Association between intake of non-sugar sweeteners and health outcomes: systematic review and meta-analyses of randomised and non-randomised controlled trials and observational studies

Ingrid Toews,1 Szimonetta Lohner,2 Daniela Küllenberg de Gaudry,1 Harriet Sommer,1,3 Joerg J Meerpohl1,4

ABSTRACT

OBJECTIVE
To assess the association between intake of non-sugar sweeteners (NSS) and important health outcomes in generally healthy or overweight/obese adults and children.

DESIGN
Systematic review following standard Cochrane review methodology.

DATA SOURCES
Medline (Ovid), Embase, Cochrane CENTRAL, WHO International Clinical Trials Registry Platform, Clinicaltrials.gov, and reference lists of relevant publications.

ELIGIBILITY CRITERIA FOR SELECTING STUDIES
Studies including generally healthy adults or children with or without overweight or obesity were eligible. Included study designs allowed for a direct comparison of no intake or lower intake of NSS with higher NSS intake. NSSs had to be clearly named, the dose had to be within the acceptable daily intake, and the intervention duration had to be at least seven days.

MAIN OUTCOME MEASURES
Body weight or body mass index, glycemic control, oral health, eating behaviour, preference for sweet taste, cancer, cardiovascular disease, kidney disease, mood, behaviour, neurocognition, and adverse effects.

RESULTS
The search resulted in 13,941 unique records. Of 56 individual studies that provided data for this review, 35 were observational studies. In adults, evidence of very low and low certainty from a limited number of small studies indicated a small beneficial effect of NSSs on body mass index (mean difference −0.6, 95% confidence interval −1.19 to −0.01; two studies, n=174) and fasting blood glucose (−0.16 mmol/L, −0.26 to −0.06; two, n=52). Lower doses of NSSs were associated with lower weight gain (−0.09 kg, −0.13 to −0.05; one, n=17 934) compared with higher doses of NSSs (very low certainty of evidence). For all other outcomes, no differences were detected between the use and non-use of NSSs, or between different doses of NSSs. No evidence of any effect of NSSs was seen on overweight or obese adults or children actively trying to lose weight (very low to moderate certainty). In children, a smaller increase in body mass index z score was observed with NSS intake compared with sugar intake (−0.15, −0.17 to −0.12; two, n=528, moderate certainty of evidence), but no significant differences were observed in body weight (−0.60 kg, −1.33 to 0.14; two, n=467, low certainty of evidence), or between different doses of NSSs (very low to moderate certainty).

CONCLUSIONS
Most health outcomes did not seem to have differences between the NSS exposed and unexposed groups. Of the few studies identified for each outcome, most had few participants, were of short duration, and their methodological and reporting quality was limited; therefore, confidence in the reported results is limited. Future studies should assess the effects of NSSs with an appropriate intervention duration. Detailed descriptions of interventions, comparators, and outcomes should be included in all reports.

SYSTEMATIC REVIEW REGISTRATION
Prospero CRD42017047668.

Introduction
Growing concerns about health and quality of life have encouraged people to adapt healthy lifestyles and avoid the consumption of food rich in sugars, salt, or fat to prevent obesity and other non-communicable diseases. With increased consumer interest in reducing energy intake, food products containing non-sugar sweeteners (NSSs) rather than simple sugars (monosaccharides and disaccharides) have become increasingly popular.1 Replacement of sugars with NSSs bears promise of health benefits primarily by reducing the contribution of sugars to daily calorie intake and thus reducing the risk of unhealthy weight...
However, evidence for health effects due to use of NSSs is conflicting. While some studies report an association between NSS use and reduced risk of type 2 diabetes, overweight, and obesity (thus suggesting a benefit for general health and the management of diabetes), other studies suggest that NSS use could increase the risk of overweight, diabetes, and cancer. Further investigations are needed to clarify the benefits and harms of NSS consumption. Therefore, the objective of our review was to investigate the health effects of NSSs in adults and children.

**Description of the exposure or intervention of interest**

Most NSSs so far have been synthesised, but through research and development in food chemistry and processing, the number of natural NSS compounds is increasing. NSSs differ from sugars not only in their taste properties, but also in how the body metabolises them and how they in turn affect physiological processes. NSSs are generally sweeter than sucrose, but contain far fewer or no calories. Each sweetener is unique in its sweetness intensity, persistence of the sweet taste, coating of the teeth, and aftertaste effect.

The definitions and terminology for NSSs vary. In some cases, the term “artificial sweeteners” is used as a synonym for NSSs, in other cases as a subcategory. In this systematic review, we use the term “NSSs” as a category including both artificial sweeteners and naturally occurring non-caloric sweeteners (fig 1). The term “NSSs” is also used by the CODEX Alimentarius (part of the Joint Food and Agriculture Organisation of the United Nations/World Health Organization Food Standards Programme), and this review was conducted in support of guidelines being developed by WHO.

The range of NSSs approved in different countries varies. In the United States, for example, the Food and Drug Administration has approved six NSSs for consumption, whereas the range of currently approved NSSs in the European Union is wider (eg, including cyclamate). In general, current evidence supports the safety of several NSSs to be used in foods. Recognised regulatory bodies have established acceptable daily intakes based on various safety studies. Other NSSs are currently declared as unsafe or have not yet been assessed.

