Technical Review 17

Nutritional Research Series

Volume 6: Concordance Between the Findings of Epidemiological Studies and Randomized Trials in Nutrition: An Empirical Evaluation and Citation Analysis

Prepared for:
Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
540 Gaither Road
Rockville, MD 20850
www.ahrq.gov

Contract No. 290-2007-10055-I

Prepared by:
Tufts Medical Center Evidence-based Practice Center
Boston, MA

Investigators:
Denish Moorthy, M.B.B.S., M.S.
Mei Chung, Ph.D.
Jounghee Lee, Ph.D.
Winifred W. Yu, Ph.D.
Joseph Lau, M.D.
Thomas A. Trikalinos, M.D.

AHRQ Publication No. 13-EHC067-EF
May 2013
Preface

The Agency for Healthcare Research and Quality (AHRQ), through its Evidence-based Practice Centers (EPCs), sponsors the development of evidence reports and technology assessments to assist public- and private-sector organizations in their efforts to improve the quality of health care in the United States. The Office of Dietary Supplements of the National Institutes of Health requested and provided funding for this report.

The reports and assessments provide organizations with comprehensive, science-based information on common, costly medical conditions and new health care technologies and strategies. The EPCs systematically review the relevant scientific literature on topics assigned to them by AHRQ and conduct additional analyses when appropriate prior to developing their reports and assessments.

To improve the scientific rigor of these evidence reports, AHRQ supports empiric research by the EPCs to help understand or improve complex methodologic issues in systematic reviews. These methods research projects are intended to contribute to the research base in and be used to improve the science of systematic reviews. They are not intended to be guidance to the EPC program, although may be considered by EPCs along with other scientific research when determining EPC program methods guidance.

AHRQ expects that the EPC evidence reports and technology assessments will inform individual health plans, providers, and purchasers as well as the health care system as a whole by providing important information to help improve health care quality. The reports undergo peer review prior to their release as a final report.

We welcome comments on this Technical Review. They may be sent by mail to the Task Order Officer named below at: Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850, or by email to epc@ahrq.hhs.gov.

Carolyn M. Clancy, M.D.  
Director  
Agency for Healthcare Research and Quality

Jean Slutsky, P.A., M.S.P.H.  
Director, Center for Outcomes and Evidence  
Agency for Healthcare Research and Quality

Stephanie Chang, M.D., M.P.H.  
Director and Task Order Officer  
Evidence-based Practice Program  
Center for Outcomes and Evidence  
Agency for Healthcare Research and Quality
Nutritional Systematic Reviews

The medical and clinical communities have effectively used systematic reviews to develop clinical and public health practice guidelines, set research agendas, and develop scientific consensus statements. However, the use of systematic reviews in nutrition applications is more recent and limited. The Office of Dietary Supplements (ODS) at the National Institutes of Health (NIH) has been proactive and developed an evidence-based review program using the EPC Program established by AHRQ, as part of a Congressional mandate to review the current scientific evidence on the efficacy and safety of dietary supplements and identify research needs (http://ods.od.nih.gov/Research/Evidence-Based_Review_Program.aspx). To date, this program has sponsored 17 evidence reports on a range of supplement-related topics including B-vitamins, ephedra, multivitamin/mineral supplements, omega-3 fatty acids, soy, and vitamin D. ODS is currently sponsoring an augmentation of the vitamin D report published in August 2007 to provide relevant information for a pending Institute of Medicine review of the current Dietary Reference Intakes for vitamin D and calcium. The completed ODS-sponsored evidence reports have resulted in numerous associated publications in scientific journals, have formed the basis for an NIH-sponsored state-of-the-science conference, and have been used to assist in setting research agendas.

To facilitate a better understanding of the challenges involved in conducting nutrition-related systematic reviews and in integrating these reviews with nutrition applications for which such reviews have not been previously used, ODS has sponsored the development of a series of technical reports via the EPC Program. The purpose of these reports was to: (a) identify the challenges, advantages, and limitations of conducting nutrition-based systematic reviews, (b) work with a panel of experts to explore approaches for integrating systematic reviews into processes associated with the derivation of nutrient intake reference values, (c) identify the breadth and quality of currently available nutrition-related systematic reviews against generally accepted quality guidelines within the context of the unique needs for nutrition topics, and (d) critically explore the consistencies and inconsistencies in results between observational and intervention studies and evaluate how the formulation of research questions may have contributed to these discrepancies.

Paul M. Coates, Ph.D.
Director, Office of Dietary Supplements
National Institutes of Health

Elizabeth A. Yetley, Ph.D.
Senior Nutrition Research Scientist (retired), Scientific Consultant
Office of Dietary Supplements
National Institutes of Health
Acknowledgments

Funding. This manuscript is based on work performed with funding from the Office for Dietary Supplements commissioned through the Agency for Healthcare Research and Quality (Contract No. 290-2007-10055-I). None other than the authors had any role in study design, collection, analysis, or interpretation of data, in the writing of the report, or the decision to submit the paper for publication. The authors are independent from the funders.

Competing interests. We declare that none of the authors has a conflict of interest in this submission. All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare that no author has support from companies for the submitted work. No other relationships or activities that could appear to have influenced the submitted work.
Concordance Between the Findings of Epidemiological Studies and Randomized Trials in Nutrition: An Empirical Evaluation and Citation Analysis

Structured Abstract

**Background.** In nutrition, there are several examples of discordance between the results of observational studies and of randomized controlled trials (RCTs).

**Objectives.** To provide empirical data on how often the summary results of epidemiological studies and of RCTs are concordant, and to explore whether the probability of concordant findings is associated with quantifiable metrics of citation maps formed between studies belonging to the evidence base of the nutrient-outcome association at hand. Citation maps are an objective representation of the translational paths in each association, and may be a surrogate of the maturity of the relevant evidence base.

**Methods.** We searched MEDLINE to identify meta-analyses of RCTs or of epidemiological studies on the association between nutrients and health outcomes. Summary findings from both research designs that were statistically significant and in the same direction were considered qualitatively concordant. We also calculated the statistical significance of the difference in the summary effects from epidemiological studies and from RCTs (a measure of quantitative concordance). For each nutrient-outcome association we defined an “evidence base” including all publications identifiable by MEDLINE searches on the nutrient and outcome of interest and constructed citation maps of all articles in the evidence base that were cited by the epidemiological studies or the RCTs in the meta-analyses, either directly, or through one or more intermediary papers. We then quantified the size of the graphs (number of vertices and citation relationships), and their connectivity (density of citation relationships, mean hub and authority scores, and mean number of citations made or received over the included papers). We tested for associations between these metrics and the probability that the summary results from epidemiological studies and from RCTs are concordant between them.

**Findings.** In 23 out of 34 associations the summary findings from meta-analyses of epidemiological studies and of RCTs were in the same direction. In 6 of 23 associations, meta-analyses of epidemiological studies and of RCTs had statistically significant findings. In the remaining 11 out of 34 associations, meta-analyses of epidemiological studies and of RCTs had summaries pointing in opposite directions. In 12 out of 34 associations the summary findings of epidemiological studies were statistically significantly different from those of RCTs, in 6 out of 12 point estimates were in the same direction, and in the other 6 in opposing directions. Despite the variation in the size and the connectivity of the citation graphs across the examined associations, we find no evidence of association between quantitative metrics of the citation graphs and the probability that the two research designs have concordant or discordant findings (using various definitions of concordance or discordance).
Conclusions. The examined quantitative characteristics of the citation maps in each association cannot predict the probability that the findings from the two designs agree or disagree. It is unclear whether there is a good way to describe the maturity of the evidence base on an association between nutrients and outcomes. At a minimum, purely bibliometric approaches are not a good way to prioritize which nutrient exposures merit further study, and for which health outcomes.
Contents

Background ..................................................................................................................................... 1
Methods .......................................................................................................................................... 3
  Overview .................................................................................................................................. 3
  Database of Associations Between Nutrients and Health Outcomes ........................................ 3
    Literature Searches To Identify Meta-Analyses on Nutrients ................................................... 3
    Matching Meta-Analyses of Epidemiological Studies With Meta-Analyses of RCTs ............... 4
    Data Extraction ...................................................................................................................... 4
    Assessment of Concordance or Discordance Between Epidemiological Studies and RCTs .... 4
Citation Analyses .......................................................................................................................... 5
  Identifying an Evidence Base for Each Association .................................................................... 5
  Citation Networks and Graphs ................................................................................................... 5
Size and Connectivity of Citation Networks ............................................................................... 6
Associations Between Citation Graph Metrics and Qualitative and Quantitative Concordance .... 7
Results ............................................................................................................................................. 8
  Database of Nutrient-Outcome Associations .............................................................................. 8
  Concordance or Discordance Between Epidemiological and RCT Data ................................... 12
  Citation Graphs .......................................................................................................................... 14
Associations Between Citation Graph Metrics and Qualitative and Quantitative Concordance .... 16
Discussion ..................................................................................................................................... 18
References ..................................................................................................................................... 20

