Author’s response to reviews

Title: Comparative Efficacy of Six Therapies for Hypopharyngeal and Laryngeal Neoplasms: a Network Meta-analysis

Authors:

Juan Che (zyyanzhenguo@126.com)
Yanlin Wang (zyysht0707@126.com)
Xiaolin Zhang (zyyyyyjm@126.com)
Jun Chen (zhry960034@bzmc.edu)

Version: 1 Date: 26 May 2018

Author’s response to reviews:

Response to reviewers:

Dear reviewer, thank you very much for your careful comment on our manuscript. We have revised the manuscript and would like to re-submit it for your consideration. Point by point responses to your comments were listed below.

Reviewer #1:

This is a well written manuscript. The methods are appropriately described and conducted, and the conclusions carefully considered and justified.

Reply: Dear reviewer, thank you very much for your comment on our manuscript.

Reviewer #2: This is an interesting network meta-analysis analyzing the effects of 6 different management strategies for laryngeal/hypopharyngeal cancers. It is quite an ambitious undertaking. However, I do have some concerns about the study design as well as the applicability of this work to clinical practice.
Major issues:

-The populations of the included studies are quite heterogeneous. Stage IV hypopharyngeal cancer is treated very differently than stage I glottic laryngeal cancer, as are many of the stages in between, and the associated prognoses are also very different as well. Ranking all the management strategies 1 through 6 for each outcome seems challenging to do in a manner that is clinically meaningful for such a heterogeneous group of diseases and patients.

Reply: Dear reviewer, thank you very much for your careful comment on our manuscript. There are several possible sources of variability or heterogeneity among studies that are included in systematic reviews and network meta-analyses. Variability in the patients, the types or timing of outcome measurements, and intervention characteristics has been termed clinical heterogeneity. Influence of cancer type/study, stage distribution of each trial's population and other potentially confounding variables can be reduced by opting for a comprehensive search for evidence and selection criteria, to maximized data yield while avoiding heterogeneity and inconsistency. Subgroup and more sensitivity analyses are effective to make the result reliable. However, among all 28 studies we included in this network meta-analysis, there were only 8 studies that the patients of which were in the same stage of hypopharyngeal cancer, which make it difficult to do a subgroup analysis and maintain the reliability of the conclusion meanwhile. Also, a random-effect Bayesian model was utilized considering the existence of heterogeneity and the influence of this heterogeneity to the analysis is limited in appropriate range.

-How (and why) are hazard ratios being separately calculated for 3-year and 5-year outcomes? Typically these are done not for discrete time points but for the hazard function overall, so that OS is one outcome and DFS is another outcome. I am also confused why 5-year overall survival rate is listed as a separate outcome from overall survival as well, and so this should be clarified. With 5 outcome measures and 6 interventions (especially when they are all quite similar), the results and conclusions certainly become more difficult to follow as well.

Reply: Dear reviewer, thank you for your comments. To fully describe the outcome index of all treatments, it is necessary to list 3-year OS, 5-year OS, 3-year DFS, 5-year DFS separately in the research. The data are obtained via software Engauge to match the overall survival curve present in included article and the Hazard Ratio was calculated by Excel formula. 5-OSR is a different index from 5-OS. 5-year overall survival rate (5-OSR) is an index that shows the percentage of people in a group who are alive after 5 years. While 5-year overall survival (5-OS) is an index that takes time factor into consideration thus shows a change process. 5-OSR is a binary variable which is compared by OR (odds ratio) while 5-OS of two treatments are
compared by HR (hazard ratio). HR represents instantaneous risk over the study time period, or some subset thereof. Hazard ratios suffer somewhat less from selection bias with respect to the endpoints chosen and can indicate risks that happen before the endpoint.

-The actual endpoints provided by each study should be comprehensively listed in Table 1 so we can see exactly what the primary data showed. This would make for a very wide table but one that is necessary during any evaluation of a meta-analysis.

Reply: Dear reviewer, thank you for your comments. The complete table that showed actual endpoints of each study has been added. We have listed all data extracted from included articles in Table S1. Please check Table S1.

-Is there room for sensitivity analyses (even the standard ones used for meta-analyses, like meta-regression, or excluding studies published before a certain year or that only had certain stages of disease) to ensure robustness of the conclusions?