Although many of the NSSs currently being used in foods have been declared safe for consumption at levels below the respective acceptable daily intakes, less is known regarding potential benefits and harms of NSSs within this range of intake, because evidence from studies and reviews is often limited and conflicting. WHO is developing guidance on the use of NSSs by adults and children based on the evidence generated by this systematic review. Following the guidance of the WHO Nutrition Guidance Expert Advisory Group Subcommittee on Diet and Health, this review seeks to comprehensively assess the association between commonly consumed NSSs and health by looking at the following research questions:

- In a general adult population, what are the effects of NSS consumption versus no consumption on relevant health outcomes?
- In a general adult population, what are the effects of higher versus lower NSS doses and more frequent versus less frequent NSS consumption on relevant health outcomes?
- In an overweight or obese adult population with explicit intentional weight loss, what are the effects of NSS consumption versus no consumption on relevant health outcomes?
- In a general child population, what are the effects of NSS consumption versus no consumption on relevant outcomes?
- In a general child population, what are the effects of higher versus lower NSS doses and more frequent versus less frequent NSSs consumption on relevant outcomes?
- In a population of overweight and obese children with explicit intentional weight loss, what are the effects of NSS consumption versus no consumption on relevant outcomes?

**Methods**

In accordance with the WHO guideline development process, we conducted a systematic review and meta-analyses according to the methodological recommendations of the Cochrane Collaboration. Ethical approval was not required for this research.

**Inclusion criteria**

The inclusion and exclusion criteria for this review were established prospectively and were based on their relevance for a WHO global guideline for NSS use by a generally healthy population. We included studies with a general, healthy population of adults (≥18 years) or children (<18 years), including those with overweight or obesity. Studies that exclusively included overweight or obese adults or children who were specifically trying to lose weight (that is, weight loss studies) were also included and analysed separately. We excluded studies including diseased populations, in vitro and animal studies. Studies with pregnant women were also excluded.

The interventions and exposures of interest included any type of NSSs, either as an individual intervention or in combination with other NSSs. Interventions or exposures described as “diet sodas,” “diet beverages,” or “diet soft drinks” were included when the sweeteners used in the products were NSSs and their...
We included non-randomised controlled trials from the second phase of crossover trials for the expected to last long enough to bias the results of the study because the effect of NSS intake is not randomised trials. In crossover randomised (quasi-)randomised controlled trials, and cluster randomised studies. We included all parallel grouped or crossover (quasi-)randomised controlled trials, and cluster randomised trials. In crossover randomised controlled trials, we considered both phases of the study because the effect of NSS intake is not expected to last long enough to bias the results from the second phase of crossover trials for the outcomes evaluated in this review. Furthermore, we included non-randomised controlled trials as well as prospective and retrospective cohort studies, case-control studies, and cross sectional studies but analysed them separately. Studies with observational design were included because the possible long term effects of NSSs—for example, on the incidence of non-communicable diseases such as cancer—are generally difficult to assess in randomised controlled trials. We included unpublished and ongoing studies.

### Search methods for identification of studies
The search strategy for this review combined electronic searches and hand searching. For the electronic searches, no date or language restrictions were applied. A systematic literature search in the following databases was conducted last on 25 May 2017 (by SL): Medline, Medline in Process and Medline Daily Update, Embase, and the Cochrane Central Register of Controlled Trials (CENTRAL). To identify ongoing or completed, but unpublished trials, the WHO International Clinical Trials Registry Platform (ICTRP) search portal was as well as ClinicalTrials.gov were searched on 23 November 2017 (by IT). Search strategies are listed in the supplementary file 1. The reference lists of relevant systematic reviews were screened manually to identify further potentially relevant citations.

### Selection of studies
All titles and abstracts of records identified in the databases and other sources above were screened for eligibility by one researcher (DKdG, SL, or IT). Two review authors independently evaluated full texts of all potentially eligible studies for appropriateness for inclusion without prior consideration of the results (DKdG, SL, IT). Any disagreements were resolved by discussion or feedback from a third author (JJM).

### Data extraction and management
Two review authors independently extracted data and cross checked the extracted information on study characteristics, and included participants, interventions, and reported outcomes using a piloted, standardised data extraction form in the online software Covidence (DKdG, SL, IT). Any differences related to the data extraction were resolved by rechecking the full text of the study or by discussion. If study data were only available from figures, data were extracted by use of the validated software Plot Digitizer (plotdigitizer.sourceforge.net). When study data were ambiguous or data were not reported in a form that could be used for formal comparison, we contacted the corresponding and first author of the original publication via email.

### Table 1 | Amount of acceptable daily intake of non-sugar sweeteners (not exhaustive) as defined by regulatory bodies for the general population

<table>
<thead>
<tr>
<th>Non-sugar sweetener</th>
<th>JECFAs17</th>
<th>European Food Safety Authority</th>
<th>US Food and Drug Administration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acesulfame K</td>
<td>15</td>
<td>918</td>
<td>15</td>
</tr>
<tr>
<td>Advantame</td>
<td>5</td>
<td>5</td>
<td>12.8</td>
</tr>
<tr>
<td>Aspartame</td>
<td>40</td>
<td>40</td>
<td>50</td>
</tr>
<tr>
<td>Brazzein</td>
<td>—</td>
<td>Not approved</td>
<td>Not approved</td>
</tr>
<tr>
<td>Cyclamate</td>
<td>11</td>
<td>7</td>
<td>Not approved Not approved</td>
</tr>
<tr>
<td>Neotame</td>
<td>0.3</td>
<td>0.218</td>
<td>0.30</td>
</tr>
<tr>
<td>Saccharin</td>
<td>15</td>
<td>5</td>
<td>15</td>
</tr>
<tr>
<td>Sucralose</td>
<td>5</td>
<td>15</td>
<td>5</td>
</tr>
<tr>
<td>Steviol glycosides</td>
<td>4</td>
<td>418</td>
<td>4</td>
</tr>
<tr>
<td>Thaumatin</td>
<td>Not approved</td>
<td>Not specified</td>
<td>Not approved</td>
</tr>
</tbody>
</table>

JECFA=Joint Food and Agriculture Organization of the United Nations/WHO Expert Committee on Food Additives.
Assessment of risk of bias
Two review authors independently assessed the risk of bias for each study. Any disagreements were resolved by discussion or a third author (JJM). For the risk of bias assessment of randomised controlled trials, we used the Cochrane risk of bias tool. For non-randomised controlled trials, we used the ROBINS-I tool (risk of bias in non-randomised studies of interventions). We planned to create funnel plots when data of 10 or more studies were available to assess the likelihood of dissemination bias. Since none of the meta-analyses included 10 studies or more, a thorough assessment of dissemination bias was not feasible.