Tables
Table 1. Eligible nutrient-outcome associations ............................................................................. 9
Table 2. Qualitative and quantitative concordance of effects in epidemiological studies and RCTs ................................................................................................................................. 13
Table 3. Summary of citation analysis of nutrient exposure-health outcome pairs studied by systematic reviews or large studies ............................................................................................................. 15
Table 4. Associations between concordance and quantitative characteristics of the citation graphs ......................................................................................................................................................... 17

Figures
Figure 1. Illustration of a hypothetical citation network of six articles ........................................ 6

Appendices
Appendix A. List of Observational Cohorts and the Screened Publications
Appendix B. Details on the Construction of Citation Graphs
Background

Many randomized controlled trials (RCTs) comparing nutrient interventions with placebo/control have failed to replicate the (usually protective) associations between nutrients and risk of chronic disease found in large-scale observational data.\(^1\)\(^-\)\(^5\) High-profile examples of such RCTs found no evidence of intervention effects for fiber and colon cancer,\(^6\) vitamin E and cardiovascular disease,\(^7\) vitamin E and lung cancer,\(^8\) β-carotene and coronary events,\(^9\) vitamin C and cardiovascular disease,\(^10\) or folate and cardiovascular disease\(^11\), and even identified adverse effects of nutrient supplementation (e.g., β-carotene and cancer\(^8\)). The discrepancies between the findings of RCTs and epidemiological studies raised serious questions about the currently used approach for determining whether the evidence base is adequate to justify launching a large-scale RCT, with hard endpoints as the outcome measure.

Deciding which specific nutrient-outcome association to further evaluate in human intervention trials is challenging. Simply having a plausible epidemiological support is probably not enough; all aforementioned negative large RCTs were motivated by hypotheses vetted in epidemiological data or in smaller RCTs examining surrogate outcomes. In previous work, we hypothesized that additional critical components pertain to the maturity and reliability of the relevant evidence base — that is, the strength of the data supporting a potential nutrient-disease association, the biological plausibility of the association, the reliability of existing data, and the likelihood of bias and systematic errors affecting interpretation of the available data. The evidence base is formed by the interplay of various translational paths, in which an initial hypothesis-forming observation supports subsequent research and is eventually “translated” to interventions for preventing or treating human disease.

Understanding the translational paths that shape the evidence base may result in insights on why epidemiological and randomized data disagree or agree. To some extent, the translational paths can be assessed with citation analysis, which is a qualitative and quantitative representation of citations between publications. In previous work, we conjectured that nutrient associations where RCT and observational data is concordant have a more extensive and mature evidence base compared with associations where the data is discordant.\(^12\) Otherwise put, differences in the observed flow of information (as captured by citations that are received or made between publications) through the various translational paths and the content of the propagated information are associated with concordance or discordance in the results of observational studies and RCT. For example, a limited evidence base and information flow may indicate inconsistency of study results and thus may be associated with topics where RCT and observational studies disagree. Reciprocally, a large evidence base with higher information flow may indicate consistency of findings and general agreement between RCT and observational studies. Of course, these are not one-to-one relationships; it is conceivable that profound inconsistencies and disagreements between studies could lead to considerable discourse among investigators, which in turn would increase information flow.

In our previous paper, we empirically explored this hypothesis by analyzing and comparing characteristics of the citation networks in two nutritional associations of disease, one where the two research designs generally agree (polyunsaturated fatty acids and cardiovascular mortality), and one where they disagree (vitamin E and cardiovascular mortality). We performed systematic reviews in each example and constructed and analyzed the respective citation networks. We identified no differences between the characteristics of the two networks. Most interestingly, we found no evidence that observational studies predated RCTs in the translational process in either example.\(^12\) Here we expand our previous work to include a large systematic sample of
associations between nutrients and health outcomes that were examined in both epidemiological studies and in RCTs. We construct and analyze citation networks to describe translational paths, and assess the relationship between metrics that describe the citation networks and the concordance of the summary findings across the two research designs.
Methods

Overview

We developed a database of nutrient associations of health outcomes that have been examined in published meta-analyses of epidemiological studies and of RCTs. We term the studies included in the meta-analyses “index studies” or “index RCTs,” respectively. For each pair of meta-analyses (each nutrient-outcome association) we defined an “evidence base” including all publications identifiable by MEDLINE® searches on the nutrient and outcome of interest. This evidence base included all index studies or RCTs. We recorded the citations made by the index studies and the index RCTs among the publications in the evidence base, and created the respective citation graphs, that is, “maps” that encode which papers in the evidence base are cited by index studies or index RCTs, either directly, or through one or more intermediary papers. We then calculated quantitative metrics that describe the connectivity of the citation graphs, and tested for associations between these metrics and the probability that the summary results from epidemiological studies and from RCTs are concordant between them.

Database of Associations Between Nutrients and Health Outcomes

We identified associations between nutrients and health outcomes that were examined in epidemiological studies and in RCTs. More specifically, we defined a nutrient as any non-energy-yielding nutrient (e.g., vitamins and minerals) or any nutrient that provides some energy but does not contribute as major source of energy intakes in a regular diet (i.e., carbohydrate, protein, and fat). An example of the latter is n-3 polyunsaturated fatty acids (hereon called “omega-3”). We accepted clinical outcomes such as mortality or cardiovascular disease, and surrogate endpoints, such as difference in blood pressure. To be eligible for inclusion in the database, the same nutrient-outcome association should be examined in epidemiological studies and in RCTs.

There is a plethora of publications on the relationship between nutrients and health outcomes. For feasibility, we prioritized associations that have been “well studied” with epidemiological studies and with RCTs. We operationally defined as well-studied associations those for which we could find at least one published meta-analysis of epidemiologic data and at least one meta-analysis of RCTs. If there was only a meta-analysis of observational data but no meta-analysis of RCTs on the same association, we accepted a large RCT (>1000 participants) instead of a meta-analysis, if any existed. Correspondingly, we accepted a single epidemiological study (>5000 participants) instead of a meta-analysis of epidemiological data, if we could identify a meta-analysis of RCTs but no meta-analysis of epidemiological data on the same association.

Literature Searches To Identify Meta-Analyses on Nutrients

We searched MEDLINE from 1996 to March 2009 to identify all systematic reviews with meta-analyses that examined our prespecified list of nutrients. We supplemented MEDLINE searches by perusing the list of systematic reviews included in the World Cancer Research Fund International’s Report on Diet, Nutrition, Physical Activity and Cancer.13
Matching Meta-Analyses of Epidemiological Studies With Meta-Analyses of RCTs

We matched meta-analyses of epidemiological studies and of RCTs (or meta-analyses and large studies) using the following algorithm:

1. First, we identified all systematic review publications that included both a meta-analysis of observational studies and a meta-analysis of RCTs. We assumed that meta-analyses of epidemiological studies and of RCTs reported in the same paper were “matched,” i.e., referred to the same nutrient-outcome association, provided that there was not clear statement suggesting otherwise.

2. Subsequently, we identified papers reporting only meta-analyses of epidemiological studies or meta-analyses of RCTs. We “matched” a meta-analysis of epidemiological studies with a meta-analysis of RCTs if both examined the same nutrient in relation to the same health outcome, and irrespective of differences in the inclusion or exclusion criteria for populations, the meta-analyses’ publication dates, or the time periods covered by their searches. If a publication contained more than one meta-analysis on different nutrient-outcome combinations, we considered each nutrient-outcome combination separately.

3. Finally, for all remaining meta-analyses of epidemiological studies that were not matched with a meta-analysis of RCTs, we tried to identify a large RCT (of more than 1000 participants) on the same association as a match. Analogously, we tried to identify large epidemiological studies (>5000 participants) to pair with each remaining meta-analyses of RCTs. To find a matching “large” observational study, we compiled a list of all large cohorts included in matched meta-analyses of steps (i) and (ii), as well as additional well-known cohorts (list provided in Appendix A), and examined whether they had publications on the same nutrient-outcome association as an yet unmatched meta-analysis of RCTs by perusing the cohort’s Web site (if one exists), or by a focused MEDLINE search.

