Low-energy sweeteners and body weight: a citation network analysis

Mie Normand, Christian Ritz, David Mela, Anne Raben

ABSTRACT

Objective Reviews on the relationship of low-energy sweeteners (LES) with body weight (BW) have reached widely differing conclusions. To assess possible citation bias, citation analysis was used to quantify the relevant characteristics of cited articles, and explore citation patterns in relation to review conclusions.

Design A systematic search identified reviews published from January 2010 to March 2020. Different characteristics (for example, type of review or research, journal impact factor, conclusions) were extracted from the reviews and cited articles. Logistic regression was used to estimate likelihood of articles with particular characteristics being cited in reviews. A qualitative network analysis linked reviews sub-grouped by conclusions with the types of articles they cited.

Main outcome measures (OR; 95% CI) for likelihood that articles with particular characteristics were cited as evidence in reviews.

Results From 33 reviews identified, 183 different articles were cited (including other reviews). Narrative reviews were 62% less likely to be cited than systematic reviews with meta-analysis (OR 0.38; 0.16 to 0.86; p=0.03). Likelihood of being cited was higher for evidence on children than adults (OR 2.27; 1.59 to 3.25; p<0.0001), and with increased journal impact factor (OR 1.15; 1.00 to 1.31; p=0.04). No other factors were statistically significant in the main analysis, and few factors were significant in subgroup analyses. Network analysis showed that reviews concluding a beneficial relationship of LES with BW cited mainly randomised controlled trials, whereas reviews concluding an adverse relationship cited mainly observational studies.

Conclusions Overall reference to the available evidence across reviews appears largely arbitrary, making citation bias likely. Differences in the conclusions of individual reviews map onto different types of evidence cited. Overall, inconsistent and selective use of the available evidence may account for the diversity of conclusions in reviews on LES and BW.

Trial registration number Prior to data analysis, the protocol was registered with the Open Science Framework (https://osf.io/9ghws).

INTRODUCTION

The relationship of low-energy sweeteners (LES) with body weight (BW) has been widely discussed. Replacing sugars (monosaccharides and disaccharides) with LES has been argued to benefit BW control by decreasing net calorie intake while satisfying a desire for sweetness. In contrast, it has also been argued that LES may promote weight gain by having the opposite effects: causing dysregulation of appetite and metabolism, and promoting intake of sweet-tasting, energy-containing foods. Widely differing views are evident in recent narrative and systematic reviews in relation to LES and BW, with some reviews concluding a beneficial effect or association, others a detrimental effect or association, and others again that the evidence is too limited or inconsistent to conclude either way. Given that the same evidence base is available to all reviewers, these different views presumably arise from the selection and importance given to different types of evidence, and potential citation bias in the reviews.

Citation bias can be defined as ‘the citation or non-citation of research findings, depending on the nature and direction of the results’. Song et al concluded that citation bias is especially known to occur for positive or significant results. Furthermore, non-systematic narrative reviews were particularly susceptible to biased citation leading to misleading conclusions. Citation bias is not unknown in nutrition research, for example, in relation to dietary fatty acid intake and the risk of cardiovascular disease (CVD). Urlings et al investigated citation bias in the literature concerning dietary trans fatty acids and serum cholesterol. They concluded that several
factors such as statistically significant results, sample size and journal impact factor (JIF) were important determinants for selective citation. Leng used a network analysis to assess the patterns of citations in reviews published before 1984 of the early randomised controlled trials (RCT) on cholesterol lowering diets for prevention of coronary heart disease. That analysis provided evidence of selective citation of RCTs, especially in the reviews supporting dietary intervention in secondary prevention of CVD. This also shows that citation bias is not a new phenomenon.

In relation to LES and BW, citation analysis could help to understand the basis for differing conclusions from reviews. Mela et al specifically raised selective citation as a general issue in the interpretation and reporting of research on LES. However, there has been no objective analysis of citation patterns or bias in the literature on LES. The aim of this analysis was therefore to assess the citation pattern in reviews on the relationship of LES with BW-related outcomes. The main analysis evaluated whether variation in the overall pattern of articles cited in reviews was quantitatively associated to specific characteristics of the cited articles, such as article type, conclusions, population, sample size, authorship, JIF and years between the review and the cited (original) article. Network analysis was used to assess qualitative relationships between individual review conclusions and the nature of cited articles.

