



STUDY PROTOCOL

**REVISED** Investigation of Risk Of Bias due to Unreported and Selectively included results in meta-analyses of nutrition research: the ROBUST study protocol [version 2; peer review: 2 approved]

Matthew J. Page <sup>1</sup>, Lisa Bero<sup>2</sup>, Cynthia M. Kroeger<sup>2</sup>, Zhaoli Dai<sup>2</sup>, Sally McDonald<sup>2</sup>, Andrew Forbes<sup>1</sup>, Joanne E. McKenzie<sup>1</sup>

<sup>1</sup>School of Public Health and Preventative Medicine, Monash University, Melbourne, VIC, 3004, Australia

<sup>2</sup>Charles Perkins Centre, School of Pharmacy, Faculty of Medicine and Health, The University of Sydney, Camperdown, NSW, 2006, Australia

**v2** First published: 16 Oct 2019, 8:1760 (<https://doi.org/10.12688/f1000research.20726.1>)  
 Latest published: 07 Feb 2020, 8:1760 (<https://doi.org/10.12688/f1000research.20726.2>)

**Abstract**

**Background:** Dietary guidelines should be informed by systematic reviews (SRs) of the available scientific evidence. However, if the SRs that underpin dietary guidelines are flawed in their design, conduct or reporting, the recommendations contained therein may be misleading or harmful. To date there has been little empirical investigation of bias due to selective inclusion of results, and bias due to missing results, in SRs of food/diet-outcome relationships.

**Objectives:** To explore in SRs with meta-analyses of the association between food/diet and health-related outcomes: (i) whether systematic reviewers selectively included study effect estimates in meta-analyses when multiple effect estimates were available; (ii) what impact selective inclusion of study effect estimates may have on meta-analytic effects, and; (iii) the risk of bias due to missing results (publication bias and selective non-reporting bias) in meta-analyses.

**Methods:** We will systematically search for SRs with meta-analysis of the association between food/diet and health-related outcomes in a generally healthy population, published between January 2018 and June 2019. We will randomly sort titles and abstracts and screen them until we identify 50 eligible SRs. The first reported meta-analysis of a binary or continuous outcome in each SR (the 'index meta-analysis') will be evaluated. We will extract from study reports all study effect estimates that were eligible for inclusion in the index meta-analyses (e.g. from multiple instruments and time points) and will quantify and test for evidence of selective inclusion of results. We will also assess the risk of bias due to missing results in the index meta-analyses using a new tool (ROB-ME).

**Ethics and dissemination:** Ethics approval is not required because information will only be extracted from published studies. Dissemination of the results will be through peer-reviewed publications and presentations at conferences. We will make all data collected from this study publicly available via the Open Science Framework.

**Open Peer Review**

**Reviewer Status** ✓ ✓

|   | Invited Reviewers |             |
|---|-------------------|-------------|
|   | 1                 | 2           |
| <b>version 2</b><br>(revision)<br>07 Feb 2020 | ✓<br>report       |             |
|   | ↑                 |             |
| <b>version 1</b><br>16 Oct 2019               | ?<br>report       | ✓<br>report |

- Robbie C.M. van Aert** <sup>1</sup>, Tilburg University, Tilburg, The Netherlands
- Joshua Wallach**, Yale School of Public Health, New Haven, USA

Any reports and responses or comments on the article can be found at the end of the article.

**Keywords**

Bias, Publication bias, Systematic reviews, Diet, Food and Nutrition

**Corresponding author:** Matthew J. Page ([matthew.page@monash.edu](mailto:matthew.page@monash.edu))

**Author roles:** **Page MJ:** Conceptualization, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing; **Bero L:** Conceptualization, Funding Acquisition, Methodology, Writing – Review & Editing; **Kroeger CM:** Methodology, Writing – Review & Editing; **Dai Z:** Methodology, Writing – Review & Editing; **McDonald S:** Methodology, Writing – Review & Editing; **Forbes A:** Methodology, Writing – Review & Editing; **McKenzie JE:** Conceptualization, Funding Acquisition, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing

**Competing interests:** No competing interests were disclosed.

**Grant information:** This research is funded by an Australian National Health and Medical Research Council (NHMRC) Project Grant [APP1139997]. MJP is supported by an Australian Research Council Discovery Early Career Researcher Award (DE200101618). JEM is supported by an NHMRC Career Development Fellowship [APP1143429]. SM is supported by the Country Women's Association (NSW) and Edna Winifred Blackman Postgraduate Research Scholarship.

*The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*

**Copyright:** © 2020 Page MJ *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**How to cite this article:** Page MJ, Bero L, Kroeger CM *et al.* **Investigation of Risk Of Bias due to Unreported and Selectively included results in meta-analyses of nutrition research: the ROBUST study protocol [version 2; peer review: 2 approved]** F1000Research 2020, **8**:1760 (<https://doi.org/10.12688/f1000research.20726.2>)

**First published:** 16 Oct 2019, **8**:1760 (<https://doi.org/10.12688/f1000research.20726.1>)

**REVISED Amendments from Version 1**

In this version of our study protocol, we have provided more details regarding our search strategy and methods for analysing whether particular characteristics modify the Potential Bias Index, and have included a link to a worked example of the potential impact of selective inclusion of results on meta-analyses.

**Any further responses from the reviewers can be found at the end of the article**

**Introduction**

Suboptimal diet is one of the leading contributors to mortality and morbidity globally<sup>1</sup>. Dietary guidelines, which provide recommendations on types and amounts of foods to consume and dietary patterns to adopt, are developed with the aim of reducing non-communicable disease attributable to diet. Systematic reviews (SRs) of the available scientific evidence often underpin dietary guidelines<sup>2</sup>. However, if the SRs are flawed in their design, conduct or reporting, the recommendations contained therein may be misleading or harmful.

Bias in the results of SRs can arise through various processes<sup>3</sup>. Systematic reviewers often face a multiplicity of results for particular outcomes in the included studies (e.g. there may be results for weight loss at multiple time points, each of which are presented unadjusted and adjusted for prognostic factors, such as age and sex)<sup>4,5</sup>. When multiple results for an outcome are available in study reports, systematic reviewers' choice about which result to include in a meta-analysis may be influenced by the P value, magnitude or direction of the result – this is known as 'selective inclusion of results'<sup>6</sup>. For example, instead of including the result for weight loss that arose from the time point considered *a priori* to be the most clinically important, or which was adjusted for the most appropriate set of prognostic factors, systematic reviewers may select a result simply because it was the largest in magnitude, or had the smallest P value.

Bias in meta-analyses can arise not only when systematic reviewers selectively include results, but also when results of some eligible studies are unavailable for inclusion<sup>7</sup>. There is extensive evidence that shows many studies are never published<sup>8</sup>, and that those studies that are published tend to have larger effect estimates than unpublished studies<sup>9</sup>. The term 'publication bias' has often been used to describe this problem. In addition, published studies often omit results for some of the outcomes that were measured<sup>10</sup>, with reported results more likely to be statistically significant than non-reported results<sup>11</sup>. The terms 'selective outcome reporting' and 'outcome reporting bias' have been used to describe this problem, but we prefer the term 'selective non-reporting bias' as it emphasises the *non-reporting* of study results. Regardless of whether an entire study report or a particular study result is unavailable selectively (e.g. because the P value, magnitude or direction of the results were considered less favourable by the investigators), the consequence is bias in a meta-analysis because available results differ systematically from missing results. The term 'bias due to missing results' has recently been coined to describe the

bias in meta-analyses that arises from non-publication or non-reporting of study results<sup>12</sup>.

There has been little empirical investigation of bias due to selective inclusion of results and bias due to missing results in SRs with meta-analyses of the association between food/diet and health-related outcomes. The only known investigation of selective inclusion of results focused on meta-analyses of randomized trials of interventions for arthritis or depressive or anxiety disorders<sup>13</sup>. Also, previous assessments of reporting biases in nutrition research have been limited to an exploration of publication bias in meta-analyses of the association between diet and cardiovascular disease or mortality<sup>14</sup>. There has been no formal investigation of selective non-reporting of results in studies included in meta-analyses of food/diet-outcome relationships.

Therefore, the aim of this research is to investigate various biases in SRs with meta-analyses of the association between food/diet and health-related outcomes. The objectives are to explore:

- (i) whether systematic reviewers selectively included study effect estimates in meta-analyses when multiple effect estimates were available;
- (ii) what impact selective inclusion of study effect estimates may have on meta-analytic effects, and;
- (iii) the risk of bias due to missing results in meta-analyses.

**Methods****Overview of the study**

We will systematically search for SRs with meta-analysis of the association between food/diet and health-related outcomes in a generally healthy population, published between January 2018 and June 2019. We will randomly sort titles and abstracts and screen them until we identify 50 eligible SRs. The first reported meta-analysis of a binary or continuous outcome in each SR (which we refer to as the 'index meta-analysis') will be evaluated. We will extract from study reports all study effect estimates that were eligible for inclusion in the index meta-analyses (e.g. from multiple instruments and time points), and will calculate a statistic – the Potential Bias Index<sup>15</sup> – to quantify and test for evidence of selective inclusion of results. The risk of bias due to missing results (arising from publication bias and selective non-reporting bias) in the index meta-analyses will also be assessed using a new tool (the Risk Of Bias due to Missing Evidence (ROB-ME) tool<sup>12</sup>).