Data Extraction

We extracted the following details from our database of matched pairs: number of studies in each meta-analysis, summary effect sizes, number of studies for different study designs, and years of recruitment for the studies (year of publication was used instead, if recruitment details were not reported in the meta-analysis).

Assessment of Concordance or Discordance Between Epidemiological Studies and RCTs

We calculated agreement between the reported meta-analysis findings of observational data and of RCTs using three definitions. We defined “qualitative” concordance based on the direction of the summary effect sizes and their statistical significance. We defined as qualitatively concordant summary effects in epidemiological studies and in RCTs that were statistically significant (at the 0.05 level) and in the same direction (e.g., both summary odds ratios suggesting lower risk of disease), as qualitatively discordant statistically significant summary effects in opposite directions, and all other cases as “unclear”.

Alternatively, we defined quantitative discordance by testing whether the difference in the summary effects of epidemiological studies and of RCTs was statistically significant or not,
based on a z-score: $z = \frac{\text{effect}_{\text{epidemiological}} - \text{effect}_{\text{RCT}}}{\sqrt{\text{variance}_{\text{epidemiological}} + \text{variance}_{\text{RCT}}}} \sim N(0, 1)$. Associations with a p-value below 0.05, were considered discordant, while those with a p-value of at least 0.05 were not discordant. Finally, we also used a third, stricter definition of discordance, where we considered as discordant associations where the z-score was significant and summary findings of RCTs and observational studies pointed in different directions. For example, an association in which the meta-analysis of RCTs and epidemiological studies point in the same direction but have sufficiently different magnitudes so that the aforementioned z-score is statistically significant would be considered qualitatively concordant (as per the first definition), quantitatively discordant (as per the second definition), and quantitatively not discordant (with the third definition).

As described in a later paragraph, we assessed relationships between quantitative characteristics of the citation graphs in each association and the concordance between evidence from epidemiological studies and from RCTs.

**Citation Analyses**

**Identifying an Evidence Base for Each Association**

For each association we searched MEDLINE to identify an evidence base of publications relevant to the nutrients and disease of interest, using proper keywords and MeSH terms. We limited searches to English-language publications and to the years searched by the original systematic reviews. When these were not reported, we searched from inception of MEDLINE to the year of publication of the systematic review. We did not use search filters for specific research designs nor limited searches to humans.

For feasibility, we revised the search strategies (by adding or removing terms for nutrients or outcomes) to return approximately 3,000 and no more than 5,000 papers. Although arbitrary, these numbers are typical for systematic reviews. We refined the searches until they missed at most three index publications. If the search missed one to three index publications, these were added manually to the evidence base.

**Citation Networks and Graphs**

For each association we constructed the citation networks of the RCTs and observational studies included in our systematic reviews (“index” articles) and represented them as citation graphs. Appendix B provides details on how we formed the citation graphs and how we checked that they possess and fulfill theoretically anticipated properties and constraints. Briefly, among the articles in the evidence base of each association, we identified all articles cited by the index studies directly or after following citation links through one or more intermediary papers.

A citation graph describes citation relationships between articles in the evidence base (see example in Figure 1). Articles are represented by vertices and citation relationships by arcs that connect pairs of vertices. The direction of each arc is from the cited towards the citing article (following the flow of information). Representing citation relationships as graphs allows us to use tools from graph theory and network analysis to characterize the evidence base and the apparent information flow in it. Without loss of clarity, we use the terms “citation network” and “citation graph” interchangeably. We then proceeded to analyze these citation networks in each association as discussed below.
Figure 1. Illustration of a hypothetical citation network of six articles

The figure shows a citation network (represented as a citation graph) in a hypothetical evidence base of 6 articles. Each article is depicted by a vertex (A through F). The horizontal placement of the articles corresponds to their year of publication, their vertical arrangement is arbitrary – its main purpose is clarity of presentation. Arcs denote citation relationships. The direction of the arc indicates flow of information from the cited paper (beginning of the arc) to the citing paper (end of the arc). In the figure 5 out of 6 papers (papers B through F) are connected with citation arcs. There are 4 observed citation arcs (B→D, C→D, C→E, and D→F) out of a total of 14 possible ones (from any earlier- to any subsequently-published article). The density of the citation relationships is defined as the ratio of observed to possible citation arcs, i.e., here it is 4/14=0.29. Citation graphs have characteristics and properties that are described in the Appendix. Briefly, citation arcs are simple directed acyclic graphs: (1) An article does not cite itself, and there is at most one citation relationship between two articles (simple graph), (2) if one follows the direction of the citation arcs along any possible path, one cannot visit the same article twice, i.e., there are no circular directed paths (directed acyclic graph). (3) In addition, there is a temporal consistency constraint, e.g., an earlier-published article cannot cite a later-published article.

Size and Connectivity of Citation Networks

As mentioned before, we treat these citation networks as operational representations of the clinical evidence base in each nutrient-outcome association. A limited evidence base (in terms of number of articles or number of citation connections between them) may indicate inconsistencies in the findings of several studies, and thus may be encountered in topics where RCT and observational studies disagree. Reciprocally, associations where there is general congruence in the study findings may have many articles and numerous citation connections.

We counted the total number of articles and the total number of citation relationships between articles in the network. These are only a subset of the total counts of citations received by an article. For example, the study on vitamin E and coronary heart disease in men by Rimm et al. received 101 citations within the network but has received more than 1,582 citations in total. We also recorded the number of citations made and number of citations received by each article (calculating density of the citation relationships as per Figure 1), as well as the article’s hub and authority scores. The latter quantify an article’s importance in the citation network. The hub score is higher for articles that cite a lot of other articles (“integrate information”) and are also connected to other articles that also cite a lot of papers. The authority score is higher for articles that receive a lot of citations (i.e., are in some sense “sources of information”) and are also connected to other articles that receive a lot of citations. These metrics convey information on the connectivity of the graph and the relative importance of articles in a network.
Associations Between Citation Graph Metrics and Qualitative and Quantitative Concordance

We explored whether quantitative characteristics of the citation graphs were associated with qualitative or quantitative concordance, as defined in an earlier paragraph. Specifically, we explored the following characteristics: total number of vertices in the connected graph, total number of citation relationships, mean hub score, mean authority score, mean number of citations made, mean number of citations received, mean number of citations made or received over all the vertices in the graph, and the corresponding mean scores or numbers over the 20 vertices with the highest respective scores or numbers in each graph.

We planned on using proportional odds ordinal logistic regression for the three-category qualitative definition of concordance, and logistic regression for the quantitative definition. However, because no association was qualitatively discordant, we used a logistic regression for the qualitative definition as well. We analyzed one variable at a time, because the small number of available associations did not allow for more complex models. We did not have a sufficiently large sample size to meaningfully evaluate the generalizability of the analyses (e.g., by developing models in a training subset of the dataset and evaluating them in a validation subset).
Results

Database of Nutrient-Outcome Associations

We identified a total of 34 eligible associations between nutrients and health outcomes (Table 1). Sixteen of 34 were based on meta-analyses of epidemiological studies and of RCTs that were reported in the same paper,12,19-27 11 were based on “matched” meta-analyses that were reported in separate papers,26,28-36 3 on a meta-analysis of observational studies matched with a large RCT (>1,000 participants),35,37-39 and 4 on a meta-analysis of RCTs matched with a large epidemiological study (>5,000 participants).40-45 Some of the included papers reported results from a meta-analyses of more than one nutrients and health outcome pair.