METHODS
The methods were originally described in a study protocol registered online prior to undertaking any analyses (https://osf.io/9ghws). Any later additions, deviations and modifications from that original protocol are noted below. The protocol is available in online supplemental material (including online supplemental tables S1, S2 and S3).

Search strategy and article selection
A systematic literature search was conducted in March 2020 using the Web of Science Core Collection. The aim was to transparently identify a comprehensive, representative and unbiased selection of reviews assessing the relationship between LES exposure and outcomes related to BW or obesity risk. Search terms included ‘low-energy sweetener(s)’ and related terms for exposure, ‘body weight’, ‘overweight’, ‘obesity’ and ‘adiposity’ for outcome, together with different types of publication, for example, ‘narrative review’, ‘systematic review’, ‘meta-analysis’, ‘scientific report’ and ‘perspective’. For a complete list of search terms, see online supplemental table S1. Articles were included if they were published in English, in refereed scientific journals (ie, excluding ‘grey’ literature). Any other potentially eligible reviews subsequently identified (for example, from reference lists in the reviews returned by our search) were not included, because this could cause an over-representation of articles cited within the network, whereas eligible articles outside the network would be ignored. As a modification of the registered protocol, the search was limited to publications within the preceding 10 years to better reflect current rather than historical practices.

Screening of review articles was done in two steps: (1) screening of title and abstract to identify potentially relevant articles, (2) full-text screening of identified articles from step 1 to confirm that relevant articles meeting the inclusion criteria were correctly identified, and that data for the required outcome measures were reported. The screening was done independently by two reviewers (DM and MN). Any uncertainties or disagreements were resolved by a third reviewer (AR). The eligibility of reviews was determined according to the predefined inclusion and exclusion criteria (online supplemental table S2). Each review was re-classified as one or more ‘evidence assessment units’ (EAU—see below) for analysis. If a single review contained multiple independent analyses and conclusions based on different evidence sets (for example, RCT vs observational evidence, evidence for adults vs children), these were treated as separate EAUs in the citation analysis. Children were defined as a study population with mean age under 18 years.

For clarity, the following terminology is used:
► Review: a published review identified by the systematic literature search. A review is a single publication, consisting of one or more EAUs.
► EAU: a single review or, where present, each of the independent evidence assessments within a review, such as wholly separate sections for evidence or meta-analyses on children or adults, with independent conclusions. For the purpose of the citation analysis, each EAU in any single review publication was treated as if they were separate publications, because each EAU cited different evidence sets and could differ in conclusions.
► Cited article: any publication cited as evidence for the LES–BW relationship in a review (EAU). Cited articles could be original research or earlier reviews (including those identified for the present citation analysis).

Inclusion criteria for articles cited in reviews
The included reviews were screened for any articles cited as evidence of possible effects or associations of LES exposure and human BW-related outcomes. Obesity or BW-related outcomes of interest included BW, body mass index, population risk of obesity or weight gain and other outcomes commonly used as indicators of relative BW or fatness (fat mass, percent body fat, waist circumference, skinfold thickness, adiposity). This was done independently by two reviewers (DM and MN). Cited articles were included or excluded from the analysis depending on the context in which they were used. Cited articles were included when they were an explicit part of the empirical evidence-base used for drawing conclusions on the effect or association of LES and BW-related outcomes.
Table 1 Continued

<table>
<thead>
<tr>
<th>Number of review authors publications in the section concerning BW</th>
</tr>
</thead>
<tbody>
<tr>
<td>n (%)</td>
</tr>
</tbody>
</table>

*From a total number of 33 included reviews. Where a review publication contained independent analyses and conclusions for randomised controlled trials and observational evidence, or adults and children, those were treated as separate evidence assessment units. This is the case for 18 papers, resulting in 51 evidence assessment units from the 33 reviews.†From evidence assessment units with meta-analysis (n=11).‡Of the 25 EAUs with support from non-profit sources, 4 were supported by primarily industry-funded non-profit organisations and the rest by grants primarily from government, independent foundations and universities.

