

Issues related to the conduct of systematic reviews: a focus on the nutrition field¹⁻³

David Moher and Andrea C Tricco

ABSTRACT

Systematic reviews (SRs) are an increasingly popular evidence-based tool and are often used to answer complex research questions across many different research domains. Early SR methodology was advanced by social scientists, and the term meta-analysis was coined by a social scientist who also conducted research in psychology. SRs have recently become popular in healthcare and are likely to be beneficial in any field. The aim of this report is to highlight issues in SR conduct with a focus on the field of nutrition and to make recommendations on improving SR conduct in this area. Development of the research question is probably the most important step in conducting an SR. The 4 main components of an answerable question are 1) the patient, population, or problem; 2) the intervention, independent variable, or exposure; 3) the comparators; and 4) the dependent variables or outcomes of interest. The question will be used to determine the optimal methods for conducting the SR. SRs often include study designs beyond randomized trials and do not always include a meta-analysis of the results. Other topics explored include understanding and interpreting discordant reviews and the importance of reporting tools [eg, QUality Of Reporting Of Meta-analyses (QUOROM Statement) or CONSolidated Standards Of Reporting of Trials (CONSORT Statement)]. Recommendations are then provided, such as developing a capacity-building program, searching the primary literature for research gaps, and extending reporting tools such as the QUOROM Statement to the field of nutrition. *Am J Clin Nutr* 2008;88:1191-9.

INTRODUCTION

Synthesizing evidence from research dates back to the early 1900s, yet improved methods for evidence synthesis have been developed only within the past few years (1). Early systematic review (SR) methods were advanced by social scientists (1), and the term "meta-analysis" was coined in 1976 by a social scientist who also did research in the area of psychology (2). Although the importance of evidence synthesis in medicine was recognized in the 1970s (3), widespread use of these techniques did not occur until 2 decades later (1). Potentially contributing to this "movement" was evidence that the judgments and opinions of experts were often biased. A pivotal study by Antman et al (4) concluded that effective therapies were not mentioned, whereas ineffective and potentially harmful therapies often were recommended by experts. The term "evidence-based medicine" was coined only ≈15 y ago (5).

Over the past 2 decades, SRs have become increasingly popular in evidence-based healthcare (6). Policy makers, clinicians,

and others have little time to keep up with the published literature. SRs offer a convenient way for them to keep up-to-date with the current evidence. Furthermore, SRs are often used in the development of clinical practice guidelines (7) and are increasingly viewed as a useful tool for health decision makers (8, 9). The increased utility of SRs has probably contributed to an increase in their publication rates. It was recently estimated that, each year, MEDLINE cites 2500 new SRs in English (10)—a rate 166 times that reported in 1991 (11).

Because of the utility of SRs, the SR approach is likely to benefit any field (12), including nutrition. The aims of this report are to highlight issues in SR conduct with a focus on the field of nutrition and to make recommendations for improving SR conduct in this area.

METHODS

Terminology

The terminology used to describe an SR and a meta-analysis has evolved over time. There is no consistent use of these terms, but we will adopt the definitions used by The Cochrane Collaboration (13). An SR consists of a clearly formulated question and explicit methods to identify, select, and critically appraise relevant research and collects and analyzes data from the studies that are included in the review. A meta-analysis is the use of statistical techniques in an SR, which integrates the results of included studies. Thus an SR does not necessarily include a meta-analysis.

Other types of reviews, such as narrative reviews, will not use the explicit methods outlined in the definition proposed by The Cochrane Collaboration (Table 1; 10). As such, these other types of reviews may be susceptible to bias (4). A properly conducted

¹ From the Children's Hospital of Eastern Ontario Research Institute, Ottawa, Canada (DM and ACT) and the Department of Epidemiology & Community Medicine, Faculty of Medicine (DM), and the Institute of Population Health, University of Ottawa, Ottawa, Canada (ACT).

² Supported by the International Life Sciences Institute Research Foundation, a University of Ottawa Research Chair (to DM), and a Canadian Institutes of Health Research Canada Graduate Scholarship and a University of Ottawa Excellence Scholarship (to ACT). The sponsor had no involvement in the writing of the manuscript or the decision to submit the manuscript for publication.

³ Reprints not available. Address correspondence to D Moher, Clinical Epidemiology Program, Ottawa Health Research Institute, Box 208, 501 Smyth Road, Ottawa, ON K1H 8L6, Canada. E-mail: dmoher@uottawa.ca. Received April 4, 2008. Accepted for publication July 30, 2008. doi: 10.3945/ajcn.2008.26255.



