ANNUAL REVIEWS

Annual Review of Nutrition

Evidence Collection and Evaluation for the Development of Dietary Guidelines and Public Policy on Nutrition

Dena Zeraatkar,[1] Bradley C. Johnston,[1,2] and Gordon Guyatt[1,3]

[1]Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Ontario L8S 4K1, Canada; email: zeraatd@mcmaster.ca

[2]Department of Community Health and Epidemiology, Faculty of Medicine, Dalhousie University, Halifax, Nova Scotia B3H 4R2, Canada; email: bjohnston@dal.ca

[3]Department of Medicine, McMaster University, Hamilton, Ontario L8S 4K1, Canada; email: guyatt@mcmaster.ca

Annu. Rev. Nutr. 2019. 39:227–47

The *Annual Review of Nutrition* is online at nutr.annualreviews.org

https://doi.org/10.1146/annurev-nutr-082018-124610

Keywords

nutrition public policy, dietary guidelines, GRADE, consumer engagement, conflicts of interest, risk of bias

Abstract

Dietary guidelines and recommendations, usually developed by government bodies or large authoritative organizations, have major downstream effects on public policy. A growing body of evidence supports the notion that there are serious deficiencies in the methods used to develop dietary guidelines. Such deficiencies include the failure to access or conduct comprehensive systematic reviews, a lack of systematic or rigorous evaluation of the quality of the evidence, a failure to acknowledge the limitations of the evidence base underlying recommendations, and insufficiently stringent management of conflicts of interest. These issues may be addressed by adhering to international standards for guideline development, including adopting systematic review methodology and using rigorous systems to evaluate the certainty of the evidence and to move from evidence to recommendations, of which the GRADE approach (Grading of Recommendations Assessment,

Development and Evaluation) is the most rigorous and fully developed. Improving the methods by which dietary guidelines are produced has considerable potential to substantially improve public policy decision-making.

Contents

INTRODUCTION

Decision-making about public policy on nutrition should be based on the best evidence and be aligned with public values and preferences. One of the primary ways in which public policy decision-making about nutrition occurs is through the development of dietary guidelines and recommendations by government bodies and authoritative organizations. These guidelines have major downstream effects and are used to inform policies on agriculture, food assistance programs, and nutrition in schools, prisons, hospitals, and nursing homes, in addition to influencing recommendations from health-care professionals and health messaging in the media (100). Evidence indicates that there are inconsistencies and limitations in the methods used to develop dietary guidelines, with subsequent adverse impacts on nutrition public policy (16, 27, 65, 87, 100). This paper addresses issues and challenges in developing dietary guidelines and discusses the implications of these issues for public policy on nutrition. We begin by describing how governments and organizations develop and use dietary guidelines to make public policy decisions about nutrition. We then discuss important issues in developing dietary guidelines, including the merits and limitations of the most common types of evidence used to inform dietary guidelines; issues related to evaluating the certainty of evidence and moving from evidence to recommendations; additional considerations such as equity, environmental impact, and feasibility that may need to be examined when developing recommendations; the engagement of consumers and stakeholders;

and the management of conflicts of interest. Finally, we describe how limitations in dietary guidelines and recommendations can lead to ineffective and harmful public policy on nutrition.

DGA: Dietary Guidelines for Americans

DGAC: Dietary Guidelines Advisory Committee

WIC: Women, Infants, and Children program

IARC: International Agency for Research on Cancer

HOW DO GOVERNMENTS AND ORGANIZATIONS DEVELOP DIETARY GUIDELINES AND TO WHAT EXTENT DO GUIDELINES AFFECT PUBLIC POLICY ON NUTRITION?

In most settings, dietary guidelines are produced by governments and authoritative organizations. These guidelines subsequently have major downstream policy effects (100). In the United States, legislation mandates that the United States Department of Agriculture (USDA) and the Department of Health and Human Services (DHHS) publish a new version of dietary guidelines every 5 years (Pub. L. No. 101-445, 101 U.S.C.). The most recent version of the guidelines was published in 2015 (109). The development of the United States Dietary Guidelines for Americans (DGA) is done in three stages: (*i*) review of the evidence; (*ii*) development of recommendations; and (*iii*) implementation of the guidelines. In the first stage, nominations from the public inform the selection of a Dietary Guidelines Advisory Committee (DGAC). The offices of the USDA and DHHS review the nominations and select the DGAC members. The 15 members of the 2015 DGAC were prominent researchers in the fields of nutrition, health, and medicine. The role of the DGAC is to provide advice and recommendations to the federal government regarding the current state of scientific evidence on nutrition and health. The committee formulates research questions and either uses existing reviews and reports or commissions additional reviews to address the questions. The evidence compiled by the DGAC is used in the second stage by the DHHS and USDA to develop recommendations. These recommendations are developed independently of the DGAC. A consequence of this is that recommendations may be more political and may not be completely aligned with the conclusions drawn in the DGAC's report. The public is subsequently invited to review and comment on drafts of the DGAC's report and the recommendations. The DGA and DGAC's report are revised based on feedback and, in the third stage, the recommendations are implemented through policies and programs and the distribution of nutrition education materials for the public. The DGA is endorsed by many government programs, including, among others, MyPlate (a nutrition guide depicting recommendations for the ratio in which different food groups should be consumed), the Women, Infants, and Children (WIC) program (a federal nutrition assistance program for pregnant women and mothers of young children that offers vouchers that can be used to purchase preapproved foods), and the National School Lunch Program (a federally assisted meal program operating in public schools).

Other countries typically follow a similar process to develop dietary recommendations. Unlike in the United States, the periodic review of research evidence and revision of dietary guidelines are not legislated in most other countries (30). For example, Canada's national dietary guidelines, *Eating Well with Canada's Food Guide*, are updated only if there is an identified need to revise guidelines (23, 40). In almost all countries, recommendations are formulated through a consensus process led by working groups or committees after a review of the evidence, which is completed with variable rigor (16). Unlike the United States, most countries do not develop de novo reviews of the evidence, instead relying on published reviews and reports (16). A minority of countries involve consumers through consultations or workshops to further discuss recommendations (16). Irrespective of the process, dietary guidelines have downstream effects on government policies and programs and on health messaging (100). However, the extent of the impact of these guidelines on various policies can vary based on the setting.

The World Health Organization's International Agency for Research on Cancer (IARC) is an example of an authoritative organization that provides guidance for governments in making

decisions about public policy. The mandate of the IARC is to provide government authorities with scientific opinions on carcinogens. It does this by publishing monographs on the association between various agents and cancers. While the IARC does not specifically focus on diet, it has investigated dietary factors (45, 46, 63), and its most striking and controversial opinion was its classification in 2015 of red meat as probably carcinogenic and processed meat as carcinogenic (33, 110).

To lead the development of its monographs, the IARC assembles a working group composed of scientific experts, who are reportedly without conflicts of interest, to provide an overview of the relevant research evidence. After reviewing the evidence, the working group meets and makes decisions through a consensus-based approach. A summary of the decision and process is published shortly after the meeting, and the full monograph is usually published within 6 months. The monograph on red meat and processed meat, however, remained unpublished until 3 years after the meeting (46).

Observational studies: studies that do not involve the random allocation of participants to interventions by an investigator

Systematic review: a search for and summary of the evidence that addresses a research question using a transparent and reproducible method

Randomized controlled trials (RCTs): RCTs assign participants at random to alternative interventions and thus produce balance between trial arms in both known and unknown prognostic factors

WHAT EVIDENCE IS USED IN DEVELOPING DIETARY GUIDELINES?

Evidence informing dietary guidelines and recommendations most often comes from randomized trials and observational studies that attempt to establish causal relationships between dietary exposures and health outcomes. Guideline developers may also consider evidence from experimental animal and mechanistic studies, as does the IARC. However, this evidence is severely limited in establishing causal relationships between dietary exposures and health outcomes and, hence, is usually used only to support evidence from randomized trials and observational studies. We now look in detail at the merits and limitations of randomized trials and observational studies and describe considerations in assessing their overall quality and ability to provide evidence supporting a causal relationship between nutritional exposures and health outcomes. We also review the importance of basing dietary guidelines and recommendations and public policy decisions on systematic reviews of the evidence.

Randomized Controlled Trials

Randomized controlled trials (RCTs) assign participants at random to alternative interventions and thus produce balance between trial arms in both known and unknown prognostic factors. Because they safeguard against potential imbalances in prognostic factors, RCTs have the potential to provide higher-certainty evidence of causation. Unfortunately, however, conducting rigorous RCTs with high applicability in nutrition poses enormous challenges (41, 91, 105). Trials of nutritional supplements typically pose the same challenges as trials of most pharmacological interventions. Hence, in this article we primarily focus on trials of dietary interventions.

