Author's response to reviews

Title: A systematic review on the effect of sweeteners on glycemic response and clinically relevant outcomes

Authors:

Natasha Wiebe (nwiebe@ualberta.ca)
Raj Padwal (rpadwal@ualberta.ca)
Catherine Field (catherine.field@ualberta.ca)
Seth Marks (smarks@ualberta.ca)
Rene Jacobs (rjacobs@ualberta.ca)
Marcello Tonelli (mtonelli@ualberta.ca)

Version: 2 Date: 25 July 2011

Author's response to reviews: see over
25 July, 2011

Dear Dr. Hogan:

Re:   MS: 1920640989562119
A systematic review on the effect of sweeteners on glycemic response and clinically relevant outcomes
Natasha Wiebe, Raj Padwal, Catherine Field, Seth Marks, Rene Jacobs and Marcello Tonelli

Thank you for your comments and those of the reviewers for this manuscript. We have taken care to address the comments from the editor and reviewers, and believe that our resubmission is a better paper as a consequence. We hope that our revised paper will be acceptable for publication in the BMC Medicine.

If you have further questions, or require further information, please don’t hesitate to contact us.

Sincerely,
Natasha Wiebe, Catherine Field, and Marcello Tonelli for the authors
Editor:
You will see referee 2 asks in particular to state the value of the estimated heterogeneity (and confidence intervals) throughout the results for all pairwise and network meta-analyses.
Done.

Furthermore, the referees recommend to improve the tables according to their suggestions.
We have improved the footnotes for the matrix tables and dropped the 2-hour blood insulin results to help readers understand the more complex network meta-analyses results as they did not add in the clinical interpretation of the overall findings.

We have also changed the flow of the network meta-analysis results. First we report the findings of analyses related to inconsistency and heterogeneity and then report the results of the mixed evidence. We have also added a box to highlight our main findings. We recognize that the many analyses we presented (due to the numerous permutations of sweetener type and the outcomes of interest) made our original submission difficult to follow. We think that our resubmission is substantially better in this regard.

All the other comments are minor or discretionary but need to be clarified.
Done.

Reviewer 1 (Statistical referee):

Discretionary Revisions
1. Authors are invited to consider adding rank-o-grams (plots of the probabilities that each intervention is the best, 2nd best etc) to show efficacy of the various interventions considered in the network meta-analysis on the multiple outcomes. For example of rank-o-grams see: Cipriani et al. (2009) Comparative efficacy and acceptability of 12 new generation antidepressants: a multiple-treatments meta-analysis. Lancet 373, 746-758 and Ades et al. (2010) Network meta-analysis with competing risk outcomes. Value in Health 13(8), 976-983.
We appreciate the suggestion and have considered it carefully. Rank-o-grams are a useful technique when one outcome is clearly preferred for ranking the merits of each intervention or treatment. As mentioned above, we have removed the results from the network meta-analysis of insulin (ie only one network analysis remains: for blood glucose). We would have considered presenting a rank-o-gram if we had been able to do a network analysis for change in weight. However, since blood glucose is not our primary outcome, we prefer not to present these results using a rank-o-gram, as we feel it would unduly emphasize this characteristic and its role in selection of a particular sweetener.

Minor Essential Revisions
2. Present the confidence intervals for the numbers quoted in the Abstract, or clarify if already doing so.
Done.

3. The choice of an active control group should be explained in the Introduction.
In the methods, we added: “Trials with placebo controls were also excluded as we aimed to investigate the comparative effectiveness of different sweeteners, as opposed to exploring the implications of avoiding sweeteners altogether.”

4. It should also be clarified that cross-over trials are not appropriate for weight-loss outcomes and are not included. Also, a (short) justification of why cross-over trials are suitable for the other outcomes would be reassuring.

We did not restrict type of RCT by outcome. To the study selection section we added that: “All outcomes selected for study (including weight change) are reversible and thus (providing that order was randomly assigned), a cross-over design should be appropriate.” Of note, we used the baseline estimates prior to each sweetener period – a comment was added to the statistical analysis section: “For weight change, the baseline value prior to the immediate period was used.” However, given that the rate of weight loss may slow down after some weight has already been lost, we modified the following text to specify the design: “One crossover trial was done in type 1 diabetic participants and found no difference in weight loss between groups over 4 weeks (0·8 kg [-3·3,4·9], 10 participants).”