**Eligibility criteria for SRs**

We will seek a sample of SRs with meta-analysis meeting the following criteria:

- includes studies of people of any ages (i.e. infants, children, adolescents, adults or elderly people) and backgrounds in the generally healthy population, including pregnant and breastfeeding women and people with common diet-related risk factors such as being overweight or having high blood pressure;

- includes randomized trials or non-randomized studies evaluating the effects of at least one type of food (e.g. whole grains, fruit) or at least one food-defined dietary pattern (e.g. high intake of processed meat) on any binary (e.g. mortality) or continuous (e.g. weight) health-related outcome;
- published from 1 January 2018 to 30 June 2019;
- written in English;
- includes citations for all studies included in the SR; and
- presents the summary statistics or effect estimate and its precision (e.g. 95% confidence interval) for each included study, and the meta-analytic effect estimate and its precision, for at least one meta-analysis of a binary or continuous outcome.

We will adopt the definition of “systematic review” used in the 2019 edition of the *Cochrane Handbook for Systematic Reviews of Interventions*: “A systematic review attempts to synthesize all empirical evidence that fits pre-specified eligibility criteria in order to answer a specific research question. It uses explicit, systematic methods that are selected with a view to minimizing bias, thus providing more reliable findings from which conclusions can be drawn and decisions made”<sup>16</sup>. We will use the criteria adopted by Page *et al.* to identify articles as SRs with meta-analysis, i.e. we will include articles with explicitly stated methods of study identification (e.g. a search strategy), explicitly stated methods of study selection (e.g. eligibility criteria and selection process), and a meta-analysis of study results<sup>17</sup>. We will not exclude articles based on the level of detail about the methods provided (e.g. articles with a line-by-line Boolean search strategy or just a list of the key words used in the bibliographic databases will both be considered to meet the criteria for an SR).

We will exclude:

- SRs that did not include any meta-analysis of a binary or continuous outcome;
- meta-analyses or pooled analyses of studies conducted outside the context of a SR (e.g. when individual participant data are combined from a set of cohort studies outside the context of a systematic review, e.g. see Zhong *et al.*<sup>18</sup>);
- SRs that focus only on nutrient-specific associations with outcomes (e.g. those examining the effects of single nutrients such as folic acid, salt).
- SRs including studies that were restricted to people with a health condition (e.g. type II diabetes), people who are obese, or frail elderly people who are at risk of malnutrition (by ‘frail’ we mean a person who, following a minor stress, experiences a large deterioration in function and does not return to baseline homeostasis<sup>19</sup>); and
- SRs that were co-authored by a member of the investigator team, because the assessment of bias may be influenced by the investigators’ prior involvement in the SR.

### Search for SRs

We will identify eligible SRs by searching PubMed and Epistemonikos (a database of SRs and other syntheses of health research that have been identified via systematic searches of the Cochrane Database of Systematic Reviews, PubMed, EMBASE, CINAHL, PsycINFO, LILACS, DARE, Campbell Library, JBI Database and EPPI-Centre Evidence Library). In PubMed, we will use the search strategy for diet- or nutrition-related studies developed and validated by Durao *et al.*<sup>20</sup>, combine this strategy with the search string, “meta-analysis[pt] OR meta-analy\*[ti]” to identify meta-analyses, and restrict the search from 1 January 2018 to 30 June 2019 (search strategy as extended data<sup>21</sup>). Various search filters designed to retrieve SRs exist, however they are designed to retrieve SRs with or without meta-analyses, and have high sensitivity given the lack of consensus on what constitutes a SR. We have decided to use a more targeted PubMed search strategy (that is, articles will need to have been classified by the PubMed indexers using the “Meta-Analysis” [Publication Type] MeSH term, or include the terms “meta-analysis”, “meta-analyses”, “meta-analytic” or other variants in the title) given we are including SRs only if they present a meta-analysis. In Epistemonikos, we will use the following search strategy: (title:(nutri\* OR diet\*)) OR abstract:(nutri\* OR diet\*)) and limit the search from 1 January 2018 to 30 June 2019.

### Selection process

**Selection of SRs.** We will export titles and abstracts into Microsoft Excel, remove all duplicate records, and randomly sort the remaining records. In the piloting phase, four investigators (MJP, CMK, ZD and SM) will independently assess 50 abstracts against the inclusion criteria (rating each as ‘Eligible’, ‘Ineligible’, or ‘Unsure’), to ensure the criteria are applied consistently. Following piloting, two investigators (MJP and one of CMK, ZD or SM) will independently screen titles and abstracts of 450 records. The full text of records rated as ‘Eligible’ or ‘Unsure’ will then be retrieved and assessed independently against the inclusion criteria by two investigators (MJP and one of CMK, ZD or SM). We will repeat this process (screening batches of 500 records) until the target sample of 50 SRs is met. Any discrepancies in screening decisions at each stage will be resolved via discussion, or by consultation with another investigator (JM or LB) where necessary.

**Selection of index meta-analyses.** From each SR one investigator (MJP) will select one meta-analysis of a binary or continuous outcome for assessment, stratifying the selection so that the final sample includes 25 binary and 25 continuous outcome meta-analyses. The selected meta-analysis will be the first meta-analytic result that is presented in the SR, and henceforth is referred to as the ‘index meta-analysis’. The index meta-analysis may be selected from the abstract, summary of

findings table, or results section of the SR, depending on where the meta-analytic result is first reported in the publication. We have opted for the first meta-analytic result reported in the review because the first meta-analysis is likely to be the most (or one of the most) important analysis in the review on which the conclusions of the review are based. We will include meta-analyses regardless of the outcome domain measured (e.g. all-cause mortality, diabetes incidence), meta-analytic effect measure (e.g. odds ratio, mean difference), meta-analytic model (e.g. fixed-effect, random-effects), approach to modelling (i.e. frequentist or Bayesian) and number and type of included studies (i.e. randomized trial or non-randomized study). We will select only standard pairwise meta-analyses of aggregate data, not dose-response meta-analyses, network meta-analyses or meta-analyses of individual participant data.

We will retrieve all reports of studies that were included in each index meta-analysis, as cited by the systematic reviewers. Study reports could include journal articles, conference abstracts, dissertations, trial results posted in trials registers, or any other reports (e.g. government reports). If more than one reference for a study was cited by the systematic reviewers (e.g. a study was reported in multiple journal articles, or in a journal article and a conference abstract), we will retrieve all references cited. If study reports are written in languages other than English, we will attempt to translate them using Google Translate; we will exclude reports if the translation is not interpretable. We will also retrieve reports of studies that were included in the review but excluded from the meta-analysis by the systematic reviewers, to explore whether any eligible outcome data may have been missed from these reports or potentially excluded because of the nature of the results (e.g. statistical non-significance). We also expect that in some systematic reviews, citations to 'near-miss' studies along with reasons for exclusion, will be provided. For these reviews, we will scan the reasons for exclusion, and where these reasons indicate there was no useable data, we will retrieve the study to confirm that no data were available for inclusion in the review.

### Data collection and management

We will collect data using a standardised form with detailed guidance created in REDCap (Research Electronic Data Capture), a secure, web-based software platform designed to support data capture for research studies<sup>22,23</sup>. Four investigators (MJP, CMK, ZD and SM) will initially pilot the data collection form on a random sample of five index meta-analyses and all of their included studies, to ensure consistency in the data collection. Any discrepancies will be discussed amongst all four investigators, and the form and guidance will be revised as necessary. Following piloting, two investigators (MJP and one of CMK, ZD or SM) will independently collect data from all remaining index meta-analyses and included studies. Any discrepancies between the data extracted will be resolved through discussion or adjudication by a third investigator (JM or LB) if necessary.

**Data items to describe the general characteristics of the SRs.** We will record the following characteristics of each of the included SRs:

- Journal it is published in;
- Year of publication;

- Country of corresponding author of the SR;
- Whether or not a registration record (e.g. PROSPERO) or protocol for the SR was mentioned in the paper;
- Source of funding of the SR, classified as: 'non-profit', 'for-profit', 'mixed', 'no funding', or 'not reported' (for funding sources classified as 'for-profit' or 'mixed' we will record the name of the for-profit funder and classify each as 'food industry' or 'other industry');
- Conflicts of interest of systematic reviewers as disclosed in the SR report, classified as: 'conflict of interest present' when at least one systematic reviewer reported a financial conflict of interest of any type, excluding current study funding or industry employment; 'no conflict of interest' if all systematic reviewers stated they had no conflicts; and 'missing' if there was no disclosure statement<sup>24</sup>; and
- Affiliation of the corresponding author of the SR, classified as 'industry' or 'non-industry' or 'mixed'.