In Table 1, 20 out of 34 associations pertained to vitamins, 7 pertained to minerals (such as calcium) and trace elements (such as selenium), and 7 to either macronutrients (fiber) or fatty acids. The examined endpoints span a wide range of clinical outcomes including mortality, stroke, and other cardiovascular outcomes, various cancers, acute macular degeneration, and fractures, and surrogate outcomes such as systolic and diastolic blood pressure. In two cases nutrient intake of the mother was examined with relation to a clinical outcome in the child (prenatal multivitamin supplementation in the mother and risk of neural tube defects in the fetus, and maternal calcium intake and blood pressure in the offspring).
Table 1. Eligible nutrient-outcome associations

<table>
<thead>
<tr>
<th>ID</th>
<th>Nutrient and Health Outcome</th>
<th>Epidemiological Studies</th>
<th>RCTs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>N Trials</td>
<td>N People</td>
</tr>
<tr>
<td>1</td>
<td>α-linolenic acid and systolic blood pressure (mmHg)</td>
<td>1</td>
<td>4,680</td>
</tr>
<tr>
<td>2</td>
<td>α-linolenic acid and diastolic blood pressure (mmHg)</td>
<td>1</td>
<td>4,680</td>
</tr>
<tr>
<td>3</td>
<td>Beta-carotene and cervical cancer</td>
<td>6</td>
<td>1537</td>
</tr>
<tr>
<td>4</td>
<td>Beta-carotene and acute macular degeneration</td>
<td>4</td>
<td>126,642</td>
</tr>
<tr>
<td>5</td>
<td>Beta-carotene and lung cancer</td>
<td>10</td>
<td>370,107</td>
</tr>
<tr>
<td>6</td>
<td>Beta-carotene and breast cancer incidence or mortality</td>
<td>13</td>
<td>210,962</td>
</tr>
<tr>
<td>7</td>
<td>Beta-carotene (±Vitamin A) and esophageal cancer</td>
<td>4</td>
<td>2331</td>
</tr>
<tr>
<td>8</td>
<td>Beta-carotene (±Vitamin A) and gastric cancer</td>
<td>3</td>
<td>29,908</td>
</tr>
<tr>
<td>9</td>
<td>Beta-carotene and CHD/CVD events</td>
<td>4</td>
<td>18,256</td>
</tr>
<tr>
<td>10</td>
<td>Beta-carotene and CHD/CVD mortality</td>
<td>4</td>
<td>3,053</td>
</tr>
<tr>
<td>ID</td>
<td>Nutrient and Health Outcome</td>
<td>Epidemiological Studies</td>
<td>RCTs</td>
</tr>
<tr>
<td>-----</td>
<td>-----------------------------</td>
<td>-------------------------</td>
<td>------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N Studies</td>
<td>N People</td>
</tr>
<tr>
<td>11</td>
<td>Calcium and hip fracture</td>
<td>8&lt;sup&gt;th&lt;/sup&gt;</td>
<td>239,597</td>
</tr>
<tr>
<td>12</td>
<td>Calcium (mother) and offspring systolic blood pressure (mmHg)</td>
<td>2&lt;sup&gt;nd&lt;/sup&gt;</td>
<td>7283</td>
</tr>
<tr>
<td>13</td>
<td>Calcium and diastolic blood pressure (mmHg)</td>
<td>17&lt;sup&gt;th&lt;/sup&gt;</td>
<td>38,950</td>
</tr>
<tr>
<td>14</td>
<td>Calcium and systolic blood pressure (mmHg)</td>
<td>17&lt;sup&gt;th&lt;/sup&gt;</td>
<td>38,950</td>
</tr>
<tr>
<td>15</td>
<td>Fiber and colorectal adenoma/cancer</td>
<td>1&lt;sup&gt;st&lt;/sup&gt;</td>
<td>33,971</td>
</tr>
<tr>
<td>16</td>
<td>Folate and cancer cervix</td>
<td>8&lt;sup&gt;th&lt;/sup&gt;</td>
<td>2835</td>
</tr>
<tr>
<td>17</td>
<td>Folate and CHD/CVD events</td>
<td>4&lt;sup&gt;th&lt;/sup&gt;</td>
<td>225,808</td>
</tr>
<tr>
<td>18</td>
<td>Folic acid and stroke</td>
<td>1&lt;sup&gt;st&lt;/sup&gt;</td>
<td>43,732</td>
</tr>
<tr>
<td>19</td>
<td>Omega-3 and all-cause mortality</td>
<td>3&lt;sup&gt;rd&lt;/sup&gt;</td>
<td>3801</td>
</tr>
<tr>
<td>20</td>
<td>Omega-3 and cancer or cancer mortality</td>
<td>7&lt;sup&gt;th&lt;/sup&gt;</td>
<td>112,454</td>
</tr>
<tr>
<td>21</td>
<td>Omega-3 and combined CVD events</td>
<td>7&lt;sup&gt;th&lt;/sup&gt;</td>
<td>69,732</td>
</tr>
<tr>
<td>22</td>
<td>Omega-3 and stroke</td>
<td>4&lt;sup&gt;th&lt;/sup&gt;</td>
<td>52,026</td>
</tr>
</tbody>
</table>
### Table 1. Eligible nutrient-outcome associations (continued)

<table>
<thead>
<tr>
<th>ID</th>
<th>Nutrient and Health Outcome</th>
<th>Epidemiological Studies</th>
<th>RCTs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>N Studies</td>
<td>N People</td>
</tr>
<tr>
<td>23</td>
<td>Omega-3 and CHD mortality</td>
<td>6</td>
<td>83,578</td>
</tr>
<tr>
<td>24</td>
<td>Selenium and prostate cancer</td>
<td>15</td>
<td>168,226</td>
</tr>
<tr>
<td>25</td>
<td>Vitamin C and breast cancer incidence or mortality</td>
<td>13</td>
<td>268,291</td>
</tr>
<tr>
<td>26</td>
<td>Vitamin C and CHD/CVD events</td>
<td>12</td>
<td>12,419</td>
</tr>
<tr>
<td>27</td>
<td>Vitamin C and CHD/CVD mortality</td>
<td>12</td>
<td>23,391</td>
</tr>
<tr>
<td>28</td>
<td>25(OH)D and total cancer</td>
<td>2</td>
<td>30,149</td>
</tr>
<tr>
<td>29</td>
<td>Vitamin E and acute macular degeneration</td>
<td>4</td>
<td>124,307</td>
</tr>
<tr>
<td>30</td>
<td>Prenatal multivitamin and neural tube defects</td>
<td>24</td>
<td>ND</td>
</tr>
<tr>
<td>31</td>
<td>Vitamin E and CHD/CVD events</td>
<td>6</td>
<td>184,594</td>
</tr>
<tr>
<td>32</td>
<td>Vitamin E and CHD/CVD mortality</td>
<td>9</td>
<td>137,237</td>
</tr>
<tr>
<td>33</td>
<td>Selenium and CHD</td>
<td>22</td>
<td>15,133</td>
</tr>
<tr>
<td>34</td>
<td>Selenium and lung cancer</td>
<td>12</td>
<td>1973 cases (total not reported)</td>
</tr>
</tbody>
</table>

**Abbreviations:** CHD = chronic heart disease; CVD = cerebrovascular disease; ND = no data reported in the systematic reviews
Concordance or Discordance Between Epidemiological and RCT Data

Table 2 shows concordance between the summary effects in the two research designs using the three definitions. Using the first definition, 6 out of 34 topics (18%) were qualitatively concordant, and the remaining 28 were classified as “unclear” (there was no topic with qualitatively discordant epidemiological and RCT data). Twelve of 34 associations (35%) had a statistically significant z-score, and using the second definition, evidence from epidemiological studies and evidence from RCTs was statistically significantly discordant. The remaining 22 associations had no evidence of significant discordance. Of the 12 association with a statistically significant z-score, in 6 the meta-analysis of the RCTs and that of the epidemiological studies pointed to different directions. Thus, six examples were quantitatively discordant using the third definition.
Table 2. Qualitative and quantitative concordance of effects in epidemiological studies and RCTs