5. Page 9, reference 19 for network meta-analysis: although Salanti et al provide a good review of methods, the appropriate reference to network meta-analysis is “Lu, G and Ades, AE. Combination of direct and indirect evidence in mixed treatment comparisons. Statistics in Medicine 23, 3105-3124. 2004” and this should be changed.

Done.

6. Page 9, sentence “Network analysis extends meta-analysis from simply pooling directly compared treatments (direct evidence) to pooling data from studies not directly compared but linked via one or more common comparators (indirect evidence)”: add “… by assuming consistency of the evidence” or similar and cite Lu & Ades, Stat Med 2004.

Done.

7. Page 9: State what was the range of the uniform prior and what were the mean and variance of the Gaussian priors.

Done.

8. State which STATA command was used to perform the pairwise meta-analysis and where the WinBUGS code was obtained from. If new code was written then this should be provided for review and published in an Appendix. Also state how convergence of the MCMC algorithm was assessed, how many burn-in iterations were used and how many samples were taken from the posterior distributions after convergence.

Done.

9. Page 12: The data in Tables 3 and 4 have different network structures and this needs to be made clear.

We have added: “Note, this network does not include non-caloric sweeteners.” to text describing the diabetic participant network.

10. Page 12 “2-hour blood insulin response”: the authors state that incoherence was detected – was this investigated further? If not, then there is no point in looking for incoherence if noting will be
done about it! There are several statements throughout the manuscript stating that incoherence and high heterogeneity were present, but no attempt is made to explain them.

We have extensively modified our presentation of the findings from network meta-analyses. Given that Reviewer 3 requested that we simplify our analyses, we have dropped analyses related to insulin levels. Additionally, we have improved the presentation of results related to 2-hour blood glucose levels:

“The network included 36 trials and 610 participants. The direct evidence from all 9 comparisons were consistent with the mixed evidence from the network. There was large heterogeneity between trials (I^2's ≥77%) for 3 of 7 multi-study direct evidence comparisons. Two of the heterogeneous comparisons included a variety of sweeteners (i.e., multiple sugar alcohols [τ^2=9.05 (95% CI 2.94,32)], or multiple other sugars [τ^2=1.72 (0.37,1.48)]) within one category. In the fructose vs glucose comparison, 6 trials were responsible for the heterogeneity [τ^2=1.40 (0.68,1.50)]. Three 36, 45, 50 were subgroups of diabetic participants; they increased the magnitude of the mean difference. The other three trials 32, 33, 46 showed important differences prior to the 2-hour timepoint (data not shown) but at 2-hours showed little or no difference between sweeteners. The single estimate of heterogeneity (τ^2) for the network meta-analysis was 0.65 (95% CI 0.35,1.10).

Using the mixed evidence, 2 comparisons: fructose vs sucrose (MD -1.12 mmol/L [-1.95,-0.27]), and fructose vs glucose (MD -1.56 mmol/L [-2.18,-1.02]) were statistically significant, all favouring fructose, but neither of the confidence limits excluded the possibility of non-clinically relevant differences (<1·15 mmol/L – calculation based on a clinical important difference of 1% for A1C). 60 The weighted regression test for publication bias was not significant.

In the subnetwork of 31 trials enrolling participants without diabetes (446 participants; τ^2=3.66 [1.66,7.31]; Appendix Table 1), the direct evidence from all 8 comparisons were consistent with the mixed evidence from the network. The heterogeneity although reduced remained large between trials (I^2's ≥60%) in both of the remaining multi-study direct evidence comparisons. Using the mixed evidence, 3 comparisons: fructose vs sucrose (-0.54 mmol/L [-1.06,-0.03]), fructose vs glucose (-0.89 mmol/L [-1.21,-0.59]), and fructose vs other sugars (-0.85 mmol/L [-1.47,-0.21]) were statistically significant, all favouring fructose, but none of the confidence limits excluded the possibility of non-clinically relevant differences.

In the subnetwork of 10 trials enrolling participants with diabetes (152 participants; Appendix Table 2), the direct evidence from all 6 comparisons were consistent with the mixed evidence from the network. Note, this network does not include non-caloric sweeteners. Because the estimate of τ^2 (224 [0.14,139]) did not converge well, we report our findings from the direct evidence. Three direct comparisons were significant and both found clinically relevant differences between agents over the entire confidence interval span: fructose vs glucose in 5 trials with 52 participants (-4·81 mmol/L [-6·34,-3·29], I^2=0%, τ^2=0 [0,7.47]), HSH vs glucose in 1 trial with 12 participants (-6·19 mmol/L [-9.78,-2·60]) and isomaltulose vs sucrose in 1 trial with 20 participants (-3·44 mmol/L [-5·31,-1·56]).”