BW, body weight; IQR, Interquartile range; LES, low-energy sweeteners; n, sample size.

Citations describing BW outcomes in animal studies were only included where they were used in this same context, and integrated into the narrative on BW or obesity risk in humans. Citations were excluded if they were only used in other contexts such as:

- Introductory descriptions of the general topic area or current public health guidance.
- Evidence limited to potential underlying mechanisms or hypotheses, for example, appetite control, energy intake or expenditure, adipogenesis, diet quality and so on.
- Cited but not used in quantitative or qualitative evidence assessments in the systematic review or meta-analysis.
- Animal studies clearly used in the context of narrative text on effects in animals.
- Evidence limited to visceral fat mass or ectopic fat as outcomes.
- Evidence limited to other health outcomes including metabolic syndrome.
- Part of an inventory (simple listing or description) of papers in a database or cited in other reviews.

Obvious errors in citations within reviews (for example, clear reference name and description in text linked to incorrect number in reference list, double-citing of the same paper in a reference list, mistakes in cited author or journal name and so on) were corrected where possible. However, if there was not an unambiguous resolution (for example, citation could not be matched to any clear source), these were treated as missing data.

Data extraction

Data on a number of characteristics were extracted from each EAU and cited article (Table 1 and Table 2). This information was independently reviewed and subsequently agreed by two authors (MN and DM), and a third author (AR) consulted where needed to reach a consensus. A number of guiding decision rules were applied to ensure greater consistency and transparency in the independent assessor judgements on the relevant citations and data for extraction. The most important guiding rules are
Table 2  Characteristics of included cited articles (n=183) in the total set of 51 evidence assessment units reported in 33 reviews

<table>
<thead>
<tr>
<th>Main message of cited article</th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decrease BW/more beneficial</td>
<td>32 (17)</td>
</tr>
<tr>
<td>Neutral (no directional effect or association)</td>
<td>39 (21)</td>
</tr>
<tr>
<td>Increase BW/less beneficial</td>
<td>54 (30)</td>
</tr>
<tr>
<td>No conclusion directly relevant to the LES–BW relationship</td>
<td>26 (14)</td>
</tr>
<tr>
<td>Evidence insufficient to draw a conclusion</td>
<td>20 (11)</td>
</tr>
<tr>
<td>Unable to draw a conclusion from the paper</td>
<td>9 (5)</td>
</tr>
<tr>
<td>Missing data</td>
<td>3 (2)</td>
</tr>
<tr>
<td><strong>Cited article type</strong></td>
<td></td>
</tr>
<tr>
<td>Randomised controlled trial</td>
<td>51 (28)</td>
</tr>
<tr>
<td>Observational study</td>
<td>72 (40)</td>
</tr>
<tr>
<td>Animal</td>
<td>13 (7)</td>
</tr>
<tr>
<td>Other</td>
<td>1 (1)</td>
</tr>
<tr>
<td>Systematic review with meta-analysis</td>
<td>16 (9)</td>
</tr>
<tr>
<td>Systematic review without meta-analysis</td>
<td>9 (5)</td>
</tr>
<tr>
<td>Narrative review</td>
<td>18 (10)</td>
</tr>
<tr>
<td>Missing data</td>
<td>3 (2)</td>
</tr>
<tr>
<td><strong>Cited article population</strong></td>
<td></td>
</tr>
<tr>
<td>Adults</td>
<td>85 (46)</td>
</tr>
<tr>
<td>Children</td>
<td>49 (27)</td>
</tr>
<tr>
<td>Both</td>
<td>32 (17)</td>
</tr>
<tr>
<td>Missing data</td>
<td>17 (9)</td>
</tr>
<tr>
<td><strong>Median (IQR)</strong></td>
<td></td>
</tr>
<tr>
<td>Sample size</td>
<td></td>
</tr>
<tr>
<td>Randomised controlled trials</td>
<td>50 (25–155)</td>
</tr>
<tr>
<td>Observational studies</td>
<td>2760 (781–15 984)</td>
</tr>
<tr>
<td>Number of authors</td>
<td>5 (3–7)</td>
</tr>
<tr>
<td>Journal impact factor, current (2018)</td>
<td>3.97 (3.05–6.57)</td>
</tr>
<tr>
<td>Journal impact factor, last 5 years</td>
<td>4.51 (3.33–7.67)</td>
</tr>
<tr>
<td>Years since cited article was published</td>
<td>5 (2–10)</td>
</tr>
</tbody>
</table>

BW, body weight; IQR, Interquartile range; LES, low-energy sweeteners; n, sample size.