Data synthesis
If not reported, we calculated the risk ratios and their respective 95% confidence intervals for randomised controlled trials, controlled clinical trials, and cohort studies, as well as odds ratios and their respective 95% confidence intervals for case-control studies. Mean differences or standardised mean differences with 95% confidence intervals were calculated for continuous outcomes. We conducted meta-analyses if comparable outcome data from two or more studies were available. In these meta-analyses, we used the random effects model. When baseline and final values were given, we computed changes from baseline. We imputed any missing standard deviation values using an imputed correlation coefficient. In this review, we used a correlation coefficient of zero. Statistical analyses were conducted by the statistical software R with the R package meta and metasens.

Sensitivity analyses
We tested the robustness of our results using sensitivity analyses. In forest plots, we reported results of analyses with the random effects model as our primary effect estimate. For all meta-analyses, we conducted sensitivity analyses using the fixed effect model. In most sensitivity analyses with the fixed effect model, the effects were more precise (narrower 95% confidence intervals) and consequently statistically significant at times, compared with analyses using the random effects model. However, given the clinical heterogeneity of the included studies, these were judged to not be appropriate, and therefore the results are not reported in detail. We found only one study with low risk of bias; thus, an analysis of studies with a low risk of bias only was not feasible. Study populations were divided into participants aged 18 years and older and those aged younger than 18 years in sensitivity analyses so that the effect of NSSs on children only and adults only could be analysed.

Assessment of the certainty of the evidence
We used the GRADE approach (grading of recommendations assessment, development, and evaluation) to assess the certainty of the evidence for the most relevant, available measures of all critical and important outcomes. According to the GRADE approach, we classified the certainty of evidence in four categories: high, moderate, low, and very low certainty of evidence. The GRADE certainty assessment per outcome was documented in GRADE evidence profiles, together with the pooled effects for the interventions. We used GRADEpro GDT online software to compile the evidence profiles. Assessments of the certainty of evidence for all outcomes were reviewed with the WHO Nutrition Guidance Expert Advisory Group Subgroup on Diet and Health as part of the WHO guideline development process.

For the outcomes with available evidence from randomised controlled trials, additional evidence from non-randomised studies and observational studies can be found in the supplementary materials (supplementary file 1, table 1). If case-control studies and cross sectional studies provided the best available body of evidence, we presented this evidence in the main text. Presentation of the results in this systematic review is primarily structured according to age group (adults or children) and outcome. Within the each outcome, we presented the results for each PICO question separately (that is, population, intervention, comparator, and outcome), describing results of randomised controlled trials first, followed by those of non-randomised and observational studies.

Patient and public involvement
No patients were involved in setting the research question or the outcome measures, nor were they involved in developing plans for design or implementation of this systematic review. No patients were asked to advise on interpretation or writing up of results. The results of this review will be disseminated to appropriate audiences. It was not evaluated whether the studies included in the review had any patient involvement.

Results
Details of the study selection are presented in figure 2. Key characteristics of all included studies are available in supplementary file 3.

Detailed results of the assessment of risk of bias in included randomised controlled trials (n=21) are summarised in supplementary file 1. Unclear reporting about random sequence generation and allocation concealment were the main reasons for unclear risk of bias in randomised controlled trials, while lack of blinding of participants and personnel was the main reason for high risk of bias. Other potential sources of bias were rarely suspected. The overall risk of bias assessment of controlled clinical trials and observational studies (n=35) was serious mainly due to suspected bias caused by confounding, and bias caused by classification of the intervention. The risk of bias assessment for individual non-randomised studies can be found in supplementary file 2.

NSS intake and health outcomes in adults
We included 17 randomised controlled trials, six controlled clinical trials, five prospective or retrospective cohort studies, 18 19 30 44 of evidence. The GRADE certainty assessment per outcome was documented in GRADE evidence profiles, together with the pooled effects for the interventions. We used GRADEpro GDT online software to compile the evidence profiles. Assessments of the certainty of evidence for all outcomes were reviewed with the WHO Nutrition Guidance Expert Advisory Group Subgroup on Diet and Health as part of the WHO guideline development process.

For the outcomes with available evidence from randomised controlled trials, additional evidence from non-randomised studies and observational studies can be found in the supplementary materials (supplementary file 1, table 1). If case-control studies and cross sectional studies provided the best available body of evidence, we presented this evidence in the main text. Presentation of the results in this systematic review is primarily structured according to age group (adults or children) and outcome. Within the each outcome, we presented the results for each PICO question separately (that is, population, intervention, comparator, and outcome), describing results of randomised controlled trials first, followed by those of non-randomised and observational studies.

Patient and public involvement
No patients were involved in setting the research question or the outcome measures, nor were they involved in developing plans for design or implementation of this systematic review. No patients were asked to advise on interpretation or writing up of results. The results of this review will be disseminated to appropriate audiences. It was not evaluated whether the studies included in the review had any patient involvement.

Results
Details of the study selection are presented in figure 2. Key characteristics of all included studies are available in supplementary file 3.

Detailed results of the assessment of risk of bias in included randomised controlled trials (n=21) are summarised in supplementary file 1. Unclear reporting about random sequence generation and allocation concealment were the main reasons for unclear risk of bias in randomised controlled trials, while lack of blinding of participants and personnel was the main reason for high risk of bias. Other potential sources of bias were rarely suspected. The overall risk of bias assessment of controlled clinical trials and observational studies (n=35) was serious mainly due to suspected bias caused by confounding, and bias caused by classification of the intervention. The risk of bias assessment for individual non-randomised studies can be found in supplementary file 2.

