GUIDING PRINCIPLES FOR
Developing Dietary Reference Intakes Based on Chronic Disease

Committee on the Development of Guiding Principles for the Inclusion of Chronic Disease Endpoints in Future Dietary Reference Intakes

Shiriki Kumanyika and Maria P. Oria, *Editors*

Food and Nutrition Board

Health and Medicine Division

A Consensus Study Report of

The National Academies of

SCIENCES • ENGINEERING • MEDICINE

THE NATIONAL ACADEMIES PRESS

Washington, DC

www.nap.edu

THE NATIONAL ACADEMIES PRESS 500 Fifth Street, NW Washington, DC 20001

This activity was supported by the Agricultural Research Service of the U.S. Department of Agriculture (Grant # 59-0204-5-001); Health Canada (Contract No. 4500358459); the National Cancer Institute, National Institute of Diabetes and Digestive and Kidney Diseases, and the Office of Dietary Supplements of the National Institutes of Health; the Office of Disease Prevention and Health Promotion of the U.S. Department of Health and Human Services; the U.S. Centers for Disease Control and Prevention (Contract No. HHSN263201200074I); and the U.S. Food and Drug Administration. Any opinions, findings, conclusions, or recommendations expressed in this publication do not necessarily reflect the views of any organization or agency that provided support for the project.

International Standard Book Number-13: 978-0-309-46256-3
International Standard Book Number-10: 0-309-46256-8
Digital Object Identifier: https://doi.org/10.17226/24828

Additional copies of this publication are available for sale from the National Academies Press, 500 Fifth Street, NW, Keck 360, Washington, DC 20001; (800) 624-6242 or (202) 334-3313; http://www.nap.edu.

Copyright 2017 by the National Academy of Sciences. All rights reserved.

Printed in the United States of America

Suggested citation: National Academies of Sciences, Engineering, and Medicine. 2017. *Guiding principles for developing Dietary Reference Intakes based on chronic disease*. Washington, DC: The National Academies Press. doi: https://doi.org/10.17226/24828.

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

The National Academy of Sciences was established in 1863 by an Act of Congress, signed by President Lincoln, as a private, nongovernmental institution to advise the nation on issues related to science and technology. Members are elected by their peers for outstanding contributions to research. Dr. Marcia McNutt is president.

The **National Academy of Engineering** was established in 1964 under the charter of the National Academy of Sciences to bring the practices of engineering to advising the nation. Members are elected by their peers for extraordinary contributions to engineering. Dr. C. D. Mote, Jr., is president.

The **National Academy of Medicine** (formerly the Institute of Medicine) was established in 1970 under the charter of the National Academy of Sciences to advise the nation on medical and health issues. Members are elected by their peers for distinguished contributions to medicine and health. Dr. Victor J. Dzau is president.

The three Academies work together as the **National Academies of Sciences, Engineering, and Medicine** to provide independent, objective analysis and advice to the nation and conduct other activities to solve complex problems and inform public policy decisions. The National Academies also encourage education and research, recognize outstanding contributions to knowledge, and increase public understanding in matters of science, engineering, and medicine.

Learn more about the National Academies of Sciences, Engineering, and Medicine at **www.nationalacademies.org**.

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

Consensus Study Reports published by the National Academies of Sciences, Engineering, and Medicine document the evidence-based consensus on the study's statement of task by an authoring committee of experts. Reports typically include findings, conclusions, and recommendations based on information gathered by the committee and the committee's deliberations. Each report has been subjected to a rigorous and independent peer-review process and it represents the position of the National Academies on the statement of task.

Proceedings published by the National Academies of Sciences, Engineering, and Medicine chronicle the presentations and discussions at a workshop, symposium, or other event convened by the National Academies. The statements and opinions contained in proceedings are those of the participants and are not endorsed by other participants, the planning committee, or the National Academies.

For information about other products and activities of the National Academies, please visit www.nationalacademies.org/about/whatwedo.

