Evaluating agreement between bodies of evidence from randomised controlled trials and cohort studies in nutrition research: meta-epidemiological study

Lukas Schwingshackl,1 Sara Balduzzi,2 Jessica Beyerbach,1 Nils Bröckelmann,1 Sarah S Werner,1 Jasmin Zähringer,1 Blin Nagavci,1 Joerg J Meerpolh1,3

ABSTRACT

OBJECTIVE
To evaluate the agreement between diet-disease effect estimates of bodies of evidence from randomised controlled trials and those from cohort studies in nutrition research, and to investigate potential factors for disagreement.

DESIGN
Meta-epidemiological study.

DATA SOURCES
Cochrane Database of Systematic Reviews, and Medline.

REVIEW METHODS
Population, intervention or exposure, comparator, outcome (PI/ECO) elements from a body of evidence from cohort studies (BoE(CS)) were matched with corresponding elements of a body of evidence from randomised controlled trials (BoE(RCT)). Pooled ratio of risk ratios or difference of mean differences across all diet-disease outcome pairs were calculated. Subgroup analyses were conducted to explore factors for disagreement. Heterogeneity was assessed through $I^2$ and $\tau^2$. Prediction intervals were calculated to assess the range of possible values for the difference in the results between evidence from randomised controlled trials and evidence from cohort studies in future comparisons.

RESULTS
97 diet-disease outcome pairs (that is, matched BoE(RCT) and BoE(CS)) were identified overall. For binary outcomes, the pooled ratio of risk ratios comparing estimates from BoE(RCT) with BoE(CS) was 1.09 (95% confidence interval 1.04 to 1.14; $I^2=68%$; $\tau^2=0.021$; 95% prediction interval 0.81 to 1.46). The prediction interval indicated that the difference could be much more substantial, in either direction. We further explored heterogeneity and found that PI/ECO dissimilarities, especially for the comparisons of dietary supplements in randomised controlled trials and nutrient status in cohort studies, explained most of the differences. When the type of intake or exposure between both types of evidence was identical, the estimates were similar. For continuous outcomes, small differences were observed between randomised controlled trials and cohort studies.

WHAT THIS STUDY ADDS

Previously, several randomised controlled trials comparing dietary with control interventions have failed to replicate the (presumably protective) associations between dietary factors and risk of non-communicable diseases observed in large scale cohort studies.
However, some consistent findings between cohort studies and randomised controlled trials have been also reported.
Systematic evaluation of the two bodies of evidence between trials and cohort studies, with an investigation on factors for disagreement, has not yet been conducted.

WHAT IS ALREADY KNOWN ON THIS TOPIC:

Non-randomised controlled trials comparing dietary with control interventions have failed to replicate the apparent associations between dietary factors and risk of non-communicable diseases observed in large scale cohort studies.

Introduction

The Global Burden of Disease study group showed that non-communicable diseases accounted for 73% of deaths worldwide.1 According to the Global Burden of Disease study, which is based on evidence from prospective cohort studies, suboptimal diet accounted for 22% of all deaths worldwide, and 15% of all disability adjusted life years.2

The Global Burden of Disease studies and dietary guidelines are predominantly based on bodies of evidence (BoE) from cohort studies,3 although evidence from randomised controlled trials exists as well. Cohort studies with patient relevant outcomes provide valuable insights into associations between diet and disease.2 4 However, nutrition research, predominantly nutritional epidemiology, has been criticised for providing potentially less trustworthy estimates of diet associated risks or benefits.5 Therefore, limitations such as residual confounding and measurement errors of cohort studies in nutrition research need to be considered in depth.3 On the other hand, randomised controlled trials, if well designed and well conducted, give robust answers to the research questions under consideration and are widely accepted as the ideal...
methodology for causal inference. However, dietary trials often have methodological limitations, such as small sample sizes, short intervention periods, as well as blinding and low compliance issues.

In the past, several randomised controlled trials comparing dietary interventions with placebo or control interventions have failed to replicate the (presumably protective) associations between dietary factors and risk for non-communicable diseases found in large cohort studies. For example, randomised controlled trials found no beneficial effect of fibre intake on colorectal cancer risk, or of vitamin E on cardiovascular diseases. In other instances, consistent findings between cohort studies and randomised controlled trials have been reported (e.g., Mediterranean diet and risk of cardiovascular disease and type 2 diabetes), but to the best of our knowledge no systematic evaluation of agreement, with an investigation on factors for disagreement, between the two BoE has ever been conducted.

This meta-epidemiological study aims to determine the extent to which estimates between diet and disease based on BoE from randomised controlled trials are in agreement with those estimates based on BoE from cohort studies, and further investigate reasons behind any disagreement. These findings will allow us to better understand and explore the possible integration of both BoE in prospective nutrition evidence syntheses.

Methods
This meta-epidemiological study was planned, written, and reported in adherence to guidelines for reporting meta-epidemiological research. The inclusion criteria (patients or population, intervention or exposure, comparator, and outcome (PI/ECO)) are described in box 1.

Identification of systematic reviews of randomised controlled trials
We searched the Cochrane Database of Systematic Reviews, for systematic reviews of randomised controlled trials, published between 1 January 2010 and 31 December 2019 (supplementary appendix 1). Titles and abstracts were screened by one reviewer (LS), and subsequently all potentially relevant full texts were screened and assessed by two reviewers independently (LS, JZ). Discrepancies were resolved by a third reviewer (JJM).

Identification of matching systematic reviews of cohort studies
After identifying all potentially relevant systematic reviews of randomised controlled trials, we searched for matching systematic reviews of cohort studies. Firstly, we screened all eligible Cochrane reviews, to determine whether they also included cohort studies. Secondly, we conducted searches for systematic reviews of cohort studies in Medline, published within the past 10 years (1 January 2010 and 31 December 2019; supplementary appendix 2). We selected a period of 10 years to ensure comparability between the two BoE.