<table>
<thead>
<tr>
<th>ID</th>
<th>Direction of Effect (Epi)</th>
<th>Significance (Epi)</th>
<th>Direction of Effect (RCT)</th>
<th>Significance (RCT)</th>
<th>Qualitative Concordance (1st definition)</th>
<th>z-Score</th>
<th>p-Value</th>
<th>Quantitative Concordance (2nd definition)</th>
<th>Quantitative Concordance (3rd definition)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Decreasing</td>
<td>Not sign</td>
<td>Decreasing</td>
<td>Not sign</td>
<td>Unclear</td>
<td>0.22</td>
<td>0.83</td>
<td>Not discordant</td>
<td>Not discordant</td>
</tr>
<tr>
<td>2</td>
<td>Decreasing</td>
<td>Sign</td>
<td>Decreasing</td>
<td>Not sign</td>
<td>Unclear</td>
<td>-0.91</td>
<td>0.36</td>
<td>Not discordant</td>
<td>Not discordant</td>
</tr>
<tr>
<td>3</td>
<td>Increasing</td>
<td>Not sign</td>
<td>Decreasing</td>
<td>Not sign</td>
<td>Unclear</td>
<td>1.36</td>
<td>0.18</td>
<td>Not discordant</td>
<td>Not discordant</td>
</tr>
<tr>
<td>4</td>
<td>Increasing</td>
<td>Not sign</td>
<td>Increasing</td>
<td>Not sign</td>
<td>Unclear</td>
<td>0.08</td>
<td>0.94</td>
<td>Not discordant</td>
<td>Not discordant</td>
</tr>
<tr>
<td>5</td>
<td>Decreasing</td>
<td>Not sign</td>
<td>Increasing</td>
<td>Not sign</td>
<td>Unclear</td>
<td>-1.50</td>
<td>0.13</td>
<td>Not discordant</td>
<td>Not discordant</td>
</tr>
<tr>
<td>6</td>
<td>Decreasing</td>
<td>Sign</td>
<td>Increasing</td>
<td>Not sign</td>
<td>Unclear</td>
<td>-1.54</td>
<td>0.12</td>
<td>Not discordant</td>
<td>Not discordant</td>
</tr>
<tr>
<td>7</td>
<td>Decreasing</td>
<td>Sign</td>
<td>Increasing</td>
<td>Not sign</td>
<td>Unclear</td>
<td>-3.08</td>
<td>0.00</td>
<td>Discordant</td>
<td>Discordant</td>
</tr>
<tr>
<td>8</td>
<td>Decreasing</td>
<td>Sign</td>
<td>Increasing</td>
<td>Not sign</td>
<td>Unclear</td>
<td>-2.00</td>
<td>0.05</td>
<td>Discordant</td>
<td>Discordant</td>
</tr>
<tr>
<td>9</td>
<td>Increasing</td>
<td>Not sign</td>
<td>Increasing</td>
<td>Not sign</td>
<td>Unclear</td>
<td>0.00</td>
<td>1.00</td>
<td>Not discordant</td>
<td>Not discordant</td>
</tr>
<tr>
<td>10</td>
<td>Increasing</td>
<td>Not sign</td>
<td>Increasing</td>
<td>Not sign</td>
<td>Unclear</td>
<td>0.00</td>
<td>1.00</td>
<td>Not discordant</td>
<td>Not discordant</td>
</tr>
<tr>
<td>11</td>
<td>Increasing</td>
<td>Not sign</td>
<td>Decreasing</td>
<td>Not sign</td>
<td>Unclear</td>
<td>1.20</td>
<td>0.23</td>
<td>Not discordant</td>
<td>Not discordant</td>
</tr>
<tr>
<td>12</td>
<td>Decreasing</td>
<td>Not sign</td>
<td>Decreasing</td>
<td>Not sign</td>
<td>Unclear</td>
<td>0.13</td>
<td>0.89</td>
<td>Not discordant</td>
<td>Not discordant</td>
</tr>
<tr>
<td>13</td>
<td>Decreasing</td>
<td>Sign</td>
<td>Decreasing</td>
<td>Sign</td>
<td>Concordant</td>
<td>2.52</td>
<td>0.01</td>
<td>Discordant</td>
<td>Not discordant</td>
</tr>
<tr>
<td>14</td>
<td>Decreasing</td>
<td>Sign</td>
<td>Decreasing</td>
<td>Not sign</td>
<td>Unclear</td>
<td>0.65</td>
<td>0.51</td>
<td>Not discordant</td>
<td>Not discordant</td>
</tr>
<tr>
<td>15</td>
<td>Decreasing</td>
<td>Sign</td>
<td>Increasing</td>
<td>Not sign</td>
<td>Unclear</td>
<td>-2.48</td>
<td>0.01</td>
<td>Discordant</td>
<td>Discordant</td>
</tr>
<tr>
<td>16</td>
<td>Decreasing</td>
<td>Not sign</td>
<td>Decreasing</td>
<td>Not sign</td>
<td>Unclear</td>
<td>-1.15</td>
<td>0.25</td>
<td>Not discordant</td>
<td>Not discordant</td>
</tr>
<tr>
<td>17</td>
<td>Decreasing</td>
<td>Sign</td>
<td>Decreasing</td>
<td>Not sign</td>
<td>Unclear</td>
<td>-2.08</td>
<td>0.04</td>
<td>Discordant</td>
<td>Not discordant</td>
</tr>
<tr>
<td>18</td>
<td>Decreasing</td>
<td>Sign</td>
<td>Decreasing</td>
<td>Not sign</td>
<td>Unclear</td>
<td>-0.98</td>
<td>0.33</td>
<td>Not discordant</td>
<td>Not discordant</td>
</tr>
<tr>
<td>19</td>
<td>Decreasing</td>
<td>Sign</td>
<td>Decreasing</td>
<td>Not sign</td>
<td>Unclear</td>
<td>-1.64</td>
<td>0.10</td>
<td>Not discordant</td>
<td>Not discordant</td>
</tr>
<tr>
<td>20</td>
<td>Increasing</td>
<td>Not sign</td>
<td>Increasing</td>
<td>Not sign</td>
<td>Unclear</td>
<td>-0.37</td>
<td>0.71</td>
<td>Not discordant</td>
<td>Not discordant</td>
</tr>
<tr>
<td>21</td>
<td>Decreasing</td>
<td>Not sign</td>
<td>Decreasing</td>
<td>Not sign</td>
<td>Unclear</td>
<td>-0.31</td>
<td>0.75</td>
<td>Not discordant</td>
<td>Not discordant</td>
</tr>
<tr>
<td>22</td>
<td>Increasing</td>
<td>Not sign</td>
<td>Increasing</td>
<td>Not sign</td>
<td>Unclear</td>
<td>0.00</td>
<td>1.00</td>
<td>Not discordant</td>
<td>Not discordant</td>
</tr>
<tr>
<td>23</td>
<td>Decreasing</td>
<td>Sign</td>
<td>Decreasing</td>
<td>Sign</td>
<td>Concordant</td>
<td>-2.07</td>
<td>0.04</td>
<td>Discordant</td>
<td>Not discordant</td>
</tr>
<tr>
<td>24</td>
<td>Decreasing</td>
<td>Sign</td>
<td>Decreasing</td>
<td>Sign</td>
<td>Concordant</td>
<td>2.14</td>
<td>0.03</td>
<td>Discordant</td>
<td>Not discordant</td>
</tr>
<tr>
<td>25</td>
<td>Decreasing</td>
<td>Sign</td>
<td>Increasing</td>
<td>Not sign</td>
<td>Unclear</td>
<td>-2.19</td>
<td>0.03</td>
<td>Discordant</td>
<td>Discordant</td>
</tr>
<tr>
<td>26</td>
<td>Decreasing</td>
<td>Sign</td>
<td>Increasing</td>
<td>Not sign</td>
<td>Unclear</td>
<td>-2.79</td>
<td>0.01</td>
<td>Discordant</td>
<td>Discordant</td>
</tr>
<tr>
<td>27</td>
<td>Decreasing</td>
<td>Not sign</td>
<td>Decreasing</td>
<td>Not sign</td>
<td>Unclear</td>
<td>0.22</td>
<td>0.83</td>
<td>Not discordant</td>
<td>Not discordant</td>
</tr>
<tr>
<td>28</td>
<td>Increasing</td>
<td>Not sign</td>
<td>Decreasing</td>
<td>Not sign</td>
<td>Unclear</td>
<td>1.46</td>
<td>0.14</td>
<td>Not discordant</td>
<td>Not discordant</td>
</tr>
<tr>
<td>29</td>
<td>Decreasing</td>
<td>Not sign</td>
<td>Increasing</td>
<td>Not sign</td>
<td>Unclear</td>
<td>-2.06</td>
<td>0.04</td>
<td>Discordant</td>
<td>Discordant</td>
</tr>
<tr>
<td>30</td>
<td>Decreasing</td>
<td>Sign</td>
<td>Decreasing</td>
<td>Sign</td>
<td>Concordant</td>
<td>-1.56</td>
<td>0.12</td>
<td>Not discordant</td>
<td>Not discordant</td>
</tr>
<tr>
<td>31</td>
<td>Decreasing</td>
<td>Sign</td>
<td>Decreasing</td>
<td>Not sign</td>
<td>Unclear</td>
<td>-1.13</td>
<td>0.26</td>
<td>Not discordant</td>
<td>Not discordant</td>
</tr>
<tr>
<td>32</td>
<td>Decreasing</td>
<td>Sign</td>
<td>Decreasing</td>
<td>Not sign</td>
<td>Unclear</td>
<td>-2.41</td>
<td>0.02</td>
<td>Discordant</td>
<td>Not discordant</td>
</tr>
<tr>
<td>33</td>
<td>Decreasing</td>
<td>Sign</td>
<td>Decreasing</td>
<td>Sign</td>
<td>Concordant</td>
<td>3.06</td>
<td>0.00</td>
<td>Discordant</td>
<td>Not discordant</td>
</tr>
<tr>
<td>34</td>
<td>Decreasing</td>
<td>Sign</td>
<td>Decreasing</td>
<td>Sign</td>
<td>Concordant</td>
<td>0.95</td>
<td>0.34</td>
<td>Not discordant</td>
<td>Not discordant</td>
</tr>
</tbody>
</table>