TABLE 1
Contrasting systematic review and the methods of other types of reviews¹

Systematic review (minimum criteria)	Other types of reviews
A protocol (ie, working plan for the systematic review) is developed, along with a clearly formulated question.	A protocol is not used, and the question may not be clear and concise.
The literature is searched broadly by using many different methods (eg, searching multiple relevant databases, contacting experts, or scanning reference lists). The search strategy, including databases and years searched and search terms employed, is transparently reported in the systematic review manuscript.	The literature may not be searched, or only the literature that the authors are aware of is searched. The search strategy is not reported or is not fully reported in the review manuscript.
To determine study eligibility, the literature is screened by using criteria set a priori. Ideally, 2 independent reviewers screen all material and resolve conflicts through discussion. The eligibility criteria, including the number of articles that were excluded and the reasons for exclusion, are transparently reported in the systematic review manuscript.	No eligibility criteria are set, and the authors are free to TAB and choose which studies should be included in the review. The inclusion criteria and the number of excluded articles are not reported in the manuscript.
The risk of bias of the included studies is assessed by using validated and applicable study appraisal instruments to determine the validity of the study results. Ideally, 2 reviewers independently appraise the quality of all included studies and resolve conflicts through discussion.	The quality of the included studies is not assessed, or unvalidated instruments are used to appraise the quality of included studies.
Data are abstracted consistently from all included studies by using a previously defined data abstraction form. Ideally, 2 reviewers independently abstract data from all included studies and resolve conflicts through discussion. The data abstraction form is transparently described in the systematic review manuscript.	The authors are free to TAB and choose which results from the included studies to report. Data may not be abstracted consistently from all included studies. The data abstraction form is not described.
The results are synthesized by using the totality of evidence. A meta-analysis may be conducted if the included studies are deemed homogenous in terms of study population, study design, exposure or intervention examined, comparators studied, and outcomes assessed.	Studies are summarized on the basis of the results that are most appealing to the authors. A meta-analysis may be performed with studies that are not homogenous.
The discussion section provides an overall summary of the strengths and weaknesses of the included studies and a summary of the strengths and weaknesses of the systematic review itself.	The discussion section may not provide an overall summary of the weaknesses of included studies or of the weaknesses of the review itself.

¹ One way to determine whether a study is a systematic review or another type of review, such as a narrative review, is to examine whether there is a methods section in the report. If a methods section is missing from the study report, chances are the study was not a systematic review, and, thus, it is susceptible to considerable bias in the results.

SR is much more resource- and labor-intensive than is a narrative review (14).

The systematic review team

Developing the SR team is a monumental step. The core SR team usually includes the funder or commissioning agency that provides the team with context for the question the review is addressing, members with content or clinical expertise (eg, clinical nutritionists), members with expertise in conducting SRs—ideally, a librarian with expertise in searching the literature in a comprehensive and unbiased manner (12)—and epidemiologists with experience in conducting primary research (eg, case-control or cohort studies of nutrition). Other relevant stakeholders or organizations may be involved, depending on the purpose of the SR, for example, if the SR will be used for decision-making or guideline development. Consultation with a statistician may also be required if statistical pooling (ie, meta-analysis) is being considered.

Developing a systematic review question

SRs are currently being conducted in many different areas and have been used to address many types of research questions (Table 2). For example, an SR was used to answer the following question within the crime and justice domain: “Does a neighborhood watch reduce crime?” (15). Another SR was used to answer the question within education: “Which approaches to parental involvement improve the academic performance of elementary school-age children?” (16). SRs have been used to answer many complex questions within the field of nutrition, including

“Which community-based nutrition and physical activity interventions are effective at preventing chronic disease in low-income populations?” (17); “What are the long-term effects of advice to restrict dietary sodium in adults with and without hypertension?” (18); and “What is the association between a diet rich in whole grain, bran, and germ and the risk of type 2 diabetes?” (19). Other real-life examples of questions answered by SRs are provided elsewhere (12).

After deciding who the SR team will comprise, developing a clear and concise question is probably the most important step in conducting an SR. The 4 components of an answerable question include 1) the patient, population, or problem (P); 2) the intervention, independent variable, or exposure (I); 3) the comparators (C); and 4) the dependent variables or outcomes of interest (O) (27). Sometimes an additional component is added, namely, study design (S), which is used to limit the SR to certain types of studies, such as cohort studies. These components are known collectively as PICOS (1). [Variations of PICOS also exist, such as one that adds a “D” for study design (PICO-D) and one that incorporates a “T” for timing and an “S” for setting (PICOTS).]

An example of the use of these components in creating a question is that, to examine the sodium intake in the United States, the question could be phrased as “What are the health-related effects (O) associated with a high-salt diet (I) compared with a low-salt diet (C) among persons living in the United States (P)?” To examine the association between fruit and vegetable intake and cardiovascular disease among diabetics, the question could be phrased as “Among persons with diabetes (P), what is the association of a high fruit and vegetable diet (I) compared



TABLE 2
Complex questions answered by systematic reviews within different domains

Domain and questions
Nutrition
Which community-based nutrition and physical activity interventions are effective at preventing chronic disease in low-income populations? (17)
What are the long-term effects of advice to restrict dietary sodium in adults with and without hypertension? (18)
What is the association between a diet rich in whole grain, bran, and germ and the risk of type 2 diabetes? (19)
Crime and justice
Does a neighborhood watch reduce crime? (15)
What are the effects of closed-circuit television on crime? (20)
What is the effectiveness and cost-effectiveness of counter-terrorism strategies? (21)
Education
Which approaches to parental involvement improve the academic performance of elementary school-age children? (16)
To what extent and in what ways does access to after-school programs affect student context (ie, student location, supervision, and safety), participation in enriching activities, behaviors, social and emotional development, and academic outcomes for youth? (22)
What is the effectiveness of volunteer tutoring programs in improving the academic skills of students from kindergarten through grade 8 in the United States? (23)
Social welfare
What is the effectiveness of behavioral and cognitive behavioral training interventions in improving placement stability, the psychological well-being and functioning of persons providing foster care, and the behavioral and relationship problems of children receiving foster care? (24)
What are the effects of cognitive behavioral therapy and similar interventions on men's physical abuse of their female partners? (25)
Does exercise alone or exercise as part of a comprehensive intervention improve self-esteem among children and young people? (26)

with a low fruit and vegetable diet (C) with respect to myocardial infarction, stroke, and cardiovascular disease-related mortality (O)?”

