Author’s response to reviews

Title: Antiviral treatment for treatment-naïve chronic hepatitis B: Systematic review and network meta-analysis of randomized controlled trials

Authors:

William WL Wong (wwlwong@uwaterloo.ca)

Petros Pechivanoglou (petros.pechivanoglou@sickkids.ca)

Josephine Wong (josephine.wong@theta.utoronto.ca)

Joanna Bielecki (joanna.bielecki@theta.utoronto.ca)

Alex Haines (alex.haines@theta.utoronto.ca)

Aysegul Erman (aysegul.erman@mail.utoronto.ca)

Yasmin Saeed (yasmin.saeed@mail.utoronto.ca)

Arcturus Phoon (rctrsphn056@gmail.com)

Mina Tadrous (TadrousM@smh.ca)

Mona Younis (mona.younis@westmountpharmacy.ca)

Noha Rayad (drnohazaki@gmail.com)

Valeria Rac (valeria.rac@theta.utoronto.ca)

Harry Janssen (harry.janssen@uhn.ca)

Murray Krahn (murray.krahn@theta.utoronto.ca)

Version: 2 Date: 26 Jun 2019

Author’s response to reviews:

Response to Reviewer #2’s comments

1. in Table 3a (and other similar tables, including Appendix F) it is not clear what the values in the top triangle represent, why they differ from the lower triangle and why there are some empty cells. Are these results obtained from simple pairwise meta-analyses? If so, how were these
computed? were random effects models used? I suspect there are not enough studies to estimate pairwise random effects models for all comparisons so are these results from a mixture of fixed and random effects? This needs to be clarified.

Response: Page# 33-36 Tables 3a and 3b and Appendix F

We thank the reviewer for this comment. We have clarified the presentations in the title legends, the sub-headings, and the footnotes as follows:

Table title: “Table 3a Relative effects on virologic response and ALT normalization of all pairs of interventions in Odds Ratios (95% credible intervals) as calculated from the network meta-analyses using random effects models”

Sub-headings: “Virologic Response: HBeAg-positive population in the left lower triangle & HBeAg-negative population in the right upper triangle”

“ALT Normalization: HBeAg-positive population in the left lower triangle & HBeAg-negative population in the right upper triangle”

Footnotes: “Empty cells in the right upper triangles are due to lack of studies comparing the treatment pairs in the HBeAg-negative networks.”

Table title: “Table 3b Relative effects on serological responses of all pairs of interventions in Odds Ratios (95% credible intervals) as calculated from the network meta-analyses using random effects models”

Sub-heading: “HBeAg seroconversion in the left lower triangle & HBeAg loss in the right upper triangle”

Footnotes: “Empty cells in the right upper triangle are due to lack of studies comparing the treatment pairs in the HBeAg loss network.”

Table title: “Appendix F: Relative effects on outcomes of all pairs of interventions in Odds Ratios (95% credible intervals) as calculated from the network meta-analyses using fixed effects models”

Sub-headings: “Virologic Response: HBeAg-positive population in the left lower triangle & HBeAg-negative population in the right upper triangle”
“ALT Normalization: HBeAg-positive population in the left lower triangle & HBeAg-negative population in the right upper triangle”

“HBeAg seroconversion in the left lower triangle & HBeAg loss in the right upper triangle”

Footnotes: “Empty cells in the right upper triangles are due to lack of studies comparing the treatment pairs in the virologic response and ALT normalization networks of the HBeAg-negative population; and the lack of studies comparing the treatment pairs in the HBeAg loss network.”

2) related to the comment above, the estimates of the between-study heterogeneity and their 95% credible intervals need to be presented for all analyses, including the node-splits. Given the small number of studies included, it needs to be noted whether or not the posterior distribution for the between-study heterogeneity has been informed by the data or if it is simply the same as the prior distribution.

Response: Page #18 Lines 401-404

We have now incorporated the estimates of the between study heterogeneity in the tables 3a and 3b. The reviewer was right to note that for some outcomes where the number of studies is small, indeed the variance is driven to some extent by the prior. We now comment on that in the discussion:

Fourthly, for some studies where information comes from fewer studies and more rare events, the estimate of between-study heterogeneity is more heavily dependent on the choice of the (vague) prior. As a consequence, the estimate of uncertainty for some of the relative effectiveness parameters is inflated. The mean estimates or relative effects however, have remained unchanged.

We have decided to proceed with the random effect estimates as conceptually we find a random effect approach is more fitting. However, we provide the fixed effect estimates in the appendix.