**Data items to describe the general characteristics of the index meta-analyses.** We will record the following characteristics of each of the index meta-analyses:

- Number of studies included in the meta-analysis;
- Number of participants included in the meta-analysis;
- Type of population investigated (i.e. participants that were eligible for inclusion in the index meta-analysis, as specified by the systematic reviewers);
- Type of interventions or exposures investigated;
- Type of studies included in the meta-analysis (classified as randomized trial or non-randomized study or both; we will also classify the type of non-randomized study as specified by the systematic reviewers, e.g. 'non-randomized trial', 'cohort study', 'case-control study');
- Outcome domain (e.g. cancer mortality, weight);
- Outcome primacy label provided by the systematic reviewers (classified as 'primary', 'secondary' or 'not labelled');
- Effect measure (e.g. odds ratio or mean difference); and
- Meta-analysis model (fixed-effect, fixed-effects or random-effects).

**Data items to evaluate selective inclusion of results in index meta-analyses.** To explore whether systematic reviewers selectively included study results in meta-analyses when multiple study results were available (e.g. included data at one time point ahead of another because the former was statistically significant), we need to determine which results were eligible for inclusion in the meta-analyses. Therefore, from each SR and its corresponding SR protocol (if available), we will extract descriptions of any eligibility criteria to select effect estimates to include in the index meta-analysis. Eligibility criteria



comprise lists of intervention groups, measurement instruments, time points and analyses that were eligible for inclusion (e.g. systematic reviewers state that results that were either unadjusted or adjusted for at least one prognostic factor would be included in meta-analyses). We will also extract any decision rules to select effect estimates to include in the index meta-analysis. Decision rules comprise strategies to either select one effect estimate or combine effect estimates when multiple were available (e.g. systematic reviewers state that if the effects of multiple levels of red meat intake were available, only the contrast between the highest and lowest intake would be included in the meta-analysis).

We will classify each eligibility criterion and decision rule as a strategy to handle results arising from:

- multiple measurement instruments;
- multiple definitions/diagnostic criteria for an event;
- multiple cut-points on a continuous outcome measure;
- multiple time points;
- multiple intervention groups;
- final and change from baseline values;
- multiple analysis samples (e.g. intention-to-treat, per-protocol and as-treated);
- unadjusted and covariate-adjusted analyses;
- period and paired analyses in crossover trials;
- multiple information sources (e.g. journal article and trials results register);
- overlapping samples of participants (e.g. men only and older adults only); or
- another source of multiplicity<sup>25</sup>.

We will also collect from each SR information about the study data included in the index meta-analysis. Such information will include the summary statistics for each group, and the effect estimate, measure of precision (e.g. standard error or confidence interval), and direction of the effect estimate for each included study, as displayed on the forest plot or in a table/text. We will also record whether systematic reviewers declared that study outcome data:

- (i) were obtained from the study investigators because the data were not reported in the study publication;
- (ii) required some algebraic manipulation of statistics in order to include the data in the meta-analysis (e.g. calculating a standard deviation from a 95% confidence interval of the mean);
- (iii) originated from a report written in a language other than English which the systematic reviewers had translated into English, or;

- (iv) required a method of imputation (such as imputing a missing standard deviation).

We will collect from study reports outcome data that could potentially be included in the index meta-analysis, according to the eligibility criteria and decision rules specified in the SR protocol, and with how the outcome was specified in the SR. By 'outcome data' we mean summary statistics (e.g. number of events, sample sizes) or an effect estimate (e.g. odds ratio) and some measure of precision (e.g. standard error, 95% confidence interval), or both if available. If no SR protocol is available, we will assume that no eligibility criteria and decision rules were pre-specified, even if some were reported in the published SR ('worst-case scenario' assumption), and extract all study outcome data based on how the outcome was specified in the SR. For example, a meta-analysis with the outcome 'reduction in cardiovascular disease risk (Framingham Risk Score)', that had no SR protocol, will have all available results for this particular outcome extracted from each study (e.g. at all time points, unadjusted and covariate-adjusted analyses, regardless of whether decision rules for these measures/analyses were stated in the published SR); however, no results for any other outcomes (e.g. using an alternative cardiovascular disease risk outcome calculated with a different algorithm) will be extracted.

When studies of more than two groups are encountered and each comparison is eligible for inclusion in a meta-analysis, systematic reviewers need to use a method that avoids multiple counting of participants. Systematic reviewers may choose to:

- (i) include data from only one of the experimental intervention/exposure groups and the control group;
- (ii) combine the two experimental intervention/exposure groups (e.g. sum the number of events across both intervention/exposure groups for binary outcomes or calculate the mean values for both experimental intervention/exposure groups for continuous outcomes), and compare this to the control group, or;
- (iii) include data from each experimental intervention/exposure group as separate comparisons in the meta-analysis by dividing the sample size of the control group by the number of comparisons.

If systematic reviewers pre-defined a method to deal with multi-arm studies, we will follow that method when extracting data. If systematic reviewers did not pre-define a method to deal with multi-arm studies, and:

- (i) selected one of the experimental intervention/exposure groups to include in the meta-analysis, we will extract the data required to calculate effect estimates for two comparisons: (a) experimental intervention/exposure A versus control, and (b) experimental intervention/exposure B versus control;
- (ii) combined the two experimental intervention/exposure groups, we will extract the data required to calculate effect estimates for three comparisons:

(a) experimental intervention/exposure A versus control, (b) experimental intervention/exposure B versus control, and (c) combination of intervention/exposure A and B versus control;

- (iii) included multiple comparisons in the meta-analysis by dividing the control group in half, we will extract the data required to calculate effect estimates for two comparisons: (a) experimental intervention A versus control, and (b) experimental intervention B versus control, where for both comparisons the control group sample size will be halved.

The three methods above will be used when dealing with three-arm studies. We will extend these methods when there are more than three arms in a study.

All study effect estimates included in mean difference meta-analyses must be in units of one particular scale, although estimates can comprise a mixture of final values and change from baseline values. In contrast, the measurement scale units of study effect estimates included in standardised mean difference (SMD) meta-analyses can vary, although it is recommended that all estimates are final values, or change from baseline values, not a mixture. Therefore, for mean difference meta-analyses we will extract only data for the particular scale included by the systematic reviewer, but extract final and change from baseline values when available. For SMD meta-analyses that included final values, we will extract final values only for any relevant measurement scale (and vice versa for SMD meta-analyses that included change from baseline values).

We will only extract study outcome data that were reported completely, defined as reporting sufficient data for inclusion in a meta-analysis (i.e. reporting of an effect estimate and a measure of precision, or summary statistics that enable calculation of these); we will not request unpublished data (e.g. missing number of events) from study authors.

### Assessment of risk of bias due to missing results in index meta-analyses

We will assess the risk of bias due to missing results in each index meta-analysis using the ROB-ME (“Risk Of Bias due to Missing Evidence”) tool, introduced in the 2019 edition of the *Cochrane Handbook for Systematic Reviews of Interventions*<sup>12</sup>. The ROB-ME tool is the first structured approach for assessing reporting biases in meta-analyses that considers both publication bias and selective non-reporting bias. Users first consider whether results known or presumed to have been generated by study investigators are unavailable for any of the included studies (e.g. by cross-checking what was pre-specified with what was reported by study authors). They then consider whether a meta-analysis is likely to be biased because of the unavailable results in the studies identified. Finally, users consider whether qualitative signals (e.g. non-comprehensive search) and the pattern of observed results suggest that additional results are likely to be missing systematically from the meta-analysis. The tool includes signalling questions,

which aim to elicit information relevant to an assessment of risk of bias. Responses to the signalling questions feed into algorithms developed to guide users of the tool to judgements about risk of bias. The possible risk-of-bias judgements are: (i) low risk of bias, (ii) some concerns, and (iii) high risk of bias.

Four investigators (MJP, CMK, ZD and SM) will independently perform ROB-ME assessments on each of the index meta-analyses. We will assign the four assessors to pairs and ask them to reach consensus on their ROB-ME assessments. Any discrepancies between the responses to signalling questions and risk-of-bias judgements that cannot be resolved via discussion will be adjudicated by a third investigator (JM or LB).

### Data analysis

**Descriptive analysis.** We will calculate descriptive summary statistics of the general characteristics of SRs and index meta-analyses. For categorical variables, we will present frequencies and percentages. For continuous variables, we will present means (with standard deviations) and medians (with interquartile ranges). We will calculate the frequencies and percentages of SR protocols and SRs reporting the different types of eligibility criteria and decision rules to select study effect estimates. We will calculate the proportion of studies that had multiple results available for inclusion in the index meta-analyses, and quantify the number of study effect estimates that were eligible for inclusion in the index meta-analyses, and the number of eligible effect estimates per study.