NSS intake and health outcomes in adults
We included 17 randomised controlled trials, six controlled clinical trials, five prospective or retrospective cohort studies, 18 19 30 44 of evidence. The GRADE certainty assessment per outcome was documented in GRADE evidence profiles, together with the pooled effects for the interventions. We used GRADEpro GDT online software to compile the evidence profiles. Assessments of the certainty of evidence for all outcomes were reviewed with the WHO Nutrition Guidance Expert Advisory Group Subgroup on Diet and Health as part of the WHO guideline development process.

For the outcomes with available evidence from randomised controlled trials, additional evidence from non-randomised studies and observational studies can be found in the supplementary materials (supplementary file 1, table 1). If case-control studies and cross sectional studies provided the best available body of evidence, we presented this evidence in the main text. Presentation of the results in this systematic review is primarily structured according to age group (adults or children) and outcome. Within the each outcome, we presented the results for each PICO question separately (that is, population, intervention, comparator, and outcome), describing results of randomised controlled trials first, followed by those of non-randomised and observational studies.

Patient and public involvement
No patients were involved in setting the research question or the outcome measures, nor were they involved in developing plans for design or implementation of this systematic review. No patients were asked to advise on interpretation or writing up of results. The results of this review will be disseminated to appropriate audiences. It was not evaluated whether the studies included in the review had any patient involvement.

Results
Details of the study selection are presented in figure 2. Key characteristics of all included studies are available in supplementary file 3.

Detailed results of the assessment of risk of bias in included randomised controlled trials (n=21) are summarised in supplementary file 1. Unclear reporting about random sequence generation and allocation concealment were the main reasons for unclear risk of bias in randomised controlled trials, while lack of blinding of participants and personnel was the main reason for high risk of bias. Other potential sources of bias were rarely suspected. The overall risk of bias assessment of controlled clinical trials and observational studies (n=35) was serious mainly due to suspected bias caused by confounding, and bias caused by classification of the intervention. The risk of bias assessment for individual non-randomised studies can be found in supplementary file 2.

NSS intake and health outcomes in adults
We included 17 randomised controlled trials, six controlled clinical trials, five prospective or retrospective cohort studies, 18 19 30 44 of evidence. The GRADE certainty assessment per outcome was documented in GRADE evidence profiles, together with the pooled effects for the interventions. We used GRADEpro GDT online software to compile the evidence profiles. Assessments of the certainty of evidence for all outcomes were reviewed with the WHO Nutrition Guidance Expert Advisory Group Subgroup on Diet and Health as part of the WHO guideline development process.

For the outcomes with available evidence from randomised controlled trials, additional evidence from non-randomised studies and observational studies can be found in the supplementary materials (supplementary file 1, table 1). If case-control studies and cross sectional studies provided the best available body of evidence, we presented this evidence in the main text. Presentation of the results in this systematic review is primarily structured according to age group (adults or children) and outcome. Within the each outcome, we presented the results for each PICO question separately (that is, population, intervention, comparator, and outcome), describing results of randomised controlled trials first, followed by those of non-randomised and observational studies.

Patient and public involvement
No patients were involved in setting the research question or the outcome measures, nor were they involved in developing plans for design or implementation of this systematic review. No patients were asked to advise on interpretation or writing up of results. The results of this review will be disseminated to appropriate audiences. It was not evaluated whether the studies included in the review had any patient involvement.

Results
Details of the study selection are presented in figure 2. Key characteristics of all included studies are available in supplementary file 3.

Detailed results of the assessment of risk of bias in included randomised controlled trials (n=21) are summarised in supplementary file 1. Unclear reporting about random sequence generation and allocation concealment were the main reasons for unclear risk of bias in randomised controlled trials, while lack of blinding of participants and personnel was the main reason for high risk of bias. Other potential sources of bias were rarely suspected. The overall risk of bias assessment of controlled clinical trials and observational studies (n=35) was serious mainly due to suspected bias caused by confounding, and bias caused by classification of the intervention. The risk of bias assessment for individual non-randomised studies can be found in supplementary file 2.
In randomised controlled trials, we saw no significant differences in change in body weight between adults receiving NSSs compared with those receiving different sugars or placebo (mean difference 1.29 kg, 95% confidence interval −2.80 to 0.21; five studies, n=229, very low certainty of evidence; fig 3). Only one study used placebo as a comparator 38 while the other studies used caloric sweeteners as a comparator. 34 37 39 40

There seemed to be no consistent difference in effect between studies using aspartame, 34 37 40 stevia, 38 or a combination of sweeteners 39 as the intervention.

Subgroup analysis by body weight status suggested that NSS use by overweight or obese individuals (that is, those not trying to lose weight, mean body weight 86.87 kg) resulted in reduced body weight of 1.99 kg (95% confidence interval −2.84 to −1.14; three studies, n=146, duration of studies, four weeks to six months) but no change in individuals of normal weight (0.03 kg, −0.03 to 0.09; two, n=110; fig 3). As assessed in randomised controlled trials, change in body mass index was 0.6 units lower in adults receiving NSSs than in those receiving sucrose (95% confidence interval −1.19 to −0.01; two studies, n=174, low certainty of evidence). Otherwise, randomised controlled trials, non-randomised controlled trials, and observational studies comparing NSS use with no use and with insufficient data for a meta-analysis indicated no consistent difference between the intervention and control group in relation to difference in body weight and other measures of overweight and obesity (supplementary material file 1, table 1).