11. Page 14, 3rd paragraph: Are cross-over trials appropriate for weight-loss outcomes? Please justify (also see comments about introduction above).

Please see above.
12. Page 16, 3rd paragraph: “Seven trials reported change in total cholesterol...”. Justify why no network meta-analysis was done on the treatments compared in these trials.
A priori we planned to perform network meta-analysis only on the two outcomes that we thought would have a very large number of trials.

Major Compulsory Revisions
13. State the value of the estimated heterogeneity (and confidence intervals) throughout the results for all pairwise and network meta-analyses. Furthermore, a statement of whether the WinBUGS...
This comment from the Reviewer appears to be incomplete. We added estimates of $\tau^2$ with corresponding 95% confidence/credible intervals to the results of pairwise and network meta-analysis.

14. In the Discussion, the authors must comment on the drawbacks of the back-calculation method used to assess incoherence. In particular, the fact that when random effects models are used for the separate MA inputs, different heterogeneity parameters are being estimated. The results are therefore hard to interpret if heterogeneity is very different between the different pairwise MA and the NMA. The node-split method recommended in Dias et al [20] is better in these circumstances, as is the new (and simpler) method for assessing inconsistency described in Dias et al 2011 TSD 4 – available from http://www.nicedsu.org.uk/Evidence-Synthesis-TSD-series%282391675%29.htm
The impact of the different heterogeneity estimates will of course depend on their values (which should be presented as mentioned above) and this needs to be discussed and justified.
We appreciate the cited documents and will use them in our future work.

We have added the following text to the limitations section of the discussion: “Our network meta-analysis has several limitations: 1) the sugar alcohol and other sugar categories contain multiple sweeteners that are likely to have different blood glucose profiles thereby causing heterogeneity, 2) power to detect inconsistency is limited by the number of trials included in the test, and 3) the back-calculation method used to detect inconsistent involved multiple tests thereby increasing the false-positive rate, however we did not detect any inconsistency.”

Reviewer 2:
This manuscript evaluates the effectiveness of intense sweeteners in controlling weight and minimizing risk factors for complications associated with obesity and type 2 diabetes by performing a systematic review and a network meta-analysis of eligible randomized clinical trials. The authors conclude from their review that little high-quality clinical research has been done to assess long term risk/benefit of non-caloric sweeteners considering their frequent consumption for weight and glycemic control. Still, despite the limited available studies, this review suggests that the evidence for a vital role for non-caloric sweeteners in an effective adult population health strategy for prevention of obesity and related co-morbidities is poor. Latest research news (albeit non-peer reviewed) show positive relationships between diet soft-drink consumption over 9 years and change in the waistlines of elderly, after adjustment for confounders and increased blood glucose with heavy aspartame exposure in diabetes-prone mice. Thus, this paper is highly pertinent in reviewing the evidence for clinically relevant outcomes from sweetener consumption to provide balance from risk and safety assessment studies, and to recognize that substituting or decreasing added sugars is
healthful. This review demonstrates that adequately powered RCTs are needed to obtain a more balanced systematic risk/benefit analysis in order to better define recommendations and provide comprehensive advice to policy-makers and consumers regarding sweeteners. The review is extensive; the criteria for selection of studies and the statistical methods used are appropriate and well described but do require careful reading, particularly tables 3 to 6.

**Discretionary Revisions:**
It would be of interest to discuss 1) the absence of brain and pancreatic responses with intense sweeteners vs. with sugar; 2) the ambiguous psychobiological signals with intense sweeteners that may confuse the body’s regulatory mechanisms leading to a loss of control over appetite and overeating; 3) that intense sweeteners are not appetite suppressants and fail to reduce motivation to eat.