Note: Sign = statistically significant at the 0.05 level
Citation Graphs

We were able to construct citation graphs for 28 topics (Table 3). For 6 topics we were unable to form reliable citation graphs (for technical reasons related to changes in the format of the citation information obtained from the ISI Web site). Briefly, there was variability in the total number of vertices in the 28 citation graphs: The median number of vertices (articles) was 253 (25th–75th percentiles: 95, 356), and the median number of edges (citation relationships) was 1181 (25th–75th percentiles: 255, 1620). The citation graphs are relatively sparsely connected with median density 0.018 (low density or low connectivity is the norm for citation graphs). The table shows also the mean hub and authority scores across all papers in a graph, and across the 20 papers with the highest respective scores. Graphs with higher mean hub scores and higher mean authority scores have more papers that “integrate” and “provide” information (as conveyed by citation relationships), respectively, compared with graphs with lower scores. When all papers in a graph were considered, the median hub score was 0.09 (25th–75th percentiles: 0.07, 0.13) and the median authority score was 0.13 (25th–75th percentiles: 0.10, 0.21). Finally, similar variability across the 28 topics was observed in the mean number of citation received, citations made, or total citations made or received. When all papers in a graph were considered the median numbers and 25th–75th percentile ranges were 4.3 (3.2, 4.7), 4.3 (3.2, 4.7) and 8.5 (6.3, 9.4), respectively. (When all papers are considered in a citation graph, the mean number of citations made equals the mean number of citations received.) The corresponding mean scores or numbers over the 20 papers with the largest respective scores or numbers are also shown (Table 3).
Table 3. Summary of citation analysis of nutrient exposure-health outcome pairs studied by systematic reviews or large studies

<table>
<thead>
<tr>
<th>ID</th>
<th>Vertices</th>
<th>Citation Relationships</th>
<th>Citation Density (*10^-3)</th>
<th>Mean Hub Score</th>
<th>Mean Authority Score</th>
<th>Mean Citations Made</th>
<th>Mean Citations Received</th>
<th>Mean Citations Made or Received</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>All</td>
<td>Top 20</td>
<td>All</td>
<td>Top 20</td>
<td>All*</td>
</tr>
<tr>
<td>1</td>
<td>91</td>
<td>226</td>
<td>27.6</td>
<td>0.11</td>
<td>0.35</td>
<td>0.10</td>
<td>0.31</td>
<td>2.48</td>
</tr>
<tr>
<td>2</td>
<td>91</td>
<td>226</td>
<td>27.6</td>
<td>0.11</td>
<td>0.35</td>
<td>0.10</td>
<td>0.31</td>
<td>2.48</td>
</tr>
<tr>
<td>3</td>
<td>62</td>
<td>264</td>
<td>69.8</td>
<td>0.22</td>
<td>0.53</td>
<td>0.21</td>
<td>0.48</td>
<td>4.26</td>
</tr>
<tr>
<td>4</td>
<td>62</td>
<td>220</td>
<td>62.1</td>
<td>0.19</td>
<td>0.45</td>
<td>0.23</td>
<td>0.53</td>
<td>3.67</td>
</tr>
<tr>
<td>5</td>
<td>254</td>
<td>1131</td>
<td>17.6</td>
<td>0.07</td>
<td>0.40</td>
<td>0.13</td>
<td>0.52</td>
<td>4.45</td>
</tr>
<tr>
<td>6</td>
<td>463</td>
<td>1784</td>
<td>8.3</td>
<td>0.13</td>
<td>0.80</td>
<td>0.04</td>
<td>0.32</td>
<td>3.85</td>
</tr>
<tr>
<td>7</td>
<td>365</td>
<td>1721</td>
<td>13.0</td>
<td>0.07</td>
<td>0.50</td>
<td>0.09</td>
<td>0.51</td>
<td>4.72</td>
</tr>
<tr>
<td>8</td>
<td>365</td>
<td>1721</td>
<td>13.0</td>
<td>0.07</td>
<td>0.50</td>
<td>0.09</td>
<td>0.51</td>
<td>4.72</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>173</td>
<td>556</td>
<td>18.7</td>
<td>0.10</td>
<td>0.48</td>
<td>0.09</td>
<td>0.50</td>
<td>3.21</td>
</tr>
<tr>
<td>12</td>
<td>183</td>
<td>565</td>
<td>17.0</td>
<td>0.25</td>
<td>0.74</td>
<td>0.04</td>
<td>0.22</td>
<td>3.09</td>
</tr>
<tr>
<td>13</td>
<td>279</td>
<td>1447</td>
<td>18.7</td>
<td>0.09</td>
<td>0.59</td>
<td>0.11</td>
<td>0.54</td>
<td>5.19</td>
</tr>
<tr>
<td>14</td>
<td>279</td>
<td>1447</td>
<td>18.7</td>
<td>0.09</td>
<td>0.59</td>
<td>0.11</td>
<td>0.54</td>
<td>5.19</td>
</tr>
<tr>
<td>15</td>
<td>60</td>
<td>189</td>
<td>53.4</td>
<td>0.17</td>
<td>0.41</td>
<td>0.27</td>
<td>0.64</td>
<td>3.15</td>
</tr>
<tr>
<td>16</td>
<td>39</td>
<td>200</td>
<td>135.0</td>
<td>0.28</td>
<td>0.50</td>
<td>0.30</td>
<td>0.50</td>
<td>5.13</td>
</tr>
<tr>
<td>17</td>
<td>252</td>
<td>1911</td>
<td>30.2</td>
<td>0.08</td>
<td>0.44</td>
<td>0.24</td>
<td>0.78</td>
<td>7.58</td>
</tr>
<tr>
<td>18</td>
<td>360</td>
<td>2602</td>
<td>20.1</td>
<td>0.05</td>
<td>0.40</td>
<td>0.21</td>
<td>0.78</td>
<td>7.23</td>
</tr>
<tr>
<td>19</td>
<td>406</td>
<td>1230</td>
<td>7.5</td>
<td>0.02</td>
<td>0.25</td>
<td>0.11</td>
<td>0.75</td>
<td>3.03</td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>392</td>
<td>2741</td>
<td>17.9</td>
<td>0.06</td>
<td>0.50</td>
<td>0.12</td>
<td>0.67</td>
<td>5.59</td>
</tr>
<tr>
<td>24</td>
<td>159</td>
<td>779</td>
<td>31.0</td>
<td>0.08</td>
<td>0.36</td>
<td>0.16</td>
<td>0.58</td>
<td>4.90</td>
</tr>
<tr>
<td>25</td>
<td>463</td>
<td>1784</td>
<td>8.3</td>
<td>0.13</td>
<td>0.80</td>
<td>0.04</td>
<td>0.32</td>
<td>3.85</td>
</tr>
<tr>
<td>26</td>
<td>98</td>
<td>255</td>
<td>26.8</td>
<td>0.09</td>
<td>0.30</td>
<td>0.18</td>
<td>0.64</td>
<td>2.60</td>
</tr>
<tr>
<td>27</td>
<td>98</td>
<td>255</td>
<td>26.8</td>
<td>0.09</td>
<td>0.30</td>
<td>0.18</td>
<td>0.64</td>
<td>2.60</td>
</tr>
<tr>
<td>28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>60</td>
<td>220</td>
<td>62.1</td>
<td>0.19</td>
<td>0.45</td>
<td>0.23</td>
<td>0.53</td>
<td>3.67</td>
</tr>
<tr>
<td>30</td>
<td>280</td>
<td>1294</td>
<td>16.6</td>
<td>0.08</td>
<td>0.57</td>
<td>0.12</td>
<td>0.49</td>
<td>4.62</td>
</tr>
<tr>
<td>31</td>
<td>351</td>
<td>1519</td>
<td>12.4</td>
<td>0.04</td>
<td>0.35</td>
<td>0.14</td>
<td>0.68</td>
<td>4.33</td>
</tr>
<tr>
<td>32</td>
<td>351</td>
<td>1519</td>
<td>12.4</td>
<td>0.04</td>
<td>0.35</td>
<td>0.14</td>
<td>0.68</td>
<td>4.33</td>
</tr>
<tr>
<td>33</td>
<td>116</td>
<td>397</td>
<td>29.8</td>
<td>0.07</td>
<td>0.30</td>
<td>0.21</td>
<td>0.70</td>
<td>3.42</td>
</tr>
<tr>
<td>34</td>
<td>302</td>
<td>1292</td>
<td>14.2</td>
<td>0.05</td>
<td>0.38</td>
<td>0.10</td>
<td>0.62</td>
<td>4.28</td>
</tr>
</tbody>
</table>

Some association topics share the same citation graph (those with identical numbers – see text).