Typically, rigorous dietary trials must possess the same characteristics as rigorous trials in other fields. To ensure prognostic balance between arms, all RCTs, including dietary RCTs, must conceal randomization—that is, prior to randomization, the person enrolling participants must be unaware of the arm to which each participant will be allocated. This is best operationalized by using central allocation through a computer or telephone. However, dietary trials are also subject to additional challenges in controlling bias, particularly in relation to blinding, which is achieved when study participants, clinicians, data collectors, and those adjudicating outcomes are all unaware of the arm to which participants are assigned in the trial. Blinding safeguards against vulnerability to placebo effects (those assigned to an intervention may do better not because of the biological effects of the intervention, but because of their belief in its merits), cointerventions (those exposed to the target intervention receive other beneficial treatments that control groups do not or vice versa), and bias in the measurement and adjudication of outcomes (personnel measuring and adjudicating

outcomes may favor one intervention over another). Blinding of study participants in dietary RCTs is, of course, not possible. People will always be aware of what they are eating. Blinding of those collecting health outcome data and adjudicating outcomes is, however, almost always possible. By contrast, and similar to most pharmacological trials, blinding of study participants, data collectors, and outcome adjudicators is almost always possible for trials of nutritional supplements.

We may be misled not only by bias in dietary trials but also by random error resulting from small sample sizes and small numbers of outcome events. This is a particular problem in studies of nutritional exposures in which effect sizes may be smaller than those of pharmacological or surgical interventions, thus increasing the number of participants required (99). Furthermore, one would anticipate that the effects of most nutritional exposures on important health outcomes will occur only after prolonged exposure, necessitating studies of long duration to ensure adequate exposure. Such long trials impose a significant burden on participants, are costly to conduct, and are subject to another serious source of bias: the loss of participants to follow-up (i.e., missing outcome data) (55). To address the need for long follow-up, RCTs may choose to focus on surrogate outcomes (see, e.g., 28). Unfortunately, there are a myriad of examples of interventions with apparently salutary effects on surrogates but no effect—or even detrimental effects—on outcomes important to patients and the public (11, 20). Hence, dietary RCTs focusing on surrogate outcomes usually provide only poor-quality evidence on outcomes of critical importance, such as mortality and major morbidity.

A final challenge for dietary RCTs is that participants often do not adhere closely to the diet to which they are assigned, and adherence may diminish over time (91, 101). This reduces differences in diet between trial arms, further reducing the size of a possible effect and increasing the number of participants required and, to the extent that one is interested in a comparison of the diets as prescribed, reducing the applicability of the results. While a lack of compliance with the intervention may be a sign that the intervention is not tolerable, it does not necessarily mean that the intervention will be completely ineffective if implemented. Although participants may not adhere to a dietary pattern within a trial, they may adhere to the dietary pattern if widely promoted by physicians, governments, the food industry, and the media, as was the case with the low-fat diet that became widespread in much of the latter half of the twentieth century (64). Additionally, even low levels of adherence across many individuals can have a large impact on health outcomes across a population.

Despite these formidable obstacles, there is reason to persist with the conduct of dietary RCTs, the most compelling of which may be (as we will show) that observational studies are unlikely to provide high-quality evidence regarding dietary impact on outcomes important to patients. Trialists may be able to ultimately offer high-quality evidence regarding the health effects of various dietary patterns if they ensure concealed randomization and blind adjudication of outcomes, implement strategies to optimize adherence to assigned diets, minimize loss to follow-up, and enroll large numbers of participants and follow them for a long enough duration so that it can be reasonably expected that the outcome will be impacted by the intervention (usually a decade or longer for outcomes important to the population, such as all-cause mortality and cardiovascular and cancer events). Although the cost of such RCTs will be high, one might argue—as some have—that they will nevertheless prove cost effective in comparison to the hundreds of conflicting and misleading observational studies published every year (48, 105).

Observational Studies

Unlike RCTs, observational studies provide evidence on the health effects of dietary exposures when they are self-selected. Hence, adherence, which is a major difficulty with conducting RCTs

in nutrition, is not an issue. Unfortunately, observational studies have other problems, which we enumerate, making strong causal inferences unlikely.

FFQ: food frequency questionnaire

NHANES: National Health and Nutrition Examination Survey

Confounding: confounding occurs when the estimated effect of an exposure on an outcome is biased due to differences in risk factors between exposure groups

The first limitation of observational studies in nutrition, and perhaps the most important, is dietary measurement. Typically, nutritional epidemiology studies measure diet by using a food frequency questionnaire (FFQ). FFQs query participants' usual frequency and quantity of consumption of a list of foods and beverages. Memory-based dietary assessments are prone to both random and systematic error because people are not good at recalling or quantifying their diet (5). For example, the National Health and Nutrition Examination Survey (NHANES) used the 24-hour recall method to measure diet, which requires participants to recall their diets during a 24-hour period. However, two out of three participants in the NHANES reported an amount of energy intake incompatible with life (5, 47). Authors of such observational studies often argue that their memory-based dietary assessments, be they based on FFQs or 24-hour recall, have been adequately validated. Typically, however, these memory-based dietary assessments are tested against other methods that also have important limitations. For example, FFQs are often validated against 24-hour recall, which is also prone to error, as the NHANES experience illustrates (see, e.g., 17, 98). Furthermore, methods for validation are diverse, and no consistent criteria are used in the literature to assess whether dietary assessment methods are adequately valid and reliable, and investigators have yet to reach a consensus on a gold standard for dietary measurement. An additional issue in dietary measurement is its timing. Even if dietary measurement is accurate at one point in time, if measured years before the onset of the disease under study (as is often the case), the measure may not capture participants' eating patterns through the key period of exposure (see, e.g., 56, 58, 102). Thus, without periodic dietary assessment, changes in diet over time may inaccurately depict diet–disease associations.

The next threat to trustworthy causal inference from observational studies is one that is not at all specific to the field of nutrition. RCTs provide stronger evidence than observational studies because if they are sufficiently large, they ensure that prognostic factors such as age, sex, and comorbidities—whether known or unknown, measured or unmeasured—are similarly distributed between trial arms. Observational studies offer no such assurance. Indeed, when, as in dietary observational studies, participants' preferences or food availability and access are the primary determinants of exposure, prognosis will almost certainly differ substantially among groups, a situation we refer to as confounding. Confounding occurs when the estimated effect of an exposure on an outcome is biased due to differences in risk factors of the outcome between exposure groups. Investigators deal with this issue by conducting adjusted analyses, primarily through regression models, which, in principle, create prognostically homogeneous groups (1). The problem that remains, however, is that adjusted analyses will yield unbiased results only if all important prognostic factors are known and measured accurately; when that is not the case, prognostic imbalance will persist. The risk of disease associated with a particular dietary exposure may also be influenced by the presence of other risk factors, many of which might be unknown and, hence, uncontrolled in studies of free-living populations.

Were residual confounding not a problem, RCTs and optimally conducted observational studies would yield similar results, and some have argued that in most instances, at least in nutrition, they do (4, 91). That may be the case, but the problem remains that in any instance, prior to evidence from RCTs, we do not know whether we are dealing with one of the majority of situations in which results from an RCT will confirm those of observational studies or the minority in which they will not. Indeed, many examples exist of the failure of RCTs to replicate the results of observational studies (38, 60, 111, 112).

A final problem with evidence from observational studies is the potential for selective reporting and publication bias. Publication bias in nutrition is likely common, primarily due to

the overabundance of data that can be used to correlate many exposures with many outcomes and so dredge the data for interesting findings (47). Safeguards against such practices, such as registration of observational studies, do not exist. Even when there are published protocols of observational studies, authors rarely prespecify exposure–outcome relationships of interest.

Meta-analysis: a statistical technique to pool results from separate studies investigating the same or similar research questions; it provides a quantitative summary across studies

Despite the myriad of limitations of observational studies, we may be more certain of their findings in three uncommon situations (36). First, when observed effect sizes are large, we can be more confident that the effect can be attributed to the intervention or exposure of interest. In such cases, residual confounding is less likely to explain the association. However, if the large effect size comes from a study with methodological limitations (i.e., those that do not comprehensively or accurately and periodically measure prognostic factors and the exposure of interest, accurately measure the target outcomes, minimize missing outcome data, and conduct appropriate adjusted analyses according to a prespecified and publicly accessible protocol), the large effect size may represent a large overestimate. Typically, our confidence in the findings of observational studies is increased when the effect estimate on the relative risk scale is less than 0.5 or greater than 2 (36).

Second, the presence of a dose–response gradient has long been recognized as an important criterion for supporting a causal relationship (93). Although dose–response relationships may increase our certainty about the findings of observational studies, they will mislead us when dose-dependent confounding exists. For instance, an apparent dose–response relationship between an exposure of interest (e.g., processed meat) and a health outcome (e.g., mortality from cardiovascular disease) may be confounded with another variable that is also correlated with the exposure (e.g., salt). Due to dense correlation networks among nutritional exposures, dose–response relationships may strengthen causal inferences to a lesser extent in nutrition than in in other fields.