RESULTS

Included articles

Out of 153 potentially eligible reviews identified from the systematic search (figure 1), 33 reviews met the criteria and were included in the analysis. Of these, 16 reviews had two independent sections (EAUs) separately reviewing evidence from RCT and observational studies. Two reviews had two independent sections (EAUs) separately assessing evidence from adults and children, whereas the remaining 15 reviews did not have such independent sections, and were thus treated as single EAUs. Thus, a more detailed discussion of the approach to the identification and selection of the literature is necessitated.
Several of the included EAUs were also cited as evidence by other EAUs. 2-4

About half of the EAUs (n=25) concluded that there was an effect or association of LES and BW, either beneficial, neutral or adverse, whereas the other half (n=26) concluded that the evidence was insufficient to draw a conclusion. Half of the EAUs were narrative reviews (n=26), 22% were systematic review with meta-analysis (n=11) and 27% systematic review without meta-analysis (n=14).

The majority (68%) of the 183 cited articles concluded that there was a relationship (beneficial, neutral or adverse) of LES with BW outcomes (n=125) (table 2). The majority of the cited articles were original research (n=136), either RCT, observational or animal, while 43 of the cited articles were reviews, either narrative or systematic with or without meta-analysis. Notably, however, almost 15% of the cited articles contained no relevant conclusion (n=26), mainly because they contained no relevant data or analyses. The article type of one article could not be classified.

For a small number of the cited articles (n=9), it was not possible to discern any conclusion regarding LES, even though they contained potentially relevant data (for example, where LES were not really a focus of the research).

**ORs for likelihood of being cited**

Only a small number of characteristics of the cited articles were significantly associated with likelihood of being cited (table 3). On average, an article was 62% less likely to be cited if it was a narrative review compared with a systematic review with meta-analysis. Articles on children were 127% more likely to be cited than articles on adults. An article was 15% more likely to be cited for every twofold increase in the JIF. No other statistically significant associations were seen. However, an article was 64% more likely to be cited when it did not contain any conclusion directly relevant to the relationship of LES with BW (p=0.08). Adjusting for number of authors, JIF and years since cited study was published did not alter the results (online supplemental table S4).

Statistically significant findings for the subgroup analyses are depicted in table 4 (full results are available in online supplemental tables S5–S12). Subgrouping by the direction of EAU conclusions (neutral, beneficial, adverse) showed limited relationships with the nature or conclusions of the cited articles. For EAUs concluding a beneficial or adverse effect or association of LES and BW, there were no significant associations with the main message (conclusions) of the cited articles (see online supplemental tables S5, S7). For EAUs concluding a neutral effect or association of LES and BW, articles from which it was not possible to draw a conclusion were 67% more likely to be cited than articles with neutral conclusions (p=0.03), and observational studies 45% more likely to be cited than systematic reviews with meta-analysis (p=0.03). For EAUs concluding that there was insufficient evidence...
evidence to draw a conclusion, articles on children were 84% more likely to be cited than articles on adults (p<0.001). For systematic reviews without meta-analysis, articles on children were 66% more likely to be cited than articles on adults (p=0.002), and articles on both children and adults were 86% more likely to be cited (p=0.004) than articles only on adults.

**Network analysis**

Figures 2 and 3, online supplemental figures 2 and 3 show the network analysis divided into subgroups based on the conclusions of the EAUs with cited articles coloured based on study type.

EAUs concluding a beneficial effect or association (figure 2) cited mainly RCTs, with the exception of Baker-Smith et al., who cited a large number of observational studies. Conversely, EAUs reporting an adverse effect or association (figure 3) cited mostly observational studies. More observational studies than RCTs were cited in EAUs concluding a neutral effect or association of LES on BW (online supplemental figure S1). EAUs concluding that there was insufficient evidence to draw a conclusion (online supplemental figure S2) cited a mix of both RCTs and observational studies.