In one cohort study, 30 researchers assessed different levels of NSS intake and reported that weight gain was 0.09 kg lower in women consuming up to 5.8 g saccharin per day compared with women consuming more than 5.8 g saccharin per day (95% confidence interval −0.13 to −0.05; one study, outcome assessed in n=17 934, very low certainty of evidence). Two randomised controlled trials 31 33 investigated the effect of NSS intake in overweight populations trying to lose weight, although they did not provide enough data to conduct meta-analysis (standard error or standard deviation not reported). One study 31 showed no difference in body weight between the study groups (mean difference 0.10 kg, 95% confidence interval −0.31 to 0.11; n=163, low certainty of evidence). The other study 33 showed no significant differences between the study groups with regard to reduction in body weight, body mass index, or body fat.

**Diabetes or glycaemic control**

In two randomised controlled trials, levels of fasting blood glucose were 0.16 mmol/L lower in the groups receiving aspartame or a combination of NSSs than in groups receiving sugar (95% confidence interval −0.26 to −0.06; two studies, n=52, very low certainty of evidence). 37 39 However, no differences were observed in plasma insulin levels (mean difference −1.60 pmol/L, 95% confidence interval −8.39 to 5.19; two, n=52) or in insulin resistance and β cell function as measured by the homeostatic model assessment of insulin resistance (HOMA-IR; −0.14, −0.38 to 0.10; two, n=66, very low certainty of evidence). 37 39 Additional markers for diabetes were reported by single studies only (supplementary material file 1, table 2).

**Eating behaviour**

*Energy intake and appetite—*Pooled data from four randomised controlled trials 18 39 41 (n=318 at baseline) showed that mean daily energy intake was 1064.73 kJ lower in people receiving NSSs than in those receiving sugar (95% confidence interval −1867.03 to −262.44; four studies, n=278, very low certainty of evidence; fig 4). Subgroup analysis by study duration and type of sweetener used as the intervention indicated that this result was largely being driven by one study that lasted for 10 weeks and used a combination of aspartame, cyclamate, acesulfame K, and saccharin (mean difference −2597.00, 95% confidence interval −3125.35 to −2068.65; n=62). Studies of short duration (lasting four weeks) using aspartame as the intervention did not show a significant reduction (−598.94, 95% confidence interval −1445.24 to 247.36; three studies, n=276). In one randomised
controlled trial, researchers reported narratively (that is, without numerical data) that there were no significant differences in energy intake between the stevia and placebo groups. Data from two non-randomised controlled trials (n=22) suggested no difference between the intervention and control groups for energy intake.

One randomised controlled trial investigated the effect of NSSs on energy intake in overweight populations trying to lose weight. In this study, mean daily energy intake was reported to be 548 kJ lower in the group receiving NSSs than in the group avoiding NSSs (95% confidence interval −692.73 to −403.27; n=128). In addition, no significant differences were observed for self control with respect to eating (mean difference −0.20, 95% confidence interval −1.03 to 0.63; n=186, low certainty of evidence) or feelings of hunger (−0.20, −1.03 to 0.63; n=186, low certainty of evidence). In another randomised controlled trial, researchers reported narratively that self reported appetite remained the same in groups receiving NSSs as well as those receiving no intervention over the study period of 12 weeks.

Sugar intake and sweet preference—The pooled effect from three randomised controlled trials showed that daily sugar intake was 89.71 g lower in adults receiving NSSs than in those receiving sugar (95% confidence interval −127.63 to −51.80; n=224). All three studies included overweight or obese participants. Both studies by Reid measured sugar intake by including the stevia and placebo groups. Data from two non-randomised controlled trials and one cross sectional study showed no differences in sugar intake between the intervention and control groups.

Two randomised controlled trials investigated the effect of NSSs on preference for sweet taste or sugar intake in overweight populations trying to lose weight. The preference for sweet taste, as assessed by desire for sweets (measured on a 0-10 scale with higher values indicating increased desire), was slightly lower in the group receiving NSSs than in the group not receiving NSSs (mean difference −0.2, 95% confidence interval −0.34 to −0.06; one study, n=186, moderate certainty of evidence). Sugar intake was similar between the groups after three years of follow-up (−0.00 g, −0.18 to 0.18; one, n=186).

Cancer
The risk for bladder or lower urinary tract cancer as assessed in meta-analysis of case control studies seemed to be similar in those exposed to sweeteners and those unexposed to sweeteners (odds ratio 1.03, 95% confidence interval 0.84 to 1.25; eight studies, n=4509, very low certainty of evidence; fig 6). The odds ratios for other types of cancer as reported in various observational studies suggested no difference in risk for different cancers except for ovarian cancer (0.61, 0.38 to 0.98; one case-control study, n=459) and pancreatic cancer (0.19, 0.08 to 0.46, one case-control study, n=978). The certainty of evidence for the risk of different types of cancers was very low.

We saw no association between consumption of higher doses of aspartame and incidence of the main subtypes of lymphoid cancers, non-Hodgkin lymphoma subtypes (P=0.69), or non-lymphoid leukaemia, in two prospective cohort studies with up to 10 years of follow-up (n=473 984). Similarly, no association was seen between consumption of higher NSS doses and lower urinary tract cancer (n=149, very low certainty of evidence) in one case-control study.

Blood pressure
Data from three randomised controlled trials showed that systolic and diastolic blood pressure were lower in people receiving NSSs than in those receiving sugar or placebo (systolic, mean difference −4.90 mm Hg, 95% confidence interval −9.78 to −0.03; diastolic, −3.27 mm Hg, −7.21 to 0.67; three studies, n=202 at baseline, very low certainty of evidence). The effect seemed stronger in studies using caloric sweeteners as comparators than in those that used a non-caloric comparator. In another randomised controlled trial, researchers reported narratively that there was no change in blood pressure in the study groups.

No significant differences in systolic and diastolic blood pressure were reported in one randomised controlled trial assessing the effect of aspartame in overweight populations trying to lose weight. After 12 weeks, the group differences in diastolic blood pressure were 6 mm Hg less in men and 1 mm Hg more in women when the aspartame group was compared with controls (not enough data for formal statistical comparison, very low certainty of evidence).