The explicit methods involved in conducting an SR are dependent on the question that is being asked. Literature can be searched according to the PICOS components. The eligibility criteria (ie, inclusion and exclusion criteria) will be based on the question. The data abstraction instrument or instruments will be developed to abstract from the included studies the data that will be used to answer the question. The synthesis of results abstracted from the included studies will also help provide an answer to the question. The methods used in an SR may vary depending on the type of question and the requirements of the organization conducting or commissioning the review. Three examples of such an organization include the Cochrane Collaboration, the Campbell Collaboration, and the Evidence-based Practice Center (EPC) program.

The Cochrane Collaboration (13) is a global entity including ≈10 000 members organized into Clinical Review Groups (eg, Schizophrenia Group), Methods Groups (eg, Bias Methods Group), and Fields (eg, Diet and Nutrition Subfield of the Cochrane Primary Health Care Field). The mission of The Cochrane Collaboration is to conduct SRs in all areas of healthcare, including nutrition. Cochrane reviews are completed by using an approach described in the Cochrane Handbook (13). A distinctive characteristic of Cochrane reviews is

that they often answer questions regarding the effectiveness or efficacy of an intervention (ie, “what works”). As such, The Cochrane Collaboration relies strongly on the synthesizing of evidence from randomized controlled trials (RCTs). However, there are some exceptions to this within Cochrane. An example is the Effective Practice and Organisation of Care (EPOC) group, which often addresses broader questions and includes other study designs.

RCTs are given high prominence within the Cochrane Collaboration because they minimize the influence of bias in their results. Whereas there is considerable evidence to support this perspective and the importance of randomized trials, there also are data indicating that such study designs account for only ≈10% of the healthcare literature (28, 29). In contrast, the Campbell Collaboration, the sister group to The Cochrane Collaboration, addresses broad questions related to crime and justice, education, and social welfare and includes all types of study designs (eg, RCTs, cross-sectional studies, case-control studies, cohort studies, and case series) (Table 2). Campbell reviews are conducted according to the guidance provided on the organization's website (Internet: www.campbellcollaboration.org/guidelines.asp).

The EPC program (30), started in 1997, is funded by the Agency for Healthcare Research and Quality and consists of 14 centers throughout North America (Internet: <http://www.ahrq.gov/clinic/epcix.htm>). The remit of this program is to complete SRs in a wide variety of healthcare and policy areas, and guides for conducting these reviews have been proposed (32). As do those of the Campbell Collaboration, these SRs address broader types of questions and thus are not limited to reports of RCTs. A more detailed discussion of the EPC program can be found elsewhere (31).

Whereas The Cochrane Collaboration, Campbell Collaboration, and EPC program are high-profile, ≥100 other groups conduct reviews, commission SR conduct, or do both (33). Examples include the Canadian Agency for Drugs and Technologies in Health (Internet: cadth.ca/index.php/en/hta/programs/health-technology-assessment/process), the Centre for Reviews and Dissemination (York University, York, United Kingdom; Internet: www.york.ac.uk/inst/crd/report4.htm), and, specific to the field of nutrition, the American Dietetic Association (Internet: www.eatright.org/cps/rde/xchg/ada/hs.xsl/education_13146_ENU_HTML.htm). All of these organizations conduct SRs by using different templates.

FINDINGS

Differences in questions or conduct may lead to differences in results and conclusions

Variations of the SR question or differences in the SR methods used may explain why SRs addressing a similar topic have different, sometimes startlingly different, results or conclusions, or both; these are called discordant reviews. An example of discordant reviews is apparent among 4 SRs examining the cardiovascular effects of vitamin E supplements (Table 3; 34–37). Although the SR questions were similar, variations are apparent. For example, one of the reviews (35) focused on effectiveness, whereas another (36) focused on efficacy. Furthermore, different methods were used to conduct the reviews. For example, one SR (36) searched multiple databases, whereas another (34) searched



TABLE 3
Differences among 4 systematic reviews examining the cardiovascular effects of vitamin E¹

