

Responses to editors' and reviewers' comments on the manuscript submitted by Imamura et al., "Consumption of sugar-sweetened beverages, artificially sweetened beverages, and fruit juice and incidence of type 2 diabetes: a systematic review, meta-analysis, and estimation of population attributable fraction" (Manuscript ID BMJ.2014.023070)

We appreciate all of the valuable comments from the reviewers of our work. We have revised our manuscript, according to the reviewers' comments, questions, and suggestions and have shortened the supplementary materials considerably. We believe that the manuscript is now much improved.

This document includes our responses to reviewers' comments:

	Page
Comments from the committee	1
Comments from Reviewer 1	5
Comments from Reviewer 2	13
Comments from Reviewer 3	15
Comments from Reviewer 4	18

Comments from the committee

1. Deadline: Your revised manuscript should be returned within one month.
2. Online and print publication: All original research in The BMJ is published with open access. The full text online version of your article, if accepted after revision, will be the indexed citable version (full details are at <http://resources.bmj.com/bmj/about-bmj/the-bmjs-publishing-model>), while the print and iPad BMJ will carry an abridged version of your article, usually a few weeks afterwards. This abridged version of the article is essentially an evidence abstract called BMJ pico, which we would like you to write using a template and then email it to papersadmin@bmj.com (there are more details below on how to write this using a template).

Response: We are pleased to resubmit our revised manuscript in time. We have also prepared an abridged version of our manuscript as suggested.

#0-1. Editors feel the title of the paper could be improved to make clearer the different beverage types included in this review.

Response: The title has now been revised as follows, with clarity about the different beverage types: "Consumption of sugar-sweetened beverages, artificially sweetened beverages, and fruit juice and incidence of type 2 diabetes: a systematic review, meta-analysis, and estimation of population attributable fraction".

#0-2. Explanation of findings could be clearer throughout the manuscript. There is a very large number of supplementary files, tables and diagrams. It would be helpful if these could be reduced to the most important and essential supplementary items.

Response: We have revised the manuscript throughout for better clarity. We also have shortened the supplementary document from 40 pages to 25 pages, keeping essential contents, those following reviewers' suggestions, and consistency with the PRISMA guideline. Briefly, the supplementary materials were shortened as follows:

- i) Forest plots for fixed effect analyses have been removed, as suggested (Comments #0-5)
- ii) A table for funding sources has been removed, as details are not essential. We have kept the text that there was no study funded by industry (Page 9, Line 5-6): "No study or publication was funded by industry".
- iii) Tables for study characteristics, potential bias, and confounding have been modified for clarity.
- iv) Supplementary tables for exposure and outcome assessment (Table S3 and Table S5 in previous version) have been merged into one table (Table S3). Non-essential details, which can be found in citations, have been shortened.
- v) In the previous version, Table S4 included validity measures of exposure assessment for different cohorts in the EPIC-InterAct consortium. The table has been removed and briefly noted in the footnote of the current Table S3.

#0-3. Our statistician team have advised they would encourage more documentation of confounders and how these were dealt with. They have advised to include these as part of the quality assessment (Cochrane has a new tool for assessing quality of non-randomised intervention studies which could help).

Response: We appreciate this suggestion. A Cochrane Risk of Bias Assessment Tool: for Non-Randomized Studies of Interventions (ACROBAT-NRSI) has now been reviewed and incorporated into our assessment of risks of bias. We have also revised the assessment and documentation of confounding. Here, we summarise the major revisions:

- i) Table 2 has now included summary estimates of relative risks (RR) without any adjustment (crude), based on the suggestion of Reviewer #2 (comment #2-5). The comparison shows the influence of confounding easily, and is useful addition.
- ii) Table S4 has listed crude RR and adjusted RR of each study and of meta-analysis. In Table S4, we have described what factors were not adjusted for, as Reviewers #1, #2 and #3 expressed concerns about confounding (#1-8, #2-5, and #3-1). Specific explanation for those factors has been added to the footnote.
- iii) The main manuscript has now included the following description of confounding in the Results section (Page 9, Line 15-22).

“Confounding was likely to exist in all of the studies. As would be expected, consumers of ASB tended to be overweight or obese and to be hypertensive.^{11,39,49,58,65} In longitudinal analysis, all studies statistically adjusted for potential confounders including socio-demographic variables, clinical factors (family history of diabetes or prevalent diseases), and lifestyle factors including diet (Table S4). None of these factors was identified as a single source of confounding, according to studies assessing influence of potential confounding in different regression models.^{11,41,48–50,53,58,60,62–64} However, a combination of multiple factors was likely to cause confounding (Table 2, Table S4). After adjustment for multiple potential confounders, RR for SSB was attenuated from 1.25 to 1.18 (32% change); and for ASB, 1.48 to 1.25 (43%). By contrast, the point estimate for fruit juice was shifted upward, from 0.97 to 1.05.”

- iv) We have also noted that an issue of confounding is related to behaviours of individuals. For example, “adults at high risk of T2D preferentially consumed more ASB.^{5,6,8,10,11,36}” (Page 14, Line 1). Then, we have now linked this point to research and clinical implications, because adults’ beverage consumption is related to health-seeking behaviours and clinical characteristics. We have now added “These observations provide research and clinical implications for better understanding of health-seeking behaviours related to beverage consumption.” (Page 14, Line 3-4) This addition is partly responding to the request of adding clinical implications (see #0-7).

We have revised our bias assessment to follow the Cochrane Collaboration’s tools (Higgins, BMJ, 2011; and ACROBAT-NRSI). This includes assessment of seven domains of bias (confounding, selection, exposure measurement, misclassification over time, missing data, outcome measurement, and selective reporting). Results from the bias assessment have been presented in Table S2, in a way visually readable as recommended by Higgins et al, 2011. We have also summarised our assessment for each of the 7 domains in the Supplementary Materials (from the last paragraph on Page 16 to the first on Page 17). Brief description in the Table 1 footnote has been kept with minor revision.

We have decided to apply one revision in process of bias assessment. In the previous manuscript, we have not considered bias due to missing data on outcomes in the Singapore Chinese Health Study that excluded participants who died during the follow-up (15% of adults). This could cause bias, because of the evidence that SSB consumption and mortality could be associated (Q Yang et al., Added sugar intake and cardiovascular disease mortality among US adults, JAMA Internal Medicine, 2014;174(4):516-524). This qualification of bias has been added. The overall quality remains the same, because, in the previous version of our manuscript, we had already rated that the study had a high risk of bias.

Our sensitivity analysis excluding studies that were rated as having a high risk of bias was revised accordingly. No difference was found as shown in Table S5.

#0-4. While the conclusions drawn from statistical analysis seem appropriate, the methods as described are overly-complicated and would benefit from clearer explanation and less supplementary material.

We appreciate this suggestion. We have deleted technical details which can be found in citations. We also have revised the manuscript for better clarity. For a method to calculate population attributable fraction, in

particular, we have now used a standard description consistent with *Chapter 12* of “the *Handbook for Systematic Reviews of Interventions*” that includes a method description to translate relative-risk estimates to risk difference in a general population. The method is similar to ours that additionally accounted for survey weights. Thus, we have revised the whole method description as follows (Page 8, Line 4-15) using the terms presented in the Handbook. The assumption that any individuals’ characteristics did not change over time has been added, in response to Reviewer #1 (#1-11).

“PAF was estimated, by applying an algorithm of the Cochrane Collaboration to survey data.^{22,43} We first estimated habitual SSB consumption based on 24-hour recalls in US and 4-day food records in UK. Then, we estimated a 10-year T2D risk based on a risk-prediction algorithm developed and validated in each country.^{46,47} The predicted T2D risk for each individual was considered as a ‘assumed control risk’ (ACR)²² if the current SSB consumption would remain constant. Then, we calculated an alternative T2D risk for each individual if the SSB consumption would become zero, calculating $ACR \times (1/RR \text{ per serving/day}) \times \text{observed SSB servings/day}$. The difference between the two risk estimates represented a risk attributable to SSB consumption. Using the risk estimates, sampling weights, and a population size, we estimated the absolute numbers of events over 10 years, events attributable to SSB consumption (‘absolute risk reduction’²²), and PAF (the proportion of events attributable to SSB consumption). The estimation assumed causality and no change in individuals’ characteristics over time. Validation of 10-year risk prediction was performed in the US survey, predicting diabetes prevalence in 2009-2010 by using data collected in 1999-2000. Further details are presented in Supplementary Text.”