Finally, occasionally all plausible confounders and biases from observational studies will result in an underestimation of an apparent effect. This situation is unlikely to occur in nutritional epidemiology because all plausible confounders and their directions of bias are only rarely understood with a high degree of confidence (83). Since we almost never see strong associations in nutritional studies, dose–response relationships provide only modest additional reassurance, and because we are rarely aware of all of the important confounding factors with certainty, seldom will we be able to make causal inferences from observational studies with high certainty.

Unfortunately, dietary guidelines have placed too heavy of an emphasis on the findings of observational studies without appropriately acknowledging their limitations, as we will show (39, 100). While observational evidence may also be supported by other lines of evidence, such as in vitro biochemical studies and animal experiments, the certainty of evidence from such sources is also severely limited and cannot be used to surmount the inherent limitations of observational studies (49, 84, 85).

Systematic Reviews

Standards for developing trustworthy guideline recommendations suggest the need for comprehensive systematic reviews of the evidence (7, 19, 54, 80). A systematic review is a search for and summary of the evidence that addresses a research question using a transparent and reproducible method. It is often, but not always, accompanied by a meta-analysis, which is a statistical method for pooling results from different studies to provide a single effect estimate for each outcome of interest. Systematic reviews have several advantages compared with single studies. Single studies, as well as reviews of the evidence without a systematic search and objective eligibility criteria, are liable to be unrepresentative of the total body of evidence and, thus, can be misleading. Also, single studies may not be adequately powered to detect important effects. This is especially true in

CRITERIA FOR RIGOROUS SYSTEMATIC REVIEWS AND META-ANALYSES

Rigorous systematic reviews and meta-analyses should:

- explicitly address a sensible research question
- conduct an exhaustive and reproducible search to select relevant studies
- assess the risk of bias using appropriate criteria
- use appropriate statistical methods to pool results from primary studies
- conduct an evaluation of the overall quality of evidence

nutrition, where effect sizes are usually small (99). Compared with single studies, systematic reviews include a greater range of study participants and so have greater external validity. The sidebar Criteria for Rigorous Systematic Reviews and Meta-Analyses presents the relevant criteria (74).

Unfortunately, many dietary recommendations are not based on systematic reviews or are based on systematic reviews that are not methodologically rigorous. In a review of guidelines on sugar intake, Erickson and colleagues (27) found that in four of nine guidelines, systematic searches for relevant studies to inform recommendations had not been conducted. Blake and colleagues (16) found that less than 20% of national dietary guidelines were based on systematic reviews. The US National Academies of Sciences, Engineering, and Medicine (NAS) has published two reports critiquing the process used to develop the DGA (77, 78). The reports identified important deficiencies in the systematic review processes and called for increased adherence to standardized processes and methods for evidence synthesis. The 2015 DGAC did not use systematic reviews for approximately half of its research questions, instead choosing to rely on unsystematic reports prepared by third-party organizations and ad hoc examinations of the evidence by experts (25). For example, the DGAC based recommendations for low-carbohydrate diets on what they referred to as "exploratory searches" rather than systematic reviews (25, p. 186). Furthermore, the DGAC undertook very few meta-analyses, instead choosing to use what they referred to as "qualitative synthesis" in order to identify "Key Trends" in the evidence (25, p. 35). Such an approach does not consider the magnitude of effect (important because very small effects may not warrant changes in dietary behavior), the precision of estimates, or the extent of inconsistency in the results of primary studies. Both imprecision and inconsistency represent important considerations in determining the certainty of evidence (37).

Although the IARC states that it conducts a review of the evidence, details of the search strategy and eligibility criteria are not reported in its monographs, raising concerns that rather than conducting a systematic search, the working group identifies relevant studies based on their experience and prior knowledge of the field. The apparent failure to conduct a comprehensive review may explain why the IARC report on red and processed meats did not include two key studies: the Women's Health Initiative Dietary Modification (WHI-DM) trial and the Polyp Prevention Trial (14, 92). The WHI-DM trial randomized nearly 49,000 participants to either a low-fat diet or their usual diet. The low-fat dietary intervention arm reduced the intake of red meat by 20% (14). The Polyp Prevention Trial tested a similar low-fat, high-fiber dietary intervention in more than 2,000 participants and found a statistically significant decreased intake of red and processed meats in the intervention group (92). Furthermore, although the IARC states that it critically reviews the evidence, it does not provide explicit judgments regarding the risk of bias in each study (46). Without an appropriate assessment of the risk of bias, conflicting results from studies of varying quality may be inappropriately given equal weight.

NAS: US National Academies of Sciences, Engineering, and Medicine

HOW IS THE CERTAINTY OF EVIDENCE EVALUATED TO INFORM DIETARY GUIDELINES?

GRADE: Grading of Recommendations Assessment, Development and Evaluation

To move from evidence to recommendations, decision-makers must consider criteria regarding the overall certainty of the body of evidence in addition to the magnitude of desirable and undesirable consequences and the values and preferences of the target population (37). The expression values and preferences refers to the importance people place on each of the outcomes associated with a decision, and it is crucial to balancing desirable and undesirable consequences. There are various systems for rating the certainty of evidence and moving from evidence to recommendations (see, e.g., 25, 26, 37, 43). Of the systems available, the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach is by far the most rigorously developed and widely used (37). GRADE was developed to address the limitations of previous systems for rating the certainty of evidence (8, 9). Unlike many of its predecessors, GRADE yields clear and understandable recommendations and is simple to apply (having undergone extensive user testing) and, hence, is the focus of this review (8, 9, 21, 62, 76, 90). The GRADE system uses four categories to rate the certainty of evidence: high, moderate, low, or very low (see the sidebar Certainty of Evidence and Definitions). Lower levels of certainty indicate a greater likelihood that the underlying true effect is substantially different from the observed effect.

Rating the certainty of evidence begins by considering study design. A body of evidence based on RCTs is initially assigned high certainty, and observational studies are assigned low certainty due to concerns about confounding. The certainty may be reduced if reviewers note limitations regarding the risk of bias, inconsistency, imprecision, indirectness, or publication bias (12). For each criterion, certainty may be reduced by one or two levels depending on the seriousness of concerns. An increase in certainty may occur when observational studies produce large effect sizes or a dose–response gradient or when all plausible confounders or biases would reduce a demonstrated effect or suggest a spurious effect when results show no effect. As we described in the previous section, it is rare to have high-certainty evidence from observational studies of nutrition because effect sizes are usually small; dose–response gradients are not as reassuring against confounding bias as they are in other fields; and we seldom ever have knowledge of all plausible confounders.

According to the GRADE system, recommendations can be either strong or weak. The direction and strength of a recommendation is determined by the balance between desirable and undesirable consequences and our certainty about the effect estimates (37). The balance between desirable and undesirable consequences is, in turn, determined by the magnitude of effect estimates for each outcome of interest and the values and preferences of the target population. The larger the gradient between the desirable and undesirable effects—based on the values and preferences of the target population—the higher the likelihood that a strong recommendation is warranted. Another determinant of the strength of recommendations is our certainty of the effect estimates. Typically, a strong recommendation is associated with high, or at least moderate, certainty about

CERTAINTY OF EVIDENCE AND DEFINITIONS

- **High certainty**: The true effect is most likely very close to the effect estimate.
- **Moderate certainty**: The true effect is likely to be close to the effect estimate, but there is a possibility that it is different.
- **Low certainty**: The true effect may be substantially different from the effect estimate.
- **Very low certainty**: The true effect is likely to be substantially different from the effect estimate.

the effect estimates for critical outcomes. The more closely balanced the trade-offs between desirable and undesirable outcomes, the more likely that low-certainty evidence underlying any critical outcome will result in a weak recommendation. A final concern that may decrease the strength of recommendations is uncertainty about or variability in the values and preferences of the target population. Uncertainty in values and preferences means that we do not confidently know whether undesirable outcomes are outweighed by desirable outcomes, according to the perspective of the target population. Variability in values and preferences means that a single recommendation would not apply uniformly across the target population. In such situations, weak recommendations are warranted. In general, making strong recommendations based on low-certainty evidence is discouraged, apart from a few paradigmatic situations (3). An alternative way of defining strong recommendations is that they indicate high- or at least moderate-certainty evidence that the benefits of a course of action clearly outweigh harms, or vice versa, according to the majority of the target population, whereas weak recommendations indicate that we are uncertain whether benefits are clearly outweighed by harms, or vice versa, for the majority of the target population.