Study component	Vivekananthan et al (34)	Eidelman et al (35)	Shekelle et al (36)	Miller et al (37)
Question posed	What effects do antioxidant vitamins (ie, vitamin E and β -carotene) have on all-cause mortality and CV death?	What is the effectiveness of vitamin E in the treatment and prevention of CVD?	What is the efficacy of vitamin E supplements in the prevention and treatment of CVD?	What effect do antioxidant vitamins (ie, vitamin E and β -carotene) have on all-cause mortality and CV death?
Methods				
Search details	Searched a database and the references of included studies. Search dates NR.	"Computerized search." Search dates NR.	Searched databases (inception to 2001 or 2002, depending on the database) and the references of included studies; contacted experts.	Searched databases (1996–August 2004), the references of included studies, and the references of previous SRs.
Databases searched	MEDLINE	NR	MEDLINE, EMBASE, MANTIS, Allied & Complementary Medicine, Biosis Previews, CAB Health, Cancerlit, The Cochrane Library, Social SciSearch, SciSearch Cited Ref Sci, TGG Health & Wellness Database	MEDLINE, Cochrane Central Register of Controlled Trials (CENTRAL)
Inclusion criteria	Studies with >1000 participants; developed countries without vitamin E deficiency; trials: study quality (trial size, randomization, ITT analysis)	Randomized trials; vitamin E therapy; treatment or prevention of CVD	Published or unpublished material; trials examining vitamin E	Random allocation; vitamin E alone or in combination; placebo group; men or nonpregnant women; study duration >1 y; ≥ 10 deaths had to occur during the trial
Exclusion criteria	Trials without all-cause mortality data	NR	NR	NR
Language limits	NR	English only	No language restrictions	No language restrictions
Screening	2 independent reviewers	NR	2 independent reviewers	NR
Data abstraction	2 independent reviewers	NR	2 independent reviewers	3 independent reviewers
Outcomes	All-cause mortality; CV death; all-cause cerebrovascular accident; nonfatal MI	CV events: MI; nonfatal stroke; CV death	All-cause mortality; CV deaths	Adjusted and unadjusted all-cause mortality for high and low dosages
Analysis	Meta-analysis: odds ratios and 95% CIs	Meta-analysis: odds ratios and 95% CIs from the difference between observed minus expected number of events	Random effects meta-analysis: log risk ratios and 95% CIs	2-level hierarchical logistic regression model analyzed by low and high doses; analysis was transformed into risk differences and risk ratios
Included studies				
Included studies (n)	7	7	84	19
Abbreviated trial name (if provided)	ATBC, CHAOS, GISSI, HOPE, HPS (MRC), AREDS, PPP	Linxian, ATBC, CHAOS, GISSI, HOPE, PPP, HPS (MRC)	Linxian, ATBC, PPP, ASAP, SPACE, HOPE, GISSI, CHAOS, MASI, HATS, MVP	MINVITAOX, Linxian, SUVIMAX, ATBC, Linqu, GISSI, PPP, HOPE, AREDS, PPS, VECAT, CHAOS, RACT, HPS (MRC), SPACE, WAVE, ADCS, DATATOP
Follow-up (y) ²	4.5 (1.4–6.3)	4.5 (1.5–6.1)	NR (2–7)	4.0 (1.4–8.2)
Participants (n) ²	4761 (1035–14 564)	11 324 (2002–29 584)	NR (NR)	3492 (256–32 704)

(Continued)



TABLE 3 (Continued)

Study component	Vivekananthan et al (34)	Eidelman et al (35)	Shekelle et al (36)	Miller et al (37)
All-cause mortality	1.02 (0.98–1.06)	NR	0.96 (0.84–1.10)	
Unadjusted results				20 IU/d: 0.98 (0.95–1.01) 50 IU/d: 0.99 (0.96–1.01) 100 IU/d: 0.99 (0.97–1.02) 200 IU/d: 1.01 (0.98–1.03) 500 IU/d: 1.03 (1.0–1.06) ³ 1000 IU/d: 1.05 (1.01–1.09) ³ 2000 IU/d: 1.07 (1.01–1.12) ³
Adjusted results				20 IU/d: 0.98 (0.95–1.02) 50 IU/d: 0.99 (0.96–1.03) 100 IU/d: 1.0 (0.97–1.04) 200 IU/d: 1.01 (0.98–1.05) 500 IU/d: 1.04 (0.99–1.08) 1000 IU/d: 1.06 (1.0–1.11) ³ 2000 IU/d: 1.08 (1.01–1.14) ³
CV event	NR	0.98 (0.94–1.03)	NR	NR
CV death	6.0% vs 6.0% control ($P = 0.94$)	1.00 (0.92–1.05)	0.97 (0.80–1.19)	NR
All-cause stroke	3.6% vs 3.5% control ($P = 0.7$)	NR	NR	NR
CV death or nonfatal MI	9.8% vs 9.8% control ($P = 0.93$)	NR	NR	NR
Nonfatal MI	NR	1.00 (0.92–1.09)	0.72 (0.51–1.02)	NR
Nonfatal stroke	NR	1.03 (0.93–1.14)	NR	NR
Ischemic stroke	NR	1.01 (0.90–1.14)	NR	NR
Hemorrhagic stroke	NR	1.24 (0.96–1.59)	NR	NR
Fatal MI	NR	NR	0.97 (0.74–1.27)	NR
Intermediate outcomes: total cholesterol, LDL, HDL	NR	NR	NS	NR
Authors' conclusions	"When used as secondary prevention, vitamin E did not reduce the risk of CV endpoints . . . [W]e do not support the continued use of vitamin E . . ."	"This overview . . . of vitamin E . . . provides conclusive evidence of a lack of statistically significant or clinically important benefit or harm regarding any important CV events or its components . . ."	"The available scientific studies offer little evidence that supplementation with vitamin E has any benefit on CVD prevention or treatment."	". . . [W]e identified a dose-response relationship between vitamin E supplementation and all-cause mortality . . . [A]ll-cause mortality progressively increased for dosages > 150 IU/d."

¹ CV, cardiovascular; CVD, CV disease; NR, not reported; ITT, intention-to-treat; MI, myocardial infarction; SR, systematic review.

² Median; range in parentheses (all such values).

³ $P < 0.05$.

only one database. Different inclusion or exclusion criteria, language limitations (eg, including only reports in English), and outcomes of interest also were evident. The number of included studies varied from 7 (34, 35) to 84 (36).

Although the results of the reviews varied, consistency was observed in 3 reviews (34–36); no association between vitamin E and any cardiovascular endpoint was observed (Table 3). However, the fourth review (37) conducted a dose-response analysis, for which high doses of vitamin E (≥ 500 IU/d) were shown to significantly increase the risk of all-cause mortality by 9–14%. Two of the SRs (34, 36) concluded that vitamin E had no benefit with respect to cardiovascular events, and one SR (35) concluded that vitamin E had neither benefit nor harm with respect to cardiovascular events. However, the fourth review (37) concluded that a dose-response relation between vitamin E and all-cause mortality was observed (ie, that vitamin E at high doses is harmful).