Matching bodies of evidence according to PI/ECO criteria
For all potentially eligible BoE of cohort studies, two reviewers judged whether each PI/ECO element matched those of the corresponding BoE of randomised controlled trials according to three definitions (supplementary table 1): more or less identical (very closely matched), similar but not identical (closely matched), or broadly similar (matched, but less close). Differences in reviewer ratings of one level disagreement were resolved by discussion (we considered the broader similarity rating for the overall PI/ECO rating); a third reviewer adjudicated the overall PI/ECO match rating for differences of more than two levels. Based on these criteria, we classified each comparison of effect estimates from randomised controlled trials and cohort studies for a given outcome according to the same three definitions: more or less identical, similar but not identical, and broadly similar. For each eligible systematic review of randomised controlled trials, we matched a maximum of six outcomes (maximum three patient relevant outcomes; and maximum three intermediate disease outcomes) for a given intervention or exposure. Selection of outcomes was based on the ranking in the summary of findings tables in the identified Cochrane reviews (from top to bottom).

Data extraction
We extracted data for every eligible BoE pair (BoE from a randomised controlled trial and matched BoE from a cohort study) related to the association between diet and disease (e.g., all cause mortality, cardiovascular disease, stroke, type 2 diabetes). These data included the name of first author, year of publication, description of population (e.g., disease status), age range, intervention or exposure (e.g., dietary pattern, food group, food, macronutrient, micronutrient), description of comparator (e.g., placebo, lowest intake or status category, control diet), definition of outcome, study design (parallel, crossover, factorial for randomised controlled trials); prospective, nested case-control studies, case cohort studies (for cohort studies); effect estimates (risk ratio, hazard ratio, odds ratio, mean differences, 95% confidence interval), type of comparison (e.g., high v low, dose-response), number of studies included, sample size, number of cases, duration of intervention or exposure (range), risk of bias or study quality ratings, and certainty of evidence rating. Data were extracted by three reviewers (LS, JB, or SSW) using a piloted data extraction form.
Box 1: Detailed description of inclusion criteria, by population

**Intervention or exposure**
- Dietary pattern: for example, Mediterranean diet, Dietary approaches to Stop Hypertension, low carbohydrate diet
- Food groups (macro-level) and foods (micro-level): for example, grains, vegetables, fruit, milk and dairy products, meat, processed meat, fish, eggs, nuts, chocolate, oils

**Macronutrients**:
- Carbohydrates: starch, fructose, glucose, sucrose
- Fat: for example, n-3 fatty acids (eicosapentaenoic acid, docosahexaenoic acid, α-linolenic acid), n-6 fatty acids (linoleic acid), monounsaturated fat
- Protein: for example, amino acids

**Micronutrients**:
- Vitamins: B carotene; vitamins A, E, C (ascorbic acid), and D (cholecalciferol, ergocalciferol); B vitamins (thiamine, riboflavin, niacin, pyridoxine, cobalamin, folic acid)
- Minerals: magnesium, calcium, selenium, sodium, potassium, iron, zinc, copper, iodine
- Other: fibre (psyllium, inulin, cellulose), probiotics, prebiotics, and synbiotics

**Control or comparison**
- Low (no) intake (status) level of the above interventions or exposure
- Placebo or usual care

**Outcomes**
- For example, all cause mortality, cardiovascular disease, coronary heart disease (myocardial infarction, ischaemic heart disease, and acute coronary syndrome), stroke, cancer, type 2 diabetes, dementia, fractures, age related macular degeneration, anthropometric outcomes, important intermediate disease markers such as systolic blood pressure, and diastolic blood pressure, fasting glucose, and low density lipoprotein cholesterol

**Study design**
- Systematic reviews of randomised controlled trials
- Matching systematic reviews of cohort studies (if available prospective cohort studies were preferred)

Where a BoE reported effect estimates based on a pool of studies of variable design (that is, case-control, cross sectional studies, retrospective cohort studies, or quasi-randomised controlled trials), we recalculated the pooled effect estimates by excluding non-cohort studies and non-randomised controlled trials. Also, if an intervention in a BoE of randomised controlled trials (eg, low v high sodium) and an exposure in a BoE of cohort studies (eg, high v low sodium) investigated opposite comparisons, we recalculated the risk estimates, respectively (eg, low v high sodium). Moreover, where a BoE reported effect estimates based on dietary intake and dietary supplements, nutrient status (eg, plasma selenium status) and dietary intake, or nutrient status and dietary supplements, we recalculated effect estimates whenever feasible to improve comparability between exposures in cohort studies and interventions in randomised controlled trials. For example, if a meta-analysis of randomised controlled trials investigated the effect of selenium supplements, and the authors of the matched meta-analysis of cohort studies combined plasma selenium status with selenium supplements, we excluded the studies with plasma selenium status and recalculated the effect estimates only based on the studies with selenium supplements.

**Statistical analysis**

If the effect estimate of the BoE from randomised controlled trials was expressed in a different measure than the effect estimate of a BoE from cohort studies, we used the appropriate conversion formulas in order to express both estimates in the same measure—that is, risk ratios for binary outcomes and mean differences for continuous outcomes. The relevant formula to transform an odds ratio to a risk ratio requires an assumed control risk:

\[
RR = (OR ÷ (1 − ACR × (1 − OR)))
\]

Where \( RR \) = risk ratio, \( OR \) = odds ratio, and \( ACR \) = assumed control risk.

Ten meta-analyses of cohort studies, included in seven systematic reviews, used an odds ratio as a summary measure (supplementary table 2); the median comparator group risk from the included studies was used for the assumed control risk required for transformation of each pooled odds ratio. If these data were not directly available in the meta-analyses of cohort studies, we used the median comparator group risk from the studies included in the corresponding meta-analyses of randomised controlled trials.

In six analyses (each including only one study) coming from two systematic reviews of randomised controlled trials, the results were expressed using hazard ratios; we did not consider the hazard ratios, but went back to the primary studies and extracted relevant data in order to obtain a risk ratio (that is, the number of randomised patients and number of patients with the outcomes of interest, in each arm; supplementary table 2).