The GRADE Evidence to Decision frameworks were developed to facilitate the process of moving from evidence to recommendations (2). They build on the GRADE method and provide a structured process through which decision-makers can be informed about the relative benefits and harms of the options being considered, provide decision-makers with a rating of the certainty of the evidence about benefits and harms, and ensure that all important factors are considered in the decision-making process. The Evidence to Decision frameworks can also support stakeholders in understanding the judgments made by decision-makers and the evidence supporting those judgments.

It must be noted that there is by necessity a considerable amount of subjectivity involved in rating the certainty of evidence and grading recommendations and that two persons evaluating the same body of evidence might reasonably come to different conclusions (75). However, the advantage of GRADE is that it provides a transparent framework through which these subjective judgments can be documented and communicated to stakeholders.

The GRADE system has been criticized by nutrition researchers for rating evidence from observational studies as having low certainty (see, e.g., 95). Such criticism is usually centered on the challenges of conducting RCTs in nutrition, challenges which our prior discussion has made clear. Picture, however, two bodies of observational studies that are identical in the rating of their certainty of evidence, but one addresses a question for which RCTs are feasible and the other does not. Should a higher rating of certainty in the context of unfeasible RCTs be provided (in other words, does the unfeasibility of RCTs increase one's certainty in the evidence)? We would argue that it should not.

There are compelling reasons for maintaining similar criteria for evaluating certainty across disciplines. Clinicians, patients, and other policy-makers must weigh the benefits and harms of many different kinds of interventions—including pharmacological, surgical, and nutritional—for a clinical or public health problem. For example, to choose between bariatric surgery or lifestyle interventions for obesity, patients and clinicians must weigh the potential benefits and harms associated with a surgical intervention and a lifestyle intervention. Using modified criteria to assess the certainty of evidence for some interventions and not others would be misleading. We have highlighted the formidable limitations of both RCTs and observational studies in nutrition. Acknowledging these limitations is preferable to pretending they do not exist.

Some have suggested that in situations in which the certainty of evidence is low or very low, guideline panels should refrain from making recommendations (18). This situation is common in nutrition. However, ultimately the public, clinicians, and policy-makers will be faced with situations in which they must make decisions. Refraining from making a recommendation is not

helpful to stakeholders. In such situations, making a weak recommendation or a recommendation to consider the values and preferences of the target population when deciding between alternative courses of action is preferable to not making a recommendation.

Turning our attention to how the certainty of evidence is evaluated and recommendations are formulated in existing dietary guidelines, there is great variability in whether decision-making bodies rate the certainty of the evidence and the methods by which they do so. In one investigation, only 5 of 32 national dietary guidelines committees explicitly evaluated the certainty of the evidence underpinning their recommendations (16). Among those committees that did, the appropriateness of the criteria used was questionable. For example, the technical advisory group for the development of New Zealand's national dietary guidelines evaluates the certainty of evidence through discussion using unspecified criteria (16). The DGAC used the USDA Nutrition Evidence Library system to rate the certainty of evidence as strong, moderate, limited, or not assignable, based on the risk of bias, numbers of studies and study participants, consistency across studies, impact, and generalizability, criteria that, although they appear reasonable and have close correspondence to GRADE criteria, still have limitations (25). For example, the criterion "impact" considers the magnitude of effect and the directness with which studies evaluate the association between the exposure and the outcomes of interest: two seemingly unrelated concepts. If the effect size is small and not clinically important or the evidence does not address the outcomes of interest, the certainty of the evidence is rated as limited. If the effect size is large with evidence directly addressing the outcomes of interest, the certainty of evidence is rated as strong, given that there are no limitations in the other domains. The framework does not, however, consider the situation in which there is high-certainty evidence that an effect is very small and of marginal importance. Moreover, the DGAC does not present explicit judgments for each of the five Nutrition Evidence Library criteria, instead presenting only the overall rating, the validity of which remains in question due to the lack of transparency in the ratings for each of the five criteria.

In their reports critiquing the development of the DGA, NAS also identified a lack of transparency in how evidence is translated into recommendations and important inconsistencies between the 2015 DGAC report and the subsequent recommendations (77, 78). Notable inconsistencies include the 2015 DGA stance on healthy dietary patterns and sodium restriction. Despite the limitations of the evidence regarding the health effects of vegetarian dietary patterns highlighted in the DGAC report, the guidelines recommend a vegetarian dietary pattern as one of three healthy dietary patterns for Americans (109). The DGA also recommends limiting daily sodium intake (to <2,300 mg) for all adults (109). The DGAC report, however, notes that evidence regarding the effects of sodium on cardiovascular outcomes is inconsistent and insufficient to conclude that lowering sodium intake to below 2,300 milligrams per day either increases or decreases the risk of cardiovascular events in the general population (25). Such inconsistencies between the evidence and the recommendations raise concerns regarding the extent to which the DGA recommendations are evidence based.

WHAT ADDITIONAL CONSIDERATIONS ARE NEEDED WHEN DEVELOPING DIETARY GUIDELINES?

In addition to the health effects associated with diet, dietary guidelines may also choose to consider issues such as resource use, equity, environmental impact, feasibility, and acceptability. **Table 1** presents these additional considerations and their relations to dietary guidelines. Whether guideline developers address these additional considerations largely depends on the perspective they assume. For example, dietary guidelines targeted at health professionals are less likely to consider costs and environmental impact compared with guidelines that take a societal perspective.

Table 1 Additional considerations for those developing guidelines in nutrition

Consideration	Details
Resource use	Healthy dietary patterns may be costly. Evaluating the cost effectiveness of programs that promote healthy dietary patterns may be important. Evaluations of cost effectiveness should ideally be specific to each setting in which the guidelines will be implemented as costs will vary between regions.
Equity	There may be reasons for anticipating differences in the relative effectiveness of dietary recommendations for disadvantaged groups. For example, recommended foods and beverages may be less accessible to disadvantaged groups. Therefore, the implementation of recommendations may result in an increase in health disparities. Recommendations may include strategies to reduce anticipated inequity.
Environmental impact	Although certain dietary patterns may optimize health outcomes, they may not be environmentally sustainable in the long term. For example, meat production requires large amounts of land and water to grow grain to feed livestock. It also contributes to methane emissions.
Feasibility	There may be important barriers to implementing dietary recommendations. For example, global and national food suppliers may not be prepared to deal with the shift in a population's dietary patterns that may follow the implementation of a new dietary guideline. The food industry may not be adequately prepared to respond to changes in demand with new or alternative products.
Acceptability	Some dietary interventions may not be acceptable to stakeholders. Stakeholders are broadly defined as individuals, groups, or organizations that may be directly affected by or interested in a proposed policy. Cultural and religious considerations may make some dietary patterns unacceptable to some groups of people. Certain dietary interventions may be burdensome and unacceptable to the general public.

If guideline developers choose to address additional considerations, they should make explicit the weight of each consideration in their decision-making process.

Ideally, additional considerations should be accompanied by a comprehensive search for and synthesis of the related evidence, conducted with a rigor similar to that used in systematic reviews addressing health outcomes. It is perhaps unsurprising that guideline panels do not always achieve this degree of rigor. For example, the Nordic Nutrition Recommendations describe the characteristics of environmentally sustainable diets without undertaking a systematic review of the relevant evidence (82). The magnitude of consideration given to environmental sustainability in relation to health effects is also not specified, raising questions about how decision-makers might balance situations in which environmentally sustainable dietary patterns are at odds with those that optimize health outcomes.

Most national dietary guidelines limit their scope to considering only health outcomes (30). Some countries, such as the Netherlands, Sweden, and Brazil, also consider environmental sustainability. The 2015 DGAC recommended considering environmental sustainability in its report to the USDA and DHHS. The DGAC conducted a systematic review of modeling studies to identify the most environmentally sustainable dietary patterns. This sparked an intense debate regarding whether environmental sustainability is within the scope of national dietary guidelines (10). Ultimately, the guidelines were published without reflecting sustainability considerations. It is possible that the recommendations in the 2015 DGA would have been different had environmental sustainability been considered. For instance, nuts (particularly almonds), which the guideline recommends, require many liters of water to produce a very small quantity (10). The production of seafood, another recommended food group, is in the midst of rapid expansion to meet growing worldwide demand, but the collapse of some fisheries due to overfishing in past decades raises concerns about sustainability (29). It remains controversial whether the consideration of environmental sustainability has a place in national dietary guidelines.

HOW SHOULD CONSUMERS BE ENGAGED IN DEVELOPING DIETARY GUIDELINES?

Consumer engagement is now internationally recognized as an important component of guideline development (19, 80, 86, 94). Involving consumers recognizes individuals as having valuable input about their health, respects the rights of citizens to participate in health policy development, and improves the implementability and uptake of guidelines. Within the GRADE framework, consumers can provide input regarding the relative importance placed on the outcomes of interest and, in turn, decide whether the benefits of a course of action are outweighed by its harms. In dietary guidelines, consumers are predominantly members of the general public. Other consumers may include health professionals and policy-makers.