#0-5. We do not want fixed effect analyses and feel the graphs on page 18 of the material give the wrong impression about the precision of estimates at the lowest category.

Response: We have removed the graphs showing estimates from fixed-effects modelling.

#0-6. Observational and reverse causality was a concern and the editors would like the authors to discuss this in more detail in the limitations of the study.

Response: We agree with this suggestion. Also, given the emphasis on this issue in ACROBAT-NRIS, we have revised the sentences as follows (Page 12, Line 11-22):

Original: “Although we examined residual confounding by adiposity and objectively indicated a non-significant result for ASB, we could not control for health consciousness and other unmeasured confounders.”

Revised: “Although this study has strength of assessing influence of confounding and providing results adjusted for potential confounders, residual confounding by many other factors could exist.³⁸ Confounding by socioeconomic and dietary factors were not detected to be strong in published studies. However, these variables are likely to have been measured with errors and have caused residual confounding in individual studies and our meta-analysis. Additionally, lifestyle factors and adiposity could change over time. The time-varying characteristics might not be random and could cause bias in an unknown direction and cause insufficient adjustment for adiposity during the follow-up. Reverse causality could also exist, because co-morbid conditions and health consciousness might alter consumption of beverages, particularly ASB, and risks of T2D.”

In addition, if confounding was plausible on the basis of prior publications, we have now described it (page 13, Line 24-27): “Residual confounding in the finding for ASB is also plausible because adults at high risk of T2D may preferentially consume more ASB.^{5,6,8,10,11,37} By contrast, confounding in the opposite direction was found in our analysis of fruit juice. This is consistent with observations that leaner, lower-risk adults consumed more fruit juice.^{10,11}”

#0-7. It would also be helpful to explain more clearly the clinical application of the findings in the discussion.

Response: We agree that clinical application deserves discussion as well as public-health implication. We have now added the following text in the Discussion (Page 14, Line 7-14).

“Although causality has not been established, our findings and available evidence indicate a benefit of lowering SSB consumption for the primary prevention of T2D. In the same context, our findings also imply that consumption of ASB or fruit juice is unlikely to reduce the T2D risk, and these should not be considered as a healthy option of beverages. However, consuming ASB to lower caloric intake and

body weight may have clinical benefit for obese or overweight adults.^{6,78} This effect on weight should be considered separately from our study that could not rule out the effect of weight on beverage consumption.⁸ In addition, clinical applications of our finding deserve further appraisal about the effects of altering beverage consumption on changes in lifestyle behaviours and on risks of other clinical outcomes.^{3,6,8,}

#0-8. Please address reviewer concerns as included at the end of this letter.

Response: We have carefully considered the helpful comments from all reviewers and have provided point by point explanations below and revised the manuscript as appropriate.

IMPORTANT

When you revise and return your manuscript, please take note of all the following points about revising your article. Even if an item, such as a competing interests statement, was present and correct in the original draft of your paper, please check that it has not slipped out during revision.

#0-9, a. In your response to the reviewers and committee please provide, point by point, your replies to the comments made by the reviewers and the editors, and please explain how you have dealt with them in the paper. It may not be possible to respond in detail to all these points in the paper itself, so please do so in the box provided

b. If your article is accepted it will then be edited, proofed, and - after your approval - published on bmj.com with open access. This open access Online First article will not be a pre-print. It will represent the full, citable, publication of that article. The citation will be year, volume, elocator (a unique identifier for that article): eg [BMJ 2008;337:a145](https://doi.org/10.1136/bmj.2008.337.a145) — and this is what will appear immediately in Medline, PubMed, and other bibliographical indexes. We will give this citation in print and online, and you will need to use it when you cite your article.

c. Please write an abridged version of the article for the print and iPad BMJ using the appropriate BMJ pico template for your study's design.

d. Please include these items in the revised manuscript to comply with BMJ style:

Response: We have understood these points. We have prepared an abridged version of our manuscript.

#0-10, Title: this should include the study design eg "systematic review and meta-analysis"

Response: We have kept the title including "systematic review", "meta-analysis", and "estimation of population attributable fraction". The revised title is "Consumption of sugar-sweetened beverages, artificially sweetened beverages, and fruit juice and incidence of type 2 diabetes: a systematic review, meta-analysis, and estimation of population attributable fraction".

#0-11, Please include in the results section of your structured abstract (and, of course, in the article's results section) the following terms, as appropriate:

For a cohort study:

- Absolute event rates over time (eg 10 years) among exposed and non-exposed groups
- RRR (relative risk reduction)

Response: We have now revised the followings:

Abstract: "Of 20.9 million events of T2D predicted to occur over 10 years in the US (absolute event rate=11.0%), 1.8 million would be attributable to SSB consumption (PAF=8.7%, 95% confidence interval=8.3–9.2%); and of 2.6 million events in the UK (absolute event rate=5.8%), 79 thousand would be attributable to SSB consumption (PAF=3.6%, 3.3–4.0%)."

Text in results (Page 11 Line 20-21): "Absolute event rate over 10 years from 2010 were estimated to be 11.0% in US (20.9 million events) and 5.8% in UK (2.6 million events)."

Figure legend (Page 22) (Figure 3) had the same information revised as appropriate.

We could not use RRR, because observational studies did not observe reduction of risks. Instead, we have revised the manuscript to describe what percent difference in risks according to higher or lower levels of exposure. For example, without using relative risks, the abstract has now included: "Higher SSB consumption was associated with higher incidence of T2D by 18% per one serving/day (95% confidence interval=8.8–28%, I^2 for heterogeneity=89%)"

#0-12, There is no need to provide a formal reference for a very widely used package that will be very familiar to general readers eg STATA, but please say in the text which version you used for articles.

Response: Thank you; we have now adopted this approach.

#0-13, include explicit statements of the quality of evidence and strength of recommendations, we prefer reporting using the GRADE system

Response: We have now provided quality of evidence in text. We note that we aimed to quantify associations of beverage consumption with incident diabetes and to estimate population attributable fraction, but not to provide a clinical or public health recommendation for beverage consumption. Thus, we have provided the quality of evidence, but not a summary table that would imply our general recommendation. Our rating of evidence has been based on overall findings, potential risks of bias, and the principles of GRADE system. As results, evidence for a positive association of SSB with T2D was considered as 'moderate' quality. For each of ASB and fruit juice, 'low' quality was assigned, based on our assessment of quality of bias. We have revised the Methods section and Results section as follows.

Page 6, Line 11-12, in the Methods section: "Overall quality of evidence was assessed based on study quality, results from sensitivity analysis, and principles of the Grades of Recommendation, Assessment, Development and Evaluation (GRADE).²³"

Page 11, Line 12-15 in the Results section: "We rated quality of evidence for each of SSB, ASB, and fruit juice. We rated 'moderate' quality for SSB, because the main findings were likely to be robust against different sources of bias, despite observational design. 'Low' quality was assigned each for ASB and fruit juice. Findings for ASB were subject to publication bias and residual confounding; and for fruit juice, concern of stability of the positive association was present."

#0-14, Please write the discussion section of your paper in a structured way, to minimise the risk of careful explanation giving way to polemic. Please follow this structure: statement of principal findings of the study strengths and weaknesses of the study strengths and weaknesses in relation to other studies, discussing important differences in results and what your study adds. Whenever possible please discuss your study in the light of relevant systematic reviews and meta-analyses (eg Cochrane reviews) meaning of the study: possible explanations and implications for clinicians and policymakers and other researchers; how your study could promote better decisions unanswered questions and future research

Response: we have revised the Discussion accordingly, with the following subheadings: Strengths and limitations (Page 12, Line 10); Interpretation in relation to other studies (Page 13, Line 10); Clinical and public health implications (Page 14, Line 6; see #0-7); and Conclusions (Page 15, Line 1). We have also made recommendations for future research on the basis of limitations of previous work and our current work. For example, on Page 13 Line 5-7, we have stated: "To address limitations typical of observational research and also needs for a policy intervention in different populations, future research should include a randomised trial examining people's health and behaviours and also a trial examining population impact."