**Analysis of selective inclusion of results.** We will follow the analyses used in a previous investigation of selective inclusion of results, as described by Page *et al.*<sup>13,15</sup>. We will calculate a statistic (called the ‘Potential Bias Index’ (PBI)) to quantify and test for evidence of selective inclusion. In brief, this index is based on ordering the effect estimates in each study based on their magnitude and direction of effect, and then determining the position within that order where the effect estimate included in the index meta-analysis sits. The PBI is the weighted average rank position of the selected effect estimates, where the weights are the inverse of the number of effect estimates available per study. This weighting system therefore attributes greater priority to the rank positions of effect estimates where there are a larger number of effect estimates to choose from. The expression for PBI is:

$$PBI = \frac{\sum_{i=1}^k \frac{n_i(X_i - 1)}{n_i - 1}}{\sum_{i=1}^k n_i}$$

where there are  $k$  studies,  $n_i$  is the number of effect estimates in study  $i$ , and  $X_i$  is the rank of the selected effect estimate in study  $i$ . Derivation of the PBI, and a worked example, are provided in Page *et al.*<sup>15</sup>.

The PBI ranges from 0 to 1. For meta-analyses comparing an experimental intervention/exposure with no

intervention or placebo control, the PBI will have the value 1 when the effect estimate that is most favourable to the experimental intervention/exposure is always selected for inclusion from each study. By “most favourable” we mean the effect estimate that suggested the most benefit or least harm of the intervention/exposure. Conversely, the PBI will have the value of 0 when the effect estimate that is least favourable to the experimental intervention/exposure is always selected. For meta-analyses comparing different levels or patterns of intake of the same food (e.g. wholegrain bread consumed 5 days per week versus wholegrain bread consumed once a week), we will determine from the text of the review whether the systematic reviewers hypothesised that the higher or lower category would have the most benefit or least harm, and rank study effect estimates based on their favourability to the category of consumption considered to be most beneficial/least harmful. If we cannot determine the hypothesis of the systematic reviewers, we will exclude the meta-analysis from the PBI analyses. For meta-analyses comparing different foods/diets (e.g. vegan versus vegetarian diet), we will determine from the text of the review which intervention/exposure the systematic reviewers were most interested in evaluating (which we will consider the experimental intervention/exposure), and rank the study effect estimates based on their favourability to the experimental intervention/exposure. We will also perform a sensitivity analysis excluding meta-analyses comparing different foods/diets to examine the impact on the PBI.

Several methods for selecting effect estimates are acceptable in terms of not introducing bias, including (i) randomly selecting effect estimates, (ii) selecting effect estimates based on some clinical or methodological rationale or (iii) selecting the median effect estimate<sup>25</sup>. If systematic reviewers employed selection methods ii and iii across the studies, we expect that the distribution of the selected effect estimates would be consistent with what we would observe under purely random selection, so on average, the selected effect estimates would be at the middle rank position and the PBI would take the value of 0.5. A PBI of 0.5 therefore suggests that there is no selective inclusion of the most or least favourable effect estimates. We will run a statistical test based on the PBI that has been constructed to test whether the observed selection of effect estimates is consistent with randomness of selection<sup>15</sup>. Confidence intervals (95%) for the PBI will be obtained by bootstrap resampling<sup>26</sup>.

For meta-analyses of binary outcomes, we will express all study effect estimates in terms of odds ratios (ORs) to enable ranking of them on the same metric. For meta-analyses of continuous outcomes, we will express all study effect estimates in terms of SMDs to enable ranking of them on the same metric. In addition, we will standardise the direction of effects so that ORs below 1 or SMDs below 0 represent effects that are more favourable to the experimental intervention/exposure. We will exclude from all PBI analyses index meta-analyses that included no studies with multiplicity of effect estimates, given there is no potential for selective inclusion of results in such meta-analyses.

We will also investigate the impact of any potential selective inclusion of study effect estimates on the magnitude of the resulting meta-analytic ORs and SMDs. For each of the meta-analyses of binary outcomes, we will calculate all possible meta-analytic ORs from all combinations of available study effect estimates. When the number of possible combinations is prohibitively large to calculate all combinations (i.e. >30,000), we will generate a random sampling distribution of 5,000 meta-analytic ORs. Each meta-analytic OR will be created by randomly selecting (with equal probability) an effect estimate for inclusion from each study comparison, and meta-analysing the chosen effects. For each distribution of generated meta-analyses, we will calculate (i) the percentile rank of the index meta-analytic OR; (ii) the median of all possible meta-analytic ORs, which represents the median of a distribution where study effect estimates were not selectively included, and (iii) the difference between the index meta-analytic OR and the median meta-analytic OR. When the difference between the index and median meta-analytic OR is minimal, we will conclude that any potential selective inclusion had a limited impact on the meta-analytic effect (worked example as extended data<sup>21</sup>). We will use non-parametric statistics to describe these differences. We will also synthesise these differences using a random-effects meta-analysis model, with the meta-analytic weights based on the variance of the index meta-analytic OR estimate and the between-study variability estimated using the restricted maximum likelihood estimator. The Hartung-Knapp-Sidik-Jonkman confidence interval method will be used to calculate uncertainty in the combined differences<sup>27,28</sup>. We will repeat the above analyses for each of the meta-analyses of continuous outcomes by calculating meta-analytic SMDs rather than ORs. We will also convert all SMDs to log ORs by multiplying the SMDs by  $\pi/\sqrt{3} = 1.814$ <sup>29</sup>, and will re-run the analyses including all 50 meta-analyses of ORs. We will quantify statistical inconsistency using the  $I^2$  statistic<sup>30</sup> along with a 95% confidence interval.

We will conduct a fixed-effect meta-analysis of the PBI obtained in the current study with that estimated in the previous study by Page *et al.*<sup>13</sup>. We will synthesise estimates of the PBI using a fixed-effect meta-analysis model because the number of included studies ( $n=2$ ) will be too small to adequately estimate the between-study variance.

We will conduct subgroup analyses to explore whether the availability of an SR protocol or registration record; the SR being funded by the food industry; and the SR having at least one author disclosing a financial conflict of interest of any type, modifies the PBI. The confidence limits and P value for the difference in PBI between subgroups will be constructed using bootstrap methods<sup>26</sup>, because statistical theory does not currently exist for the distribution of the difference between two PBIs. Specifically, the steps will be to (1) draw a bootstrap sample of trials from each subgroup (e.g. separate samples from ‘SRs funded by the food industry’ and ‘SRs not funded by the food industry’), (2) calculate the PBI for each subgroup, and (3) calculate the difference in the PBIs across the subgroups. This process will be repeated 2000 times. The confidence limits for the



difference in PBI between subgroups will be the 2.5<sup>th</sup> and 97.5<sup>th</sup> percentiles of the bootstrap distribution of differences. The P value will be identified through an iterative search for the confidence level of the differences that touches the null value, where the P value will be calculated as 1 – confidence interval level.

A similar approach will be used to examine whether the PBI is modified by the number of available effect estimates. The approach will only deviate at step 3 where a regression of PBI on the number of available effect estimates will be fitted. The regression coefficient for the estimate of linear association will be stored. The bootstrap distribution of these regression coefficients will be used to calculate the confidence limits and P value as described above.

We will undertake a series of sensitivity analyses to investigate whether the PBI is robust to certain assumptions. For SRs without protocols or registration records, we will not be able to determine whether the eligibility criteria and decision rules to select results in the methods section of the SR were developed prior to or while undertaking the SR. Therefore, in these SRs, our primary calculation of the PBI will be based on the set of study effect estimates that were compatible with the assumption of ‘no pre-specified eligibility criteria or decision rules’. However, we will perform a sensitivity analysis where study effect estimates that were compatible only with the eligibility criteria and decision rules in the methods sections of the SR are included, to examine if this affected the PBI.

In our primary analysis of meta-analyses of continuous outcomes, we will convert all study effect estimates to SMDs to allow us to calculate the PBI in circumstances where multiple effect estimates were available for the same outcome domain, but measured on different scales. However, there is not necessarily a one-to-one relationship between the rank positions of effect estimates based on the mean difference and SMD (because the SMD additionally depends on the pooled standard deviation). Therefore, in a sensitivity analysis we will calculate the PBI based on the rank positions of the mean difference for the subset of study effect estimates that were measured on the same scale as the effect estimate included in the index meta-analysis. This will allow us to assess more accurately whether systematic reviewers had selectively included study effect estimates based on the magnitude of the mean difference in raw measurement scale units.

We anticipate that in some study reports, only an effect estimate and its standard error or 95% confidence interval will be presented (that is, the number of events or means and standard deviations per group will not be available). In this circumstance, to include the result in a meta-analysis, algebraic manipulation will be required. Algebraic manipulation may be considered challenging by some systematic reviewers, so effect estimates requiring algebraic manipulation may not have been considered by reviewers in the set of effect estimates to potentially include in the meta-analysis. For the primary calculation of the PBI, we will exclude study effect estimates that required algebraic manipulation; however, we will undertake a sensitivity

analysis to explore whether the PBI is modified when we include these study effect estimates.