Online supplemental figures S3-S6 show network analysis divided into subgroups based on the conclusions of the EAUs with cited articles coloured based on conclusion of the article. EAUs concluding a beneficial effect or association of LES with BW tended to cite articles which also concluded a beneficial effect or association (online supplemental figure S3). Similarly, EAUs concluding an adverse effect or association tended to cite articles which also concluded an adverse effect or association of LES with BW (online supplemental figure S4). EAUs concluding a neutral effect or association (online supplemental figure...
Table 4 Statistically significant findings for the subgroup analysis based on evidence assessment unit conclusions and type of review. Data from 51 evidence assessment units reported in 33 reviews

<table>
<thead>
<tr>
<th>Evidence assessment units concluding a neutral effect or association of LES on BW (n=7)</th>
<th>OR (95% CI)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main message of cited articles</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neutral (no directional effect or association)</td>
<td>1 (ref)</td>
<td></td>
</tr>
<tr>
<td>Unable to draw a conclusion from the article</td>
<td>1.67 (1.07 to 2.54)</td>
<td>0.03</td>
</tr>
<tr>
<td>Cited article type</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Systematic review with meta-analysis</td>
<td>1 (ref)</td>
<td></td>
</tr>
<tr>
<td>Observational study</td>
<td>1.45 (1.06 to 2.02)</td>
<td>0.03</td>
</tr>
<tr>
<td>Evidence assessment units concluding insufficient evidence to draw a conclusion about the effect of LES on BW (n=26)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cited article type</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Systematic review with meta-analysis</td>
<td>1 (ref)</td>
<td></td>
</tr>
<tr>
<td>Systematic review without meta-analysis</td>
<td>1.97 (1.12 to 3.45)</td>
<td>0.02</td>
</tr>
<tr>
<td>Cited article population</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adults</td>
<td>1 (ref)</td>
<td></td>
</tr>
<tr>
<td>Children</td>
<td>1.84 (1.43 to 2.37)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Cited article journal impact factor, current (2018)*</td>
<td>1.10 (1.00 to 1.20)</td>
<td>0.049</td>
</tr>
<tr>
<td>Systematic reviews (evidence assessment units) without meta-analysis (n=14)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cited article type</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Systematic review with meta-analysis</td>
<td>1 (ref)</td>
<td></td>
</tr>
<tr>
<td>Randomised controlled trial</td>
<td>0.61 (0.38 to 1.00)</td>
<td>0.04</td>
</tr>
<tr>
<td>Cited article population</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adults</td>
<td>1 (ref)</td>
<td></td>
</tr>
<tr>
<td>Both</td>
<td>1.86 (1.20 to 2.82)</td>
<td>0.004</td>
</tr>
<tr>
<td>Children</td>
<td>1.66 (1.20 to 2.29)</td>
<td>0.002</td>
</tr>
</tbody>
</table>

Logistic mixed-effects regression. The analysis of neutral reviews is additionally adjusted for overdispersion. Bold value indicates result is statistically significant with p<0.05 or lower.

*S Journal impact factor was base 2 log-transformed, so OR is the change per twofold change in journal impact factor.

BW, body weight; CI, confidence interval; LES, low-energy sweeteners; n, sample size; OR, odds ratio; ref, reference variable.

S5) and EAUs reporting insufficient evidence to draw a conclusion (online supplemental figure S6) cited articles with various conclusions.

Comparing figure 2 with online supplemental figure S3 shows that most of the cited RCTs reported a beneficial effect of LES on BW. Similarly, comparing figure 3 with online supplemental figure S4 shows that most of the cited observational studies reported either a neutral or an adverse association of LES on BW.

**DISCUSSION**

This study assessed the pattern of citations in reviews on the relationship of LES with BW-related outcomes. Surprisingly, across all reviews only a few consistent determinants of likelihood of citation were evident, favouring systematic reviews with meta-analysis, studies on children and publications in higher impact journals. In the overall data set, there was little clear quantitative association between the direction of review conclusions on the LES–BW relationship (neutral, adverse, beneficial) and the conclusions of the cited articles. The network analysis indicated that individual reviews concluding a beneficial relationship of LES with BW cited mainly RCTs, whereas the reviews concluding an adverse relationship cited mainly observational studies.