For an understanding of discordant reviews, an assessment of the risk of bias (ie, study quality) of the SRs is likely to be beneficial (38). Such an exercise was completed for the 4 SRs discussed above by

using the Oxman and Guyatt (39) instrument (Table 4). This validated instrument consists of 9 main criteria for assessing the scientific quality of review articles. The final item asks the assessor to rate the overall scientific quality of the SR by using a score ranging from 1 (ie, extensive flaws) to 7 (ie, minimal flaws). When the Oxman-Guyatt tool was applied to these 4 SRs, the scores ranged from 1 [indicating extensive flaws in the SR (35)] to 7 [indicating minimal flaws in the SR (36)]; the other 2 SRs fell between the extremes.

Relation between the conduct and reporting of systematic reviews

The example of discordant reviews highlights an important issue—that an intertwined relation between the conduct and reporting of SRs is likely. For example, if SR authors do not report an assessment of the risk of bias (ie, study quality), given the integral part of a review that such an assessment is, it is unlikely that they conducted their review appropriately. Describing methods to enhance the quality of reporting of SRs also is likely to be important to any efforts to improve the conduct of such reviews.



TABLE 4
Risk of bias in 4 systematic reviews of vitamin E¹

Item	Vivekananthan et al (34)	Eidelman et al (35)	Shekelle et al (36)	Miller et al (37)
1. Were the search methods used to find evidence (original research) on the primary questions(s) stated?	Y	P	Y	Y
2. Was the search for evidence reasonably comprehensive?	N	N	Y	Y
3. Were the criteria used for deciding which studies to include in the overview reported?	Y	Y	P	Y
4. Was bias in the selection of studies avoided?	Y	C	Y	C
5. Were the criteria used for assessing the validity of the included studies reported?	P	C	Y	C
6. Was the validity of all of the studies referred to in the text assessed with the use of appropriate criteria (either in the selection of studies for inclusion or in analysis of the studies cited)?	C	C	Y	C
7. Were the methods used to combine the findings of the relevant studies (to reach a conclusion) reported?	P	Y	Y	Y
8. Were the findings of the relevant studies combined appropriately with respect to the primary question the overview addressed?	Y	Y	Y	Y
9. Were the conclusions made by the author or authors supported by the data or the analysis (or both) reported in the overview?	Y	Y	Y	Y
Score ²	3	1	7	4

¹ Y, yes; N, no; P, partially; C, not clear.

² The maximum score is 7. A score of 1 means the review has extensive flaws; a score of 2–3 represents major flaws; a score of 4–5 represents minor flaws; and a score of 6–7 represents minimal flaws. For further information on the scoring scheme, see Oxman and Guyatt (39).

As it would be for any type of research, reporting SRs to the highest possible standard is critical, because the report is one of the most important ways of disseminating the results to the broader research community and beyond. Dissemination is likely to be more widespread if the report accurately and transparently describes the research methods and results. It is important that the report of a research endeavor is, in many instances, the only public record of the existence of the research itself. High-quality reports enable readers to critically interpret and use the results. There also are emerging data indicating that the use of reporting guidelines improves the quality of the research reports (40).

To help improve the quality of reporting of all research, the UK National Health Service National Library for Health and the UK National Institute for Health Research recently provided funds to set up the EQUATOR Network (Internet: www.equator-network.org). The objective of this network is to improve the quality of scientific publications by promoting transparent and accurate reporting of health research and by supporting the development of high-quality reporting guidelines. An early project completed by the network executive was an SR that identified existing reporting guidelines and a survey that identified issues faced by those developing such guidelines (41). More than 50 sets of reporting guidelines were identified; they can be accessed via the EQUATOR Network website.

Many studies have evaluated the quality of SR reports. In 1987, Mulrow (42) examined 50 review articles published in 1985 and 1986 in 4 leading medical journals; he found that none met all 8 explicit scientific criteria that they examined, such as a quality assessment of included studies. Also in 1987, Sacks et al (43) evaluated the adequacy of reporting on 23 characteristics in

6 domains in 83 meta-analyses. They found that reporting generally was poor; between 1 and 14 characteristics ($\bar{x} \pm SD: 7.7 \pm 2.7$) were adequately reported. A 1996 update of this study found little improvement (44). To address the suboptimal reporting of meta-analyses, an international group developed guidance on the quality of reporting of meta-analyses of RCTs, called the QUOROM Statement.

The QUOROM Statement consists of a 21-item checklist that documents the process of completing a meta-analysis and a flow diagram that details the number and status of included articles at each stage of the meta-analysis process (45). Since the QUOROM Statement was published in 1999, the evidence base underlying the conduct of SRs has matured. In addition, a more comprehensive understanding of some conceptual issues, advances in methodology, practical innovations in the conduct and reporting of SRs, and changes in terminology have occurred. It is for these reasons that the QUOROM Statement was updated in June 2005. It is now called the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Statement, and it applies not only to SRs of RCTs but also to any SR assessing the effectiveness of any intervention.