To compare the two BoEs (that is, from randomised controlled trials and cohort studies), we synthesised the differences in the results coming from all eligible outcome pairs. Binary outcomes were expressed as ratio of risk ratios, while continuous outcomes were expressed as differences of mean differences. By using the BoE of cohort studies as the reference group, we examined the pooled estimate to determine a relatively larger or smaller estimate from the BoE of randomised controlled trials (that is, effect of BoE of trials > effect of BoE of cohort studies, or effect of BoE of trials < effect of BoE of cohort studies). For example, a risk ratio from randomised controlled trials of 0.95 and a risk ratio from cohort studies of 0.90 would result in a ratio of risk ratios of 1.06; whereas a risk of 1.00 in cohort studies compared with a risk ratio of 1.06 in randomised controlled trials would also result in a ratio of risk ratios of 1.06. Therefore, the ratio of risk ratios should not be interpreted as larger or smaller estimate from the BoE of randomised controlled trials, but only as differences between the two BoEs; and the direction of difference depends on direction of effect of the underlying BoEs.

We conducted a priori planned subgroup analyses: type of dietary intervention or exposure, outcome, and PI/ECD similarity degree (more or less identical, similar but not identical, and broadly similar). We also conducted two post hoc sensitivity analyses.
excluding highly correlated outcomes. Firstly, we did a conservative sensitivity analysis including only one outcome per comparison (that is, the outcome with the largest number of randomised controlled trials) from each Cochrane review. Secondly, we did a sensitivity analysis including outcomes based on their ranking in the summary of findings tables in the identified Cochrane reviews (from top to bottom). For example, for the α-linolenic acid intervention or exposure, the outcomes of coronary heart disease, cardiovascular disease, and cardiovascular mortality are likely to be highly correlated. Because cardiovascular mortality was mentioned first in the summary of findings table, cardiovascular mortality was chosen to be included, while the other two outcomes were excluded (supplementary table 3). Finally, a sensitivity analysis was also performed for Cochrane reviews that included both randomised controlled trials and cohort studies.

We obtained pooled estimates through a random effects meta-analysis model. We assessed heterogeneity through the $I^2$ and $τ^2$ statistics. The $τ^2$ statistic was estimated by the Paule and Mandel method, which is the recommended method for binary outcomes and performs well also with continuous ones. Furthermore, we calculated 95% prediction intervals to show the range of possible values for the difference between BoEs of randomised controlled trials and those of cohort studies that might be observed in future comparisons. We conducted all the meta-analyses using the R package meta.

Patient and public involvement

We did not involve patients or members of the public when we selected the research question, designed the study, interpreted the results, or wrote the manuscript. Although there was no direct patient and public involvement in this paper owing to the methodological design of our study, we asked a member of the public to read our manuscript after submission.

Results

The literature search identified 333 systematic reviews (Cochrane reviews) of randomised controlled trials, of which 65 full texts were assessed for inclusion (supplementary fig 1, and supplementary table 4), and 33 were included in this study.26 28 29 36-65 We found 3318 systematic reviews of cohort studies, from which 46 systematic reviews of cohort studies (with matching systematic reviews of randomised controlled trials) were included (supplementary fig 2, and supplementary table 4).17 18 21 25 27 66-103 Two of the Cochrane reviews contained also cohort studies and were therefore included in this study.26 166

Overall, we included 97 diet-disease outcome pairs of randomised controlled trials and cohort studies (that is, estimates based on BoE from trials matched with those based on BoE from cohort studies related to the association between diet and disease; supplementary table 5). We recalculated 34 pooled estimates from 21 systematic reviews.21 24-27 40 48 56 61 71 72 77 80 83 86 89 93 96 101 102 The number of primary studies contributing to the 97 diet-disease outcome pairs ranged from 1 to 64 (median 6) for BoE from randomised controlled trials, and from 1 to 68 (median 7) for BoE from cohort studies (overall >950 trials and >750 cohort studies). The total number of participants ranged from 56 to 211 957 for BoE from randomised controlled trials, and from 2563 to 1 797 670 for BoE from cohort studies. Of the identified 97 diet-disease outcome pairs, 83 were included in the meta-analysis (71 binary, 12 continuous). We could not include 14 diet-disease outcome pairs in the meta-analysis (reasons in supplementary table 6).

The interventions or exposures investigated in the identified systematic reviews could be categorised into micronutrients (n=47), dietary approach (n=19), fatty acids (n=17), food groups (n=5), fibre (n=4), phytonutrients (n=3), and food (n=2). Across the BoE of randomised controlled trials, the intervention was either given in the form of dietary supplements (n=43), dietary intake (n=38), or both (n=16). Interventions on intake were mainly attempts to modify dietary intake via dietary advice or dietary counselling to reduce, for example, fat or sodium intake, but dietary adherence to these interventions was mainly not assessed in the primary systematic reviews. Across the BoE of cohort studies, the exposure measured was dietary intake (n=69), nutrient status (n=16), dietary supplements (n=8), dietary intake and dietary supplements (n=2), or dietary intake and nutrient status (n=2).

The type of intake or exposure between both BoEs was the same for dietary intake across 36 diet-disease outcome pairs and for dietary supplements across eight diet-disease outcome pairs, respectively. The diseases clusters included cardiovascular disease (n=22), intermediate disease markers (n=22), pregnancy outcomes (n=17), all cause mortality (n=15), cancer (n=12), eye disease (n=3), neurodegenerative disease (n=3), bone health (n=2), and type 2 diabetes (n=1). All Cochrane reviews evaluated risk of bias, whereas the Newcastle-Ottawa scale was the most often used instrument to evaluate study quality for BoE of cohort studies (n=48; mean rating 7.5). Certainty of evidence was rated for 48 BoE of randomised controlled trials using GRADE: very low (n=5), low (n=16), moderate (n=14), and high (n=13). For 10 BoE of cohort studies rated (for two outcomes NutriGrade was used), the certainty of evidence was measured: very low (n=8), low (n=1), moderate (n=1). Detailed study characteristics including effect estimates, description of population, age, description of intervention or comparator, outcomes, range study length, and risk of bias or study quality of primary studies included in each diet disease pair are given in supplementary tables 7-12.