We have now also noted (Page 14 Line 18-21: "For future implementation of a policy-based intervention to reduce SSB consumption,^{12,13} our estimate of efficacy should be extended to estimates of effectiveness of interventions of reducing SSB, accounting for practical issues in interventions and effects on obesity, T2D risk, and lifestyle change associated with reduction of SSB consumption."^{8,78})

#0-15, Footnotes and statements (competing interests statement, contributorship statement + guarantor, copyright statement/ licence for publication, a data sharing statement, a statement describing the role of the study sponsors).

Response: We have kept these statements as instructed.

Comments from Reviewer 1

Comments:

#1-1: This is a rigorously conducted systematic review and meta-analysis that examined the role of sweet beverage consumption in the incidence of type 2 diabetes. Several earlier meta-analyses have summarized observational studies that examined sugar-sweetened beverage consumption and type 2 diabetes (e.g.,

Greenwood et al. Br J Nutr. 2014; Malik et al. Diabetes Care 2010). However, the paper adds important policy-relevant measures such as population attributable fraction. The main strengths include a range of sensitivity analyses to examine the robustness of their findings as well as an estimation of population attributable fraction. Major limitations include high heterogeneity in estimates across studies and potentially a high risk of bias because of the nature of observational studies.

Response: We appreciate the comments on the strengths of our work. We agree on the limitations owing to observational studies. We have extended our description about confounding (Page 12, Line 15-22) in response to the comment from the editors' committee (#0-3) and the other reviewers (see #2-5, #3-1, #4-2). As the reviewer #2 recommended (comment #2-5), we have now added results from meta-analysis evaluating crude (unadjusted) estimates (Table 2 and Table S3). We have also revised our assessment of bias for better transparency of the available information (Table S2).

There are points that should be better addressed:

Major points:

Introduction

#1-2. (Page 4, lines 21 to 35) The authors considered adiposity as a potential confounder between sweet beverage consumption and type 2 diabetes risk. However, as sweet beverage consumption can elevate the risk of obesity/overweight, adiposity could also be a mediator between sweet beverage consumption and type 2 diabetes risk. The authors should explicitly state that adjustment for adiposity may control confounding, but also remove the mediation effect and potentially bias the estimate.

Response: We agree that adiposity could be a mediator for the association of interest. We have now revised the sentences in the Introduction to describe this explicitly.

Original text in the second paragraph of the Introduction: "The influence of adiposity is crucial to better characterize, because obesity directly causes T2D and also relates to dietary habits, which confounds an association of beverage consumption with incident T2D"

Revision (Page 4, Line 10-13): "The influence of adiposity is crucial to better characterise, because obesity can be a mediator by directly causing T2D and, thus, mediating an association of SSB consumption with T2D; and because obesity can be a confounder by altering dietary habits and confounding an association of beverage consumption with incident T2D.^{8,10}"

We have also revised the statement of aims to highlight the same point.

Original text in the Introduction: "Second, after concerted effort to obtain unpublished estimates we evaluated the associations before and after adjustment for measured adiposity separately."

Revised (Page 4 Line 19-20): "We specifically aimed to meta-analyze the associations with and without adjustment for adiposity, because the association may be both mediated and also confounded by adiposity."

We have considered that the estimates adjusted for adiposity were not biased results because these are valid as estimates of net biological effects not including an effect on obesity per se. The estimates after adjustment for obesity are important because an effect of obesity on T2D has been well established and evidence is needed to inform that SSB can exert a non-obesogenic effect (separate to an obesogenic effect).

Nevertheless, our manuscript should be clear about separate effects on obesity and effects on T2D risk independent of obesity. Thus, in the Discussion, we have now discussed this point for a clinical application based on our findings (Page 14, Line 7-14), as also stated in response to editorial comment #0-7:

"Although causality has not been established, our findings and available evidence indicate a benefit of lowering SSB consumption for the primary prevention of T2D. In the same context, our findings also imply that consumption of ASB or fruit juice is unlikely to reduce the T2D risk, and these should not be considered as a healthy option of beverages. However, consuming ASB to lower caloric intake and body weight may have clinical benefit for obese or overweight adults.^{6,7,8} This effect on weight should be considered separately from our study that could not rule out the effect of weight on beverage consumption.⁸ In addition, clinical applications of our finding deserve further appraisal about the effects of altering beverage consumption on changes in lifestyle behaviours and on risks of other clinical outcomes.^{3,6,8}"

#1-3. (Page 4, lines 27-35) The authors stated that no review has quantified population attributable fraction due to sugar-sweetened beverages for type 2 diabetes. A recently published modelling study reported the population preventable fraction for type 2 diabetes by taxing sugar-sweetened beverages in India (PLoS Med 11(1): e1001582. doi:10.1371/journal.pmed.1001582).

As the paper published in PLoS Med provides relevant information, I suggest the authors mention the findings in the manuscript.

Response: We appreciate this citation for the paper by Basu et al. We have now included it as one of citation and provided discussion. We identified the publication previously (during the update of electronic search), but did not include it in the manuscript, because we considered that we were unable to compare our study with Basu et al. in a way beneficial for readers. While the study demonstrated a great approach to combine economic, nutritional, and health metrics, we had several concerns as summarised further below (for information only). However, we have now cited the publication as follows in our manuscript.

In the Discussion section (Page 12, Line 24-28), we have added the following sentences (assigning Reference #12 to Basu et al.).

“As population-based measures, effects of taxation of SSB on obesity and T2D incidence have been modelled previously.¹²⁻¹⁴ No study combined SSB consumption observed in multiple populations, T2D risk predicted by a validated algorithm, and quantitative evidence on association of SSB consumption with T2D incidence. Future comparison between available estimates is worthwhile to characterise efficacy and effectiveness of policy interventions in different settings.”

Here, we have described our considerations of the findings by Basu et al. Based on the considerations, we decided not to compare with our study in details.

- 1) Basu et al used household survey data which provided mean intakes of households, including children and adults. Thus, data from individual adults were not used.
- 2) Using means of households, the population mean severely lost between-individual variations of SSB consumption. Mean±standard deviation (SD) of energy from SSB consumption was estimated to be 46±1 kcal/day (Table 2 of the article) (Table 2 also includes errors in presentation. For example, means of BMI were greater than 100). Very small SD values might reflect failure to account for between-individual variation. In addition, by calculating the mean of household means, the statistics are likely to be regressed toward the mean and to follow the central limit theorem, getting close to normal distribution. Overall, the authors were unlikely to account for skewed distribution of SSB consumption. Although the authors misspelled standard errors as SD and accounted for skewed distribution by using unique probability functions, the approach was overall not clear.
- 3) Basu et al did not use the association of SSB consumption with incident T2D. They used a measure of an association between levels of glycaemic load (GL, a summary measure of impact of carbohydrates on insulin secretion) and T2D. In addition, the authors argued that the effect of GL on T2D was independent of obesity, but the estimate itself (Livesey et al., Am J Clin Nutr, 2013) was based on meta-analysis that did not separate estimates adjusted for adiposity and estimates not adjusted for adiposity. Also, even though Livesey reported significant heterogeneity of the estimate across populations (US/Europe and others), Basu et al did not account for it in their analysis of effectiveness of taxation in India.
- 4) Estimation and its validation of absolute incidence of T2D were based only on ecological trends. Comparison between two ecological estimates would not support validity.

Methods:

#1-4. The authors stated that they followed the PRISMA guidelines but they did not state whether the protocol of the systematic review was registered in database(s) such as the PROSPERO database.

Response: We apologise that we did not indicate it in addition to availability of the protocol. Now the first sentence of the method section has been revised as follows (Page 5):

Original: “Following the PRISMA guidelines, relevant studies were identified through hand searches and systematic searches of four databases on May 31, 2013 (updated on 10 February 2014).”

Revised: “Following the PRISMA guidelines¹⁵ and the protocol (not registered, available upon request), relevant studies were identified through hand searches and systematic searches of four databases on May 31, 2013 (updated on 10 February 2014)”

We did not register our study, partly because we anticipated post hoc decisions based on increase in the literature on the same topics (as we found Xi et al., PLoS One, 2014; Greenwood et al. Br J Nutr. 2014), needs for extensive bias assessment, and our multi-stage procedure including a review of reviews (the first paragraph of the Supplementary Text); meta-analysis involving contacts to authors; and estimation of a population measures based on contemporary surveys. We recognised availability of registration, but we did not recognise that we could modify a registration for post hoc decisions by specifying dates of modification.