Other outcomes
In studies comparing NSS intake with no intake, we found an increased risk of depression in one cohort study (odds ratio 1.14, 95% confidence interval 1.02 to 1.27; n=263 923). We also found no effects on the incidence of kidney disease (very low certainty of evidence), mood (moderate certainty of evidence), neurocognition (low certainty of evidence), or risk of

---

**Table 1** Effect of non-sugar sweetener intake on weight change (kg) in adults

<table>
<thead>
<tr>
<th>Study</th>
<th>Mean difference (95% CI)</th>
<th>Weight %</th>
<th>Mean difference (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal weight</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kuzma 2015</td>
<td>15.2 -0.43 (-2.53 to 1.66)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maki 2008</td>
<td>21.5 0.03 (-0.03 to 0.09)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pooled estimate</td>
<td>36.7 0.03 (-0.03 to 0.09)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterogeneity: τ²=0, P=0.67, I²=0%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overweight and obese</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maerak 2012</td>
<td>20.6 -1.00 (-1.66 to -0.32)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reid 2014</td>
<td>21.3 -2.02 (-2.35 to -1.69)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raben 2001</td>
<td>21.3 -2.80 (-3.11 to -2.49)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pooled estimate</td>
<td>63.3 -1.99 (-2.84 to -1.14)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterogeneity: τ²=0.51, P=0.01, I²=93%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall effect</td>
<td>100.0 -1.29 (-2.80 to 0.21)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Fig 3** Effect of non-sugar sweetener intake on weight change (kg) in adults.
adverse events (eg, skin reactions, loss of appetite, and headaches; risk ratio 0.65, 95% confidence interval 0.16 to 2.59; three studies, n=167, low certainty of evidence). We identified no studies investigating the incidence of asthma or the incidence of allergies.

In studies comparing different doses of NSS intake, evidence from one crossover randomised controlled trial\(^\text{16}\) indicated a significant increase in depression in people consuming the higher aspartame dose compared with those consuming the lower dose (low certainty of evidence). The study reported significantly better results in participants receiving lower doses of aspartame with respect to neurocognition (low certainty of evidence), but no difference in adverse events for higher intake versus lower intake of aspartame (low certainty of evidence).\(^\text{36}\) Similarly, in two randomised controlled trials,\(^\text{31, 32}\) no significant differences in the risk for adverse events were observed between individuals receiving NSSs and those not receiving NSSs in overweight populations trying to lose weight (risk ratio 1.38, 95% confidence interval 0.58 to 3.28; n=204, low certainty of evidence). Detailed results on all outcomes are reported in supplementary file 1.

**NSS intake and health outcomes in children**

Overall, we identified four randomised controlled trials,\(^\text{54-87}\) two non-randomised controlled trials,\(^\text{83, 88}\) one case-control study,\(^\text{89}\) and one cross sectional study\(^\text{90}\) that contributed data to our review regarding the association between NSS intake and health outcomes in children. We identified one ongoing study in children.\(^\text{90}\)

**Body weight**

Two randomised controlled trials\(^\text{85, 91}\) found a similar weight gain in children receiving sucralose and acesulfame K\(^\text{31}\) or aspartame\(^\text{85}\) and children receiving sucrose (mean difference -0.60 kg, 95% confidence interval -1.33 to 0.14; two studies, n=467, low certainty of evidence; fig 7). After exclusion of the oldest age group (13-21 years) from one study\(^\text{85}\) in a sensitivity analysis, we saw no difference in effect (-0.50 kg, -1.43 to 0.42; two, n=722). Two randomised controlled trials\(^\text{87, 92}\) reported a significantly smaller increase in body mass index z score in children receiving sucralose and acesulfame K\(^\text{21}\) or sucralose alone,\(^\text{87}\) compared with children receiving sucrose (-0.15, -0.17 to -0.12; n=528, moderate certainty of evidence).

One randomised controlled trial\(^\text{92}\) (n=641) reported no group differences in body fat measured by electrical impedance (mean difference -0.83% body fat, 95% confidence interval -2.12% to 0.46%), waist circumference (-0.50 cm, -1.73 to 0.73), skinfold thickness (-1.5 mm, -4.71 to 1.71), and waist-to-height ratio (-0.50%, -1.73 to 0.73). In one randomised controlled trial including overweight or obese children involved in a weight loss programme,\(^\text{86}\) researchers reported a lower weight gain in children receiving aspartame than in children receiving placebo (-0.75 kg, -1.08 to -0.43; one study, n=57, low certainty of evidence).

**Dental health**

In one non-randomised controlled trial,\(^\text{88}\) mouth rinses with chlorhexidine were more effective than stevioside in decreasing plaque volume. Plaque volume was similar in the groups using water or stevioside (low certainty of evidence).

**Eating behaviour**

**Satiety, appetite, and energy intake**—In one randomised controlled trial (n=141), children receiving NSSs versus those receiving sucrose had similar self reported satiety one minute after intake (odds ratio 0.77, 95% confidence interval 0.46 to 1.29) and 15 minutes after intake (1.44, 0.86 to 2.40).\(^\text{84}\) Self reported appetite increase (risk ratio 0.86, 95% confidence interval 0.22 to 3.29) or appetite decrease (1.08, 0.44 to 2.63) were similar between the study groups in another randomised controlled trial (n=126).\(^\text{85}\) According to evidence from a third randomised controlled trial, energy intake was lower in the sucrose group compared with those consuming sucrose in overweight children.\(^\text{31, 32}\) No significant differences in the risk for adverse events were observed in two randomised controlled trials,\(^\text{84, 94}\) one non-randomised controlled trial,\(^\text{83}\) or mouth rinses with chlorhexidine.\(^\text{85}\) Energy intake was 6711, 6640, or 7728 kJ daily with aspartame, saccharin, or sucrose in the preschool group, respectively, and 8100, 8284, and 9293 kJ for school age children, respectively. In one randomised controlled trial with overweight children involved in active weight loss, researchers assessed change in appetite as self reported adverse events, which were reported to be no different between the study groups (incidence rate ratio 0.94, 95% confidence interval 0.35 to 2.49; one study, n=55, very low certainty of evidence).\(^\text{86}\)

**Preference for sweet taste**—One crossover non-randomised controlled trial\(^\text{83}\) (n=47) reported significantly lower sugar intake in children receiving aspartame or saccharin than in children receiving sucrose (not enough data for formal statistical comparison, very low certainty of evidence). The effect seemed to be strongly related to the sugar content of the experimental diets.