Similarities

Of 97 diet-disease outcome pairs, none was rated as more or less identical, 57 (59%) were similar but not identical, and 40 (41%) were broadly similar. Interventions or exposures rated as broadly similar accounted for most PI/ECO dissimilarities overall.
Statistical heterogeneity
Across individual meta-analyses of randomised controlled trials, the mean \( \tau^2 \) was 21\% (\( \tau^2=0.018 \)), whereas the median \( \tau^2 \) was 2\% (\( \tau^2=0.018 \)). The heterogeneity (\( \tau^2 \)) was lower for binary outcomes (mean \( \tau^2=19\% \); median \( \tau^2=0\% \)) than for continuous outcomes (\( \tau^2=31\% \); \( \tau^2=23\% \)). Across individual meta-analyses of cohort studies, the mean \( \tau^2 \) was 47\% (\( \tau^2=0.023 \)), whereas the median \( \tau^2 \) was 54\% (\( \tau^2=0.011 \)). The heterogeneity was lower for binary outcomes (mean \( \tau^2=44\% \); median \( \tau^2=48\% \)) than for continuous outcomes (\( \tau^2=81\% \); \( \tau^2=86\% \); supplementary table 14).

Pooled estimate
Overall, 83 diet-disease outcome pairs were included in the meta-analysis. For binary outcomes, 71 pairs were included. The treatment effects were more often larger in the BoE of cohort studies (n=44) than in the BoE of randomised controlled trials (n=25), and for two outcome pairs the treatment effects were of similar magnitude (supplementary table 5). The risk ratio was <1 across 64 BoE from cohort studies, whereas the risk ratio was ≥1 across seven. The risk ratio was <1 across 48 BoE from randomised controlled trials, whereas it was ≥1 in 23 BoE from randomised controlled trials. For continuous outcomes, the treatment effects were more often larger in the BoE of randomised controlled trials (n=7) than in the BoE of cohort studies (n=5). For eight outcomes, we observed a risk ratio difference greater than 0.25 between the BoE from randomised controlled trials compared with the BoE from cohort studies (but only two instances showed a strong difference >0.5).

The pooled estimate, using the BoE of cohort studies as the reference group, showed that on average the BoE of randomised controlled trials had slightly different estimates compared to that of cohort studies (ratio of risk ratios 1.09 (95\% confidence interval 1.04 to 1.14)); 95\% prediction interval 0.81 to 1.46; fig 1 and table 1). The prediction interval indicated that the difference could be much more substantial, in either direction. Substantial heterogeneity (\( \tau^2=68\% \); \( \tau^2=0.021 \)) was observed. When the BoE from cohort studies with a risk ratio <1 versus ≥1 were analysed separately, the ratios of risk ratios were 1.12 (95\% confidence interval 1.07 to 1.17); \( \tau^2=60\% \); \( \tau^2=0.016 \); 95\% prediction interval 0.87 to 1.45; n=64; supplementary fig 3) and 0.89 (0.79 to 1.00; 44\%; 0.013; 0.64 to 1.24; n=7; supplementary fig 4), respectively.

For continuous outcome pairs (n=12), we observed no differences between randomised controlled trials and cohort studies, apart from smaller systolic and diastolic blood pressure estimates in the BoE of randomised controlled trials. The pooled difference of mean differences was −1.95 mm Hg (95\% confidence interval −3.84 to −0.05; \( \tau^2=59\% \); \( \tau^2=1.64 \); 95\% prediction interval −22.33 to 18.43) for systolic blood pressure estimates and −2.36 mm Hg (−3.16 to −1.57; \( \tau^2=0\% \); \( \tau^2=0 \); −3.16 to −1.57) for diastolic blood pressure estimates (fig 2).

Sensitivity analyses excluding highly correlated outcomes
The first sensitivity analysis, where only one outcome (with the largest number of randomised controlled trials) was chosen from each Cochrane review (n=31), confirmed the findings of the primary analysis (ratio of risk ratios 1.14 (95\% confidence interval 1.06 to 1.22); \( \tau^2=72\% \); \( \tau^2=0.027 \); 95\% prediction interval 0.81 to 1.61; supplementary fig 5). In the second sensitivity analysis, 50 diet-disease outcome pairs were included in the meta-analysis and showed also similar results (1.12 (1.06 to 1.18); \( \tau^2=68\% \); \( \tau^2=0.023 \); 0.82 to 1.52; supplementary fig 6).

Subgroup analyses
Subgroup analyses showed that estimates were marginally different in BoE of randomised controlled trials compared to BoE of cohort studies for PI/ECO matched outcomes pairs that were similar but not identical (ratio of risk ratios 1.05 (95\% confidence interval 1.00 to 1.10); \( \tau^2=61\% \); \( \tau^2=0.016 \); 95\% prediction interval 0.81 to 1.36) and substantially in disagreement for those pairs that were broadly similar (1.20 (1.10 to 1.30); \( \tau^2=62\% \); \( \tau^2=0.020 \); 0.88 to 1.63; fig 3 and fig 4). Regarding specific PI/ECO components, the dissimilarity in intervention or exposure explained most of the differences. The broadly similar category showed substantial disagreement (1.29 (1.18 to 1.41); \( \tau^2=52\% \); \( \tau^2=0.015 \); 0.97 to 1.71), whereas the more or less identical category led to estimates highly in agreement (0.98 (0.91 to 1.04); \( \tau^2=7\% \); \( \tau^2=0.00 \); 0.88 to 1.09; supplementary fig 7).

Subgroup analyses by type of dietary intervention or exposure showed different results between the two BoEs for micronutrient comparisons (mainly dietary supplements; ratio of risk ratios 1.14 (95\% confidence interval 1.06 to 1.22); \( \tau^2=69\% \); \( \tau^2=0.031 \); 95\% prediction interval 0.79 to 1.63), whereas no differences for all other types of intervention were observed (supplementary fig 8). After observing substantial heterogeneity for several types of comparison for intervention or exposure (dietary approaches, \( \tau^2=0.01 \); \( \tau^2=61\% \); micronutrients, \( \tau^2=0.03 \); \( \tau^2=69\% \)), we further explored it, by considering the type of intake or exposure (supplementary fig 9). We noticed that when the type of intake of the interventions and exposures was the same in both BoE, the estimates were similar (for dietary intake, ratio of risk ratios 0.98 (95\% confidence interval 0.93 to 1.04); \( \tau^2=4\% \); \( \tau^2=0.00 \); 95\% prediction interval 0.90 to 1.07; for dietary supplements, 1.08 (0.98 to 1.20; \( \tau^2=0\% \); \( \tau^2=0.00 \); 0.95 to 1.23); in both cases, no heterogeneity was observed). The comparison of dietary intake and...
Fig 1 | Forest plot of comparisons between bodies of evidence from randomised controlled trials versus those from cohort studies for binary outcomes as pooled ratio of risk ratios. CS=cohort study; RCT=randomised controlled trial; RR=risk ratio; omega 3=omega 3 fatty acid; omega 6=omega 6 fatty acid; Abdelhamid 2018a=reference 37; Abdelhamid 2018b=reference 38; Bjelakovic 2014a=reference 43; Bjelakovic 2014b=reference 42; Chowdhury 2014a=reference 67; Chowdhury 2014b=reference 75; Hooper 2015b=reference 51; Rees 2013b=reference 58; Zhang 2016a=reference 93.
Table 1 | Overview of main results for binary outcomes (n=71 diet-disease outcome pairs)