#1-5. Eligibility criteria included follow-up of at least two years. Two years of follow-up may be too short to examine the risk of type 2 diabetes. Please provide relevant information to support this criterion.

Response: We recognised that the Diabetes Prevention Program (DPP) (Knowler et al., N Engl J Med, 2002) indicated that two years or more would be sufficient to vary risk of diabetes by lifestyle modification. A natural experiment (ecological study) in Cuba also indicated that lifestyle change was related to change in diabetes incidence in 2 years later (Franco et al., BMJ, 2013). Thus, if we had identified a study (trial or observational study) that recruited adults in a general population or a high-risk population (which DPP targeted) and followed up participants at least for 2 years on average, we would have included it. Although we did not find any studies concerned with duration of follow-up (≥ 3.7 years for all studies, Table 1), a follow-up for 2 years or more would be enough to detect a causal role of dietary factors in altering a risk of T2D.

We have now revised the sentence as follows (Page 5 Line 8-11):

Original: “Eligibility criteria were a prospective design, assessment of beverage consumption and incident T2D, recruitment of adults free of diabetes and aged 18 years or older, and follow-up of two years or longer.”

Revised: “Eligibility criteria were a prospective design, assessment of beverage consumption and incident T2D, and recruitment of adults free of diabetes and aged 18 years or older. Follow-up of at least two years on average was also considered, because diabetes incidence could alter approximately two years after lifestyle modification.^{16,17}”

#1-6. In the estimation of population attributable fraction, the authors modelled if the sugar-sweetened beverage consumption would become zero. It might be difficult to eliminate sugar-sweetened beverage consumption from the UK or US. Alternatively, it may be practical to reduce the contribution of sugar-sweetened beverages to <5% of total energy intake (or tax sugar-sweetened beverages). Thus, I would suggest the authors examine the fraction under different scenarios.

Response: We appreciate the suggestion. We agree on difficulty in eliminating SSB consumption and on benefits of estimating PAF based on different scenarios. Yet, we have not decided to do that, because of the following reasons:

- 1) We have been considering that PAF is an efficacy estimate, whereas estimates based on different scenarios are efficacy estimates. With this distinction, we would like to keep the efficacy estimates of PAF, as this provides an important estimate of how much risk reduction could be achieved potentially.
- 2) Efficacy estimates would add too much to the current manuscript and would stand alone as a single research publication (as Reference #14 Briggs et al., Overall and income specific effect on prevalence of overweight and obesity of 20% sugar sweetened drink tax in UK: econometric and comparative risk assessment modelling study. *BMJ*, 2013).
- 3) PAF based on <5% of total energy intake would be conceivable. However, estimates of associations per 1% of total energy intake are not available, because information needed for the analysis has been limited. For example, we have not identified any studies assessing an association of percent of energy with incident diabetes. In addition, we recognised that the World Health Organisation drafted the guideline proposing that sugars should be less than 10% and be <5% for additional benefits. Use of the cutpoint proposed for total sugar intake, not SSB, would add extra complexity to our study. Thus, we would like not to investigate its use in our work.

Given these points, we would like to keep the current presentations of PAF treated as a measure of efficacy about SSB consumption. We have now considered that we should use the terms of ‘efficacy’ and

‘effectiveness’, as we believe these terms aid clinical readers to interpret PAF without confusion with effective estimates. Based on this clarification, we have further noted that effectiveness is of future interest, requiring a formal review of existing evidence on trials of dietary and/or policy interventions. Accordingly, we have provided the following revisions (partly responding to the other comment, #1-11, about lifestyle changes associated with reduction of SSB consumption):

Page 13 Line 1-3: “Nonetheless, limitations of PAF should be appraised. First, causality was assumed, although it has not been established. Second, estimates were under assumptions of no change in lifestyle potentially associated with reducing SSB consumption. Thus, our estimates should be considered as efficacy of reducing SSB, rather than effectiveness.”

Page 14 Line 18-21: “For future implementation of a policy-based intervention to reduce SSB consumption,^{12,13} our estimate of efficacy should be extended to estimates of effectiveness of interventions of reducing SSB, accounting for practical issues in interventions and effects on obesity, T2D risk, and lifestyle change associated with reduction of SSB consumption.^{8,78,}”

After this description, we have continued documenting the following (Page 14 Line 21-25). This highlights one of key inferences based on efficacy (partly in response to Reviewer #3; see #3-2):

“Despite PAF of no more than 15%, efficacy estimates are crucial, as 535 million adults are estimated to have T2D in 2035.⁶⁵ Additionally, the PAF informs that an intervention reducing SSB only would not reduce a large amount of events, and thus confirms importance of multiple modifiable risk factors, rather than a single dietary component, for the prevention of T2D.”

We have also considered that we should not imply the impractical scenario of eliminating SSB consumption. Rather, we should describe PAF as an estimate of how many events are attributed to the exposure. For example, the text for the results has been revised as follows (Page 11 Line 24-27):

Original: “Assuming a causal effect of SSB consumption independent of adiposity (adiposity adjusted), eliminating SSB consumption would show fewer T2D events by 1.8 millions in US (PAF=8.7%; 8.3-9.2%) and by 79 thousands in UK (PAF=3.6; 3.3-4.0). Younger adults and men would have greater numbers of preventable T2D than older adults and women, respectively (Figure 3, Table S9).”

Revised: “Assuming a causal effect of SSB consumption independent of obesity status (adiposity adjusted), SSB consumption would cause 1.8 million excess events in US (PAF=8.7%; 8.3 to 9.2%) and 79 thousands excess events in UK (PAF=3.6%; 3.3 to 4.0%). Younger adults and men would have greater numbers of T2D events related to SSB consumption than older adults and women, respectively (Figure 3, Table S9).”

#1-7. (Page 33, lines 15 to 18) In the supplemental text, the authors provided procedures to estimate population attributable fraction in the UK and US. They stated that (3) estimated separate ideal 10-year risk (R_i) if all adults reduced their SSB consumption to zero. Please describe a more detailed explanation to derive R_i .

Response: We are pleased to be able to clarify this point. We have provided a simple example as below, for RR adjusted for adiposity=1.13. Of note, the description has now been revised in reference to similar documentation by Cochrane’s group (see chapter 12.5.4.2 of http://handbook.cochrane.org/chapter_12/) (per-1,000 estimate was replaced with that per an actual number of non-diabetic adults in a population). Also, following the suggestion by the editors committee (#0-4), the description has been revised for simplicity.

Page 19 5th and 6th paragraphs of Supplementary Materials:

“Overall, in each survey, we (1) estimated habitual consumption of SSB among adults; (2) predicted 10-year risk (‘assumed control risk’, ACR^{113}) of developing T2D for each adult; (3) estimated separate ideal 10-year risk (R_i) of each adult if SSB consumption was reduced to zero; and (4) estimated ($ACR - R_i$) for each adult and $\Sigma(ACR - R_i) \times$ population size as a number of cases attributable to SSB consumption in a population. PAF was derived as $\Sigma(ACR - R_i) / \Sigma(ACR)$.

As a simple example, if one adult consumed 1 serving/day of SSB and had ACR of 0.10 and if SSB consumption became zero, his or her risk (R_i) would be $0.10/1.13=0.088$, where 1.13 is RR adjusted for adiposity. This calculation was applied to all adults, and pooling them as $\Sigma(ACR)$ and $\Sigma(R_i)$, the

population-based estimates were obtained.¹¹⁰ This estimation has advantage that there is no need of assumption in exposure distribution.”

Results

#1-8. (Page 9, lines 13 to 25) According to Table S6, many studies, including CARDIA, EPIC-Inter Act, SCHS, and MESA, did not control for a family history of diabetes. Because a family history of diabetes is a strong risk factor for type 2 diabetes, risk of bias is potentially high in these studies. Moreover, age was not listed as a covariate in studies, including NHS I, NHS II, HPFS, EPIC-InterAct, E3N; I believe these studies controlled for age. Thus, please check whether covariates are accurately listed. If such important covariates were not controlled in the included studies, the risk of bias of this meta-analysis is high; I suggest that the authors conduct additional sensitivity analysis (or bias analysis) to examine the impact of residual confounding due to these factors.

Response: We appreciate the attention to detail and apologise for incomplete/inappropriate description. We have revised the information on confounding in Table S4 in the current manuscript. Of note, we did not describe in details for any of particular potential confounders, because no study indicated substantial confounding by an individual factors (Please see our responses to the editors' committee, #0-3, and the other reviewers, #2-5, #3-1, #4-2).