We have added some text to the discussion with respect to satiety and sweeteners: “First, use of hypocaloric sweeteners might not induce weight loss even in the short term. For example, if reductions in calories due to sweeteners are offset by increases in caloric intake from other sources, or offset by decreases in caloric expenditure. Although our data suggest that non-caloric sweeteners may lead to clinically relevant weight loss through reduced energy consumption, this conclusion was driven by a single trial with a total of 41 participants. Unlike caloric sweeteners (which may partially compensate added calories with reduced energy intake from other sources), non-caloric sweeteners are not known to suppress appetite, and therefore would not reduce the motivation to eat. Furthermore, it has been suggested that the psychobiological signals with non-caloric sweeteners may directly influence physiological regulatory mechanisms and thus further reduce their potential for reducing net energy intake.

**Minor essential revisions:**
Page 13, line 17-19: terms ADI and CPG recommendations need clarification; Acceptable daily intake (ADI) of sweeteners are defined as safe amounts per kg body weight such that reducing energy intake would not change its value. However for sucrose and fructose, there are no ADI but Clinical Practice Guidelines’ recommendations of 10% of energy such that intake of 60 grams/day could exceed 10% with energy restriction but not necessarily a concern for safety except displacing nutritious foods.

Page 15 line 3-8: text should follow the same order as in the table (general before overweight)

Page 15 line 10 and 11: should specify that the study in ref 61 is in participants with type 1 diabetes.

Tables 3 to 6: explanations given with table 3 could be repeated under tables 4 to 6 for clarity such that each table stands on its own.

The discussion is well balanced and conclusions are supported by the results obtained. The authors could mention also the lack of good data to determine, if sugar is displaced by intense sweeteners,
whether there is incomplete compensation from other foods, a necessity if long term benefits are to be achieved.

In the discussion, we do highlight the limitation due to short-term and small trials: “Although the identified trials were numerous, they were very small and largely short-term. We found 13 trials with participant follow-up greater than 1 week and group sizes ≥10: 3 that compared non-caloric sweeteners to sucrose, and 10 that were head-to-head comparisons of saccharides. Ten of 13 trials had a Jadad score of 1 ...” We have also added some text with regards to incomplete compensation: “Unlike caloric sweeteners (which may partially compensate added calories with reduced energy intake from other sources), non-caloric sweeteners are not known to suppress appetite, and therefore would not reduce the motivation to eat.”

This review is limited to adult population and the conclusions may not be applied to sub-populations such as high soda consuming adolescents who could be better responders to intense sweeteners.

We agree.

The manuscript needs some language corrections before being published: abbreviation for kilocalories is kcal (low case c); conclusions in abstract: first word: use considering instead of despite; 2nd line of conclusions .....in this condition; diabetics (p. 10) and diabetic patient should be replaced by diabetic people or individuals or participants; p. 8 line 1: oral antihyperglycemic agents; in discussion page 17, line 11: control group was asked; page 18, line 19: whether substituting a non-caloric; page 19, line 1: delete the second consumed; line 5: most importantly; line 13: improvements in the health of the population; page 20 line 3: will improve public health. Done.

Reviewer 3:
Major Compulsory Revisions: None

Minor Essential Revisions and Discretionary Revisions:

Methods:
A Bayesian network analysis was used in addition to the standard meta-analysis. This method is beyond the typical statistical skill of a clinical investigator, which leads to the question whether there is too little material to be analyzed due to heterogeneity and lack of suitable studies. Rather than performing an in-depth exploration of a restricted set of sweeteners, we wanted to do an overview of the evidence supporting the use of any particular sweetener. Given the variety of sweeteners available, we decided a priori to use network meta-analysis to summarize the totality of evidence across multiple classes. Despite the tremendous interest in the potential effect of sweeteners on health, we found only short-term, low-quality evidence to support the widespread use of hypocaloric sweeteners, and choice between caloric sweeteners. This is the main finding we want to communicate, and we believe that the use of network meta-analysis is appropriate for this purpose.

In addition, this report looks at a large mix of sweeteners and comparisons that it seems to confuse and dilute the authors’ message. Could one draw more concrete conclusions if one compared fewer categories?
We wanted to summarize the large evidence base of almost entirely small trials. Our conclusions are limited by the limited nature of the available evidence rather than the design of our meta-analysis. Despite this, we have taken steps to make our review more concise and easy to follow. Please see responses to Reviewers 1 and 2 above.

Results:
Page 11, 3rd paragraph (2-hour blood glucose response):
For the following sentence please specify the “expected direction of effect” and specify a reference and meaning for the term of “non-clinically relevant differences”: “For the direct evidence, 3 comparisons were statistically significant, all in the expected direction of effect, but none of which excluded the possibility of non-clinically relevant differences”
We have added a sentence in the methods describing and citing our expected directions of effect. We also added a reference and described how we derived a minimum clinically important difference for serum glucose.