| Ratio of risk ratios (95% CI) | Heterogeneity (I² (%); τ²) | 95% prediction interval |
|-----------------------------|-----------------------------|------------------------|
| Main analysis               | 1.09 (1.04 to 1.14)         | 68; 0.021 (0.81 to 1.46) |
| Stratified by overall PI/ECO similarity degree | | |
| More or less identical | —                           | —                      | |
| Similar but not identical | 1.05 (1.00 to 1.10)         | 61; 0.016 (0.81 to 1.36) |
| Broadly similar             | 1.20 (1.10 to 1.30)         | 62; 0.020 (0.88 to 1.63) |
| Stratified by type of dietary intervention/exposure | | |
| Fatty acids                 | 1.05 (1.00 to 1.10)         | 62; 0.020 (0.88 to 1.63) |
| Micronutrients              | 1.14 (1.06 to 1.22)         | 69; 0.031 (0.79 to 1.63) |
| Dietary approach            | 0.99 (0.90 to 1.09)         | 61; 0.010 (0.77 to 1.27) |
| Stratified by type of intake/exposure (randomised controlled trials v cohort studies) | | |
| Intake v intake             | 0.98 (0.93 to 1.04)         | 6; 0.00 (0.90 to 1.07)   |
| Supplements v supplements   | 1.08 (0.98 to 1.20)         | 0                      |
| Intake and supplements v intake | 1.06 (0.99 to 1.14)     | 62; 0.007 (0.86 to 1.30) |
| Supplements v vintake       | 1.07 (0.93 to 1.21)         | 74; 0.049 (0.65 to 1.73) |
| Supplements v status        | 1.29 (1.17 to 1.42)         | 54; 0.018 (0.94 to 1.77) |
| Stratified by type of outcomes | | |
| Cardiovascular disease      | 1.05 (0.99 to 1.12)         | 57; 0.010 (0.85 to 1.31) |
| All cause mortality         | 1.17 (1.11 to 1.23)         | 75; 0.006 (0.99 to 1.39) |
| Cancer                      | 1.07 (0.98 to 1.16)         | 20; 0.007 (0.86 to 1.31) |
| Pregnancy outcomes          | 0.93 (0.75 to 1.15)         | 70; 0.092 (0.46 to 1.88) |

I²=inconsistency; PI/ECO=population, intervention or exposure, comparator, outcome.

Reference pair | Intervention in RCTs | Exposure in CSs | Outcome | Difference of mean differences (95% CI) | Difference of mean differences (95% CI) |
|---------------|---------------------|----------------|--------|----------------------------------------|----------------------------------------|
| Systolic blood pressure | | | | | |
| Adler 2014+Leyvraz 2018 | Low sodium | Low sodium | Systolic blood pressure | -0.59 (-2.06 to 0.88) | -3.41 (-5.50 to -1.32) |
| Rees 2013a+Kastorini 2011 | Healthy diet | Mediterranean diet | Systolic blood pressure | -2.30 (-5.22 to -0.62) | -1.95 (-3.84 to -0.06) |
| Rees 2019+Kastorini 2011 | Mediterranean diet | Mediterranean diet | Systolic blood pressure | -2.37 (-3.31 to -1.43) | -2.35 (-3.84 to -0.86) |
| Random effects model | | | | | |
| Diastolic blood pressure | | | | | |
| Adler 2014+Leyvraz 2018 | Low sodium | Low sodium | Diastolic blood pressure | -2.36 (-3.16 to -1.57) | -3.16 (-3.16 to -1.57) |
| Rees 2013a+Kastorini 2011 | Healthy diet | Mediterranean diet | Diastolic blood pressure | -2.36 (-3.16 to -1.57) | -3.16 (-3.16 to -1.57) |
| Random effects model | | | | | |
| Body weight | | | | | |
| Kelly 2017+Ye 2012 | Whole grains | Whole grains | Body weight | -0.11 (-0.75 to 0.53) | -0.11 (-0.75 to 0.53) |
| Random effects model | | | | | |
| Birth length | | | | | |
| Palacios 2019+Tous 2020 | Vitamin D | Vitamin D | Birth length | 0.08 (-0.23 to 0.39) | 0.08 (-0.23 to 0.39) |
| Random effects model | | | | | |
| Head circumference at birth | | | | | |
| Palacios 2019+Tous 2020 | Vitamin D | Vitamin D | Head circumference at birth | -51.59 (-104.25 to 1.07) | 5.20 (51.85 to 62.25) |
| Tieu 2017+Chia 2019 | Healthy diet | Vitamin D | Birth weight | -24.30 (-79.91 to 31.32) | -103.51 to 54.92 |
| Random effects model | | | | | |
| High density lipoprotein | | | | | |
| Rees 2019+Kastorini 2011 | Mediterranean diet | Mediterranean diet | High density lipoprotein | 0.01 (-0.05 to 0.07) | 0.01 (-0.05 to 0.07) |
| Random effects model | | | | | |
| Triglycerides | | | | | |
| Rees 2019+Kastorini 2011 | Mediterranean diet | Mediterranean diet | Triglycerides | -0.39 (-1.05 to 0.27) | -0.39 (-1.05 to 0.27) |
| Random effects model | | | | | |