The issue of family history: We had confirmed that adjustment for family history was unlikely to change the estimate. EPIC-InterAct did the sensitivity analysis before and after the adjustment in a multivariable-adjusted condition (The InterAct Consortium, *Diabetologia*, 2013;56(7):1520-30) and showed very similar results. This information has been added to the footnote of Table S4.

We have now confirmed similar results after excluding the four studies without adjustment for family history of diabetes. For the association of SSB consumption with T2D incidence after adjustment for adiposity, RR=1.13 (1.03-1.21) before the exclusion and 1.13 (1.05-1.21) after the exclusion of family history. For ASB consumption, RR=1.08 (1.02-1.15) and 1.08 (1.02-1.15), respectively. For fruit juice consumption, 1.10 (1.01-1.20) and 1.07 (1.01-1.14), respectively. We would agree that family history of diabetes is a strong risk factor for T2D, but unless this was strongly associated with beverage consumption after adjustment for other covariates, confounding by this variable would not occur. There was no indication that this happened.

The issue of age: NHS I, NHS II, HPFS, EPIC-InterAct, and E3N did use age as underlying time-scale of a Cox proportional hazard model, thus averaging proportional hazards across age, i.e. adjusted for age. In the previous manuscript, we failed to include this important information. We have now included it in the footnote of the Table S4.

#1-9. (Page 10, lines 3 to 4) A possible non-linear association was assessed using a cubic spline model. However, only 13 studies were included in this analysis because category-specific estimates of the other four studies were not available. Did the authors contact the authors of those studies to request the category-specific estimates? In addition, because the evaluation of the dose-response relationship is quite important in the meta-analysis of nutritional epidemiologic studies, it may be more informative to show the spaghetti plots proposed by Dr. Eric Ding to graphically illustrate the shape and direction of the dose-response relationship of individual study.

Response: We did not request additional estimates based on categorical variables, if a publication included estimates based on continuous variables before and after adjustment for adiposity measures. These two types of information have been sufficient to estimate a dose-response relationship in our primary meta-analysis.

We have now clarified this point as follows (Page 5, Line 24-28):

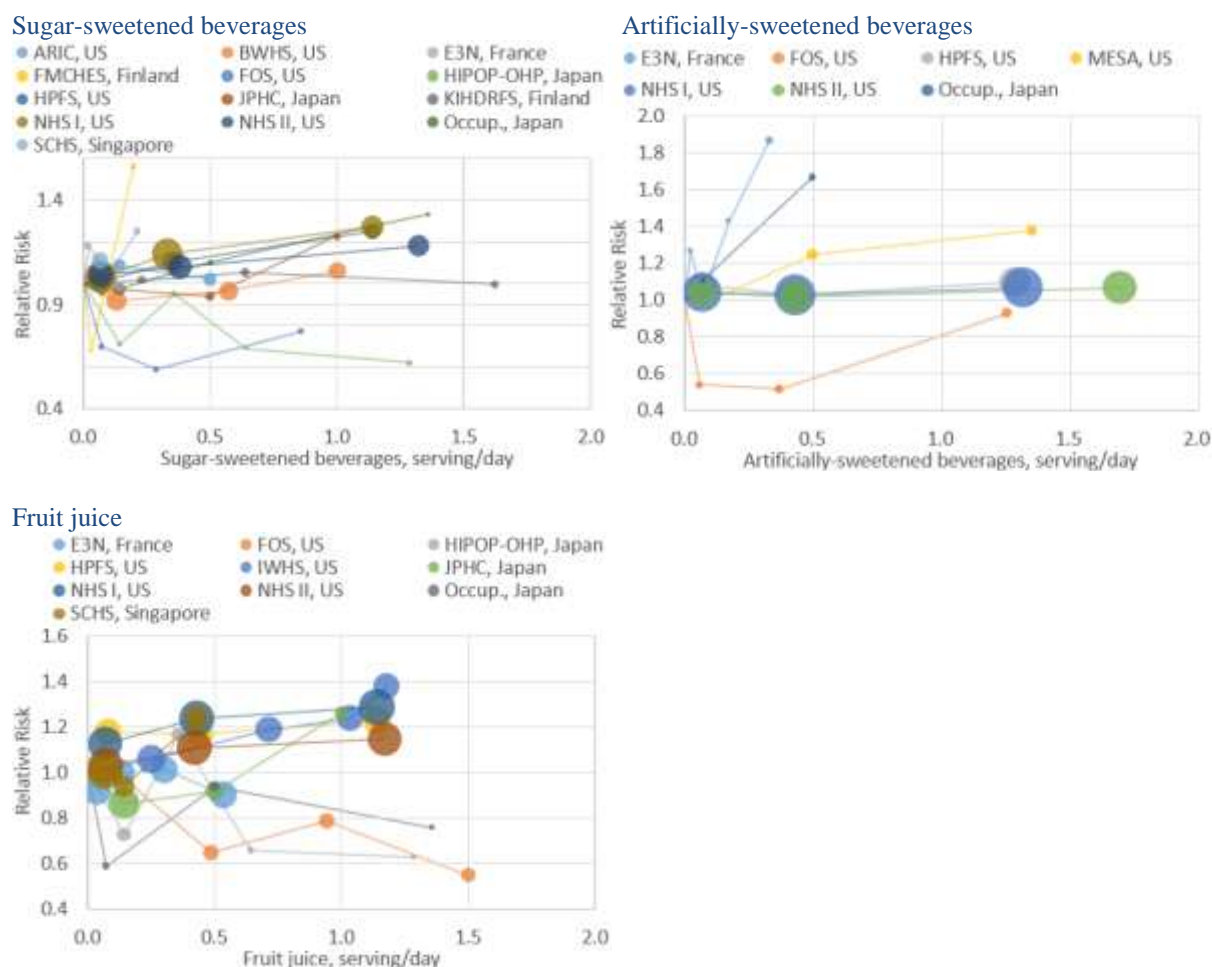
“We requested additional information if an article did not report two types of estimates before and after adjustment for adiposity, based on either categorical or continuous analysis for SSB, ASB, and fruit juice separately. When we contacted authors, we requested estimates before and after adjustment for adiposity of estimates based on both continuous and categorical variables of each beverage consumption; and requested estimates based on longer follow-up if available.”

We agree that a dose-response relationship is important in nutritional epidemiology. However, based on the available evidence, we had not primarily aimed to test a non-linear relationship. There has been no strong

biological rationale that there would be a ceiling or threshold effect in a relationship between SSB consumption, for example, and incidence of diabetes. Instead, our primary interest has been to identify and quantify a dose-response relationship. In addition, adults consuming SSB more than 1.0 servings/day, for example, are of a small segment of the population, as we have noted that more than half were zero consumers (the end of introduction, Page 4). Thus, a potential ceiling or threshold effect of high consumption of SSB would be biologically interesting, but less important for public health, compared to other types of exposure (e.g. coffee or tea, which are consumed more frequently habitually). Therefore, we did not collect categorical estimates, if a cohort reported a linear dose-response relationship in both models adjusted for adiposity and unadjusted for adiposity (e.g. EPIC-InterAct).

We have generated spaghetti plots to identify if they are informative. The figures (below) may be informative to present scatter of available information, but seem not informative to identify an average non-linear relationship. We have decided not to use them in the manuscript. The information from a subset of cohorts is not essential and each figure does not give enough information about differences in weights assigned to studies in meta-analysis.

Figure. Spaghetti plots of categorical data points available for meta-analysis of sweet beverages and incidence of type 2 diabetes. The differences in sizes of circles are proportional to differences in 1/standard error



#1-10. (Page 11, lines 9 to 11) The authors described the results of examining potential influence of residual confounding by measured adiposity with a focus on “statistical significance”. As the degree of bias (or confounding) matters, we should focus instead on the magnitude and direction of the corrected estimates. Based on the displayed results, the magnitude of the estimate was attenuated after incorporating residual confounding by measured adiposity.

Response: We appreciate the comment. We agree that the magnitude of bias is important, as we attempted to do quantitative assessment. We have now revised the relevant sentences as follows, to be more informative (Page 11 Line 8-11):

“When we examined potential influence of residual confounding by measured adiposity, bias toward the null would appear substantial for ASB.³⁷ Under realistic assumption of correlation=0.80 between measured and true adiposity, the association of SSB was attenuated by 26% to be RR 1.20 (1.04-1.38); of ASB, by 96%, 1.01 (0.81-1.25) (Figure S4); and fruit juice, strengthened by 19%, 1.12 (1.03-1.22).”