**Diabetes**

In one crossover non-randomised controlled trial,\(^\text{83}\) researchers found a significantly higher increase in blood glucose in children of preschool age receiving aspartame compared with sucrose (mean difference 0.24 mmol/L, 95% confidence interval 0.09 to 0.39; n=25), a significantly higher increase in blood glucose in children of school age receiving saccharin compared with sucrose (0.65 mmol/L, 0.44 to 0.86; n=23), and a significantly lower increase in blood glucose in children of preschool age receiving aspartame compared with saccharin (-0.75 mmol/L, -0.95 to -0.64; n=23, very
In one randomised controlled trial, a higher risk of adverse effects in overweight children involved in active weight loss not receiving NSSs versus those receiving NSSs was observed (incidence rate ratio 1.28, 95% confidence interval 0.86 to 1.91; n=126, low certainty of evidence).85 However, in one randomised controlled trial, a higher risk of adverse effects in overweight children involved in active weight loss not receiving NSSs versus those receiving NSSs was observed (incidence rate ratio 1.37, 95% confidence interval 1.05 to 1.79; n=55, low certainty of evidence).86 Overall, 103 adverse effects were noted involving the occurrence of adverse events between children receiving NSSs and children not receiving NSSs (risk ratio 1.28, 95% confidence interval 0.86 to 1.91; n=126, low certainty of evidence).85 In one non-randomised controlled trial, we found no difference in effect between children receiving NSSs and those not receiving NSSs on self rated mood (very low certainty of evidence).83 In one non-randomised controlled trial, we found no difference in effect between children receiving NSSs and those not receiving NSSs on self rated mood (very low certainty of evidence).83 In one non-randomised controlled trial, we found no difference in effect between children receiving NSSs and those not receiving NSSs on self rated mood (very low certainty of evidence).83 In one non-randomised controlled trial, we found no difference in effect between children receiving NSSs and those not receiving NSSs on self rated mood (very low certainty of evidence).83 In one non-randomised controlled trial, we found no difference in effect between children receiving NSSs and those not receiving NSSs on self rated mood (very low certainty of evidence).83 In one non-randomised controlled trial, we found no difference in effect between children receiving NSSs and those not receiving NSSs on self rated mood (very low certainty of evidence).83 In one non-randomised controlled trial, we found no difference in effect between children receiving NSSs and those not receiving NSSs on self rated mood (very low certainty of evidence).83 In one non-randomised controlled trial, we found no difference in effect between children receiving NSSs and those not receiving NSSs on self rated mood (very low certainty of evidence).83 In one non-randomised controlled trial, we found no difference in effect between children receiving NSSs and those not receiving NSSs on self rated mood (very low certainty of evidence).83
this review were often only measured indirectly with intermediate markers. Lastly, most included studies had small sample sizes and their study duration was often too short to infer any meaningful results in the long term.

Several other systematic and narrative reviews have examined the effects of NSSs on various health outcomes.\textsuperscript{14, 56, 94–97} The methodological and clinical inclusion and exclusion criteria used in these systematic reviews differed substantially from our criteria in the present study, resulting in a different pool of included studies. The data synthesis methods also differed from the ones used in the present review. Still, the reviews found similar results to our results: Brown and colleagues\textsuperscript{4} found no strong clinical evidence for an effect of artificial sweeteners on metabolic effect in youths, whereas Cheungpasitporn and colleagues\textsuperscript{3} found no effect of artificially sweetened soda on chronic kidney disease. Greenwood and colleagues\textsuperscript{5} reported no consistent association between artificially sweetened soft drinks and diabetes risk. Onakpoya and Henegham\textsuperscript{95} reported a non-significant reduction in systolic blood pressure and significant reductions in diastolic blood pressure and fasting blood glucose with steviol glycoside compared with placebo, but indicated that the evidence was not robust due to heterogeneity.

Wieber and colleagues\textsuperscript{9} reported a decrease in body mass index in people consuming foods and drinks containing non-caloric sweeteners compared with an increased body mass index in those consuming foods and drinks containing sucrose. The researchers further highlighted the lack of high quality research regarding non-caloric sweeteners. A systematic review by Azad and colleagues\textsuperscript{97} found no statistically significant effect of non-nutritive sweeteners on body mass index, body weight, fat mass, waist circumference, and HOMA-IR. Overall, published systematic reviews rarely drew firm conclusions. Main methodological concerns were limitations in the literature search and the data analyses. By contrast to our review, most meta-analyses were not planned and conducted, and the authors summarised the individual study results narratively instead.

A few large prospective cohort studies\textsuperscript{98–102} with long term follow-up investigated the association between NSS intake and different health outcomes. However, the NSSs being investigated were not sufficiently specified to match the inclusion criteria of this review. Still, their results indicate an increased risk of higher body mass index and type 2 diabetes with higher NSS consumption, or lower risk of cardiovascular disease with intake of artificially sweetened sodas compared with sugar sweetened sodas. These results partly conflict with the ones from the findings of this systematic review. Included studies investigated long term health outcomes for a relatively short duration—for example, cardiovascular health outcomes\textsuperscript{99, 100} and diabetes\textsuperscript{98, 101} investigated for six months or less. Long term studies with sufficient statistical power are key to investigating long term health outcomes such as incidence of diabetes or other long term outcomes.
cardiovascular health. Hence, results of large, long term cohort studies should be verified by studies that specify the type of sweeter used.