Fig 2 | Forest plot of comparisons between bodies of evidence from randomised controlled trials versus those from cohort studies for continuous outcomes as pooled difference of mean differences. CS=cohort study; RCT=randomised controlled trial; MD=mean difference; Rees 2013a=reference 57.
dietary supplements in randomised controlled trials versus dietary intake in cohort studies also showed similar estimates (ratio of risk ratios 1.06 (95% confidence interval 0.99 to 1.14); I^2=62%; t^2=0.007; 95% prediction interval 0.86 to 1.30). Heterogeneity was present when considering low fat dietary approaches. Focusing on dietary fatty acids only (n-3, n-6, and polyunsaturated fatty acids), we observed no heterogeneity (supplementary fig 10). Moreover, the comparisons between dietary supplements in randomised controlled trials versus dietary intake in cohort studies showed similar estimates but substantial heterogeneity and a wide prediction interval (ratio of risk ratios 1.07 (95% confidence interval 0.95 to 1.21); I^2=74%; t^2=0.049; 95% prediction interval 0.65 to 1.75). By excluding pregnancy outcomes (because all other comparisons focused on non-communicable diseases), and β

| Reference pair | Intervention in RCTs | Exposure in CSs | Outcome | Ratio of risk ratios (95% CI) | Ratio of risk ratios (95% CI) |
|----------------|----------------------|-----------------|---------|-------------------------------|-------------------------------|
| Abdelhamid 2018a+Chowdhury 2014a | Omega 3 | Omega 3 | Cardiovascular disease | 1.14 (1.01 to 1.28) |  |
carotene and vitamin A comparisons (for the outcome of mortality), which are known to increase mortality at higher doses in randomised controlled trials,\textsuperscript{105} heterogeneity disappeared (supplementary fig 11). The comparisons of dietary supplements versus nutrient status was judged to have the lowest similarity degree for intervention or exposure and also showed substantial differences between randomised controlled trials and cohort studies (1.29 (1.17 to 1.42); $I^2=56\%$; $\tau^2=0.018$; 0.94 to 1.77). Heterogeneity was driven by vitamin D comparisons (supplementary fig 9). After stratifying the analysis by outcome type, we observed differences for overall mortality (1.17 (1.11 to 1.23); $I^2=75\%$; $\tau^2=0.006$; 0.99 to 1.39), bone health (1.46 (1.16 to 1.84); $I^2=67\%$; $\tau^2=0.019$; 1.02 to 2.08), and eye disease (1.14 (1.03 to 1.26); $I^2=36\%$; $\tau^2=0.003$; 0.44 to 2.96; supplementary fig 12).

The findings of the subgroup analyses are supported by sensitivity analyses excluding outcomes that are likely to be highly correlated (supplementary figs 13-17), and when BoE from cohort studies with a risk ratio <1 were analysed separately (supplementary figs 18-22). The subgroup analyses for BoE from cohort studies with a RR >1, need to be interpreted with caution due to the very small number of comparisons (n=7) (supplementary figs 23-27).

Additional analyses

We also performed a multi-level meta-analysis, considering the pairs as grouping factor, and the findings of the primary analysis were confirmed (ratio of risk ratios 1.08 (95% confidence interval 1.02 to 1.13); $I^2=68\%$). The sensitivity analysis comparing BoEs from randomised controlled trials versus cohort studies of the two Cochrane reviews (based on six outcomes) also confirmed the findings of the primary analysis (supplementary fig 28).

Discussion

Summary of findings

This large meta-epidemiological study identified and compared empirical data to determine the extent to which diet-disease association estimates of BoE from randomised controlled trials and cohort studies are in agreement. Overall, 97 diet-disease outcome pairs were identified and 83 were suitable for meta-analysis. No outcome pair was rated as more or less identical, according to PI/ECO similarity. On average, the difference in the pooled results between the two BoEs was small, but given that prediction intervals are wide and statistical heterogeneity was in part substantial in cohort studies, important differences or potential bias in individual comparisons or individual studies cannot be excluded.

We investigated possible factors for the observed heterogeneity, finding that PI/ECO dissimilarities, in particular the comparisons of dietary supplements in randomised controlled trials and nutrient status in cohort studies, explained most of the differences. When the type of intake or exposure between both BoE was identical, the estimates were similar (and the analysis showed low statistical heterogeneity). For pooled
estimate of continuous outcomes, no differences were observed between randomised controlled trials and cohort studies, except for smaller systolic and diastolic blood pressure estimates in the BoE of trials.

Comparison with other studies

**Nutrition field**

A technical review published in 2013 identified 34 diet-disease outcome pairs of systematic reviews of randomised controlled trials and large, single randomised controlled trials (>1000 participants) versus systematic reviews of case-control or cohort studies and one large observational study (>5000 participants). Similar to our findings, 22 (65%) of 34 diet-disease outcome pairs were in the same direction, and had no evidence of significant disagreement (z score not statistically significant). By comparison, our study included a larger sample of outcome pairs and a larger number of participants. We also thoroughly matched PI/ECO criteria, pooled the effect estimate to generate a ratio of risk ratios and difference of mean differences, and also investigated the possible factors of disagreement.

Trepakowski and Ioannidis recently argued that many prominent epidemiological associations (including highly cited studies on α-tocopherol, β-carotene, vitamin C, vitamin D, selenium, calcium, and low fat diets) have not been corroborated by large randomised controlled trials or meta-analyses. Their statement, however, is not based on a systematic evaluation, and does not accord with our findings, where pooling BoEs of randomised controlled trials and cohort studies showed on average minor differences. On the contrary, Satija and colleagues argued that, when randomised controlled trials are able to successfully examine diet-disease relations, their results are more often in line with those of cohort studies. Our findings seem to accord with Satija and colleagues’ conclusions, although the pooled estimate showed some differences between both BoEs, and the prediction intervals were wide.

**Medical field**

Anglemyer et al conducted a methodological Cochrane review, including systematic reviews and overviews of reviews in different medical fields, which showed little difference between the results obtained from randomised controlled trials and observational studies (cohort and case-control studies). Their result when comparing BoE from randomised controlled trials with BoE from observational studies (ratio of odds ratios 1.08 (95% confidence interval 0.96 to 1.22)) is similar to our findings (ratio of risk ratios 1.09 (1.04 to 1.14)). The difference in the estimates, in terms of point estimate, was more in disagreement with pharmacological studies (ratio of odds ratios 1.17 (0.95 to 1.43)). This difference corresponds to our findings regarding micronutrient interventions (mainly as dietary supplements), where randomised controlled trials showed differences compared with cohort studies (ratio of risk ratios 1.14 (1.06 to 1.22)).