After this description, the following paragraph has now also been newly added to aid interpretation (Page 11, Line 12-16). As the editors’ committee confirmed (#0-4), we believe our interpretation was reasonable, given the magnitude of potential bias.

“We rated quality of evidence for each of SSB, ASB, and fruit juice (Table 2). We rated ‘moderate’ quality for SSB, because the main findings were likely to be robust against heterogeneity and residual confounding, although all studies were observational. ‘Low’ quality was assigned each for ASB and fruit juice. Findings for ASB were subject to publication bias and residual confounding; and for fruit juice, concern of stability of the positive association was present.”

Discussion

#1-11. The population attributable fraction is a useful measure but we need to assume that all other lifestyle factors remain constant after eliminating sugar-sweetened beverages. From a public health perspective, the possible impact of eliminating sugar-sweetened beverages on other lifestyle factors such as added sugar consumption (not from beverages) or fat consumption may deserve discussion. What are the authors’ expectations?

Response: We agree about the limitation of PAF under assumption of no change in any other variables. If SSB consumption were eliminated (or reduced), caloric intake would go down, and energy from protein and fat could relatively increase. One clinical trial showed that replacing SSB with ASB or water, with general dietary guidelines, altered the dietary pattern (Piernas et al., Am J Clin Nutr, 2013) along with possible increase in awareness of a healthier diet pattern. This highlight needs for effectiveness research on reducing SSB consumption. This concern is thus in line with our aforementioned view about difference between efficacy and effectiveness (#1-6). We have copied the sentences here to highlight the limitation of PAF based on a single dietary exposure and needs for future research on effectiveness.

Page 13 Line 1-6: “Nonetheless, limitations of PAF should be appraised. First, causality was assumed, although it has not been established. Second, estimates were under assumptions of no change in lifestyle potentially associated with reducing SSB consumption. Thus, our estimates should be considered as efficacy of reducing SSB, rather than effectiveness. Third, generalisability remains unknown, for example, to Central and South America where the highest per-capita sales of SSB in the world have been recorded; and China and India where the highest prevalence of T2D is expected.^{1,66} To address limitations typical of observational research and also needs for a policy intervention in different populations, future research should include a randomised trial examining people’s health and behaviours and also a trial examining population impact.”

Page 14 Line 18-21: “For future implementation of a policy-based intervention to reduce SSB consumption,^{12,13} our estimate of efficacy should be extended to estimates of effectiveness of interventions of reducing SSB, accounting for practical issues in interventions and effects on obesity, T2D risk, and lifestyle change associated with reduction of SSB consumption.^{6,78”}

Please also see our reply to the comments made by the reviewer #3 (#3-1), who also commented importance to consider multiple dietary factors in estimation of PAF.

Minor point:

#1-12. (Page 7, lines 11 to 12) Should two-sided $\alpha=0.1$ be revised to two-sided $\alpha=0.2$?

Response: We apologise for the confusion. In the analysis of potential sources of heterogeneity, we set $p=0.20$ as a criterion to retain variables in a meta-regression model (forward-variable selection). Then, in a multivariable-adjusted meta-regression, variables with $p<0.10$ were considered as significant sources of heterogeneity. We could have tested potential sources of heterogeneity one by one, but this would fail to identify independent sources of heterogeneity. We have now revised the description in the main text (Page 7 Line 9-12):

Original: “We assessed potential sources of between-study heterogeneity. Pre-specified study-specific characteristics were evaluated in multivariable meta-regression with a procedure of forward variable selection ($P < 0.2$ for entry) (two-sided $\alpha = 0.1$) (Supplementary Text). Using each variable with $P < 0.1$ and prespecified variables (age, sex, BMI, and location), stratified analyses were performed.”

Revised: “We assessed if pre-specified study characteristics explained heterogeneity of estimates of associations across studies, using meta-regression. Publication status (published or not) was also assessed *post hoc*. Stratified meta-analysis was performed by each variable identified as a significant source of heterogeneity ($P < 0.01$, see Supplementary Text for variable selection), age, sex, body-mass index (BMI), and location of study.”

We have also revised the related supplementary text as following:

Page 18, third paragraph: “Independent sources of heterogeneity were selected by meta-regression with forward-variable selection. If variables in meta-regression showed $P < 0.20$, the variable with the lowest P -value was retained in the model. Then, adjusting for the variable retained, meta-regression was repeated for remaining variables. If any of the additional variables did not meet $P < 0.20$, the model was considered best fitted. A variable with $P < 0.10$ was considered as a significant source of heterogeneity and meta-analysis stratified by the factor was performed.¹⁰³”

Comments from Reviewer 2

Comments:

Thank you for giving me the opportunity to review this interesting article. The review which examines the association between consumption of sugar sweetened beverages, artificially sweetened beverages, fruit juice and incidence T2DM has been conducted and reported to a high standard. I believe the report will be of interest to the BMJ readership, and comes at a time when interest in this topic is growing. I would recommend a statistician review the methods.

Abstract

#2-1. The conclusion should be reworded to make it clearer that the estimation of 2million events in the USA and 80 thousand events in the UK of T2DM are based on assumptions.

Response: We appreciate the suggestion. In the abstract, we have now concluded as the following. In this concluding sentence, the estimated numbers of events have also be reworded to avoid redundancy of the results and keep generalisability.

“Under assumption of causality, SSB consumption over ten years may be related to a substantial number of cases of new-onset diabetes.”

#2-2. In addition the phrase “but still support not consuming ASB or fruit juice as alternatives to SSB” is confusing to the reader, please re-phrase.

Response: We have now included the following sentences for ASB and fruit juice, partly responding also to the reviewer #4: “Although ASB and fruit juice also showed positive associations with T2D incidence, the findings were likely to involve bias. Nonetheless, both ASB and fruit juice were unlikely to be healthy alternatives to SSB for the prevention of T2D.”

Introduction

#2-3. Some re-wording of sentences is required, please check the introduction, and the whole document for grammar and sentence structure. For example, page 4, line 11: “T2DM have not be well established”

Response: Thank you. We have now extensively reviewed and revised our documents for improved English usage, including attention to correct grammar.

Methods

#2-4. Please make it clearer what volume your definition of “serving/day” is equivalent to.

Response: We have now provided the additional information on “serving/day”.

Page 6 Line 19-22: “Measures of associations were standardised to relative risk (RR) per one serving/day of beverage consumption, after we confirmed this unit was used most frequently across studies. Because volume per serving was specific to a population, ranging from 237 ml (1 cup) to 355 ml (12 oz) (median across publications=250 ml/day), we repeated the meta-analysis to estimate RR per 250 ml/day.”

#2-5. It is unclear what confounders have been adjusted for; I would prefer to see an unadjusted RR presented, followed by the adjusted RR (with a clearer presentation of what confounders have been included) and then the adiposity adjusted RR.

Response: We appreciate the request to better present characteristics of confounding. We have created Table S4 to provide unadjusted RR of each study and pooled estimate of the RRs. The summary estimates have also been added to the main Table 2. We believe that readers would now be able to see the magnitude of confounding easily and understand the covariates adjusted for. As a result, we have been able to infer that confounding for ASB estimates was substantial.

If we described all covariates in each study, the list would be too long. In response to editors' request (#0-2) to shorten the material, we have now provided the information on covariates succinctly in Table S4 and its footnote. Of note, we found that any single factor was unlikely to cause substantial confounding, except for adiposity. Although a single factor appeared not to be a substantial confounder, a net effect of confounding would be noticeable, particularly for ASB, where crude RR was 1.48 (95% CI, 1.35-1.62); adjusted RR, 1.25 (1.18-1.33); and RR further adjusted for adiposity, 1.08 (1.02-1.15). To describe the characteristics of confounding, we have now added a paragraph about confounding in the Results section (Page 9 Line 15-22):

“Confounding was likely to exist in all of the studies. As would be expected, consumers of ASB tended to be overweight or obese or hypertensive.^{11,39,49,58,65} In longitudinal analysis, all studies statistically adjusted for potential confounders including socio-demographic variables, clinical factors (family history of diabetes or prevalent diseases), and lifestyle factors including a diet (Table S4). None of these factors was identified as a single cause of confounding, according to studies assessing influence of potential confounding in different regression models.^{11,41,48–50,53,58,60,62–64} However, a combination of multiple factors was likely to cause confounding (Table 2, Table S4). After adjustment for multiple potential confounders, RR for SSB was attenuated from 1.25 to 1.18 (32% change); and for ASB, 1.48 to 1.25 (43%). By contrast, the point estimate for fruit juice was shifted upward, from 0.97 to 1.05.”