Quality of reporting of systematic reviews within nutrition

We completed an extensive search to identify studies that have examined the general quality of reporting of SRs within nutrition; we found that few studies exist, perhaps because there are few individual studies (eg, cohort studies or RCTs) to be synthesized. More likely the reasons are multifactorial, such as a lack of capacity to conduct SRs on nutrition topics and possibly a lack of clarity as to how best to conduct and report such reviews.



There are limited data on the specific issues that systematic reviewers in the nutrition domain should specifically address. Gibson et al (46) noted that the following steps were involved in nutrition assessment during the conduct of RCTs and meta-analyses: measuring food intakes via qualitative or quantitative means (eg, food diary or food recalls); converting food intakes to observed intakes of nutrients by using food-composition tables; and evaluating nutrient adequacy. Challenges encountered during a series of reviews of omega-3 fatty acids were also documented (47), including difficulties in sorting through the numerous endpoints reported in the included studies and variations in the source of dietary intake (eg, supplement or dietary component), study design, and study duration of the included studies. These challenges highlight how important it is for those completing individual studies to ensure that these points are adequately integrated into the conduct of such studies.

A systematic review is dependent on the quality of included studies

The successful conduct of any SR is dependent to a large degree on the quality of the included studies. Historians may well view the first 50 y of reporting of RCTs with some surprise. They will encounter what might be described as a cognitive dissonance: a disconnect between the increasing sophistication of the design and the swelling cost of these studies and the apparent lack of care—a disastrous lack in some cases—with which they have been reported. A series of studies beginning in 1995 found empirical evidence that results may be biased when trials use inferior methods or are reported without adequate description of the methods; notably, failure to conceal the allocation process is associated with a $\geq 30\%$ exaggeration in the effectiveness of an intervention (48). The cause for concern is obvious: if the conduct or reporting of RCTs is poor, treatments may be introduced that are less effective than was thought or that may even be ineffective, and these reports are included in SRs.

This concern led to the Consolidated Standards of Reporting of Trials (CONSORT) Statement recommendations originally published in 1996 and updated in 2001; another update is currently in progress (Internet: www.consort-statement.org). The CONSORT Statement consists of a 22-item evidence-based checklist that authors can use to guide the reporting of their trials and a flow diagram to account for the flow of participants throughout the trial process. The CONSORT Statement has been extended in several directions, such as cluster trials, equivalence and noninferiority trials, and nonpharmacologic treatments. These extensions have been made because of inadequate reporting of important study aspects not covered by the original CONSORT Statement. The CONSORT statement for herbal interventions was developed to help improve the quality of reporting of such studies, which may also have implications for how those studies, as well as observational ones, are designed and conducted. Currently, the CONSORT Statement has not been widely implemented in the field of nutrition.

SUMMARY AND RECOMMENDATIONS

SR techniques have been used to answer complex research questions across many different research domains. SRs have recently become popular within healthcare and are likely to be

beneficial in any field, including nutrition. Developing the research question is probably the most important step in conducting an SR. SRs often include study designs beyond randomized trials and do not always include a meta-analysis of the results. As with any other type of study, the optimal conduct and reporting of SRs is necessary.

To strengthen the scientific rigor of nutrition science research endeavors, we propose the following ideal recommendations, recognizing that their adoption may be difficult without the involvement of an expert panel akin to the Institute of Medicine (Table 5).

First, we recommend the development of a capacity-building program to conduct and report nutrition-related SRs, including a comprehensive training portfolio. This program could be implemented within universities and colleges that have nutrition programs or within organizations in the nutrition domain.

Second, to ascertain which nutrition topics are most likely to benefit from synthesis (ie, an SR), we recommend that a comprehensive search of the primary literature (eg, cohort studies) be conducted and that the resulting reports be categorized according to the nutrition-related topic. When there is a sufficient evidence base, we recommend that this priority list form the basis of SRs that should be commissioned for immediate completion. Such an exercise is also likely to provide evidence for locating the gaps in the evidence base and thus to give a rationale for the areas in which primary research could be earmarked for investment. To identify gaps in the SR literature and generate new SR questions, we also recommend the development of a repository of SRs within the field of nutrition.

Third, SRs are most likely to be useful if the quality of reporting of the individual studies included is sufficiently high; this is a longer-term undertaking than some of the other recommendations. We recommend that the CONSORT Statement be implemented for nutrition-related clinical trials (Internet: www.consort-statement.org). Such an initiative for herbal interventions could build on the CONSORT Statement. Similarly, where appropriate, we recommend that other reporting guides be extended or implemented for this field (Internet:

TABLE 5
Recommendations with respect to systematic reviews (SRs)

Number	Recommendation
1	Develop a capacity-building program to conduct and report nutrition-related SRs, including a comprehensive training portfolio.
2	Conduct a comprehensive search to develop a list of priority topics on which sufficient evidence exists to warrant an SR, and identify gaps in primary research. Develop a repository of SRs in nutrition to identify gaps and generate new SR questions.
3	Implement CONSORT and other reporting guidelines within the field of nutrition research. Obtain endorsement of reporting guidelines by journals publishing nutrition-related research.
4	Extend PRISMA Statement to reporting of SRs evaluating nutrition. Obtain endorsement of SR reporting guidelines by journals publishing SRs on nutrition-related topics.
5	Develop a standard guide to the conduct of SRs in nutritional science.
6	Review the quality of nutrition literature every year.



www.equator-network.org). We also recommend that journals publishing nutrition-related research endorse evidence-based reporting guidelines to improve the quality of nutrition research in their journals.