3) Appendix E states that p<0.05 will indicate evidence of inconsistency. However, a value of p=0.099 is also highlighted in bold. When checking for inconsistency, the values of the between-
study heterogeneity for each node-split model should also be checked as a reduction of heterogeneity compared to the main NMA analysis can also indicate inconsistency.

Response: Page# 12 Lines 270-274

Thank you for the comment. The p = 0.099 was an error from our end is now fixed. Regarding the check of inconsistency as a reduction in heterogeneity, we compared the variances across all the node split models for each outcome against the consistent model. There were no significant fluctuations of the variance and its confidence interval across all models, with large overlap across the confidence intervals. We have added the following in the body of the manuscript:

We compared the variances across all the node split models for each outcome against the consistent model. There were no significant fluctuations of the variance and its confidence interval across all models, with large overlap across the confidence intervals. This finding further indicated limited evidence of inconsistency.

Given the volume of information already in the manuscript, we decided to not include the tables with the variances from the node split modes in the appendix. We are happy to include it if editor deemed is necessary.

4) The odds ratios and 95% credible intervals for the main outcomes should be given in the abstract.

Response: Page# 2 Lines 59

We thank the reviewer for this helpful suggestion. We have revised the results in the Abstract to include this important information:

“Treatment strategies were ranked by the probability of achieving outcomes, and pairwise comparisons calculated from NMA were reported in odds ratios (OR). For HBeAg-positive population, tenofovir disoproxil fumarate (TDF) and tenofovir alafenamide (TAF) were the best for VR; OR versus adefovir = 14.29, 95% CI 7.69-25 and 12.5, 95% CI 4.35-33.33 respectively. TAF was the best for achieving ALT norm (OR versus placebo = 12.5, 95% CI 4.55-33.33), HBeAg loss and seroconversion (OR versus entecavir/TDF combination = 3.03, 95% CI 1.04-8.84; and 3.33, 95%CI 1.16-10 respectively). In the HBeAg-negative population, TDF and TAF
were the best for VR (OR versus adefovir = 9.79, 95% CI 2.38-42.7 and 11.71, 95% CI 1.03-150.48 respectively."

5) line 192: "Inconsistency evaluates..." should be "Inconsistency check evaluate..."

Response: Page# 10 Line 205

We thank the reviewer for pointing this out and have corrected the text as suggested:

“Inconsistency check evaluates the validity of a network…”

6) line 184: "evidence conflict against" should be "evidence conflict with"

Response: Page# 10 Line 207

We have corrected the text:

“If the results from direct evidence conflict with the results from indirect evidence …”

7) line 195: "inconsistency in the network. In this case one could argue that the direct and indirect evidence..." should be changed to "inconsistency in the network; that is the direct and indirect evidence...

Response: Page# 10 Line 208-209

We have revised the text as advised:

“In this case one could argue that inconsistency in the network, that is the direct and indirect evidence, may not be compatible.”
8) line 291: note that choice of best treatment should also take the relative effect into account.

Response: Page# 14 Line 311-312

We have edited the text as advised:

“The choice of the most desirable treatment should not be based solely on the ranking, but should take the relative effect into account.”

9) line 312: finding inconsistency does not mean different classes of treatment cannot be compared in an NMA - it means that for the studies of the different treatment classes identified, there were effect modifiers that suggested the evidence should not be combined. If the studies had been conducted in similar ways and in similar populations then they could be compared. The reasons for this heterogeneity were clinical and mentioned earlier in the manuscript. These should be identified here with a note that for this assessment the studies identified were too heterogeneous to combine, which manifested as inconsistency.

Response: Page# 15 Line 333-340

We thank the reviewer for this comment, and have revised the text as follows:

“Among these studies, PEG-IFN was used in different dosages; the duration of treatment was shorter than the oral nucleos(t)ides, and PEG-INF was combined in different order with oral nucleos(t)ides in different combination therapies. This means that there were effect modifiers in the identified studies that suggested the evidence should not be combined. For the purpose of this NMA, the studies identified were too heterogeneous to combine, which manifested as inconsistency in the network. On the basis of this clinical heterogeneity, we decided to further exclude studies that included the PEG-IFN treatment.”

10) line 368: I do not understand the comments on lack of connectivity of the networks. Maybe specify which outcomes these comments refer to.

Response: Page#17 Line 397-399

We have edited the text to make this clearer. We were referring to the number of closed loops within the network.
“These small sample sizes produced wide credible intervals and reduced the number of closed loops in the network, especially for the HBeAg-negative population.”