The findings of our review might be biased by the fact that only one reviewer assessed inclusion of studies in the initial title and abstract screening phase. Hence, relevant references could have inadvertently not been included in this review. However, this possibility is unlikely because only clearly irrelevant references were excluded at this stage. Furthermore, we did not seek clarification with the study authors about whether our assessment of risk of bias in the individual studies was correct. In the statistical analyses, missing standard deviations for change in outcomes were imputed, and in some cases, approximation was used for the analyses. Therefore, the reliability of analyses of changes in outcomes might have been weakened by the unavailability of data and the use of imputed values and approximation.

Implications for clinicians and policy makers, unanswered questions, and future research
This review was prepared to inform a WHO guideline on NSS use. The guideline will provide information on implications for actions by health experts and policy makers. So far, several studies on the effects of NSSs on different health outcomes have been conducted. However, their methodological or reporting quality is mostly limited and often not sufficiently detailed to include their results in meta-analyses. Moreover, included studies differed substantially in their design (that is, choice of population, intervention, comparator, and outcome measures). Given these relevant differences between studies, a reliable review of the effects by type of sweeter or of the caloric effects versus non-caloric effects is challenging. Type of intervention and comparator might affect health outcomes differently and should be considered in future research.

We also recommend that future studies assess the effects of NSS use on health outcomes with an appropriate study duration. Study planning should consider the duration necessary for plausible, relevant effects to occur in the different outcomes of interest. Longer term studies are needed to assess effects on overweight and obesity, risk for diabetes, cardiovascular disease, and kidney disease. Type and dose of sweeter use should be reported precisely and transparently in all studies. Precise reporting of sweeter content (that is, type and amount of sweeteners) in ready-to-consume foods and beverages is highly desirable and could be helped by more detailed information on ingredients as provided by manufacturers. Consistent use of core outcome measures and consensus on timing and mode of assessment would further help researchers pool data across studies. In addition to studying the effects on NSS use in a general healthy population of adults and children, research should focus on diseased populations and other subgroups, including pregnant women and their offspring and people who use NSSs in amounts higher than average (such as those with diabetes).

Most of the studies identified for this review used single sweeteners and the use patterns of sweeteners in the studies might differ from that in real life practice. Therefore, the certainty in the evidence presented in this review might further be affected by indirectness. For example, NSSs can be consumed in different ways, including as a table top sweeter (that is, added to tea or coffee as a replacement for sugar) where the dose is freely determined by users themselves and might be higher than in that recorded the studies. Moreover, by contrast to many of our included studies that used a single NSS only, many food items have different types of NSSs that are combined to cover different bitter or metallic aftertastes of individual sweeteners and provide an adequate sweetness. Future research might consider exploring the effects of different combinations of sweeteners in doses similar to real life use patterns and compare the effects of higher versus lower NSS doses. Development and research on NSSs is ongoing, and new alternatives to sugar are presented on a regular basis. Therefore, we also need data on the safety and benefits and harms of other sweeteners not assessed in this review for a comprehensive overview of the health effects of NSSs.

Results of observational studies on the health effects of NSSs should be interpreted with caution, and attention should focus on plausible residual confounding as well as reverse causality (such as a higher consumption of NSSs by overweight or obese populations aiming at weight management). Appropriate long term studies that consider baseline consumption of sugar and NSSs and have an appropriate comparator should investigate whether NSSs are a safe and effective alternative to sugar, and results should be interpreted in light of these study design characteristics.

The WHO Nutrition Guidance Advisory Group (NUGAG) Subgroup on Diet and Health provided valuable insight on aims and objectives of this review. WHO agreed to the publication of this systematic review in a scientific journal because it serves as the background evidence review for WHO guidelines on non-sugar sweeteners and should therefore be available widely.

Contributors: SL, IT, and JIM conceived and designed the review. JIM coordinated the review. SL, IT, DKdG, and JIM screened the papers against eligibility criteria. SL, IT, and DKdG appraised the quality of the papers. SL, IT, DKdG, and HS extracted data from the papers. HS analysed the data. SL, JIM, IT, DKdG, and JIM wrote the review and its protocol. JIM, SL, and DKdG provided general advice on the review. JIM secured funding for the review. SL, JIM, and IT performed previous work that was the foundation of the current review. JIM is the guarantor of this manuscript. The questions guiding the review were discussed and developed by the WHO NUGAG Subgroup on Diet and Health, and the study protocol was approved, by the NUGAG Subgroup on Diet and Health. Neither WHO nor the WHO NUGAG Subgroup on Diet and Health played a role in data collection or analysis. All authors, external and internal, had full access to all of the data (including statistical reports and tables) in the study and can take responsibility for the integrity of the data and the accuracy of the data analysis. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

**Funding:** The research was funded by WHO. The research was conducted independently from the funder, and researchers are independent from the funder.

**Competing interests:** All authors have completed the CIJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare support from WHO for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years, no other relationships or activities that could appear to have influenced the submitted work.

**Ethical approval:** Ethical approval was not required for this research.

**Data sharing:** Full datasets can be obtained from the corresponding author Meerpohl@cochrane.de.

The lead author affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.


**Supplementary file 1:** Supplementary materials

**Supplementary file 2:** Results of the assessment of risk of bias in included observational studies

**Supplementary file 3:** Details of included studies (RCT-randomised, controlled trial; non-RCT-non-randomised controlled trial; AS-artificial sweetener, CVD-cardiovascular disease); *For profit funding includes sponsoring of study material, i.e. intervention substances, as well as financial sponsoring for conducting the study.