While the value of consumer engagement is well recognized, there is great variability in how consumers are engaged in the development of both medical and dietary guidelines, with no accepted gold standard approach (6). Ideally, consumers should be engaged throughout the entire process of guideline development (6). A comprehensive framework on consumer engagement that we find helpful was published by Armstrong and colleagues in 2017 (6). The framework summarizes strategies for ensuring meaningful consumer engagement at various steps of the guideline development process. These strategies include, but are not limited to, soliciting priority topics from the public, including consumers on guideline development panels, surveying consumers regarding the importance of proposed outcomes for guideline questions, consulting consumers on decisions related to the analysis and presentation of data, involving consumers in drafting recommendations or reviewing drafts of recommendations, and querying consumers regarding expected barriers to the dissemination and implementation of guidelines. Additionally, consumers can be engaged in two ways: passively (whereby consumers are solicited for their ideas or views through public forums) and actively (whereby consumers are involved in meetings and focus groups and a dialogue between guideline developers and consumers is established), with active engagement being preferred to passive engagement. Barriers to consumer engagement include ensuring recruitment of a representative spectrum of consumers, determining how to facilitate meaningful engagement with consumers who have insufficient understanding of the relevant content or guideline development process, and handling situations in which consumers may hold scientifically indefensible views.

Unfortunately, current dietary guidelines fall short of ensuring adequate consumer engagement (27, 87). In most cases, consumer engagement is passive and restricted to one or two steps within the guideline development process or completely overlooked. For example, while the DGA process engages the public through nominations of members to the DGAC, the ultimate selection process for members of the DGAC is a black box (77). Furthermore, members of the public do not directly participate in the development of DGA recommendations. Rather, they are restricted to opportunities for public comment on the DGA and DGAC report. It is unclear the extent to which and how public comments subsequently shape recommendations. Similarly, the 2012 Nordic Nutrition Recommendations included a period of public comment without any detail regarding how recommendations were modified based on public comments (82).

HOW SHOULD CONFLICTS OF INTEREST BE ADDRESSED?

Interpreting a body of evidence involves a level of subjectivity that may lead to discordant conclusions by different individuals (97). This can be problematic when conflicts of interest influence the interpretation of the evidence and the resulting recommendations. Financial conflicts of interest and industry sponsorships have received substantial attention in nutrition (81). Financial conflicts include ownership of stocks or shares, paid employment, or receipt of research grants.

Nonfinancial conflicts, including intellectual and personal conflicts, may, however, also be important (35, 50). Intellectual conflicts include previous involvement in research that bears on the recommendation. Personal conflicts can include previous advocacy, personal preferences, and culture and religion.

To date, the most common strategies proposed to limit the influence of conflicts of interest on recommendations have been to exclude individuals with conflicts from the guideline group or limit the number of individuals with conflicts to a minority of the group (35, 69, 70). It has also been proposed that important conflicts of interest should prohibit panelists from leadership roles and active participation in the final decision-making process for the recommendations for which they have conflicts (35). Some guideline groups have used an external standing committee to screen candidates by reviewing their conflicts of interest and evaluating the likelihood of their conflicts influencing recommendations (69, 70). Additionally, it has been suggested that information on conflicts in guideline groups should be monitored and periodically updated, given that guideline development is a lengthy process with some guidelines taking many years to produce (54).

A previous report suggests that less than 15% of national dietary guidelines disclose the conflicts of interest of working members, and less than 10% have a policy for managing conflicts (16). In 2017, Erickson and colleagues (27) found problems in the reporting and handling of conflicts of interest in most published guidelines on sugar consumption. Among dietary guidelines that have a policy regarding the handling of conflicts of interest, simple disclosure is the most common (16, 27). Most dietary guidelines do not have policies for dealing with intellectual or personal conflicts or managing conflicts beyond disclosure.

The 2015 DGA also fell short on appropriately managing conflicts of interest. Relevant conflicts were not disclosed in either the DGAC report or the accompanying guideline document (25, 109). Yet the research history of DGAC members reveals many relevant conflicts (104), including the receipt of funding from the California Walnut Commission, the International Tree Nut Council, and various food companies, including Bunge, Unilever, General Mills, and PepsiCo (13, 42, 61, 72, 96). Some members of the DGAC have also been closely involved with large cohort studies that have been used to support the guideline's recommendations (24, 32, 89). The DGAC relied heavily on nonsystematic reports produced by third-party organizations, such as the American Heart Association and the American College of Cardiology, that receive substantial funding from the vegetable oil industry (104). Although the report recommends a high consumption of vegetable oils and nuts and places undue emphasis on findings from particular observational studies with a high potential for confounding, the extent to which conflicts influenced recommendations remains uncertain.

In contrast to the DGA, the IARC reports that it bars experts with conflicts of interest from participating in the working group. The panel may consult experts with conflicts as invited specialists who may contribute expertise to noninfluential issues but who cannot serve in leadership positions, draft texts that pertain to the interpretation of data, or participate in deliberations. Whether the IARC takes a comprehensive approach and also includes intellectual and personal conflicts in its policy is not explicitly reported: It is most likely that the IARC's definition of conflicts of interest is restricted to financial conflicts.

Organizations producing guidelines, in addition to considering conflicts of interest among guideline developers, might also consider the extent to which primary studies being used to guide recommendations are industry sponsored. Evidence regarding the extent to which industry sponsorships can bias the results of studies in nutrition is mixed. A systematic review found there was no statistically significant difference between conclusions from industry-sponsored studies and nonindustry-sponsored studies in nutrition (22). Other research has found evidence for substantial bias in industry-sponsored systematic reviews of sugar-sweetened beverages and artificial

sweeteners (15, 66). If partnerships between researchers and the food industry are transparent and provide open access to protocols and verifiable data, records indicating the intent to publish results, and the eventual publication of results, methodologists participating in guideline development should be able to use the results of industry-sponsored studies to guide recommendations without being affected by potentially misleading reporting (68).

WHAT ARE THE CONSEQUENCES OF USING POOR EVIDENCE TO GUIDE PUBLIC POLICY DECISIONS ABOUT NUTRITION?

Starting in the 1970s, the US government, along with other authoritative organizations, recommended a shift to a low-fat diet. In 1980, the first DGA was published with the recommendations to "avoid too much fat, saturated fat, and cholesterol" (103; 107, p. 11). The 1985 DGA went further and recommended that fats be restricted to less than 30% of total caloric intake (108). Other governments shortly followed with their own dietary recommendations for low-fat diets. For example, in 1983, the United Kingdom's National Advisory Committee on Nutritional Education published dietary guidance echoing the 1980 DGA (79). The food industry responded to increased consumer demand for low-fat foods by developing thousands of new products that were lower in fat. In an effort to maintain palatability, fat-replacement strategies often required increasing the carbohydrate content, and fat was replaced by a combination of sugars and starches (65). In the following years, the intake of dietary fats decreased from approximately 36% of total caloric intake in the late 1970s to approximately 33% in 2000, and carbohydrate intake increased by approximately 10% (67).

Recent research suggests that the focus on reducing dietary fat may have directly contributed to the growing burden of chronic diseases. Increased rates of obesity and heart disease have been found to coincide with the introduction of recommendations to consume a low-fat diet (44). Although a causal relation cannot be concluded—given the many limitations of observational research already discussed—the temporal association is striking between reduced fat intake and increased carbohydrate intake and the rise in the prevalence of overweight, obesity, and cardiometabolic conditions (73). At the very least, dietary recommendations to reduce fat intake did not reduce the prevalence of cardiometabolic diseases as they were intended to. The health effects of low-fat dietary patterns remain a contentious issue and, given the magnitude of scientific uncertainty that exists even today, further highlight that past recommendations to consume a low-fat diet were largely inappropriate (31, 34, 88, 103).

The case of dietary fats illustrates the importance of using strong evidence to guide recommendations. In general, the evidence used to guide recommendations is now considered to be poor quality and potentially misleading (59, 71, 106). Much of the evidence came from ecological studies with a high potential for confounding. Other studies addressed the effects of short dietary interventions on surrogate outcomes. At the time, data from RCTs addressing patient-important outcomes did not support the effectiveness of low-fat diets, but observational evidence was viewed as sufficient to implicate dietary fats in the pathogenesis of heart disease (39). At the time of the inception of the recommendations to consume low-fat diets, the techniques of epidemiology and causal inference were much different from today (53). The RCT had not yet attained its hegemonic gold standard status, and the hazards of observational evidence had not yet been widely recognized (53). It has also been argued that recommendations to consume low-fat diets were advanced by the sugar industry, which reportedly employed various spin tactics to implicate fats and not sugar in the pathogenesis of cardiovascular disease (57). In 2018, Johns & Oppenheimer (53) presented an alternative and equally compelling account of how both the sugar and fat industries engaged in attempts to influence dietary recommendations. However, the hypothesis implicating

IOM: Institute of Medicine

dietary fats in heart disease was supported by more evidence, albeit very weak, compared with the antisugar hypothesis, and, hence, gained more endorsement and traction (53). The example of dietary fats illustrates that failure to acknowledge the limitations of the evidence base may lead to recommendations that are ineffective, at best, or harmful, at worst.