*In each topic, when all papers are considered, the mean citations made and received are the same. The corresponding numbers differ when only the subset of 20 papers with the highest number of citations made or received are considered.*
Associations Between Citation Graph Metrics and Qualitative and Quantitative Concordance

There was no association between the characteristics of the citation graphs and the observed qualitative or quantitative concordance in the 28 analyzable topics (Table 4). The table shows analyses in which the predictors (numbers or mean scores) were log transformed using the base 2. Thus, the odds ratios in the table represent the change in the odds of finding concordance per doubling of the value of the predictor. The p-values remain identical for any other logarithmic transformation of the predictors. Analyses using other transformations (square root, no transformation) were qualitatively similar. It appears that the examined connectivity metrics of the citations maps do not correlate with the likelihood that the summary findings from meta-analyses of epidemiological studies and of RCTs are in qualitative or quantitative agreement.
<table>
<thead>
<tr>
<th>Citation Graph Metric</th>
<th>Qualitative (First Definition) OR (95% CI)*</th>
<th>Qualitative (Second Definition) OR (95% CI)*</th>
<th>Qualitative (Third Definition) OR (95% CI)*</th>
<th>Concordance P-value</th>
<th>Concordance P-value</th>
<th>Concordance P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among all papers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Papers (vertices)</td>
<td>1.57 (0.60, 4.10)</td>
<td>0.82 (0.40, 1.68)</td>
<td>1.15 (0.38, 2.00)</td>
<td>0.33</td>
<td>0.59</td>
<td>0.74</td>
</tr>
<tr>
<td>Citation relationships (edges)</td>
<td>1.65 (0.74, 3.70)</td>
<td>0.78 (0.44, 1.40)</td>
<td>1.21 (0.42, 1.64)</td>
<td>0.19</td>
<td>0.40</td>
<td>0.59</td>
</tr>
<tr>
<td>Graph citation density</td>
<td>0.83 (0.32, 2.18)</td>
<td>1.03 (0.48, 2.21)</td>
<td>1.01 (0.39, 2.51)</td>
<td>0.70</td>
<td>0.95</td>
<td>0.99</td>
</tr>
<tr>
<td>Hub score (mean)</td>
<td>0.54 (0.16, 1.86)</td>
<td>1.19 (0.46, 3.02)</td>
<td>0.62 (0.51, 5.12)</td>
<td>0.32</td>
<td>0.72</td>
<td>0.41</td>
</tr>
<tr>
<td>Authority score (mean)</td>
<td>1.15 (0.34, 3.91)</td>
<td>0.70 (0.26, 1.88)</td>
<td>1.14 (0.30, 2.60)</td>
<td>0.82</td>
<td>0.47</td>
<td>0.81</td>
</tr>
<tr>
<td>Citations made (mean)</td>
<td>3.11 (0.28, 34.59)</td>
<td>0.44 (0.07, 3.00)</td>
<td>2.25 (0.04, 4.41)</td>
<td>0.35</td>
<td>0.40</td>
<td>0.48</td>
</tr>
<tr>
<td>Citations received (mean)</td>
<td>3.11 (0.28, 34.59)</td>
<td>0.44 (0.07, 3.00)</td>
<td>2.25 (0.04, 4.41)</td>
<td>0.35</td>
<td>0.40</td>
<td>0.48</td>
</tr>
<tr>
<td>Total citations (mean)</td>
<td>3.11 (0.28, 34.59)</td>
<td>0.44 (0.07, 3.00)</td>
<td>2.25 (0.04, 4.41)</td>
<td>0.35</td>
<td>0.40</td>
<td>0.48</td>
</tr>
<tr>
<td>Among the 20 papers with the highest number or score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hub score (mean)</td>
<td>0.74 (0.08, 7.35)</td>
<td>1.13 (0.19, 6.74)</td>
<td>0.52 (0.24, 15.83)</td>
<td>0.80</td>
<td>0.89</td>
<td>0.54</td>
</tr>
<tr>
<td>Authority score (mean)</td>
<td>3.70 (0.22, 62.12)</td>
<td>0.26 (0.03, 2.04)</td>
<td>1.11 (0.12, 6.70)</td>
<td>0.32</td>
<td>0.16</td>
<td>0.92</td>
</tr>
<tr>
<td>Citations made (mean)</td>
<td>1.73 (0.43, 6.98)</td>
<td>0.70 (0.25, 1.98)</td>
<td>1.26 (0.24, 2.63)</td>
<td>0.42</td>
<td>0.50</td>
<td>0.71</td>
</tr>
<tr>
<td>Citations received (mean)</td>
<td>1.97 (0.62, 6.27)</td>
<td>0.74 (0.32, 1.68)</td>
<td>1.47 (0.26, 1.80)</td>
<td>0.22</td>
<td>0.46</td>
<td>0.43</td>
</tr>
<tr>
<td>Total citations (mean)</td>
<td>1.94 (0.50, 7.53)</td>
<td>0.70 (0.26, 1.89)</td>
<td>1.50 (0.21, 2.12)</td>
<td>0.31</td>
<td>0.48</td>
<td>0.49</td>
</tr>
</tbody>
</table>

*per doubling of the predictor (score or number).
Discussion

In this work we provide a large-scale empirical exploration of the hypothesis that the connectivity (in terms of citation relationships) between observational studies and RCTs is associated with the likelihood that the summary findings of these two research designs would agree or disagree. In previous hypothesis forming work we compared quantitative characteristics of the citation maps in one example where the two research designs agree (n-3 polyunsaturated fatty acids and cardiovascular mortality), and in one where they disagree (vitamin E and cardiovascular mortality). This larger evaluation shows that depending on the definition of concordance, the summary findings from meta-analyses of observational studies and of RCTs were in qualitative agreement approximately one out of five times (6 out of 34) and were not quantitatively discordant about two thirds of the time (22 out of 34). Despite the variation in the size and the connectivity of the citation graphs across the examined associations, we find no evidence of association between quantitative metrics of the citation graphs and the probability that the two research designs have concordant findings.

The definitions of concordance or discordance used in this work are based on findings from published meta-analyses. Meta-analyses of RCTs and of observational studies are principled systematic syntheses of the totality of the evidence in a given topic, and therefore, it is reasonable and practical to use them as such in meta-epidemiological research. Meta-analyses were used as an operational way to assess a large body of knowledge, but this operationalization necessitates some caveats. Any observed discordance between the results of RCTs and epidemiological studies should not be automatically attributed to the effect of bias, within each study or across studies. Differences in the populations, levels and form of nutrient intakes, and in the exact definitions of the outcomes are alternative explanations, especially if RCTs and epidemiological studies have systematic differences with respect to factors that affect the strength of the association. Conversely, large methodological and epidemiological diversity can lead to substantial statistical heterogeneity, and excessive between-study variability can in make it more difficult to identify discorndances between meta-analyses of RCTs and of epidemiological studies.

Research builds on previous findings, and knowledge advances as scientific data and its interpretation are communicated to inform subsequent studies. The generally accepted vehicle for this communication of knowledge is the scientific publication. Therefore, in some sense, the citation networks formed by scientific papers are a representation of the flow of biomedical knowledge through various translational paths, and the analysis of citation networks may provide insights on the size and maturity of the relevant evidence base.