COMMITTEE ON THE DEVELOPMENT OF GUIDING PRINCIPLES FOR THE INCLUSION OF CHRONIC DISEASE ENDPOINTS IN FUTURE DIETARY REFERENCE INTAKES

SHIRIKI KUMANYIKA (*Chair*), Emeritus Professor of Epidemiology, University of Pennsylvania Perelman School of Medicine, Research Professor, and Chair, African American Collaborative Obesity Research Network, Dornsife School of Public Health, Drexel University, Philadelphia, Pennsylvania
CHERYL ANDERSON, Associate Professor, Department of Family and Preventive Medicine, University of California, San Diego
SUSAN I. BARR, Professor Emeritus of Food Nutrition and Health, University of British Columbia, Vancouver
KATHRYN G. DEWEY, Distinguished Professor, Department of Nutrition, University of California, Davis
GORDON GUYATT, Distinguished Professor, Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Ontario
JANET C. KING, Professor Emeritus, University of California, Berkeley and Davis, Senior Scientist, Children's Hospital Oakland Research Institute, Oakland, California
MARIAN L. NEUHOUSER, Full Member, Cancer Prevention Program, Fred Hutchinson Cancer Research Center, Seattle, Washington
ROSS L. PRENTICE, Member, Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Washington
JOSEPH RODRICKS, Founding Principal, Ramboll Environ, Arlington, Virginia
PATRICK J. STOVER, Professor and Director, Division of Nutritional Sciences, Cornell University, Ithaca, New York
KATHERINE L. TUCKER, Professor of Nutritional Epidemiology, University of Massachusetts Lowell
ROBERT B. WALLACE, Irene Ensminger Stecher Professor of Epidemiology and Internal Medicine, The University of Iowa

Study Staff

MARIA P. ORIA, Study Director
ALICE VOROSMARTI, Research Associate
RENEÉ GETHERS, Senior Program Assistant
FAYE HILLMAN, Financial Officer
ANN L. YAKTINE, Director, Food and Nutrition Board

NOTE: See Appendix F, Disclosure of Conflict of Interest.

Consultants

WEIHSUEH A. CHIU, Professor, College of Veterinary Medicine and Biomedical Sciences, Texas A&M University
ANNE RODGERS, Science Writer

Reviewers

This Consensus Study Report was reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise. The purpose of this independent review is to provide candid and critical comments that will assist the National Academies of Sciences, Engineering, and Medicine in making each published report as sound as possible and to ensure that it meets the institutional standards for quality, objectivity, evidence, and responsiveness to the study charge. The review comments and draft manuscript remain confidential to protect the integrity of the deliberative process.

We thank the following individuals for their review of this report:

John F. Aloia, New York University Winthrop-University Hospital
Lawrence J. Appel, Johns Hopkins Medical Institutions
Alicia Carriquiry, Iowa State University
Mei Chung, Tufts University School of Medicine
Nancy R. Cook, Brigham and Women's Hospital and Harvard Medical School
Robert J. Cousins, University of Florida
David L. Eaton, University of Washington
John W. Erdman, Jr., University of Illinois at Urbana-Champaign
Bertha Hidalgo, University of Alabama at Birmingham
Frank Hu, Harvard T.H. Chan School of Public Health
Samuel Klein, Washington University School of Medicine
JoAnne E. Manson, Brigham and Women's Hospital and Harvard Medical School

Bernadette P. Marriott, Medical University of South Carolina
Suzanne P. Murphy, Emeritus, University of Hawaii Cancer Center
Sydne Jennifer Newberry, RAND Corporation
Cheryl L. Rock, University of California, San Diego
Elizabeth A. Stuart, Johns Hopkins Bloomberg School of Public Health
Connie M. Weaver, Purdue University

Although the reviewers listed above provided many constructive comments and suggestions, they were not asked to endorse the conclusions or recommendations of this report nor did they see the final draft before its release. The review of this report was overseen by **Diane F. Birt,** Iowa State University, and **David B. Allison,** University of Alabama at Birmingham. They were responsible for making certain that an independent examination of this report was carried out in accordance with the standards of the National Academies and that all review comments were carefully considered. Responsibility for the final content rests entirely with the authoring committee and the National Academies.

Preface

Identifying minimum recommended intakes of food substances known to be essential for preventing life-threatening nutritional deficiency diseases has been a prominent and longstanding stream in nutrition research, with a complementary focus on identifying high intakes likely to result in adverse health effects. Adequate and safe (non-toxic) intakes are expressed as quantitative thresholds around the lower and upper ends, respectively, of a distribution of possible intakes of the nutrient or food substance in question. Having such thresholds—termed Dietary Reference Intakes (DRIs)—is critical for a variety of food and nutrition policy uses. DRIs are intended for application to the apparently healthy population rather than those with medical conditions requiring specialized diets. However, within healthy populations, DRIs also must account for potential differences in nutritional needs and vulnerabilities by age, developmental stage, gender, reproductive status, and other population characteristics that may influence the adequacy and safety of a given nutrient intake.