Finally, in our primary analysis of investigating the impact of any potential selective inclusion of study effect estimates on meta-analytic effect estimates, we will use the random-effects meta-analysis model to pool effect estimates when calculating the distribution of possible meta-analytic effects. We will also perform a sensitivity analysis to explore whether our primary analysis is modified when the distribution of meta-analytic effect estimates are calculated using a fixed-effect model.

**Analysis of risk of bias due to missing results.** We will calculate the agreement in responses to signalling questions and risk-of-bias judgements for the ROB-ME tool for consensus assessments across the pairs of reviewers using the weighted Kappa statistic and percentage agreement (both metrics will be presented with 95% confidence intervals)<sup>31</sup>. We will interpret Kappa values as poor ( $\leq 0.00$ ), slight (0.01–0.20), fair (0.21–0.40), moderate (0.41–0.60), substantial (0.61–0.80), or almost perfect (0.81–1.00)<sup>32</sup>.

After reaching consensus on ROB-ME judgements across the reviewer pairs, we will calculate the frequency and percentage of index meta-analyses rated at ‘low risk of bias’, ‘high risk of bias’ or ‘some concerns’. We will conduct subgroup analyses to explore whether the SR being funded by a for-profit company, and the SR having at least one author reporting a conflict of interest of any type, were associated with the index meta-analysis being rated at high risk of bias due to missing results.

### Software

We plan to use **Stata** version 15 software<sup>33</sup> to conduct all analyses.

### Sample size

The sample size of 50 SRs with meta-analysis was primarily selected for feasibility reasons given our available resources. This was informed by the time taken to search, screen, extract data and undertake the analysis in a previous similar study<sup>13</sup>. This sample size will allow estimation of the PBI to within a margin of error of  $\pm 0.05$ , assuming each meta-analysis includes an average of 8.1 studies, with 2.2 effect estimates per study (as observed in Page *et al.*<sup>13</sup>).

### Study status

We have run the searches, piloted the screening form and started screening titles and abstracts against the eligibility criteria.

### Dissemination of information

Dissemination of the results will be through peer-reviewed publications and presentations at conferences. We will make all data collected from this study publicly available via the Open Science Framework.

### Discussion

To our knowledge, this is the first study to investigate bias due to selective inclusion of results and bias due to missing results

in SRs with meta-analyses of the association between food/diet and health-related outcomes. Our study will address several aspects of selective inclusion of results that have not yet been explored, including the extent to which it occurs in meta-analyses of binary outcomes and in meta-analyses of non-randomized studies, and whether the practice is associated with funding source and conflicts of interest of the systematic reviewers. Our study will also provide the first evaluation of the measurement properties of the recently developed ROB-ME tool, which should provide insight into any possible revisions that may need to be made to the tool.

Our study has several strengths. We have pre-specified methods to identify, select and collect data from eligible SRs and studies, and will declare any modifications to the protocol in the final report. We will use systematic methods to identify eligible SRs with meta-analyses, including use of explicit inclusion criteria, sensitive search strategies, duplicate selection and collection of data from SRs and studies, and standardised and pilot-tested data collection forms.

There are also some limitations of our planned methods. Most of the studies included in the index meta-analyses are unlikely to have been registered or have analysis plans available, which will make it challenging to detect selective non-reporting of results reliably. For example, in a cross-sectional analysis of 264 randomized trials of nutrition and dietetics interventions published in 2016, only 62 (24%) were registered prospectively<sup>34</sup>; the proportion of non-randomized nutrition studies that are prospectively registered is likely to be far lower. In addition, our investigations of whether selective inclusion of results and risk of bias due to missing results is associated with conflicts of interest of systematic reviewers relies on systematic reviewers declaring such interests in the SR report; however, previous research suggests that the level of conflict of interest disclosure in nutrition research articles is low<sup>35,36</sup>.

## Conclusion

Meta-analyses of nutrition research underpin the recommendations of dietary guidelines. Therefore, it is essential that the findings of such meta-analyses are robust. Our study will examine previously underexplored sources of bias in meta-analyses of nutrition research. The findings may have implications for the design, conduct and reporting of future SRs with meta-analyses of the association between food/diet and health-related outcomes.

## Ethics

Ethics approval is not required because information will only be extracted from published studies.

## Data availability

### Underlying data

No data are associated with the article.

### Extended data

Open Science Framework: Investigation of Risk Of Bias due to Unreported and Selectively included results in meta-analyses of nutrition research: the ROBUST study. <https://doi.org/10.17605/OSF.IO/TCVJQ><sup>21</sup>

This project contains the following extended data:

- ROBUST\_protocol\_appendix\_1\_20190920.pdf (PDF containing study search strategy)
- ROBUST\_protocol\_appendix\_2\_20191219.pdf (PDF containing worked example of the potential impact of selective inclusion of results on meta-analyses)

Data are available under the terms of the [Creative Commons Zero “No rights reserved” data waiver](#) (CC0 1.0 Public domain dedication).

## References

1. GBD 2017 Diet Collaborators: **Health effects of dietary risks in 195 countries, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017.** *Lancet.* 2019; **393**(10184): 1958–72. [PubMed Abstract](#) | [Publisher Full Text](#)
2. Blake P, Durão S, Naude CE, *et al.*: **An analysis of methods used to synthesize evidence and grade recommendations in food-based dietary guidelines.** *Nutr Rev.* 2018; **76**(4): 290–300. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
3. Whiting P, Savović J, Higgins JP, *et al.*: **ROBIS: A new tool to assess risk of bias in systematic reviews was developed.** *J Clin Epidemiol.* 2016; **69**: 225–34. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
4. Page MJ, McKenzie JE, Chau M, *et al.*: **Methods to select results to include in meta-analyses deserve more consideration in systematic reviews.** *J Clin Epidemiol.* 2015; **68**(11): 1282–91. [PubMed Abstract](#) | [Publisher Full Text](#)
5. Mayo-Wilson E, Fusco N, Li T, *et al.*: **Multiple outcomes and analyses in clinical trials create challenges for interpretation and research synthesis.** *J Clin Epidemiol.* 2017; **86**: 39–50. [PubMed Abstract](#) | [Publisher Full Text](#)
6. Page MJ, McKenzie JE, Forbes A: **Many scenarios exist for selective inclusion and reporting of results in randomized trials and systematic reviews.** *J Clin Epidemiol.* 2013; **66**(5): 524–37. [PubMed Abstract](#) | [Publisher Full Text](#)
7. Li T, Mayo-Wilson E, Fusco N, *et al.*: **Caveat emptor: the combined effects of multiplicity and selective reporting.** *Trials.* 2018; **19**(1): 497. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
8. Schmucker C, Schell LK, Portalupi S, *et al.*: **Extent of non-publication in cohorts of studies approved by research ethics committees or included in trial registries.** *PLoS One.* 2014; **9**(12): e114023. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
9. Dechartres A, Atal I, Riveros C, *et al.*: **Association Between Publication Characteristics and Treatment Effect Estimates: A Meta-epidemiologic Study.** *Ann Intern Med.* 2018; **169**(6): 385–93. [PubMed Abstract](#) | [Publisher Full Text](#)
10. Goldacre B, Drysdale H, Dale A, *et al.*: **COMPare: a prospective cohort study correcting and monitoring 58 misreported trials in real time.** *Trials.* 2019; **20**(1): 118. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
11. Chan AW, Hróbjartsson A, Haahr MT, *et al.*: **Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles.** *JAMA.* 2004; **291**(20): 2457–65. [PubMed Abstract](#) | [Publisher Full Text](#)