However, the methodological review by Anglemyer et al did not conduct PI/ECO matching, did not calculate 95% prediction intervals, and did not differentiate various types of intervention and outcomes.

When comparing our results with findings from meta-epidemiological studies investigating the impact of design features of randomised controlled trials, the magnitude of differences were similar. For example, lack of reporting of adequate random sequence generation, allocation sequence concealment, and double blinding tend to overestimate intervention effects (ratio of odds ratios ranging from 0.87 to 0.93). The extent of overestimation was lower for objective outcomes (eg, mortality) and therefore unlikely to be influenced by knowledge of the intervention received. A recent meta-epidemiological study of 142 meta-analyses found no evidence for difference in treatment effect between randomised controlled trials with and without patients, healthcare providers, or outcome assessors blinded to treatment. The impact of design features of randomised controlled trials and cohort studies has not yet been explored in the field of nutrition using meta-epidemiological methods.

Potential implications

What constitutes best evidence in nutrition research has been debated extensively, and whether it comes from randomised controlled trials, which are considered the ideal methodology for causal inference and in which the effects of a dietary change on disease or intermediate disease markers are evaluated experimentally. However, most randomised controlled trials of dietary interventions are short and do often not target patient relevant outcomes such as morbidity or mortality. Further limitations are the difficulty of inducing and maintaining dietary changes in the long term, and the low adherence to a specific dietary regimen that often occurs. Moreover, although several long term trials have been conducted (eg, the Women’s Health Initiative Dietary Modification trial, or the PREDiMED study), the costs of such large scale dietary trials are challenging. Cohort studies, on the other hand, provide methodologically less robust information regarding causality, but are usually considered more applicable for nutrition research.

In general, the two BoE (trials vs cohort studies) often differ in terms of study populations (inclusion and exclusion criteria, comparison group), different exposure levels (dose, duration, sources), different outcomes, and different sample sizes and follow-up durations, as shown in our study. For example, in randomised controlled trials of supplements such as selenium, participants might already have an adequate selenium status. Observational studies, on the other hand, usually include participants with a broader range of selenium status. Therefore, the comparison of risk ratios from trials and cohort studies might not be a perfect match.

Randomised controlled trials are experimental studies, where participants are usually given fixed intake levels in the form of dietary supplements or...
where investigators try to modify dietary intake via dietary advice or dietary counselling.\textsuperscript{112} Cohort studies are observational and do not actively intervene in the behaviour of study participants, and participant groups are based on reported intake (or status) of dietary advice (eg, higher \textit{v} lower sodium intake), thus implying a variable difference in intake levels.\textsuperscript{112} In each case, the estimated risk ratio reflects the direction of the effect and an indication of the strength of the association. In our study, BoE of randomised controlled trials were predominantly based on dietary supplements (n=43) and dietary intake (n=38), whereas BoE of cohort studies investigated mainly dietary intake (n=69). In randomised controlled trials of dietary supplements, which are similar to drug trials, study participants are randomly assigned into study arms, thus balancing measured and unmeasured potential confounders across the comparison groups, allowing differences in the outcome measure to be attributed to the dietary intervention.\textsuperscript{20} By contrast, cohort studies are prone to residual confounding, and the direction and magnitude of risk ratio is influenced by the variables included in the statistical models built to estimate the effect, and by the potential measurement error of dietary factors (which is also a problem in long term randomised controlled trials on dietary intake) and all other factors.\textsuperscript{112}

Despite these circumstances, our study matched PI/ECO similarities between the two most important study designs in nutrition research. Because of the above described differences, no diet-disease outcome pair was rated as more or less identical. However, when the type of intake or exposure between both BoEs was identical, the estimates were similar (and the analysis showed low statistical heterogeneity). Such PI/ECO dissimilarities are often present even between studies with the same design, which contributes to statistical heterogeneity in the primary meta-analyses.\textsuperscript{20} A meta-epidemiological study of meta-analyses such as ours could further increase the complexity of heterogeneity, but the exploration of statistical heterogeneity among comparisons between different types of dietary intake or exposure yielded plausible explanations in our study.

At the systematic review level, the established approach to evaluate the credibility of results from primary studies is risk-of-bias assessment. In our study, all Cochrane reviews used the Cochrane risk-of-bias tool, whereas for cohort studies the Newcastle-Ottawa scale was mainly used, as reported elsewhere.\textsuperscript{106} Risk-of-bias assessment is an integral part of the GRADE approach, which rates the certainty of evidence based on a BoE.\textsuperscript{119,120} According to the GRADE approach, the certainty of evidence is initially determined by study design: a BoE from randomised controlled trials starts with high certainty, whereas a BoE from observational studies starts with low certainty owing to confounding and selection bias (if the ROBINS tool is used, both BoEs start as high certainty).\textsuperscript{119,120} Use of the GRADE approach, especially in relation to the risk-of-bias assessment, is challenging and could lead to excessive downgrading. For example, GRADE users might inappropriately double count the risk of confounding and selection bias by downgrading the initial certainty of the BoE to low, followed by further downgrading due to unknown confounders.\textsuperscript{119,121} When GRADE was used, very low and high certainty of evidence ratings accounted for 10\% and 27\% of ratings for BoE of randomised controlled trials, respectively, compared to 80\% and 0\% for BoE of cohort studies, respectively. In this regard, we could show in a recent methodological survey that very low and high certainty of evidence ratings accounted for 61\% and 1\% of ratings in systematic reviews of observational studies, respectively, compared to 16\% and 5\% in systematic reviews of randomised controlled trials, respectively.\textsuperscript{122}

A recent cross sectional study has shown that very few Cochrane nutrition reviews include observational studies (2\%),\textsuperscript{123} which has been criticised.\textsuperscript{124} BoE from cohort studies can strengthen or complement BoE from randomised controlled trials, and vice versa, so our meta-epidemiological study provides support for integration based on thorough assessment of PI/ECO similarities, and could be a starting point for future work.