Please also find our replies to the editors' committee (#0-3) and the reviewer #3 (#3-1) for discussion about confounding.

#2-6. RR and OR are not the same, although converted this should be noted as a limitation in the discussion.

Response: We agree that our meta-analysis has a limitation of approximation. We have now revised the description of limitations as follows (Page 12 Line 19-22):

“Our meta-analysis included statistical approximation that might involve errors. For example, we derived dose-response estimates partly from categorical estimates and odds ratios. Without such approximations, analysis standardised across different cohorts is of future interest to characterise associations of different beverages with risks of T2D.”

Results

#2-7. On page 10, line 31, it states after stratification none of the beverages were significantly associated with T2DM. Please can you provide more information on this.

Response: We appreciate the attention to detail. We have now clarified (Page 10 Line 22-23) that:

“Demographic variables and BMI were not significant sources of heterogeneity ($P>0.14$ each), whereas each of SSB, ASB, and fruit juice was not significantly associated with T2D in studies recruiting more men than women or in Asia, with a fewer number of studies than the main analysis (Table S5).”

We have pre-specified stratified analysis by age, sex (% men), BMI, and location (US, Europe, and Asia). These variables were selected for stratification, because each variable was important for public health perspective. Any of these were not identified as a significant source of heterogeneity. After the pre-specified stratified meta-analysis, no significant association of any beverage with T2D was found in stratum of a fewer number studies with men>women and of Asian studies. The null findings were partly based on loss of

statistical power due to stratification. Therefore, because heterogeneity by each of these factors was not significant, we have not drawn main conclusion based on the stratified analysis. We have intended that these results will be helpful for readers who are interested in our findings in a specific population.

This is related to an issue of generalisability, and has been kept in the revised manuscript (Page 13 Line 4-8): “generalisability remains unknown, for example, to Central and South America where the highest per-capita sales of SSB in the world have been recorded; and China and India where the highest prevalence of T2D is expected.^{1,66} To address limitations typical of observational research and also needs for a policy intervention in different populations, future research should include a randomised trial examining people’s health and behaviours and also a trial examining population impact.”

Discussion

#2-8. The discussion is nicely presented, however please note a number of limitations; the data is from observational studies, interventions would be of interest and add to the findings.

Response: We have extended our discussion about limitations of this study. We have revised the descriptions of limitations about meta-analysis and estimates of population attributable fraction (Page 12 Line 10 to Page 13 Line 8). At the end of the limitations section in the Discussion, we have now commented needs for a trial as follows (Page 13 Line 6-8):

“To address limitations typical of observational research and also needs for a policy intervention in different populations, future research should include a randomised trial examining people’s health and behaviours and also a trial examining population impact..”

#2-9. Have you adjusted for lifestyle factors? As stated earlier, this is unclear. Those who drink SSB may have other behaviours which are detrimental to their health.

Response: We agree that confounding due to lifestyle factors would be concerning. Fortunately, most of the studies reported estimates of associations after adjustment for lifestyle factors. Thus, as far as variables measured, confounding was unlikely to exist in a substantial manner. Table S4 and its footnote now include information on confounding which each study attempted to adjust for.

As we have stated above, confounding by a single set of confounders appeared not to be substantial. Nurses Health Study I and II, Health Professionals’ Follow-up Study, EPIC-InterAct, and Framingham Offspring Study have informed influence of dietary covariates and other lifestyle factors on estimates of associations, but no substantial influence was observed. This has been highlighted as written in the above reply (see response at comment 5 above: #2-5).

#2-10. Table S1; a number of studies were excluded but the table does not make it clear if they were excluded because they were ineligible due to study design or ineligible because the authors did not respond to your contact? If this is the case this is unfortunate.

Response: We have now revised the Table (S1) to clarify which studies were ineligible and which studies were eligible but not included. We identified five cohorts were eligible, but not included in our study, because the authors of each cohort could not respond to our request. We agree that exclusion of these is unfortunate. This is limitation of our meta-analysis, so we have kept the sentence on Page 12 Line 19: “Weakness of meta-analysis includes exclusion of eligible cohorts by lack of information.”

Comments from Reviewer 3

Comments:
TO EDITOR:

This is a manuscript on a very interesting topic: the relationship between consumption of sweet beverages and incidence of type 2 diabetes, one of the epidemics of the XXI century. No previous review or It is especially relevant the estimation of the population attribute fraction for type 2 diabetes due to sweet beverages. However, in addition to adiposity, I think that the authors should take into account in the statistical analysis the consumption of other foods that have been related to protection against diabetes (whole-grain cereals, coffee, tea,

skimmed milk, fruit, vegetables, legumes, nuts and moderate alcohol intake) or induction of diabetes (red meat, processed meat products, eggs and high alcohol intake). Other lifestyle factor related to diabetes is physical activity. In addition, these dietary and lifestyle factors should be included in the discussion.

TO THE AUTHORS:

This is a very interesting systematic review and meta-analysis on the relationship between consumption of sweet beverages and incidence of type 2 diabetes. It is especially relevant the estimation of the population attributable fraction (PAF) for type 2 diabetes due to sweet beverage intake. No previous review or individual study has evaluated PAF. They concluded that habitual consumption of sweet beverages is associated with a greater type 2 diabetes incidence, independently of adiposity. However, the authors should take into account the following in order to improve the quality of the manuscript:

#3-1. According to the conclusions of a landmark study (Nurses') on key factors that may influence the appearance of diabetes mellitus in 84,941 nurses followed up a mean of 16 years (N Engl J Med 2001;345:790), the main factors that induced diabetes were: body weight, physical activity, smoking, moderate consumption of alcohol and four dietary factors (intake of fiber, PUFA/SFA, trans fatty acids and low glycemic load foods). In addition to adiposity, these other factors should be taken into account in the statistical analysis and also included in the discussion.

Response: We appreciate the positive appraisal of our meta-analysis. We agree that the association of beverage consumption with T2D risk should be carefully evaluated with regards to confounding by the other dietary and lifestyle components. Also, we agree that, ideally, our PAF estimation should account for the multiple modifiable risk factors. Here, we have commented for meta-analysis and for PAF separately.

Meta-analysis:

In our meta-analysis, we have largely relied on published estimates. Thus, because each publication did not test confounding effect of each lifestyle component, we would not logistically be able to test each of the possible modifiable risk factors for incident T2D in the way we did for adiposity measures. Instead, given available information in the papers we assessed, we have confirmed the following:

- i) All studies examined the association of beverage consumption with T2D risk after adjustment for socio-demographic factors, lifestyle factors (alcohol, physical activity, and different dietary factors or dietary patterns).
- ii) In the studies examining influence of confounding by different factors, including Nurses' Health Study I, Nurses' Health Study II, and Health Professionals' Follow-up Study, there has been no indication that the associations of beverages with T2D risk were substantially confounded by dietary factors and by the other measured factors.
- iii) However, there was a substantial impact of confounding if all variables were taken into consideration (comparison between crude RR and RR after adjustment for sociodemographic variables, lifestyle factors, and other variables). This information has been now presented in Table 2 of the main manuscript; and Table S4 of Supplementary Materials.

We have now added overall characteristics of potential confounders in the main text (Page 9 Line 15-19) as the following. We believe that this text, Table 2 and Table S4 together capture the relevant information to minimise the concern of confounding in our study.

“Confounding was likely to exist in all of the studies. As would be expected, consumers of ASB tended to be overweight or obese or hypertensive.^{11,39,49,58,65} In longitudinal analysis, all studies statistically adjusted for potential confounders including socio-demographic variables, clinical factors (family history of diabetes or prevalent diseases), and lifestyle factors including a diet (Table S4). None of these factors was identified as a single cause of confounding, according to studies assessing influence of potential confounding in different regression models.^{11,41,48–50,53,58,60,62–64} However, a combination of multiple factors was likely to cause confounding (Table 2, Table S4). After adjustment for multiple potential confounders, RR for SSB was attenuated from 1.25 to 1.18 (32% change); and for ASB, 1.48 to 1.25 (43%). By contrast, the point estimate for fruit juice was shifted upward, from 0.97 to 1.05.”