12. Page MJ, Higgins JPT, Sterne JAC: **Chapter 13: Assessing risk of bias due to missing results in a synthesis.** In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al., editors. *Cochrane Handbook for Systematic Reviews of Interventions*. 2nd Edition. Chichester (UK): John Wiley & Sons; 2019. [PubMed Abstract](#) | [Publisher Full Text](#)
13. Page MJ, Forbes A, Chau M, et al.: **Investigation of bias in meta-analyses due to selective inclusion of trial effect estimates: empirical study.** *BMJ Open*. 2016; 6(4): e011863. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
14. de Rezende LFM, Rey-López JP, de Sá TH, et al.: **Reporting bias in the literature on the associations of health-related behaviors and statins with cardiovascular disease and all-cause mortality.** *PLoS Biol*. 2018; 16(6): e2005761. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
15. Page MJ, McKenzie JE, Green SE, et al.: **An empirical investigation of the potential impact of selective inclusion of results in systematic reviews of interventions: study protocol.** *Syst Rev*. 2013; 2: 21. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
16. Higgins JPT, Thomas J, Chandler J, et al.: **Cochrane Handbook for Systematic Reviews of Interventions**. 2nd Edition. Chichester (UK): John Wiley & Sons; 2019. [Reference Source](#)
17. Page MJ, Shamseer L, Altman DG, et al.: **Epidemiology and Reporting Characteristics of Systematic Reviews of Biomedical Research: A Cross-Sectional Study.** *PLoS Med*. 2016; 13(5): e1002028. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
18. Zhong VW, Van Horn L, Cornelis MC, et al.: **Associations of Dietary Cholesterol or Egg Consumption With Incident Cardiovascular Disease and Mortality.** *JAMA*. 2019; 321(11): 1081–95. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
19. Clegg A, Young J, Iliffe S, et al.: **Frailty in elderly people.** *Lancet*. 2013; 381(9868): 752–62. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
20. Durão S, Kredt T, Volmink J: **Validation of a search strategy to identify nutrition trials in PubMed using the relative recall method.** *J Clin Epidemiol*. 2015; 68(6): 610–6. [PubMed Abstract](#) | [Publisher Full Text](#)
21. Page MJ: **Investigation of Risk Of Bias due to Unreported and Selectively included results in meta-analyses of nutrition research: the ROBUST study.** 2019. <http://www.doi.org/10.17605/OSF.IO/TCVJQ>
22. Harris PA, Taylor R, Thielke R, et al.: **Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support.** *J Biomed Inform*. 2009; 42(2): 377–81. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
23. Harris PA, Taylor R, Minor BL, et al.: **The REDCap consortium: Building an international community of software platform partners.** *J Biomed Inform*. 2019; 95: 103208. [PubMed Abstract](#) | [Publisher Full Text](#)
24. Grundy Q, Dunn AG, Bourgeois FT, et al.: **Prevalence of Disclosed Conflicts of Interest in Biomedical Research and Associations With Journal Impact Factors and Altmetric Scores.** *JAMA*. 2018; 319(4): 408–9. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
25. López-López JA, Page MJ, Lipsey MW, et al.: **Dealing with effect size multiplicity in systematic reviews and meta-analyses.** *Res Synth Methods*. 2018; 9(3): 336–351. [PubMed Abstract](#) | [Publisher Full Text](#)
26. Efron B, Tibshirani RJ: **An introduction to the bootstrap.** New York, NY: Chapman & Hall; 1993. [Reference Source](#)
27. Hartung J, Knapp G: **A refined method for the meta-analysis of controlled clinical trials with binary outcome.** *Stat Med*. 2001; 20(24): 3875–89. [PubMed Abstract](#) | [Publisher Full Text](#)
28. Sidik K, Jonkman JN: **A simple confidence interval for meta-analysis.** *Stat Med*. 2002; 21(21): 3153–9. [PubMed Abstract](#) | [Publisher Full Text](#)
29. Chinn S: **A simple method for converting an odds ratio to effect size for use in meta-analysis.** *Stat Med*. 2000; 19(22): 3127–31. [PubMed Abstract](#) | [Publisher Full Text](#)
30. Higgins JP, Thompson SG, Deeks JJ, et al.: **Measuring inconsistency in meta-analyses.** *BMJ*. 2003; 327(7414): 557–60. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
31. Hartling L, Hamm MP, Milne A, et al.: **Testing the risk of bias tool showed low reliability between individual reviewers and across consensus assessments of reviewer pairs.** *J Clin Epidemiol*. 2013; 66(9): 973–81. [PubMed Abstract](#) | [Publisher Full Text](#)
32. Landis JR, Koch GG: **The measurement of observer agreement for categorical data.** *Biometrics*. 1977; 33(1): 159–74. [PubMed Abstract](#) | [Publisher Full Text](#)
33. StataCorp: **Stata Statistical Software: Release 15.** College Station, TX: StataCorp LLC. 2017. [Reference Source](#)
34. Azar M, Riehm KE, Saadat N, et al.: **Evaluation of Journal Registration Policies and Prospective Registration of Randomized Clinical Trials of Nonregulated Health Care Interventions.** *JAMA Intern Med*. 2019; 179(5): 624–32. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
35. Chartres N, Fabbri A, McDonald S, et al.: **Association of industry ties with outcomes of studies examining the effect of wholegrain foods on cardiovascular disease and mortality: systematic review and meta-analysis.** *BMJ Open*. 2019; 9(5): e022912. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
36. Mandrioli D, Kearns CE, Bero LA: **Relationship between Research Outcomes and Risk of Bias, Study Sponsorship, and Author Financial Conflicts of Interest in Reviews of the Effects of Artificially Sweetened Beverages on Weight Outcomes: A Systematic Review of Reviews.** *PLoS One*. 2016; 11(9): e0162198. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

# Open Peer Review

Current Peer Review Status:  

---

## Version 2

Reviewer Report 14 February 2020

<https://doi.org/10.5256/f1000research.24643.r59686>

© 2020 van Aert R. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Robbie C.M. van Aert** 

Department of Methodology and Statistics, Tilburg University, Tilburg, The Netherlands

I am satisfied with the revisions of the study protocol.

One remaining minor comment is that the authors write on page 8:

- "When the difference between the index and median meta-analytic OR is minimal, we will conclude that any potential selective inclusion had a limited impact on the meta-analytic effect."

What is considered here as a "minimal difference"? I recommend the authors to define what they see as a minimal difference.

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Statistics

**I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

---

## Version 1

Reviewer Report 30 October 2019

<https://doi.org/10.5256/f1000research.22794.r55246>

© 2019 Wallach J. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Joshua Wallach**

Department of Environmental Health Sciences, Yale School of Public Health, New Haven, CT, USA

In this study protocol, Page et al. outline a project focused on exploring potential biases due to selective inclusion of results and missing results in systematic reviews of food/diet-outcome relationships. In particular, the authors have 3 main objectives:

1. To explore whether systematic reviews selectively include study effect estimates in meta-analyses when multiple effects are available.
2. To explore what impact selective inclusion of study effect estimates may have on meta-analytic effects.
3. To explore the risk of bias due to missing results in meta-analyses.

In the introduction of the protocol, the authors clearly justify the rationale for, and objectives of, the study. As the authors outline, if systematic reviews are flawed, the recommendations that could inform dietary guidelines could be misleading or harmful. Nutrition/dietary exposures are difficult to evaluate in observational studies, and numerous prior empirical evaluations have suggested concerns related to various biases.

The study design outlined by the authors is appropriate. Furthermore, most of the methods are clear and are sufficient to allow replication by others. A few minor suggestions:

- It would be helpful if the authors provided more information about their search for SRs. In particular, the authors note that they will use a PubMed to identify meta-analyses. Previous studies have outlined difficulties searching for meta-analyses. Is there a validated search strategy that the authors considered for systematic reviews/meta-analyses?

Other comments:

- In the introduction, the authors highlight that systematic reviews evaluating the association between food/diet and health-related outcomes could inform dietary guidelines. Have the authors considered limiting their sample to reviews that actually were used to inform guidelines?
- When selecting one meta-analysis, why focus on the first identified analysis? Could the authors have considered the primary outcome? Or, if the goal is identify potentially selectively reported outcomes, only focusing on those with results highlighted in the abstract?
- When looking at decision rules to select effect estimates: what if a meta-analysis includes multiple analyses based on different decision rules? It seems like this would make it difficult to evaluate selective inclusion?
- I am a bit confused by this statement "If no SR protocol is available, we will assume that no eligibility criteria and decision rules were pre-specified, even if some were reported in the published SR ('worst-case scenario' assumption), and extract all study outcome data based on how the outcome was specified in the SR." Why not use the information in the published SR?
- Would it be helpful if the authors discussed the potential bias index in the introduction? This is an important part of the study, but it is not discussed until the data analysis.

**Is the rationale for, and objectives of, the study clearly described?**

Yes



**Is the study design appropriate for the research question?**

Yes

**Are sufficient details of the methods provided to allow replication by others?**

Yes

**Are the datasets clearly presented in a useable and accessible format?**

Not applicable

**Competing Interests:** In the past 36 months, I have received support from the Collaboration for Research Integrity and Transparency (CRIT), which is funded by Arnold Ventures.

**Reviewer Expertise:** Meta-research; meta-analysis

**I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

Author Response 05 Feb 2020

**Matthew Page**, Monash University, Melbourne, Australia

It would be helpful if the authors provided more information about their search for SRs. In particular, the authors note that they will use a PubMed to identify meta-analyses. Previous studies have outlined difficulties searching for meta-analyses. Is there a validated search strategy that the authors considered for systematic reviews/meta-analyses?

**AUTHOR RESPONSE: We decided to search PubMed and Epistemonikos for meta-analyses of food/diet-outcome relationships. In PubMed, we have used the following terms to identify meta-analyses: “meta-analysis[pt] OR meta-analy\*[ti]”. That is, articles will need to have been classified by the PubMed indexers using the “Meta-Analysis” [Publication Type] MeSH term, or include the terms “meta-analysis”, “meta-analyses”, “meta-analytic” or other variants in the title. Various search filters designed to retrieve systematic reviews exist (see <https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/filters-to-identify-systematic-reviews>). However, these are designed to retrieve systematic reviews with or without meta-analyses, and typically have high sensitivity given the lack of consensus on what constitutes a systematic review. We decided to opt for a more targeted PubMed search strategy for articles classified by database indexers or the authors themselves as meta-analyses, given we are including systematic reviews only if they present a meta-analysis. We believe this search strategy will be sufficiently comprehensive given the consensus on what constitutes a meta-analysis. The other database we are searching, Epistemonikos, has been populated by conducting systematic searches for systematic reviews (with or without meta-analyses) indexed in 10 databases (listed at [https://www.epistemonikos.org/en/about\\_us/methods#](https://www.epistemonikos.org/en/about_us/methods#)). We have provided this additional information under “Search for SRs”.**

In the introduction, the authors highlight that systematic reviews evaluating the association between food/diet and health-related outcomes could inform dietary guidelines. Have the authors considered limiting their sample to reviews that actually were used to inform guidelines?