Turning to another issue, nutrition as a field is more susceptible to influence from industry and those with conflicts of interest because the evidence is usually of low certainty. This is illustrated by the controversy surrounding the WIC program in the United States and its stance on white potatoes. In 2014, WIC removed white potatoes from its list of approved foods, based on a report from the Institute of Medicine (IOM) that concluded that low-income families already consumed large quantities of white potatoes and that potatoes have low nutritional value (51). Potato lobbyists, spearheaded by the National Potato Council, worked to overturn this change, arguing that potatoes are just as nutritious as other fruits and vegetables. Recently, WIC added white potatoes back to the list of approved foods, based on a more recent report from the IOM that concluded white potatoes are a cost-effective source of fiber and potassium and are nutritionally valuable (52). While the reversal of the decision may be attributed to scientific developments in potato research, in our opinion, that is unlikely. Rather, low-certainty evidence is generally more open to undue influence from those with conflicts of interest. Inconsistent and alternating dietary recommendations are arguably more likely when the evidence is of low certainty, given that uncertain evidence is generally more open to interpretation.

Inaccurate and conflicting dietary guidance is detrimental to the public's understanding of nutrition and people's ability to build healthy diets. Nutritional guidance, once adopted, appears frequently in the media. At a time when consumers are already subjected to an overabundance of nutrition and health information, governments and organizations should be held accountable for developing policies that are rooted in strong science and free from influence by those with conflicts of interest (54).

HOW CAN DECISION-MAKING ABOUT PUBLIC POLICY ON NUTRITION BE IMPROVED?

Although current dietary guidelines suffer from shortcomings, they have a large impact on public policy. Improving the methods for developing dietary guidelines has the potential to improve public policy decision-making about nutrition. To make this improvement, we must base all recommendations on methodologically rigorous systematic reviews and should primarily rely on higher quality evidence to guide recommendations. If only low-certainty evidence is available, guideline panels must acknowledge the limitations of the evidence base and refrain from making strong recommendations.

All decision-making is complex. Decision-makers may not have clear criteria, may sometimes neglect important criteria, or may give inappropriate importance to select criteria. Hence, we recommend the use of a transparent and structured system to evaluate the certainty of the evidence and to move from evidence to recommendations, such as the GRADE system, which is the most widely used system and an improvement over its predecessors (8, 9, 21, 62, 76, 90).

Norms for handling conflicts of interest in developing dietary guidelines are inadequate. While disclosure is an important step, further measures should be taken to protect recommendations from undue influence. Such measures should include prohibiting those with conflicts of interest from participating in the final decision-making process for the recommendation for which they have conflicts and excluding those with conflicts of interest from leadership positions.

Most dietary guidelines do not meet standards for adequate stakeholder engagement (87). Current dietary guidelines would benefit from further stakeholder engagement, including

engagement of the general public, academic researchers, and members of professional organizations. While broad participation in the process should be proactively sought, participation needs to be incorporated thoughtfully. Ideally, such engagement should occur throughout the entire process of developing recommendations and should include some active components, such as involving consumers as part of guideline panels.

Finally, decision-makers can make use of several resources that set standards for the development of guidelines. While many of these resources target the development of medical guidelines, most, if not all, are relevant to nutrition. Such resources include the IOM's standards for developing trustworthy guidelines (80); guidance from the GRADE working group, which sets the gold standard for grading the certainty of evidence and strength of recommendations (**http://www.gradeworkinggroup.org/**); the Guidelines International Network–McMaster Guideline Development Checklist (94); tools for measuring the quality of guidelines, such as the AGREE (Appraisal of Guidelines, Research and Evaluation) tool (19); and a recent initiative to adhere to the best international standards in the development of nutritional recommendations (54; **https://nutrirecs.com**).

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

LITERATURE CITED

1. Agoritsas T, Merglen A, Shah ND, O'Donnell M, Guyatt GH. 2017. Adjusted analyses in studies addressing therapy and harm: users' guides to the medical literature. *JAMA* 317:748–59
2. Alonso-Coello P, Schünemann HJ, Moberg J, Brignardello-Petersen R, Akl EA, et al. 2016. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. *BMJ* 353:i2016
3. Andrews JC, Schünemann HJ, Oxman AD, Pottie K, Meerpohl JJ, et al. 2013. GRADE guidelines: 15. Going from evidence to recommendation—determinants of a recommendation's direction and strength. *J. Clin. Epidemiol.* 66:726–35
4. Ankarfeldt MZ. 2014. Comment on "Limitations of observational evidence: implications for evidence-based dietary recommendations." *Adv. Nutr.* 5:293
5. Archer E, Hand GA, Blair SN. 2013. Validity of US nutritional surveillance: National Health and Nutrition Examination Survey caloric energy intake data, 1971–2010. *PLOS ONE* 8:e76632
6. Armstrong MJ, Rueda JD, Gronseth GS, Mullins CD. 2017. Framework for enhancing clinical practice guidelines through continuous patient engagement. *Health Expect.* 20:3–10
7. Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, et al. 2004. Grading quality of evidence and strength of recommendations. *BMJ* 328:1490
8. Atkins D, Briss PA, Eccles M, Flottorp S, Guyatt GH, et al. 2005. Systems for grading the quality of evidence and the strength of recommendations II: Pilot study of a new system. *BMC Health Serv. Res.* 5:25
9. Atkins D, Eccles M, Flottorp S, Guyatt GH, Henry D, et al. 2004. Systems for grading the quality of evidence and the strength of recommendations I: Critical appraisal of existing approaches. The GRADE Working Group. *BMC Health Serv. Res.* 4:38
10. Aubrey A. 2015. New dietary guidelines will not include sustainability goal. *NPR*, Oct 6. **https://www.npr.org/sections/thesalt/2015/10/06/446369955/new-dietary-guidelines-will-not-include-sustainability-goal**
11. Ball JR, Micheel CM. 2010. *Evaluation of Biomarkers and Surrogate Endpoints in Chronic Disease*. Washington, DC: Natl. Acad. Press