Taken together, this shows a very diverse and inconsistent pattern of citations, suggesting that the citation of evidence across reviews overall is somewhat arbitrary, which may contribute to the diversity in review conclusions. For individual reviews, conclusions mapped onto different patterns of cited evidence providing support for those conclusions.

This is the first citation analysis of its kind in relation to LES, and we believe it has several strengths. First, the approach was systematic and pre-planned for the search and selection of review articles, extraction of data and analyses. This reduced the potential for bias and strengthened the likely reproducibility of the study. Second, citations were only extracted from sections used as evidence for the LES–BW relationships and conclusions of the reviews. This approach ensured that only relevant articles cited as evidence were included.

The analysis also has a number of potential limitations. We were limited to using a database (Web of Science) from which it was possible to readily extract citations. The risk of only using one database is that relevant reviews may have been missed, although it would still provide an unbiased, reasonably comprehensive and representative sampling of reviews in this field. Future analyses may benefit from extraction of citations from several databases to extend the citation network with additional relevant reviews. Second, some of the guiding criteria and principles for the cited article selection and data extraction, as described in the methods, had to be operationalized by the authors specifically for this research, because there were no pre-existing recommended guidelines. Although this approach provided an objective and transparently replicable basis for decision-making, there is always some possible subjectivity and thus potential bias. This was minimised by requiring that the entire process of study selection and data extraction was agreed by two independent assessors against our defined criteria. Lastly, quality assessment of the included reviews was not undertaken, since this was not of primary interest in the present study.
study. However, quality assessment of included reviews can potentially add an extra element, providing information about a possible association between the quality of reviews and citation of different types of primary studies. From a broader perspective, formulation of standard approaches and guidelines for citation analyses should be encouraged, similar to the ‘Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines’ for systematic reviews and meta-analyses.

The present findings for LES and BW differ somewhat from other citation analyses in nutrition research. Leng examined how the first RCTs evaluating the efficacy of cholesterol-lowering diets in the secondary prevention of coronary heart disease were interpreted in reviews of the literature. They concluded that reviews supporting dietary interventions underutilised the available RCTs to a greater extent than other reviews. However, in contrast to the present study, their research used only network analysis as methodology and did not include a statistical analysis of the specific factors underlying citation of studies in reviews. Urlings et al. found a higher likelihood of being cited for studies with statistically significant results in the literature of dietary trans fatty acids and serum cholesterol. Sample size, JIF and authority of the author were

Figure 2  Network analysis based on evidence assessment units concluding a beneficial effect or association of low-energy sweeteners with body weight (n=11). Cited articles (n=33) are marked based on the type of study.

Figure 3  Network analysis based on evidence assessment units concluding an adverse effect or association of low-energy sweeteners with body weight (n=7). Cited articles (n=22) are marked based on the type of study.
also important determinants of citation. One significant finding from our analysis was that OR for being cited was higher for articles on children compared with articles in adults. This was also evident in several of the subgroup analyses. However, a simple explanation for this consistent finding may be that the number of cited articles on children (n=49) was considerably smaller than the number of cited articles on adults (n=89), suggesting that the total available evidence for children is smaller, thus making those articles more frequently cited in reviews.

A surprising result was the limited evidence for quantitative relationships between conclusions of reviews and the citing of articles with corresponding conclusions. A possible explanation for this can be that, in the absence of a quantitative (meta-) analysis, review authors selectively interpret some of the articles they cite in reviews. This is supported by the surprisingly large percentage (almost 15%) of cited articles which had no relevant conclusion at all in relation to LES and BW. Subgroup analysis showed that articles which contained information about the relationship of LES with BW, but from which it was not possible to draw a conclusion, were significantly more likely to be cited in reviews concluding a neutral effect or association. Thus, variable interpretations of the LES–BW relationship may arise from differences in the qualitative ‘weights’ authors assign to different specific parts of the same cited evidence base, possibly influenced by their prior beliefs or published views on this relationship. It can also arise if qualitative conclusions are being influenced by other less direct evidence, such as mechanistic studies (which were excluded from this analysis).