**Strengths and limitations**

Our study had several strengths. Firstly, we included a large sample of diet-disease outcome pairs (n=97), based on more than 950 randomised controlled trials and 750 cohort studies, with both study designs considered as the most reliable in nutrition research.\textsuperscript{6} Secondly, the conducted PI/ECO matching process was novel, and provided important insights in our understanding of which factors are associated with disagreement. Thirdly, the data extraction was extensive, retrieving information on the description of interventions or exposures and comparators, population, study design, risk of bias of the primary studies, and certainty of the evidence for each diet-disease pair. Fourthly, we conducted various statistical analyses, such as recalculating 34 pooled estimates, converting odds ratios and hazard ratios into risk ratios, including binary and continuous outcomes, and pooling the estimates across all diet-disease outcome pairs. Finally, the exploration of factors potentially associated with disagreement through a priori planned subgroup analyses for PI/ECO dissimilarities, types of intervention, intake, exposure, and outcomes was an additional strength of this study.

The present study also had several limitations. Firstly, we only searched the Cochrane Database of Systematic Reviews to identify systematic reviews of randomised controlled trials, after which we screened all eligible Cochrane reviews to see if they also included cohort studies. It would have been ideal if Cochrane reviews also included cohort studies to ensure better comparability in terms of systematic review methodological approaches (eg, search strategy, risk-of-bias assessment, or GRADE rating). However, only two Cochrane reviews contained also
cohort studies. Therefore, we searched and matched systematic reviews of cohort studies retrieved from Medline, which might have affected the validity of our findings. However, a sensitivity analysis comparing BoE from randomised controlled trials versus cohort studies of the two Cochrane reviews (based on six outcomes) confirmed the findings of the primary analysis. Secondly, the meta-analyses in the present research could have themselves had limitations, from the primary data and how the evidence has been summarised; therefore, readers should consider the original studies for more detailed information.

Thirdly, although the PI/ECO matching process was conducted by two reviewers, subjectivity cannot be ruled out completely. A quantitative PI/ECO matching approach might have been more objective but has yet to be developed. Another limitation, particularly for the BoE from cohort studies, is that some studies were included multiple times, and from the systematic reviews, the same original studies were used with the same exposure but for different outcomes. The sensitivity analysis where only one outcome (with the largest number of randomised controlled trials) was chosen from each Cochrane review (n=31) confirmed the findings of the primary analysis. Moreover, sensitivity analyses excluding outcomes that were likely to be highly correlated showed similar findings as the primary analysis (ratio of risk ratio 1.12 (95% confidence interval 1.06 to 1.18)). We also did a multi-level meta-analysis considering the pairs as grouping factor, which confirmed the findings of the primary analysis. A further limitation was that we did not explore other potential factors for disagreement, such as dietary adherence in the included primary randomised controlled trials.

Finally, the impact of potential bias in cohort studies should be considered at three levels: generally causing a systematic bias; causing bias in individual comparisons; and causing bias in individual studies, hence leading to heterogeneity in individual meta-analyses. Considering these three levels in order, our findings implied no strong indication for systematic bias overall (meta-analytical ratio of risk ratios close to 1), but this did not exclude an important risk for bias in individual comparisons, or in individual studies. The ratio of risk ratios in individual comparisons was often not equal to 1; in fact, we observed a different direction (eg, ratio of risk ratios <1 in 24 of 71 comparisons) and magnitude (eg, four comparisons showed a ratio of risk ratios <0.75, and 13 comparisons showed a ratio of risk ratios >1.25). Furthermore, the prediction intervals indicated that the difference between the two BoEs as shown in our study (ratio of risk ratios 1.09) could be much more substantial in either direction (95% prediction interval 0.81 to 1.46). By exploring factors associated with those differences, we found that PI/ECO dissimilarity (mainly in type of intake) was the main driver. Finally, statistical heterogeneity was higher in the individual meta-analyses of cohort studies (mean I²=47%; τ²=0.023) than in those of randomised controlled trials (mean I²=21%; τ²=0.018), possibly due potential bias in individual cohort studies.

Conclusion

On average, the difference in the pooled results between the two BoEs was small, but with wide prediction intervals and some substantial statistical heterogeneity in cohort studies, important differences or potential bias in individual comparisons or individual studies cannot be excluded. We investigated possible factors for the observed heterogeneity, finding that PI/ECO dissimilarities, especially for the comparisons of dietary supplements in randomised controlled trials and nutrient status in cohort studies. When the type of intake or exposure between both BoEs was identical, the estimates were similar (and the analysis showed low statistical heterogeneity). Nevertheless, the comparison between randomised controlled trials and cohort studies should be interpreted very carefully.

These findings provide valuable insights towards better understanding the integration of both BoEs of randomised controlled trials and cohort studies in prospective nutrition evidence syntheses. Considering that few Cochrane nutrition reviews include cohort studies, and that most of the evidence in nutrition actually comes from cohort studies, evidence based guidance is urgently needed on how to incorporate and possibly integrate both BoEs in nutrition evidence syntheses. We consider that this information is needed not only for research purposes but also for decision making processes and tailoring better evidence based dietary guidelines.

Contributors: LS, SB, JB, NB, SSW, JZ, and JIM designed the research. LS and SB analysed the data and wrote the first draft of the paper. LS, SB, JB, NB, SSW, JZ, BN, and JIM interpreted the data, read the manuscript, and approved the final version. LS and JIM are guarantors. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Funding: Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Projektnummer 459430615 and Forschungskommission der Medizinischen Fakultät Freiburg. The funders had no role in considering the study design or in the collection, analysis, interpretation of data, writing of the report, or decision to submit the article for publication.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/disclosure-of-interest/ and declare: support from the Deutsche Forschungsgemeinschaft and Forschungskommission der Medizinischen Fakultät Freiburg for the submitted work; LS is a member of the GRADE working group; JIM is a member of the GRADE working group and director of the Freiburg GRADE centre; no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: Not required.

Data sharing: Data were extracted from published meta-analyses, all of which are available and accessible. The lead authors (the manuscript’s guarantors) affirm 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 have been explained.

Dissemination to participants and related patient and public communities: Results from our meta-epidemiological study will be disseminated through a press release (Cochrane Germany), the website (eg, Cochrane Germany, Cochrane Nutrition), blog (eg, wissenschaftskritik.org), and social media (Twitter). Results dissemination will also be targeted to health professionals, including dietitians, nutritionists, physicians, guidelines developers, and scientists.

Provenance and peer review: Not commissioned; externally peer reviewed.