12. Balshem H, Helfand M, Schünemann HJ, Oxman AD, Kunz R, et al. 2011. GRADE guidelines: 3. Rating the quality of evidence. *J. Clin. Epidemiol.* 64:401–6
13. Bao Y, Han J, Hu FB, Giovannucci EL, Stampfer MJ, et al. 2013. Association of nut consumption with total and cause-specific mortality. *N. Engl. J. Med.* 369:2001–11
14. Beresford SA, Johnson KC, Ritenbaugh C, Lasser NL, Snetselaar LG, et al. 2006. Low-fat dietary pattern and risk of colorectal cancer: the Women's Health Initiative Randomized Controlled Dietary Modification Trial. *JAMA* 295:643–54
15. Bes-Rastrollo M, Schulze MB, Ruiz-Canela M, Martinez-Gonzalez MA. 2013. Financial conflicts of interest and reporting bias regarding the association between sugar-sweetened beverages and weight gain: a systematic review of systematic reviews. *PLOS Med.* 10:e1001578
16. Blake P, Durão S, Naude CE, Bero L. 2018. An analysis of methods used to synthesize evidence and grade recommendations in food-based dietary guidelines. *Nutr. Rev.* 76:290–300
17. Block G, Hartman AM, Naughton D. 1990. A reduced dietary questionnaire: development and validation. *Epidemiology* 1:58–64
18. Bretthauer M, Kalager M. 2018. When no guideline recommendation is the best recommendation. *Lancet* 392:898–99
19. Brouwers MC, Kho ME, Browman GP, Burgers JS, Cluzeau F, et al. 2010. AGREE II: advancing guideline development, reporting and evaluation in health care. *Can. Med. Assoc. J.* 182:E839–42
20. Bucher HC, Guyatt GH, Cook DJ, Holbrook A, McAlister FA. 1999. Users' guides to the medical literature: XIX. Applying clinical trial results A. How to use an article measuring the effect of an intervention on surrogate end points. *JAMA* 282:771–78
21. Carrasco-Labra A, Brignardello-Petersen R, Santesso N, Neumann I, Mustafa RA, et al. 2016. Improving GRADE evidence tables part 1: A randomized trial shows improved understanding of content in summary of findings tables with a new format. *J. Clin. Epidemiol.* 74:7–18
22. Chartres N, Fabbri A, Bero LA. 2016. Association of industry sponsorship with outcomes of nutrition studies: a systematic review and meta-analysis. *JAMA Intern. Med.* 176:1769–77
23. Colapinto CK, Ellis A, Faloon-Drew K, Lowell H. 2016. Developing an evidence review cycle model for Canadian dietary guidance. *J. Nutr. Educ. Behav.* 48:77–83.e1
24. Colditz GA, Hankinson SE. 2005. The Nurses' Health Study: lifestyle and health among women. *Nat. Rev. Cancer* 5:388–96
25. Diet. Guidel. Advis. Comm. 2015. *Scientific Report of the 2015 Dietary Guidelines Advisory Committee: Advisory Report to the Secretary of Health and Human Services and the Secretary of Agriculture*. Washington, DC: USDA (U.S. Dep. Agric.)
26. Ebell MH, Siwek J, Weiss BD, Woolf SH, Susman J, et al. 2004. Strength of Recommendation Taxonomy (SORT): a patient-centered approach to grading evidence in the medical literature. *Am. Fam. Physician.* 69:548–56
27. Erickson J, Sadeghirad B, Lytvyn L, Slavin J, Johnston BC. 2017. The scientific basis of guideline recommendations on sugar intake: a systematic review. *Ann. Intern. Med.* 166:257–67
28. Estruch R, Martínez-González MA, Corella D, Salas-Salvadó J, Ruiz-Gutiérrez V, et al. 2006. Effects of a Mediterranean-style diet on cardiovascular risk factors: a randomized trial. *Ann. Intern. Med.* 145:1–11
29. FAO (Food Agric. Organ. U.N.). 2018. *The State of World Fisheries and Aquaculture: Meeting the Sustainable Development Goals*. Rome: FAO
30. FAO (Food Agric. Organ. U.N.). 2019. Food-based dietary guidelines. *Food and Agriculture Organization of the United Nations*. **http://www.fao.org/nutrition/education/food-dietary-guidelines/regions/en**
31. Forouhi NG, Krauss RM, Taubes G, Willett W. 2018. Dietary fat and cardiometabolic health: evidence, controversies, and consensus for guidance. *BMJ* 361:k2139
32. Fung TT, Rexrode KM, Mantzoros CS, Manson JE, Willett WC, Hu FB. 2009. Mediterranean diet and incidence of and mortality from coronary heart disease and stroke in women. *Circulation* 119:1093–100
33. Gallus S, Bosetti C. 2016. Meat consumption is not tobacco smoking. *Int. J. Cancer* 138:2539–40
34. Grundy SM. 2001. Dietary fat: at the heart of the matter. *Science* 293:801–4
35. Guyatt GH, Akl EA, Hirsh J, Kearon C, Crowther M, et al. 2010. The vexing problem of guidelines and conflict of interest: a potential solution. *Ann. Intern. Med.* 152:738–41

36. Guyatt GH, Oxman AD, Sultan S, Glasziou P, Akl EA, et al. 2011. GRADE guidelines: 9. Rating up the quality of evidence. *J. Clin. Epidemiol.* 64:1311–16
37. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, et al. 2008. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 336:924–26
38. Guyatt G, Rennie D, Meade MO, Cook DJ, eds. 2015. *Users' Guides to the Medical Literature: Essentials of Evidence-Based Clinical Practice*. New York: McGraw-Hill. 3rd ed.
39. Harcombe Z, Baker JS, Cooper SM, Davies B, Sculthorpe N, et al. 2015. Evidence from randomised controlled trials did not support the introduction of dietary fat guidelines in 1977 and 1983: a systematic review and meta-analysis. *Open Heart* 2:e000196
40. Health Can. 2011. *Eating Well with Canada's Food Guide*. Ottawa: Health Can.
41. Hébert JR, Frongillo EA, Adams SA, Turner-McGrievy GM, Hurley TG, et al. 2016. Randomized controlled trials are not a panacea for diet-related research. *Adv. Nutr.* 7:423–32
42. Herman J. 2010. Saving U.S. dietary advice from conflicts of interest. *Food Drug Law J.* 65:285–316
43. Howick J, Chalmers I, Glasziou P, Greenhalgh T, Heneghan C, et al. 2011. Explanation of the 2011 Oxford Centre for Evidence-Based Medicine (OCEBM) levels of evidence: background document. *Oxf. Cent. Evid.-Based Med.* **http://www.cebm.net/index.aspx?o=5653**
44. Hu FB, Stampfer MJ, Manson JE, Grodstein F, Colditz GA, et al. 2000. Trends in the incidence of coronary heart disease and changes in diet and lifestyle in women. *N. Engl. J. Med.* 343:530–37
45. IARC (Int. Agency Res. Cancer). 2010. *Alcohol Consumption and Ethyl Carbamate*. Lyon, France: IARC
46. IARC (Int. Agency Res. Cancer). 2010. *Red Meat and Processed Meat*. Lyon, France: IARC
47. Ioannidis JPA. 2013. Implausible results in human nutrition research. *BMJ* 347:f6698
48. Ioannidis JPA. 2016. We need more randomized trials in nutrition—preferably large, long-term, and with negative results. *Am. J. Clin. Nutr.* 103:1385–86
49. Ioannidis JPA. 2017. Nonreproducibility of preclinical research—reply. *JAMA* 317:2453
50. Ioannidis JPA, Trepanowski JF. 2018. Disclosures in nutrition research: why it is different. *JAMA* 319:547–48
51. IOM (Inst. Med.). 2006. *WIC Food Packages: Time for a Change*. Washington, DC: Natl. Acad. Press
52. IOM (Inst. Med.). 2015. *Review of WIC Food Packages: An Evaluation of White Potatoes in the Cash Value Voucher*. Washington, DC: Natl. Acad. Press
53. Johns DM, Oppenheimer GM. 2018. Was there ever really a "sugar conspiracy"? *Science* 359:747–50
54. Johnston BC, Alonso-Coello P, Bala MM, Zeraatkar D, Rabassa M, et al. 2018. Methods for trustworthy nutritional recommendations NutriRECS (Nutritional Recommendations and accessible Evidence summaries Composed of Systematic reviews): a protocol. *BMC Med. Res. Methodol.* 18:162
55. Johnston BC, Guyatt GH. 2016. Best (but oft-forgotten) practices: intention-to-treat, treatment adherence, and missing participant outcome data in the nutrition literature. *Am. J. Clin. Nutr.* 104:1197–201
56. Kappeler R, Eichholzer M, Rohrmann S. 2013. Meat consumption and diet quality and mortality in NHANES III. *Eur. J. Clin. Nutr.* 67:598–606
57. Kearns CE, Schmidt LA, Glantz SA. 2016. Sugar industry and coronary heart disease research: a historical analysis of internal industry documents. *JAMA Intern. Med.* 176:1680–85
58. Kelemen LE, Kushi LH, Jacobs DR Jr., Cerhan JR. 2005. Associations of dietary protein with disease and mortality in a prospective study of postmenopausal women. *Am. J. Epidemiol.* 161:239–49
59. Keys A, Menotti A, Aravanis C, Blackburn H, Djordevic BS, et al. 1984. The seven countries study: 2,289 deaths in 15 years. *Prev. Med.* 13:141–54
60. Knekt P, Reunanen A, Jävinen R, Seppänen R, Heliövaara M, Aromaa A. 1994. Antioxidant vitamin intake and coronary mortality in a longitudinal population study. *Am. J. Epidemiol.* 139:1180–89
61. Kris-Etherton PM, Hu FB, Ros E, Sabate J. 2008. The role of tree nuts and peanuts in the prevention of coronary heart disease: multiple potential mechanisms. *J. Nutr.* 138:1746S–51S
62. Langendam M, Carrasco-Labra A, Santesso N, Mustafa RA, Brignardello-Petersen R, et al. 2016. Improving GRADE evidence tables part 2: A systematic survey of explanatory notes shows more guidance is needed. *J. Clin. Epidemiol.* 74:19–27
63. Loomis D, Guyton KZ, Grosse Y, Lauby-Secretan B, El Ghissassi F, et al. 2016. Carcinogenicity of drinking coffee, mate, and very hot beverages. *Lancet Oncol.* 17:877–78