Across systematic reviews, important differences may arise from choices relating to setting inclusion and exclusion criteria, prior to (or even while) performing the review (see below).

Taken together, these findings show a very arbitrary pattern of citation across the overall body of reviews, which could be explained by selective choice and interpretation of cited articles by review authors. A well-known weakness of narrative reviews is the lack of a systematic search process, leaving space for selective and potentially biased citing of evidence. However, systematic reviews (with or without meta-analysis) are also potentially subject to citation bias through subjective choices made in the inclusion and exclusion criteria, and how studies are grouped or compared. In the case of the LES–BW relationship, recent systematic reviews have differed markedly in, for example, arbitrary restrictions on the types of LES exposures included, duration of RCTs or follow-up of prospective cohort studies, differences in comparators (water or sugar, for example) and so on, all leading to conclusions being based on quite different evidence sets. Over half of the EAUs concluded that evidence was insufficient to draw a conclusion about the relationship of LES with BW. It is a subjective judgement as to when evidence is believed to be sufficient to draw a conclusion within a specific area. However, the network analysis showed that differing conclusions are qualitatively associated with differing degrees of citation of RCTs vs observational evidence. If both types of evidence are cited and assigned equal weight in the overall interpretation, it is perhaps not surprising if authors deem it impossible to draw a conclusion.

Several approaches are suggested to address citation bias in future reviews of this topic. First, it has been recommended that new research is always placed in the context of the totality of different types of evidence, with consideration of their relative strengths and weaknesses. Looking at the totality of evidence will per definition lead to avoidance of citation bias. Furthermore, it can potentially close anticipated gaps in the literature, and focus the resources on the true gaps of evidence. Citation of the totality of evidence in reviews must be encouraged independent of the direction of the results. Arguably, this is what systematic reviews already should do, so perhaps greater emphasis needs to be given to better justification and consensus on the criteria for nature and quality of evidence. Second, when concluding that evidence is insufficient to draw a conclusion, more attention should be given to elaborating the specific needs for research that would resolve the gaps. This also raises the question of whether the topic really suffers from gaps in the evidence itself, or gaps in the consistent understanding and use of the existing evidence. These approaches are suggested to limit the number of reviews concluding simply that ‘evidence is insufficient to draw a conclusion’, and to focus instead on where there is consensus and remaining gaps in the literature on the effects of LES in relation to BW.

One of the more interesting findings of this study is from the network analyses showing that reviews concluding a beneficial relationship of LES with BW cited mainly RCTs, whereas reviews concluding an adverse relationship cited mainly observational studies. These findings represent a possible source of citation bias in the included reviews. However, from this analysis it is not possible to identify whether differences in cited articles arise due to ‘neutral’ processes used to select the literature or whether this is potentially (with intent or not) biased by review authors’ view of the relationship of LES with BW. Variation in the literature cited across reviews can only be explained by a few characteristics of the cited articles, suggesting that citation across reviews assessing the relationship of LES with BW is overall inconsistent and arbitrary. Inconsistent use of the available evidence may allow and account for the diversity of conclusions in the currently available reviews on LES and BW. Replication of the current analyses with further expansion or more types of analyses would be useful to confirm or refute the observations and suggested explanations given here.

**Contributors** The project was initiated by DM. All authors (MN, DM, AR and CR) made substantial contributions to the development of idea and methods. The protocol was drafted by MN with support from DM, AR and CR. Extraction of data was done by MN and DM and discussed with AR, where needed. The manuscript draft was written by MN, and carefully revised by DM, AR and CR. All authors were involved in both the analysis and interpretation of the data. MN and CR were
REFERENCES


Maki KC, Curry LL, Reeves MS, et al. Chronic consumption of rebaudioside a, a steviol glycoside, in men and women with type 2 diabetes mellitus. Food and Chemical Toxicology 2008;46:547–53.


99 Fowler SPG. Low-Calorie sweetener use and energy balance: results from experimental studies in animals, and large-scale prospective studies in humans. *Physiol Behav* 2011;118:517–23.