The process of developing DRIs involves deliberations of expert panels convened under the rubric of the Food and Nutrition Board (FNB) of the Health and Medicine Division of the National Academies of Sciences, Engineering, and Medicine (the National Academies) and involves collaboration between the United States and Canada. This is both a policy-driven and science-driven process. The policy relevant questions about nutrition and health issues are developed by federal agencies. Scientists on DRI committees consider relevant evidence and draw conclusions about intake levels appropriate as DRI values, which are then translated by the

agencies for policy purposes. This consensus study reflects a new set of challenges that has entered into the realm of DRI development for which the study sponsors identified a need for clearer guidance to DRI committees. The challenges relate to how to formally incorporate chronic disease considerations.

Chronic disease considerations have already been included in some DRI committee deliberations to enhance the evidence picture for making judgments about adequacy and safety or because of the DRI committee's awareness of extant chronic disease issues in their particular sphere of interest. The statement of task for this consensus study asks the committee to develop guidance for making judgments about reference intakes for reducing chronic disease risk per se. As explained in detail in the committee's Consensus Study Report, nutrient-chronic disease risk questions, concepts, and methods are qualitatively different from those for adequacy and safety issues. Nutrient-related chronic disease questions address risk within the range already determined to be safe and adequate and address probable, long-term effects on multifactorial pathways instead of the rapid-onset effects of specific nutrients that are associated with traditional DRI questions. The multiple U.S. federal agencies collaborating to sponsor this study indicate the breadth of policy interests and considerations for which chronic disease DRIs may be relevant: the Agricultural Research Service of the U.S. Department of Agriculture; the National Cancer Institute, National Institute of Diabetes and Digestive and Kidney Diseases, and the Office of Dietary Supplements at the National Institutes of Health; the Office of Disease Prevention and Health Promotion of the U.S. Department of Health and Human Services; the U.S. Centers for Disease Control and Prevention; and the U.S. Food and Drug Administration. Health Canada is also a sponsor of this study, reflecting a longstanding tradition of U.S.–Canadian collaboration on DRIs.

The statement of task was based on results of prior, substantive deliberations over a 2-year period on the question of how to develop chronic disease DRIs. These deliberations resulted in the publication *Options for Basing Dietary Reference Intakes (DRIs) on Chronic Disease Endpoints: Report from a Joint US-/Canadian-sponsored Working Group*[1] (i.e., the Options Report), which articulated the scientific and policy context for such an undertaking. It identified a series of issues to be resolved and set

[1] Yetley, E. A., A. J. MacFarlane, L. S. Greene-Finestone, C. Garza, J. D. Ard, S. A. Atkinson, D. M. Bier, A. L. Carriquiry, W. R. Harlan, D. Hattis, J. C. King, D. Krewski, D. L. O'Connor, R. L. Prentice, J. V. Rodricks, and G. A. Wells. 2017. Options for basing Dietary Reference Intakes (DRIs) on chronic disease endpoints: Report from a joint US-/Canadian-sponsored working group. *Am J Clin Nutr* 105(1):249S-285S.

out specific options for resolving them, including addressing the question of how a chronic disease DRI process would relate to the existing process. The study committee was asked to consider these options and develop guiding principles that would serve future DRI committees. The 12 members and 1 consultant named to the committee included scholars whose combined expertise covers a spectrum of relevant areas: nutrient requirements and metabolism, maternal and child nutrition, methods in nutrition epidemiology and chronic disease epidemiology, preventive medicine, biostatistics, methods of systematic literature review and use of evidence in guideline development, toxicology and risk analysis, the DRI process, and other areas of dietary guidance and nutrition policy. At a day-long workshop early in its deliberations, the committee also obtained vital input from top experts on the topics it was to address, particularly about emerging methodological advances that could improve the ability of future DRI committees to make judgments about nutrient-chronic disease relationships and intake-response relationships.