Fourth, we recommend that the PRISMA Statement be extended to the reporting of SRs that evaluate nutrition (45). We also recommend that journals publishing SRs on nutrition-related topics endorse the PRISMA Statement to improve the quality of reporting of this research in their journals.

Fifth, because the conduct and reporting of SRs are so intertwined, we recommend that an evidence-based approach to developing a standard guide for conducting SRs within the nutritional science domain be undertaken. Preexisting SR conduct guides providing some insight into the nutrition field, such as the Cochrane Handbook (13), can be used as a starting point in the development of this standard guide. It would also be beneficial to include persons with expertise in many different areas to develop evidence-based guidelines for the best reporting and conduct of nutrition-related SRs. These persons may include a broad base of funders and those who commission nutrition-related SRs, those actively involved in obtaining empirical evidence to guide the conduct of nutrition-related SRs, journal editors who publish SRs, and persons with experience and expertise in conducting SRs in the nutrition field. Such a consortium is likely to have a monumental effect in improving the reporting and conduct of nutrition-related SRs.

Sixth, we recommend that an annual review of the quality of nutrition literature be conducted. Such a review will provide important stakeholder groups with an evidence base as to the current quality of this important literature.

We thank the attendees of the Evidence-based Approaches to Developing Dietary Guidance Workshop, hosted by the International Life Sciences Institute Research Foundation in December 2007, for their useful feedback on a preliminary version of the manuscript. We also thank members of Chalmers Research Group for their feedback on Table 1.

The authors' responsibilities were as follows—DM: conceptualized the study and wrote the manuscript; ACT: helped write the manuscript and edited it. Neither author had a personal or financial conflict of interest.

REFERENCES

- Chalmers I, Hedges LV, Cooper H. A brief history of research synthesis. *Eval Health Prof* 2002 March;25(1):12–37.
- Glass G. Primary, secondary, and meta-analysis of research. *Educational Research* 1976;10:3–8.
- Cochrane AL. Effectiveness and efficiency. Random reflections on health services. London, United Kingdom: Nuffield Provincial Hospitals Trust, 1972.
- Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC. A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts. Treatments for myocardial infarction. *JAMA* 1992;268(2):240–8.
- Evidence-based medicine. A new approach to teaching the practice of medicine. *JAMA* 1992;268(17):2420–5.
- Lavis J, Davies H, Oxman A, Denis JL, Golden-Biddle K, Ferlie E. Towards systematic reviews that inform health care management and policy-making. *J Health Serv Res Policy* 2005;10(suppl):35–48.
- Cook DJ, Greengold NL, Ellrodt AG, Weingarten SR. The relation between systematic reviews and practice guidelines. *Ann Intern Med* 1997;127(3):210–6.
- Lavis JN, Posada FB, Haines A, Osei E. Use of research to inform public policymaking. *Lancet* 2004;364(9445):1615–21.
- Lavis JN, Davies HTO, Gruen RL, Walshe K, Farquhar CM. Working within and beyond the Cochrane Collaboration to make systematic reviews more useful to healthcare managers and policymakers. *Healthcare Policy* 2006;1(2):21–33.
- Moher D, Tetzlaff J, Tricco AC, Sampson M, Altman DG. Epidemiology and reporting characteristics of systematic reviews. *PLoS Med* 2007;4(3):e78.
- Chalmers TC. Problems induced by meta-analyses. *Stat Med* 1991;10(6):971–9.
- Petticrew M. Systematic reviews from astronomy to zoology: myths and misconceptions. *BMJ* 2001;322(7278):98–101.
- Higgins JPT, Green S, eds. *Cochrane handbook for systematic reviews of interventions*. Chichester, United Kingdom: Wiley-Blackwell, 2008.
- Allen IE, Olkin I. Estimating time to conduct a meta-analysis from number of citations retrieved. *JAMA* 1999;282(7):634–5.
- Bennett T, Holloway K, Farrington DP. Does neighborhood watch reduce crime? A systematic review and meta-analysis. *J Exp Criminol* 2006;2(4):437–58.
- Nye C, Turner H, Schwartz J. Approaches to parent involvement for improving the academic performance of elementary school age children. Campbell Collaboration, Campbell Library of Systematic Reviews. Internet: http://www.campbellcollaboration.org/campbell_library/index.php (accessed 17 Mar 2008).
- Chaudhary N, Krieger N. Community-based nutrition and physical activity interventions for chronic disease prevention aimed at low-income populations: a systematic review. *Can J Dietet Pract Res* 2007;68(4):201–6.
- Hooper L, Bartlett C, Davey SG, Ebrahim S. Systematic review of long term effects of advice to reduce dietary salt in adults. *BMJ* 2002;325(7365):628.
- de Munter JS, Hu FB, Spiegelman D, Franz M, van Dam RM. Whole grain, bran, and germ intake and risk of type 2 diabetes: a prospective cohort study and systematic review. *PLoS Med* 2007;4(8):e261.
- Farrington DP, Gill M, Waples S, Argomaniz J. The effects of closed-circuit television on crime. *J Exp Criminol* 2007;3(1):21–38.
- Lum C, Kennedy LW, Sherley AJ. The effectiveness of counter-terrorism strategies. Campbell Collaboration, Campbell Library of Systematic Reviews. Internet: http://www.campbellcollaboration.org/campbell_library/index.php (accessed 17 Mar 2008).
- Zief SG, Lauver S, Maynard RA. Impacts of after-school programs on student outcomes: a systematic review for the Campbell Collaboration. Campbell Collaboration, Campbell Library of Systematic Reviews. Internet: http://www.campbellcollaboration.org/campbell_library/index.php (accessed 17 Mar 2008).
- Ritter G, Denny G, Albin G, Barnett J, Blankenship B. The effectiveness of volunteer tutoring programs: a systematic review. Campbell Collaboration, Campbell Library of Systematic Reviews. Internet: http://www.campbellcollaboration.org/campbell_library/index.php (accessed 17 Mar 2008).
- Turner W, MacDonald GM, Dennis JA. Behavioural and cognitive behavioural training interventions for assisting foster careers in the management of difficult behaviour. Campbell Collaboration, Campbell Library of Systematic Reviews. Internet: http://www.campbellcollaboration.org/campbell_library/index.php (accessed 17 Mar 2008).
- Smedslund G, Dalsbo TK, Steiro AK, Winsvold A, Clench-Aas J. Cognitive behavioural therapy for men who physically abuse their female partner. Campbell Collaboration, Campbell Library of Systematic Reviews. Internet: http://www.campbellcollaboration.org/campbell_library/index.php (accessed 17 Mar 2008).
- Ekeland E, Heian F, Hagen KB, Abbott J, Nordheim L. Exercise to improve self-esteem in children and young people. Campbell Collaboration, Campbell Library of Systematic Reviews. Internet: http://www.campbellcollaboration.org/campbell_library/index.php (accessed 17 Mar 2008).
- Stone PW. Popping the (PICO) question in research and evidence-based practice. *Appl Nurs Res* 2002;15(3):197–8.
- Funai EF, Rosenbush EJ, Lee MJ, Del PG. Distribution of study designs in four major US journals of obstetrics and gynecology. *Gynecol Obstet Invest* 2001;51(1):8–11.
- Scales CD Jr, Norris RD, Peterson BL, Preminger GM, Dahm P. Clinical research and statistical methods in the urology literature. *J Urol* 2005;174(4 Pt 1):1374–9.
- Atkins D, Fink K, Slutsky J. Better information for better health care: the Evidence-based Practice Center program and the Agency for Healthcare Research and Quality. *Ann Intern Med* 2005;142(12 Pt 2):1035–41.
- Helfand M, Morton S, Guallar E, Mulrow C, eds. Challenges of summarizing better information for better health: the evidence-based practice center experience. *Ann Intern Med* 2005;142:1032–126, W-262.