Reply: Dear reviewer, thank you for your comments. Sensitivity analyses are often used in meta-analysis to access the between-study heterogeneity, which is done by showing the influence of detaching one trial from the whole analysis. An important aspect in meta-analysis is to investigate statistical heterogeneity, which is also important in network meta-analysis. If heterogeneity exists, then the possible sources should be explored and implementation of random-effects modeling, sensitivity analyses, subgroup analyses should be considered if sufficient data are available.

Here in our study, we used heat plot to show the influence of detaching a pair of direct evidence to the network evidence, which play a role of sensitivity analysis in our network meta-analysis. In network meta-analysis inconsistency arises as another aspect of heterogeneity. The colors in heat plot are associated with the change in inconsistency between direct and indirect evidence in design shown in the row after detaching the effect of design d’ shown in the column.

-Is there any other way to quantitatively examine heterogeneity other than heat plots in NMAs?
Reply: Dear reviewer, thank you for your comments. An important aspect in network meta-analysis is to investigate statistical heterogeneity. If heterogeneity exists, then the possible sources should be explored and implementation of random-effects modeling, sensitivity analyses, subgroup analyses should be considered if sufficient data are available. Here we used random-effect Bayesian model to reduce the influence of heterogeneity between included studies. However the data are not sufficient enough for us to conduct a subgroup analysis. In network meta-analysis inconsistency arises as another aspect of heterogeneity. Therefore, we used heat plot to show the influence of detaching a pair of direct evidence to the network evidence. The colors in heat plot are associated with the change in inconsistency between direct and indirect evidence in design shown in the row after detaching the effect of design d’ shown in the column.

Minor issues:

-RT should be specified as "RT alone" when discussed as a management strategy (same with surgery alone and TLM alone); the conclusions of the abstract make it sounds like RT is detrimental in general, whereas it is really just RT alone performed poorly overall.

-"Chemotherapy radiotherapy" in the abstract should called "chemoradiotherapy"

-OS, DFS, and OSR should be defined in the captions for Table 2 and Table3.

Reply: Dear reviewer, we have revised the manuscript and RT has been specified as "RT alone". Also, "Chemotherapy radiotherapy" in the abstract has been changed to "chemoradiotherapy". OS, DFS, and OSR were defined in the captions for Table 2 and Table3.

Reviewer #3: The manuscript provides an interesting network meta-analysis of 28 clinical trials published over the previous ~30 years to compare approaches for treating hypopharyngeal and laryngeal neoplasms. While the article presents a potentially interesting finding, it is difficult to interpret with the limited analysis that has been completed thus far. For example, information regarding the classification of responses across trials is important to understand with increased clarity. How the outcome of the meta-analysis is being driven by the input data as well as how some of the potentially confounding variables affect the model are not discussed in the context of effect on the model. Some of these variables include cancer type/study, stage distribution of each trial's population and other potentially confounding variables. Overall, the conclusion is too strongly stated given the limited depth of analysis completed so far.
Reply: Dear reviewer, thank you so much for your serious evaluation on our manuscript. Our responses and explanations to your comments are as follows.

This network meta-analysis (NMA) was conducted to investigate effectiveness of six therapies being utilized in clinical practice nowadays. The outcome index include 3/5 year overall survival, which shows the percentage of people in a group who are alive in 3/5 years. 3/5-DFS shows the percentage of people survives without any signs or symptoms of that cancer in a group 3/5 years after primary treatment. Our study discusses the HRs of each index to compare the efficiency of six therapies. The data are obtained via software Engauge to match the overall survival curve present in included article and the Hazard Ratio was calculated by Excel formula. A relative ranking of the six regimens was calculated by the surface under the curve ranking area (SUCRA). To be sure, there are several possible sources of variability or heterogeneity among our studies that are included in systematic reviews and meta-analyses. For example, influence of cancer type/study, stage distribution of each trial's population and other potentially confounding variables can be reduced by opting for a comprehensive search for evidence and selection criteria, to maximize data yield while avoiding heterogeneity and inconsistency. Subgroup and more sensitivity analyses are effective to make the result reliable. However, among all 28 studies we included in this network meta-analysis, there were only 8 studies that the patients of which are in the same stage of hypopharyngeal cancer, which make it difficult to do a subgroup analysis and maintain the reliability of the conclusion meanwhile. Also, a random-effect Bayesian model was utilized considering the existence of heterogeneity and the influence of this heterogeneity to the analysis is limited in appropriate range.