**AUTHOR RESPONSE: We decided not to limit our study to systematic reviews that were used to inform dietary guidelines for two reasons. First, investigators have found that**

most dietary guidelines are not underpinned by systematic reviews. For example, Blake et al. found that of 32 food-based dietary guidelines published from 2010 to 2016, only 10 (31%) were informed by previously published systematic reviews (<https://www.ncbi.nlm.nih.gov/pubmed/29425371>). Therefore, relying on dietary guidelines to identify systematic reviews may result in a smaller than desired sample. Second, the systematic reviews used to inform dietary guidelines are not likely to be current because of the time-lag between development and publication of guidelines. By limiting our focus to recently published systematic reviews (Jan 2018 to Jul 2019), we will be able to comment on current issues with transparency and risk of bias in reviews.

When selecting one meta-analysis, why focus on the first identified analysis? Could the authors have considered the primary outcome? Or, if the goal is identify potentially selectively reported outcomes, only focusing on those with results highlighted in the abstract?

**AUTHOR RESPONSE:** We have decided not to focus on the primary outcome of the systematic review because many reviews do not specify a primary review outcome. For example, Bassani et al. found that of 480 systematic reviews in dentistry indexed in PubMed in 2017, only 151 (32%) specified a primary outcome (<https://www.ncbi.nlm.nih.gov/pubmed/30716451>). Also, in a sample of reviews not restricted by topic that were indexed in MEDLINE in Feb 2014, 136/288 (47%) specified a primary outcome (<https://www.ncbi.nlm.nih.gov/pubmed/27218655>). Therefore, assuming that many recent reviews will not specify a primary outcome, we need an alternative rule to select meta-analyses that can be applied across all reviews. We have opted for the first meta-analytic result reported in the review because the first meta-analysis is likely to be the most (or one of the most) important analysis in the review on which the conclusions of the review are based.

When looking at decision rules to select effect estimates: what if a meta-analysis includes multiple analyses based on different decision rules? It seems like this would make it difficult to evaluate selective inclusion?

**AUTHOR RESPONSE:** In our previous study investigating selective inclusion of results (<https://www.ncbi.nlm.nih.gov/pubmed/27121706>), we found that it was always clear which decision rules applied to which meta-analysis, because the rule was applicable only to a particular outcome (e.g. authors specified a hierarchy of measurement scales for pain, and a hierarchy of scales for anxiety), or authors made it clear that the rule applied to all meta-analyses (e.g. stated that for all meta-analyses of continuous outcomes, final values were preferred over change from baseline values). We assume that the same level of specificity will occur in the systematic reviews examined in current study.

I am a bit confused by this statement "If no SR protocol is available, we will assume that no eligibility criteria and decision rules were pre-specified, even if some were reported in the published SR ('worst-case scenario' assumption), and extract all study outcome data based on how the outcome was specified in the SR." Why not use the information in the published SR?

**AUTHOR RESPONSE:** In our primary analysis, we will ignore the eligibility criteria and decision rules to select results that were specified in the published SR, because there is a risk that such criteria and rules were created by systematic reviewers post-hoc. For example, systematic reviewers may have examined the study results for different scales measuring satiety, and crafted a post-hoc decision rule for scales that prioritises inclusion of the result with the largest effect estimate or smallest P value. However, as noted in paragraph eight of the section on "Analysis of selective inclusion of results",

**“...we will perform a sensitivity analysis where study effect estimates that were compatible only with the eligibility criteria and decision rules in the methods sections of the SR are included, to examine if this affected the PBI”.**

Would it be helpful if the authors discussed the potential bias index in the introduction? This is an important part of the study, but it is not discussed until the data analysis.

**AUTHOR RESPONSE: We believe it is best to leave the description of the Potential Bias Index until the data analysis section, because of the extensive detail that is required to describe the index. Including information about the index in the background may detract from the rationale for the study and the objectives.**

**Competing Interests:** No competing interests were disclosed.

Reviewer Report 21 October 2019

<https://doi.org/10.5256/f1000research.22794.r55247>

© 2019 van Aert R. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Robbie C.M. van Aert** 

Department of Methodology and Statistics, Tilburg University, Tilburg, The Netherlands

The reviewed manuscript is a study protocol for studying the risk of bias due to unreported and selectively included results in meta-analyses of nutrition research. I really appreciate the effort the authors put in this detailed study protocol. Below I list some major and minor comments to improve the protocol.

#### Major comments:

1. If there is no systematic review protocol available, there is assumed that there were no eligibility criteria or decision rules pre-specified. I believe this is way too restrictive, because meta-analysts will most of the time have explained what their inclusion and exclusion criteria were in the paper. This is also what is acknowledged in the proposed study protocol, but there is still decided to not use this information.  
Why are the systematic reviews without study protocol not included as a separate category? It would be interesting to see whether risk of bias is smallest in the systematic reviews with study protocols, larger in the ones with a clear description in the paper, and largest in the ones without a clear description in a study protocol or the paper.
2. There is planned to also include study reports that were eligible for inclusion in the index meta-analysis but excluded in the systematic review. Do the authors of the study protocol expect that a list is included in any systematic review with the study reports that were excluded or are they planning to redo the literature search? If the authors will be relying on a list of excluded study reports, how are they planning to deal with situations were such a list is unavailable?
3. The first meta-analysis is the selected meta-analysis that will be included in the study. Is it expected that the first meta-analysis is equivalent to any other meta-analysis in the systematic

review? Do the authors expect that the first meta-analysis is representative for the other meta-analyses in the systematic reviews? I can imagine that the first meta-analysis is on the main effect under study of a systematic review and that the risk of bias is expected to be the largest for this meta-analysis.

**Minor comments:**

1. P.3: The authors may want to consider to change the order of the second and third paragraph of the introduction section. I believe that it is more logical to start with explaining that not all studies are usually accessible for being included in a meta-analysis and then explain that systematic reviewers themselves can also purposefully select which studies to include.
2. P.3: An eligibility criterion of a systematic review to be included is that it can be either a randomized trial or a non-randomized study. Please consider using stratification here such that 50% of the included systematic reviews are randomized trials and 50% non-randomized studies. I believe that this may yield relevant insights as randomized trials are usually seen as of higher quality.
3. P.8: I have a hard time understanding what the planned analysis is for studying the impact of any potential selective inclusion of study effect estimates on the magnitude of resulting meta-analytic ORs and SMDs (last paragraph, left column). Are the authors planning to compute odds ratios for each possible constellation of the 2x2 table of each individual study and then meta-analyzing these odds ratios? I do not see how this will yield insight into the effect of selective inclusion of study effect estimates on the magnitude of the meta-analytic estimate. Moreover, did the authors already write syntax/code for this? If yes, please add this to the study protocol as well.
4. P.8: Related to my previous minor comment, I also do not understand why converting of SMDs to log ORs is necessary. Why not conducting such an analysis based on SMDs?
5. P.8: A fixed-effect meta-analysis is planned to be used to combine the PBI that will be observed in the planned study with a PBI obtained in an earlier study. I doubt whether it is useful to meta-analyze these to PBI-values. Why not interpreting the PBI-value that will be obtained independently from the previously obtained PBI-value? It is, of course, good to relate this PBI-value to the previously obtained one, but I do not really see the need for meta-analyzing the two. Especially not since there are only two PBI-values that can be combined, so it is perfectly possible to interpret the two independently of each other.  
If the authors still decide to use a fixed-effect meta-analysis for combining the PBI values, please report the  $I^2$  statistic together with its confidence interval. The computed  $I^2$  statistic will be very uncertain in case of two studies and it is important to acknowledge this uncertainty.
6. P.8: The effect of several independent variables on PBI will be tested by means of bootstrap methods. Why are not more conventional methods used for testing these independent variables such as regression analysis?
7. P.9: The impact of any potential selective inclusion of study effect estimates on meta-analytic effect estimates will be studied using random-effects meta-analysis. Please also report what estimator will be used for estimating the between-study variance in the random-effects meta-analysis.