Page 12, 1st paragraph (2-hour blood insulin response):
It is previously mentioned that “I2’s from 3 of 7 multi-study comparisons indicate very large heterogeneity (#77%) between the trials” and then once again in this section “I2’s from 3 of 6 multi-study comparisons indicated large homogeneity between the trials (#48%). Are these 3 multi-study comparisons with large heterogeneity the same? If so, please make note of this somewhere within the report.
The “#” should be “>=”. The former refers to blood glucose outcomes; the latter refers to blood insulin outcomes.

Page 15, 2nd paragraph:
Please specify which type of diabetes is meant in the statement “The trial in 10 diabetic participants found no effect...”
Done.

Conclusion:
Page 20, 1st paragraph:
The authors express the need for “long-term, high-quality, adequately powered trials”. Maybe one should specify that these trials should be interventional and controlled. However, while this point is theoretically absolutely correct, the practical aspect is rather difficult. It might be a real bonus if the authors made a concrete suggestion how such a trial should look like, what the sample size should be and how long the observation period should last.
We have added: “The longest trial was only 10 weeks – not long enough to determine whether substituting a non-caloric sweetener for a caloric sweetener is sustainable in daily practice. To detect an important reduction in weight over one year such as 2.5 kg/m$^2$(less than 0.05 kg/m$^2$/week) in a RCT would require a minimum of 85 participants (assumptions: 25% loss-to-follow-up, $\alpha=0.05$, power=90%, SD=3 kg/m$^2$).”

Tables:
It would be beneficial to the readers’ understanding of the report to include a table outlining the inclusion and exclusion criteria. Thus, some information from the appendix should be summarized or at least indicated as a source of information.
We have added a box to the appendix to summarize our selection criteria succinctly.

**Tables 3-6:**
Please make these tables more accessible for a general reader. These are non-intuitive and difficult to understand.

We have added text in the table footnotes to describe what statistics each cell contains, and have rearranged the order that the information is presented in the footnote: “The mixed evidence of the Bayesian network analysis are in the shaded upper triangle and the direct evidence calculated using the REML estimate of $\tau^2$ are in the lower triangle. Sweeteners are reported in the expected order of efficacy (with the exception of other sugars) from the expected lowest to highest 2 hour glucose response, with the estimated probability (or rank) listed in the diagonal. Each table cell contains the mean difference (MD) with the accompanying 95% confidence intervals. In the cells with direct evidence, we also list the number of studies, the $I^2$ (percent of heterogeneity due to between-study heterogeneity) and $\tau^2$ (the between-study variance). Blank cells in the lower triangle indicate that no direct evidence was available. In the cells with mixed evidence, we list whether the mixed evidence was consistent with the available direct evidence. Also, in the first cell of the mixed evidence, we list the single $\tau^2$ estimate for the mixed evidence. Results are the MD of the expected higher-ranked sweeteners compared to the expected lower-ranked sweeteners (e.g., MD of sugar alcohols vs sucrose is 0.41 and is in column 2, row 5 for the direct results, and is -0.93 and is in column 5, row 2 for the network analysis results). MDs less than zero favour the expected higher-ranked sweetener (smaller glucose response). For example, sugar alcohols show an increased serum glucose response by 0.41 mmol/L compared to sucrose using the direct evidence. However, sugar alcohols show a decreased serum glucose response by 0.93 mmol/L using the mixed evidence. However since both confidence intervals include zero, neither analysis allows a confident judgment about which sweetener is preferable. Pooled evidence significant at $p<0.05$ are presented in bold font. All 9 mixed and direct results are consistent.”

The text in the methods section first discusses heterogeneity and consistency of the findings, and then presents the mixed evidence in detail. The text referring to the direct evidence has been revised to focus only on supporting the findings from the consistent networks. Hopefully this revised presentation will make it easier for readers to follow our findings. We have also added a box that summarizes our key findings.

**Table 3:** This table shows that there is no difference in blood glucose levels when non-caloric sweeteners are compared to sucrose. What does this mean?
Since the confidence intervals include zero, both the direct and mixed evidence are not statistically significant. We have added an example that explains the interpretation of these findings to the footnote.