We hypothesized that nutrient associations of disease where RCTs and observational studies agree in the direction and significance of their findings may differ in the size of their evidence base, and the patterns of information flow in it. However, we found no association between the quantitative characteristics of the citation networks and the likelihood of concordance in the examined examples. This negative finding is easy to explain: Only the most dramatic aberrations in a translational path would have been identified in an analysis of citation relationships. An example would be a whole body of literature generated from an unwarranted extrapolation of previous findings.

There are limitations to using citation analysis to understand translational paths in a given topic. Citing previous research is a complex process. As much as we would like citation practices to be impartial and scientific, they are influenced by personal beliefs, biases and preferences, and are subject to citation distortion. The latter includes citation bias, i.e., when one systematically
ignores articles that contain content at odds with ones claims, citation amplification where a lot of citations propagate a belief without any evidence support, and citation invention, that includes citing content but ascribing different meaning to it and converting an hypothesis to fact through citation alone \(^{48}\). The operational searches used to define the evidence base in each association may have missed early publications that potentially influenced subsequent work, either observational or experimental. We did not assess the methodological quality or the content of the publications in the citation networks, nor did we examine whether any citations made were supportive or dismissive, factual or mistaken, or just providing general context. Publication bias, which occurs as a result of researchers publishing only studies with significant positive results and ignoring studies that report non-significant results, also affects the citation network.

Despite the above limitations, we believe that citation analysis is one of the few representations of the translational process that can be objectively quantified. It merits attention as it can provide a framework to analyze the series of research steps that led to RCTs in humans, and perhaps identify unwarranted extrapolations in the translational process, if any exist \(^{12,48}\). The examined quantitative characteristics of the citation maps in each association cannot predict the probability that the findings from the two designs agree or disagree. It is unclear whether there is a good way to describe the maturity of the evidence base on an association between nutrients and outcomes. At a minimum, purely bibliometric approaches are not a good way to prioritize which nutrient exposures merit further study, and for which health outcomes.


Appendix A. List of Observational Cohorts and the Screened Publications

Japan Public Health Center-Based Prospective Study


**Nurses’ Health Study**


**Seven-Day Adventist Study**


**The Health Professional Follow-up Study 2009**


The Framingham Heart Study


**The Iowa Women’s Health Study**


The Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study


Netherlands Cohort Study on Diet and Cancer


The Singapore Chinese Health Study


**Norwegian Mother & Child Cohort Study**


**European Prospective Investigation into Diet and Cancer Study**


**NIH-AARP Diet and Health Study**


**Dietary Intake of Folic Acid and Colorectal Cancer Risk in a Cohort of Women**


**The Danish Diet, Cancer and Health Study**


Women’s Health Initiative


**Swedish Mammography Cohort Study**


**Multiethnic Cohort Study**


**Women’s Antioxidant Cardiovascular Study**


**Northern Swedish Health and Diseases Study**


**Black Women’s Health Study**

Appendix B. Details on the Construction of Citation Graphs

For each of the $N$ papers in each topic (in each corpus), we queried Thompson ISI to obtain a list of its references. We examined these lists to identify which papers are cited by other papers in the corpus, ignoring citations by papers that were not included in the corpus. We organized this information into an adjacency matrix $M$, i.e., an $N \times N$ matrix whose elements $m_{ij}$ code the number of citations (0 or 1) from the $j$-th to the $i$-th paper. This matrix contains all the information that is necessary to create the citation graph of the $N$ papers.

In graph theory terminology, citation graphs are directed and acyclic. Directed, because the direction of the arcs is always from the paper that is being cited towards the paper that is making the citation. Acyclic means that there are no closed “loops” in a citation network, because a paper cannot cite itself, and, generally, two or more papers do not cite each other. There are however rare instances where two papers that are, e.g., published in the same issue, cite each other, resulting in a non-acyclic graph. We can transform this graph into a directed graph by assuming that such papers essentially cite each other’s preprints (Appendix Figure 1). This transformation does not affect important characteristics of the networks such as the distributions of in-degrees, authority scores, hub scores, and main path scores (see Glossary).

Appendix Figure 1. Resolving mutual citations to ensure that the citation graph is a simple directed acyclic graph.

After correcting for papers that cite each other, we verified that the resulting networks were acyclic, and that there was temporal consistency, i.e., that there were no citations from earlier-published to later-published papers (a paper published in 1979 cannot cite a paper published in 2000).
Practicalities and Coherency Assessment

We matched citations by exact string matching of titles, after basic preprocessing. Subsequently, we used fuzzy string-matching algorithms (algorithms that tolerate small discrepancies between two title strings) to identify titles that did not match exactly, but pertained to the same paper. This can happen especially for older papers that were entered manually in the Thompson ISI database, because of typos or alternative spelling of title words (e.g., “Randomized trial of …” in MEDLINE may become “Randomized trial of …” in ISI). A human manually reviewed all title pairs that had a Levenshtein edit distance of 5 or less. The results of the manual review were taken into account when forming the final citation graph.

Main Path Articles

Main paths go from a source vertex to a sink vertex in a citation network, and include vertices and arcs with the highest traversal weights. A source vertex is a vertex that has only outgoing arcs (indegree=0, outdegree>0) and a sink vertex is a vertex that has only incoming arcs (indegree>0, outdegree=0). The traversal weight of a vertex or an arc expresses the proportion of paths from all sources to all sinks in the entire network that include the particular vertex or arc. We calculated transversal weights using the Search Path Link Count (SPLC) method implemented in the Pajek software. The results of the SPLC algorithm were very similar to those of alternative methods (vertex pair projection count, VPPC, and search path vertex pair, SPVP). For details on these algorithms and a comparison of their relative performance see the technical report by Batagelj 2003. One may consider main path articles as central in a field because they integrate information from previous articles (vertices) and propagate information to other articles (vertices).

Glossary of graph theory and network terms

<table>
<thead>
<tr>
<th>Term</th>
<th>Description</th>
<th>Example in citation network</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graph</td>
<td>Mathematical construct consisting of vertices or nodes, and edges that connect pairs of vertices. If the edges are directed, the graph is called a directed graph. A directed graph is acyclic when one cannot return to the same vertex following any combination of directed edges.</td>
<td>A citation network graph is a simple directed acyclic graph.</td>
</tr>
<tr>
<td>Subgraph</td>
<td>A part of a graph that includes a subset of the vertices and all the edges between them.</td>
<td></td>
</tr>
<tr>
<td>Vertex or node</td>
<td>The fundamental unit of a graph.</td>
<td>In a citation network vertices represent papers</td>
</tr>
<tr>
<td>Arc or directed edge</td>
<td>A connection between a pair of nodes. It is directed when the order in which the nodes are connected is important.</td>
<td>Arcs go from cited papers (A) to citing papers (B) to denote that some information is flowing from the former to the latter: A→B</td>
</tr>
<tr>
<td>Path</td>
<td>A sequence of vertices such that from each vertex there is an edge to the next vertex in the sequence, as if one were visiting vertices by walking a long the edges. In a directed path, the direction of the edges matters, one would be allowed to walk only in the direction of the arcs.</td>
<td></td>
</tr>
<tr>
<td>Term</td>
<td>Description</td>
<td>Example in citation network</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Indegree, outdegree, hub and authority scores</td>
<td>These are measures of the centrality (&quot;importance&quot;) of a paper in a citation network. Indegree is the number of incoming arcs (number of papers cited). Outdegree is the number of outgoing arcs (citations received). Hub and authority scores, is the relative importance of a vertex in a network. A vertex has higher hub or authority score if it has a higher indegree or outdegree respectively, and if it is connected to other vertices with high hub or authority scores respectively.</td>
<td>The distribution of these measures can characterize the connectivity of the citation network, and potentially the amount of information that flows through citations. These measures may identify &quot;key&quot; papers in the corpus of citations.</td>
</tr>
<tr>
<td>Temporal consistency</td>
<td>A citation graph is temporally consistent when the cited articles have been published earlier than the corresponding citing articles.</td>
<td>A temporally consistent graph must be acyclic.*</td>
</tr>
</tbody>
</table>

*In theory two articles can cite each other. This can happen e.g., in articles appearing in the issue. Such an instance would render the graph non-acyclic (for two papers A and B that cite each other, there is a directed cyclic path: $A \rightarrow B \rightarrow A$).

**References**