**Is the rationale for, and objectives of, the study clearly described?**

Yes

**Is the study design appropriate for the research question?**

Yes

**Are sufficient details of the methods provided to allow replication by others?**

Partly

**Are the datasets clearly presented in a useable and accessible format?**

Not applicable

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Statistics

**I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.**

Author Response 05 Feb 2020

**Matthew Page**, Monash University, Melbourne, Australia

Major comments:

If there is no systematic review protocol available, there is assumed that there were no eligibility criteria or decision rules pre-specified. I believe this is way too restrictive, because meta-analysts will most of the time have explained what their inclusion and exclusion criteria were in the paper. This is also what is acknowledged in the proposed study protocol, but there is still decided to not use this information. Why are the systematic reviews without study protocol not included as a separate category? It would be interesting to see whether risk of bias is smallest in the systematic reviews with study protocols, larger in the ones with a clear description in the paper, and largest in the ones without a clear description in a study protocol or the paper.

**AUTHOR RESPONSE: In our primary analysis, we will ignore the eligibility criteria and decision rules to select results that were specified in the published SR, because there is a risk that such criteria and rules were created by systematic reviewers post-hoc. For example, systematic reviewers may have examined the study results for different scales measuring satiety, and crafted a post-hoc decision rule for scales that prioritises inclusion of the result with the largest effect estimate or smallest P value. However, as noted in paragraph eight of the section on “Analysis of selective inclusion of results”, “...we will perform a sensitivity analysis where study effect estimates that were compatible only with the eligibility criteria and decision rules in the methods sections of the SR are included, to examine if this affected the PBI”. We also state in paragraph seven of this section that “We will conduct subgroup analyses to explore whether the availability of an SR protocol or registration record...modifies the PBI”. In other words, we will compare selective inclusion in SRs with a protocol/registration entry versus SRs without.**

There is planned to also include study reports that were eligible for inclusion in the index meta-analysis but excluded in the systematic review. Do the authors of the study protocol expect that a list is included in any systematic review with the study reports that were excluded or are they planning to redo the literature search? If the authors will be relying on a list of excluded study reports, how are they planning to deal with situations where such a list is unavailable?

**AUTHOR RESPONSE: We have clarified in paragraph two of the section, “Selection of**



index meta-analyses”, that “We will also retrieve reports of studies that were included in the review but excluded from the meta-analysis by the systematic reviewers...”). We expect to be able to identify which of the studies included in the review were omitted from the meta-analysis given that all (or nearly all) systematic reviews cite every study included in the review. We have added the following text to this paragraph: “We also expect that in some systematic reviews, citations to ‘near-miss’ studies along with reasons for exclusion, will be provided. For these reviews, we will scan the reasons for exclusion, and where these reasons indicate there was no useable data, we will retrieve the study to confirm that no data were available for inclusion in the review.” We will not redo the literature search.

The first meta-analysis is the selected meta-analysis that will be included in the study. Is it expected that the first meta-analysis is equivalent to any other meta-analysis in the systematic review? Do the authors expect that the first meta-analysis is representative for the other meta-analyses in the systematic reviews? I can imagine that the first meta-analysis is on the main effect under study of a systematic review and that the risk of bias is expected to be the largest for this meta-analysis.

**AUTHOR RESPONSE: We do not expect the first meta-analysis will necessarily be representative of all other meta-analyses in the systematic review. However, the first meta-analysis is likely to be the most (or one of the most) important analysis in the review on which the conclusions of the review are based. For this reason, we have chosen to focus our assessment on the first meta-analysis.**

Minor comments:

P.3: The authors may want to consider to change the order of the second and third paragraph of the introduction section. I believe that it is more logical to start with explaining that not all studies are usually accessible for being included in a meta-analysis and then explain that systematic reviewers themselves can also purposefully select which studies to include.

**AUTHOR RESPONSE: We would prefer to retain the order of the second and third paragraph, as we think it is more logical to start with explaining how meta-analyses may be biased because of actions of review authors (i.e. selective inclusion of results), followed by actions beyond review authors’ control (i.e. non-reporting by study authors).**

P.3: An eligibility criterion of a systematic review to be included is that it can be either a randomized trial or a non-randomized study. Please consider using stratification here such that 50% of the included systematic reviews are randomized trials and 50% non-randomized studies. I believe that this may yield relevant insights as randomized trials are usually seen as of higher quality.

**AUTHOR RESPONSE: As noted under “Selection of index meta-analyses”, we are already stratifying the sample so that it includes 25 meta-analyses of binary outcomes and 25 meta-analyses of continuous outcomes. The screening of titles and abstracts conducted thus far suggests there is a strong association between outcome type and type of included study in meta-analyses of nutrition research. That is, meta-analyses of continuous outcomes (such as weight) are based typically on data from randomized trials, and meta-analyses of binary outcomes (such as all-cause mortality) are based typically on data from non-randomized studies. Therefore, we believe our current stratification approach (by outcome type) will lead to a roughly equal sample of meta-analyses of randomized trials and meta-analyses of non-randomized studies.**

P.8: I have a hard time understanding what the planned analysis is for studying the impact of any

potential selective inclusion of study effect estimates on the magnitude of resulting meta-analytic ORs and SMDs (last paragraph, left column). Are the authors planning to compute odds ratios for each possible constellation of the 2x2 table of each individual study and then meta-analyzing these odds ratios? I do not see how this will yield insight into the effect of selective inclusion of study effect estimates on the magnitude of the meta-analytic estimate. Moreover, did the authors already write syntax/code for this? If yes, please add this to the study protocol as well.

**AUTHOR RESPONSE: We plan to generate all possible meta-analytic effects that could be generated from the various combinations of study effect estimates available. For some studies included in an index meta-analysis, there will only have been one possible effect estimate available from the study report. For these studies, that sole effect estimate will be included in every iteration of the meta-analysis. For studies that have multiple effect estimates that are eligible for inclusion in the index meta-analysis (e.g. based on multiple scales or time points), we will randomly select a study effect estimate to include in each iteration of the meta-analysis. We will repeat this process until every possible meta-analytic effect estimate is computed (unless the number of possible combinations is prohibitively large to calculate all combinations (i.e. >30,000)). After calculating all possible meta-analytic effect estimates that the systematic reviewers could have generated, we will compare the meta-analytic effect estimate that was reported by the systematic reviewers with the median of all possible meta-analytic effect estimates that could have been generated. If the meta-analytic effect estimate presented by the systematic reviewers is much greater in magnitude than the median of all possible meta-analytic effect estimates, this suggests that systematic reviewers' decision to selectively include study effect estimates had a considerable impact on the meta-analytic effect estimate reported. We have uploaded a worked example to the Open Science Framework (<https://osf.io/pn38b/>).**

P.8: Related to my previous minor comment, I also do not understand why converting of SMDs to log ORs is necessary. Why not conducting such an analysis based on SMDs?

**AUTHOR RESPONSE: To undertake the analysis including all 50 meta-analyses, we will have to use the same effect measure. Of the 50 included meta-analyses, 25 will be of continuous outcomes using the SMD, and 25 will be binary outcomes using the OR. We therefore need to either convert the SMDs to (log)ORs or the (log)ORs to SMDs. Given that we have an equal number of effect estimates to transform, we had no particular reason to adopt transforming (log)ORs to SMDs ahead of SMDs to (log)ORs.**

P.8: A fixed-effect meta-analysis is planned to be used to combine the PBI that will be observed in the planned study with a PBI obtained in an earlier study. I doubt whether it is useful to meta-analyze these two PBI-values. Why not interpreting the PBI-value that will be obtained independently from the previously obtained PBI-value? It is, of course, good to relate this PBI-value to the previously obtained one, but I do not really see the need for meta-analyzing the two. Especially not since there are only two PBI-values that can be combined, so it is perfectly possible to interpret the two independently of each other. If the authors still decide to use a fixed-effect meta-analysis for combining the PBI values, please report the I<sup>2</sup> statistic together with its confidence interval. The computed I<sup>2</sup> statistic will be very uncertain in case of two studies and it is important to acknowledge this uncertainty.

**AUTHOR RESPONSE: We believe there is value in meta-analysing the two PBI estimates. Doing so will yield a single estimate with greater precision. We plan to present a forest plot with estimates of PBI from the two studies, along with the combined PBI. Given there are only two estimates, we are not convinced that the I<sup>2</sup> statistic will provide a useful**

**quantification of inconsistency, and not more than can be obtained from visually inspecting the overlap in confidence intervals of the PBIs from each study.**

P.8: The effect of several independent variables on PBI will be tested by means of bootstrap methods. Why are not more conventional methods used for testing these independent variables such as regression analysis?

**AUTHOR RESPONSE: We plan to use bootstrap methods to estimate the confidence limits and P values since the PBI is a complex metric for which statistical theory concerning its distribution does not currently exist. We now provide more specific detail about our proposed bootstrap methods.**

P.9: The impact of any potential selective inclusion of study effect estimates on meta-analytic effect estimates will be studied using random-effects meta-analysis. Please also report what estimator will be used for estimating the between-study variance in the random-effects meta-analysis.

**AUTHOR RESPONSE: We stated in paragraph five of the section “Analysis of selective inclusion of results”, that the between-study variability will be “...estimated using the restricted maximum likelihood estimator”.**

**Competing Interests:** No competing interests were disclosed.

---

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact [research@f1000.com](mailto:research@f1000.com)

**F1000Research**