64. Ludwig DS. 2016. Lowering the bar on the low-fat diet. *JAMA* 316:2087–88
65. Maki KC, Slavin JL, Rains TM, Kris-Etherton PM. 2014. Limitations of observational evidence: implications for evidence-based dietary recommendations. *Adv. Nutr.* 5:7–15
66. Mandrioli D, Kearns CE, Bero LA. 2016. Relationship between research outcomes and risk of bias, study sponsorship, and author financial conflicts of interest in reviews of the effects of artificially sweetened beverages on weight outcomes: a systematic review of reviews. *PLOS ONE* 11:e0162198
67. MMWR (Morb. Mortal. Wkly. Rep.). 2004. Trends in intake of energy and macronutrients—United States, 1971–2000. *MMWR* 53:80–82
68. Montori VM, Jaeschke R, Schünemann HJ, Bhandari M, Brozek JL, et al. 2004. Users' guide to detecting misleading claims in clinical research reports. *BMJ* 329:1093–96
69. Morciano C, Basevi V, Faralli C, Boon MH, Tonon S, Taruscio D. 2016. Policies on conflicts of interest in health care guideline development: a cross-sectional analysis. *PLOS ONE* 11:e0166485
70. Morciano C, Faralli C, Basevi V. 2017. Managing conflicts of interest in practice guidelines panels. *JAMA* 318:867–68
71. Mozaffarian D. 2016. Dietary and policy priorities for cardiovascular disease, diabetes, and obesity: a comprehensive review. *Circulation* 133:187–225
72. Mozaffarian D, Hao T, Rimm EB, Willett WC, Hu FB. 2011. Changes in diet and lifestyle and long-term weight gain in women and men. *N. Engl. J. Med.* 364:2392–404
73. Mozaffarian D, Ludwig DS. 2015. The 2015 US dietary guidelines: lifting the ban on total dietary fat. *JAMA* 313:2421–22
74. Murad MH, Montori VM, Ioannidis JP, Jaeschke R, Devereaux PJ, et al. 2014. How to read a systematic review and meta-analysis and apply the results to patient care: users' guides to the medical literature. *JAMA* 312:171–79
75. Murad MH, Mustafa R, Morgan R, Sultan S, Falck-Ytter Y, Dahm P. 2016. Rating the quality of evidence is by necessity a matter of judgment. *J. Clin. Epidemiol.* 74:237–38
76. Mustafa RA, Santesso N, Brozek J, Akl EA, Walter SD, et al. 2013. The GRADE approach is reproducible in assessing the quality of evidence of quantitative evidence syntheses. *J. Clin. Epidemiol.* 66:736–42.e5
77. Natl. Acad. Sci. Eng. Med. 2017. *Optimizing the Process for Establishing the Dietary Guidelines for Americans: The Selection Process*. Washington, DC: Natl. Acad. Press
78. Natl. Acad. Sci. Eng. Med. 2017. *Redesigning the Process for Establishing the Dietary Guidelines for Americans*. Washington, DC: Natl. Acad. Press
79. Natl. Advis. Comm. Nutr. Educ. 1983. *A Discussion Paper on Proposals for Nutritional Guidelines for Health Education in Britain*. London, U.K.: Health Education Council
80. Natl. Res. Counc. 2011. *Clinical Practice Guidelines We Can Trust*. Washington, DC: Natl. Acad. Press
81. Nestle M. 2013. *Food Politics: How the Food Industry Influences Nutrition and Health*. Berkeley, CA: Univ. Calif. Press
82. Nordic Counc. Minist. 2012. *Nordic Nutrition Recommendations*. Copenhagen: Nordic Counc. Minist.
83. Patel CJ, Ioannidis JP. 2014. Placing epidemiological results in the context of multiplicity and typical correlations of exposures. *J. Epidemiol. Community Health* 68:1096–100
84. Pound P, Bracken MB. 2014. Is animal research sufficiently evidence based to be a cornerstone of biomedical research? *BMJ* 348:g3387
85. Pound P, Ebrahim S, Sandercock P, Bracken MB, Roberts I. 2004. Where is the evidence that animal research benefits humans? *BMJ* 328:514–17
86. Qaseem A, Forland F, Macbeth F, Ollenschläger G, Phillips S, van der Wees P. 2012. Guidelines International Network: toward international standards for clinical practice guidelines. *Ann. Intern. Med.* 156:525–31
87. Rabassa M, Ruiz SG-R, Solà I, Pardo-Hernandez H, Alonso-Coello P, García LM. 2018. Nutrition guidelines vary widely in methodological quality: an overview of reviews. *J. Clin. Epidemiol.* 104:62–72
88. Ravnskov U, Allen C, Atrens D, Enig MG, Groves B, et al. 2002. Studies of dietary fat and heart disease. *Science* 295:1464–65
89. Rimm EB, Stampfer MJ, Ascherio A, Giovannucci E, Colditz GA, Willett WC. 1993. Vitamin E consumption and the risk of coronary heart disease in men. *N. Engl. J. Med.* 328:1450–56

90. Santesso N, Carrasco-Labra A, Langendam M, Brignardello-Petersen R, Mustafa RA, et al. 2016. Improving GRADE evidence tables part 3: Detailed guidance for explanatory footnotes supports creating and understanding GRADE certainty in the evidence judgments. *J. Clin. Epidemiol.* 74:28–39
91. Satija A, Stampfer MJ, Rimm EB, Willett W, Hu FB. 2018. Are large, simple trials the solution for nutrition research? *Adv. Nutr.* 9:378–87
92. Schatzkin A, Lanza E, Corle D, Lance P, Iber F, et al. 2000. Lack of effect of a low-fat, high-fiber diet on the recurrence of colorectal adenomas. *N. Engl. J. Med.* 342:1149–55
93. Schünemann HJ, Hill S, Guyatt G, Akl EA, Ahmed F. 2011. The GRADE approach and Bradford Hill's criteria for causation. *J. Epidemiol. Community Health* 65:392–95
94. Schünemann HJ, Wiercioch W, Etxeandia I, Falavigna M, Santesso N, et al. 2014. Guidelines 2.0: Systematic development of a comprehensive checklist for a successful guideline enterprise. *Can. Med. Assoc. J.* 186:E123–42
95. Schwingshackl L, Knuppel S, Schwedhelm C, Hoffmann G, Missbach B, et al. 2016. NutriGrade: a scoring system to assess and judge the meta-evidence of randomized controlled trials and cohort studies in nutrition research. *Adv. Nutr.* 7:994–1004
96. Shi Y, Hu FB. 2014. The global implications of diabetes and cancer. *Lancet* 383:1947–48
97. Shrier I, Boivin JF, Platt RW, Steele RJ, Brophy JM, et al. 2008. The interpretation of systematic reviews with meta-analyses: an objective or subjective process? *BMC Med. Inform. Decis. Mak.* 8:19
98. Singh PN, Fraser GE. 1998. Dietary risk factors for colon cancer in a low-risk population. *Am. J. Epidemiol.* 148:761–74
99. Siontis GC, Ioannidis JP. 2011. Risk factors and interventions with statistically significant tiny effects. *Int. J. Epidemiol.* 40:1292–307
100. Slavin JL. 2015. The challenges of nutrition policymaking. *Nutr. J.* 14:15
101. Stampfer M. 1997. Observational epidemiology is the preferred means of evaluating effects of behavioral and lifestyle modification. *Control. Clin. Trials* 18:494–99
102. Takata Y, Shu X-O, Gao Y-T, Li H, Zhang X, et al. 2013. Red meat and poultry intakes and risk of total and cause-specific mortality: results from cohort studies of Chinese adults in Shanghai. *PLOS ONE* 8:e56963
103. Taubes G. 2001. The soft science of dietary fat. *Science* 291:2536–45
104. Teicholz N. 2015. The scientific report guiding the US dietary guidelines: Is it scientific? *BMJ* 351:h5686
105. Trepanowski JF, Ioannidis JP. 2018. Limiting dependence on nonrandomized studies and improving randomized trials in human nutrition research: why and how. *Adv. Nutr.* 9:367–77
106. Turpeinen O, Karvonen MJ, Pekkarinen M, Miettinen M, Elosuo R, Paavilainen E. 1979. Dietary prevention of coronary heart disease: the Finnish Mental Hospital Study. *Int. J. Epidemiol.* 8:99–118
107. USDA (US Dep. Agric.), USDHHS (US Dep. Health Hum. Serv.). 1980. *Nutrition and Your Health: Dietary Guidelines for Americans*. Washington, DC: U.S. Gov. Print. Off.
108. USDA (US Dep. Agric.), USDHHS (US Dep. Health Hum. Serv.). 1985. *Nutrition and Your Health: Dietary Guidelines for Americans*. Washington, DC: U.S. Gov. Print. Off.
109. USDA (US Dep. Agric.), USDHHS (US Dep. Health Hum. Serv.). 2015. *Dietary Guidelines for Americans 2015–2020*. Washington, DC: USDA, USDHHS
110. Vineis P, Stewart BW. 2016. How do we judge what causes cancer? The meat controversy. *Int. J. Cancer* 138:2309–11
111. Young SS, Karr AF. 2011. Deming, data and observational studies: a process out of control and needing fixing. *Significance* 8:116–20
112. Yusuf S, Dagenais G, Pogue J, Bosch J, Sleight P. 2000. Vitamin E supplementation and cardiovascular events in high-risk patients. *N. Engl. J. Med.* 342:154–60