The committee members and consultant, three of whom had participated in developing the above-mentioned Options Report, exhibited an outstanding depth of knowledge, collegiality, and commitment to achieving the assigned tasks. Over the course of three in-person meetings and several teleconferences, it became clear that they understood that this next phase of DRI development would require bridging across disciplinary perspectives and blending various areas of knowledge and experience, and that doing so would be necessary for nutrition guidance to evolve along with changing public health contexts.

This report would not have been possible without the energy, patience, dedication, and expertise of the FNB staff. Maria Oria, the study director, together with Ann Yaktine, Alice Vorosmarti, and Reneé Gethers, have worked tirelessly and creatively to help us with our task within the relatively short, 10-month, period available to complete this fast-track study. On behalf of the committee, I express my utmost appreciation for their efforts.

Finally, I remind readers that this guiding principles report points to pathways for finding answers rather than providing answers themselves. Its value will be found only as it helps future DRI committees draw conclusions that do justice to the unique aspects of nutrition—as a universal and fundamental exposure affecting the health of all people—while also using the most rigorous methodologies available when making judgments about how a nutrient or other food substance contributes to health risks and benefits. Some level of uncertainty will be inherent in any judgments about nutrient-chronic disease associations, especially when substantial

gaps in the relevant evidence exist. The committee sought to minimize this uncertainty wherever possible by facilitating use of a comprehensive and rigorous process for evaluating the evidence that is available. We hope the report will have a positive influence on standards for evidence generation in this critical area of public health.

Shiriki Kumanyika, *Chair*
Committee on the Development of Guiding Principles for the Inclusion of Chronic Disease Endpoints in Future Dietary Reference Intakes

Contents

APPENDIXES

Summary

Since 1938 and 1941, nutrient intake recommendations have been issued to the public in Canada and the United States, respectively. Currently defined as the Dietary Reference Intakes (DRIs),[1] these values are a set of standards established by consensus committees (DRI committees) under the National Academies of Sciences, Engineering, and Medicine (the National Academies) and used for planning and assessing diets of apparently healthy individuals and groups. The *Dietary Guidelines for Americans* and the Canadian Food Guide, for example, draw partly on the DRI reports. Other important uses are establishing reference points for monitoring dietary intake of populations, evaluating government food assistance

[1] Dietary Reference Intakes are a group of six different standards: (1) *Estimated Average Requirement (EAR):* The average daily intake of a nutrient that is expected to meet the requirement of half of the healthy individuals in a group defined by sex and life-stage. (2) *Recommended Dietary Allowance (RDA):* The average daily intake level that is sufficient to meet the nutrient requirements of 97 to 98 percent of healthy individuals in a specified sex and life-stage group. (3) *Adequate Intake (AI):* When available evidence is not sufficient to determine the EAR for a nutrient, an AI is set. The AI is the average daily nutrient intake observed in apparently healthy individuals in a specified sex and life-stage group. (4) *Tolerable Upper Intake Level (UL):* The highest average daily nutrient intake level that is likely to pose no risk of adverse effects to nearly all healthy individuals in the specified sex and life-stage sex group. (5) *Acceptable Macronutrient Distribution Range (AMDR):* A range of usual intakes for a macronutrient that is associated with reduced risk of chronic disease while providing adequate intakes of essential nutrients. An AMDR is expressed as a percentage of total energy intake. (6) *Estimated Energy Requirement (EER):* A calculated level of energy intake that is estimated to maintain energy balance (or as appropriate, normal growth), that incorporates weight, height, physiological state (i.e., pregnancy), and level of energy expenditure.

programs, planning diets for military personnel, and guiding nationwide health programs.