32. Systems to rate the strength of scientific evidence. Rockville, MD: Agency for Healthcare Research and Quality, 2008. (Evidence Report/Technology Assessment no. 47.)
33. Garrity C, Tricco A, Tsertsvadze A, Sampson M, Moher D. Updating systematic reviews: an international survey. *Ger J Evidence Qual Health Care* 2008;102(suppl):78.
34. Vivekananthan DP, Penn MS, Sapp SK, Hsu A, Topol EJ. Use of antioxidant vitamins for the prevention of cardiovascular disease: meta-analysis of randomised trials. *Lancet* 2003;361(9374):2017–23.
35. Eidelman RS, Hollar D, Hebert PR, Lamas GA, Hennekens CH. Randomized trials of vitamin E in the treatment and prevention of cardiovascular disease. *Arch Intern Med* 2004;164(14):1552–6.
36. Shekelle PG, Morton SC, Jungvig LK, et al. Effect of supplemental vitamin E for the prevention and treatment of cardiovascular disease. *J Gen Intern Med* 2004;19(4):380–9.
37. Miller ER III, Pastor-Barriuso R, Dalal D, Riemersma RA, Appel LJ, Guallar E. Meta-analysis: high-dosage vitamin E supplementation may increase all-cause mortality. *Ann Intern Med* 2005;142(1):37–46.
38. Jadad AR, Cook DJ, Browman GP. A guide to interpreting discordant systematic reviews. *Can Med Assoc J* 1997;156(10):1411–6.
39. Oxman AD, Guyatt GH. Validation of an index of the quality of review articles. *J Clin Epidemiol* 1991;44(11):1271–8.
40. Plint AC, Moher D, Morrison A, et al. Does the CONSORT checklist improve the quality of reports of randomised controlled trials? A systematic review. *Med J Aust* 2006;185(5):263–7.
41. Simera A, Altman DG, Moher D, Schulz KF, Hoey J. Guidelines for reporting health research: the EQUATOR Network's survey of guideline authors. *PLoS Med* 2008;5(6):e139.
42. Mulrow CD. The medical review article: state of the science. *Ann Intern Med* 1987;106(3):485–8.
43. Sacks HS, Berrier J, Reitman D, Ancona-Berk VA, Chalmers TC. Meta-analyses of randomized controlled trials. *N Engl J Med* 1987;316(8):450–5.
44. Sacks HS, Reitman D, Pagano D, Kupelnick B. Meta-analysis: an update. *Mt Sinai J Med* 1996;63(3-4):216–24.
45. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. *Lancet* 1999;354(9193):1896–900.
46. Gibson RS, Sazawal S, Peerson JM. Design and quality control issues related to dietary assessment, randomized clinical trials and meta-analysis of field-based studies in developing countries. *J Nutr* 2003;133(suppl):1569S–73S.
47. Balk EM, Horsley TA, Newberry SJ, et al. A collaborative effort to apply the evidence-based review process to the field of nutrition: challenges, benefits, and lessons learned. *Am J Clin Nutr* 2007;85(6):1448–56.
48. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;273(5):408–12.