Please also see our reply to the comments made by the editors (#0-3) and reviewer #2 (#2-5), who also requested us to better describe confounding in our study.

PAF and other modifiable risk factors:

As the reviewer states, we estimated PAF based solely on SSB consumption. We wish to estimate PAF based on multiple dietary exposures. However, to conduct the estimation, we need to identify existing meta-analyses evaluating each lifestyle component before adjustment for adiposity and after adjustment for adiposity and for other lifestyle factors. This would be particularly difficult to achieve. For example, most of past studies evaluating smoking status and alcohol consumption did not adjust for other dietary factors, indicating severe confounding in each of the factors. In the meta-analysis of alcohol consumption and T2D (Balunas et al, Diabetes Care, 2009), 20 studies were used in the meta-analysis, but only 2 studies adjusted for multiple dietary components as potential confounders; and only 4 studies adjusted for physical activity; and only 2 studies adjusted for socioeconomic indicators.

Therefore, unfortunately we would not be able to provide estimates of PAF based on appropriate observational evidence. For our estimation of PAF for T2D owing to SSB consumption, we would like to keep the assumption that other lifestyle factors would be held consistent over time.

In addition, while we strongly agree on the importance of multiple lifestyle factors, we also think that policy interventions will be different between factors. For example, taxation on SSB is considered separately from a policy intervention for physical activity. Thus, PAF for a single dietary component (SSB) is informative as it is, for future consideration of effectiveness research and policy debate.

We would agree that multiple modifiable risk factors should be accounted for in future work on PAF. Considering that PAF was approximately 10%, we have rephrased this point and considered we should also focus on the other 90% of T2D risk. Thus, we have now added the following statement in the Discussion section (Page 14 Line 21-25):

“Despite PAF no more than 15%, efficacy estimates are crucial, as 535 million adults are estimated to have T2D in 2035.⁶⁵ Additionally, the PAF can be interpreted that an intervention reducing SSB only would not reduce a large amount of events, and thus confirms importance of multiple modifiable risk factors, rather than a single dietary component, for the prevention of T2D.”

Please also see our reply to the comments made by the reviewer #1 (#1-6, #1-11), who also requested us to clarify assumption and interpretation of PAF.

#3-2. Diabetes has also been related to high consumption of several foods such as red meat, processed meat products and eggs. On the other hand, consumption of other foods may protect against the development of diabetes such as whole-grain cereals, coffee, tea, skimmed milk, fruit, vegetables, legumes and moderate consumption of alcohol. These other key foods should be included in the analysis.

Response: We agree that, in addition to the factors which Nurses’ Health Study reported, there are many different lifestyle components which we would have to take into consideration. As we replied to the first comment (#3-1) (and #2-5), there is no indication that whether or not we accounted for these dietary factors was likely to influence our results and alter our conclusion.

Moreover, it is logistically infeasible to compile evidence from meta-analysis for each different dietary item for both estimates before adjustment for adiposity and after adjustment for adiposity, because individual studies contributing this meta-analysis did not report such information. Furthermore, it would be challenging to keep high quality of meta-analysis and high quality of estimates of PAF, because quality of some prior publications are of concern (see our reply to #3-1). Therefore, we would like to keep the current presentation in which we have demonstrated the translation of the finding from meta-analysis to estimation of PAF.

Importance of multiple risk factors has been now added to our discussion about clinical implications (recommended by the editors’ committee, #0-7) as follows (Page 14, Line 7-14):

“Although causality has not been established, our findings and available evidence indicate a benefit of lowering SSB consumption for the primary prevention of T2D. In the same context, our findings also imply that consumption of ASB or fruit juice is unlikely to reduce the T2D risk, and these should not be considered as a healthy option of beverages. However, consuming ASB to lower caloric intake and body weight may have clinical benefit for obese or overweight adults.^{6,78} This effect on weight should be considered separately from our study that could not rule out the effect of weight on beverage consumption.⁸ In addition, clinical applications of our finding deserve further appraisal about the

Comments from Reviewer 4

Recommendation:

Comments:

This meta-analysis used standard methodology and was thoroughly conducted.

#4-1. My only comment is about presentation of the ASB and fruit juice results as the authors have mentioned the issues of heterogeneity and potential reverse causation. I can see that authors tried to be careful about this by mentioning multiple times the caveats and caution needed for interpreting the ASB and fruit juice results, but this study has the potential to see a lot of media attention, so one can't be too careful (throughout the entire manuscript). In particular, the first sentence of the discussion lumps all three beverages together. Although the authors did specify that it was about the summary estimates, it is confusing then to see another sentence later saying that the estimates for ASB and fruit juice could be questionable. I think it is better not to mention ASB and fruit juice in the same sentence as SSB, and maybe use something like "Although ASB and fruit juice also shown....., however,(caveats, potential bias).....".

Response: We very much appreciate the suggestion. We have now revised the conclusion of the abstract: "Habitual SSB consumption was associated with greater T2D incidence, independently of adiposity. Although ASB and fruit juice also showed positive associations with T2D incidence, the findings were likely to involve bias. Nonetheless, both ASB and fruit juice were unlikely to be healthy alternatives to SSB for the prevention of T2D. Under assumption of causality, SSB consumption over ten years may be related to a substantial number of cases of new-onset diabetes."

#4-2. Regarding the reverse causation, could the authors check the "table 1" of the original papers to see if those consumed high amounts of ASB were also at high risk (e.g. obesity, physical activity, etc) than those with low ASB consumption?

Response: We agree that, for ASB, descriptive statistics are informative. We have repeated review of the publications that assessed ASB consumption. The key findings were the following:

- i) In the Coronary Artery Risk Development in Young Adults (CARDIA) Study and in Multiethnic Study on Atherosclerosis (MESA), consumption of ASB was significantly higher among college graduated than those with shorter education history (Duffey, *Am J Clin Nutr*, 2012; Nettleton et al., *Diabetes Care*, 2009).
- ii) The opposite trend was present in EPIC-InterAct (The InterAct Consortium, *Diabetologia*, 2013).
- iii) Higher consumption of ASB was associated with higher BMI or waist circumference or both in all studies in which descriptive statistics were available (Health Professionals' Follow-up Study; EPIC-InterAct; occupational cohort in Japan; CARDIA; and MESA).
- iv) EPIC-InterAct and HPFS clearly showed that ASB consumers tended to have history of hypertension. Comparing the group of highest ASB consumption and the lowest, EPIC-InterAct had 20.1% and 17.5%, respectively; and HPFS had 23% and 16%, respectively.
- v) Despite seemingly adults with overweight, obesity or hypertension, adults consuming ASB tended to be younger (except for CARDIA which recruited only young adults).
- vi) Trends of physical activity across ASB consumption were not clear, except that HPFS showed clear positive trend.

We have now added the following statement in the result section describing confounding (Page 9 Line 15-16): "As would be expected, consumers of ASB tended to be overweight or obese and to be hypertensive."^{11,39,49,58,65,}

As descriptive statistics are subject to confounding and would cause biased inference, we have not described more than this statement. However, we have now included comparison between crude RR and multivariable-adjusted RR (Table 2; Table S4). The Methods section has now added "We additionally estimated crude RR

without any adjustment to assess a magnitude of overall confounding.” (Page 7 Line 2-3). The paragraph about description of confounding has now included the following (Page 9 Line 15-22):

“In longitudinal analysis, all studies statistically adjusted for potential confounders including socio-demographic variables, clinical factors (family history of diabetes or prevalent diseases), and lifestyle factors including a diet (Table S4). None of these factors was identified as a single cause of confounding, according to studies assessing influence of potential confounding in different regression models.^{11,41,48–50,53,58,60,62–64} However, a combination of multiple factors was likely to cause confounding (Table 2, Table S4). After adjustment for multiple potential confounders, RR for SSB was attenuated from 1.25 to 1.18 (32% change); and for ASB, 1.48 to 1.25 (43%). By contrast, the point estimate for fruit juice was shifted upward, from 0.97 to 1.05.”

Please also see our responses made in relation to comments from the editors’ committee and the other reviewers (#0-3, #1-1, #2-5, #3-1).