assumptions of network meta-analysis will be violated when an effect modifier is heterogeneously distributed throughout an evidence base. [18] Prior to conducting the network meta-analysis, analysts can use the data.plot() function to examine the distribution of an effect modifier within the network. The determination of whether or not a variable is an effect modifier and whether or not the observed differences in its distribution are clinically meaningful is determined according to expert opinion and prior evidence. To demonstrate this function, we have simulated a patient characteristic that may modify the treatment effect (i.e. the age of participants). To mimic a lack of reporting, we have omitted the standard deviation for a few of the studies. As observed in Figure 2, the mean age of participants within each treatment arm (the individual bars) is similar to the overall mean age of participants within the evidence base (the red dotted line). According to the standard deviation (the +/- error bars), the variability of ages within each treatment arm appear to be similar as well (where available). Based on this analysis, one would conclude that there is no meaningful heterogeneity in the distribution of age. This analysis would be repeated for all potentially important effect modifiers identified a priori by clinical opinion and a review of previous studies. If no heterogeneity is detected, then one may proceed to conducting the network meta-analysis. If heterogeneity is detected, one can attempt to adjust for imbalances by using meta-regression (if there are an adequate number of studies) or by using alternative statistical techniques that leverage individual patient data (e.g. matching-adjusted indirect comparison or simulated treatment comparison). [18]”