Traditionally, for each nutrient deemed essential for normal physiological functioning (e.g., vitamins, minerals, protein), the scientific literature is reviewed to determine the most appropriate indicator of adequacy (e.g., calcium balance for adequacy of calcium and vitamin D) and toxicity (e.g., liver damage for excessive copper intake) with the objective to establish standards for 22 groups defined by life-stage and sex. In some cases, associations of nutrients with indicators of chronic diseases have been used to establish DRIs where indicators of adequacy or toxicity were lacking or where the chronic disease considerations were critical. At present, a more general need to explore whether specific levels of nutrients or other food substances (NOFSs) can ameliorate the risk of chronic disease is recognized, given changes in diets and the availability of a larger body of evidence about potential roles of NOFSs in causal pathways leading to chronic diseases. Before chronic disease DRIs are determined, however, various conceptual and methodological differences with traditional DRIs should be considered. A major difference reflects the nature of the goal: while traditional DRIs are required to guide efforts to ensure that populations meet essential nutritional needs for normal physiological functioning, chronic disease DRIs are desirable but not essential. Other differences relate to the slow progression of a chronic disease leading to long latency of any effects, as well as the multi-factorial etiology of a chronic disease, the types of data available, and the analytical methodologies required to evaluate such data. Stakeholders have reflected on approaches for how to address chronic disease DRIs but no agreement yet exists on a process that can be consistently applied. A recent effort to identify and articulate challenges associated with evaluating NOFS-chronic disease relationships occurred under the leadership of a working group sponsored by the Canadian and U.S. governments, resulting in the publication *Options for Basing Dietary Reference Intakes (DRIs) on Chronic Disease Endpoints: Report from a Joint US-/Canadian-Sponsored Working Group* (i.e., the Options Report) (Yetley et al., 2017, Appendix B). The Options Report is the primary reference resource for this consensus study, which aims to develop guiding principles and recommendations to develop chronic disease DRIs.

THE TASK AND APPROACH

The statement of task (see Chapter 1, Box 1-1) directs an ad hoc committee of the National Academies to assess the options (see Table S-1) presented in the Options Report and to determine guiding principles for including chronic disease endpoints for NOFSs that will be used by future National Academies committees to develop DRIs. In the Options Report,

TABLE S-1 Options for Addressing Questions Related to Establishing Chronic Disease (from the Options Report in Appendix B)

Question/Issue	Options
How should chronic disease outcomes be selected?	Option 1: Endpoint (outcome) is the incidence of a chronic disease or a qualified surrogate disease marker.
	Option 2: Endpoint (outcome) may include nonqualified disease markers.
What are acceptable levels of confidence that the relation is causal?	Option 1: Require a high level of confidence. This level of confidence likely requires at least some evidence from high-quality randomized controlled trials (RCTs) in which the measured outcome is a chronic disease event or qualified surrogate disease marker.
	Option 2: Use level B evidence. This level of evidence suggests a moderate degree of confidence that the relation of interest is causal, but new findings could change the DRI decision.
	Option 3: Use actual level of certainty.
	Option 4: Make decisions on a case-by-case basis.
What is the approach to selecting indicators and specifying intake-response relations?	Option 1: Choose a single outcome indicator on the causal pathway.
	Option 2: Use multiple indicators of a chronic disease.
	Option 3: Use of multiple indicators for multiple diseases. This option may be necessary when a single food substance has different intake-response relations with multiple chronic diseases. In this situation, the DRI committee might need to develop criteria for selecting appropriate disease indicators to establish multiple intake-response relations, methods to integrate multiple endpoints, and approaches to account for the inevitable inter-individual variability in the relations of interest.
When should intake-response data be extrapolated?	Option 1: Establish reference intake values only for similar populations.
	Option 2: Allow extrapolation when sufficient evidence is available.

continued

TABLE S-1 Continued

Question/Issue	Options
What should be the different types of DRIs associated with benefit?	Option 1: Establish chronic disease risk-reduction intake values, by modifying the traditional EAR/RDA approach to estimate the mean intakes of individuals and the inter-individual variability associated with specified disease risk reductions. This option is conceptually very similar to the traditional EAR/RDA approach but uses relative risks and requires knowledge of baseline disease prevalence.
	Option 2: Identify ranges of beneficial intakes.
What should be the different types of DRIs associated with reduction in chronic disease risk?	Option 1: Base ULs on the traditional threshold model when UL values based on chronic disease endpoints are higher than those based on traditional adverse effects.
	Option 2: Base UL_{CD} on intakes associated with chronic disease risk.
What are acceptable levels of confidence in the intake-response data?	Option 1: Require a high confidence level by, for example, using RCTs with a chronic disease event or a qualified surrogate disease marker as the outcome measure.
	Option 2: Accept a moderate confidence level.
	Option 3: Piecemeal approach. This option pieces together different relations in which the biomarker of interest is a common factor between the food substance and a chronic disease.
What approaches can be taken to make decisions when benefits and harms overlap?	Option 1: Avoid overlap between beneficial intakes and intakes associated with adverse events.
	Option 2: Establish criteria related to severity and risk of chronic disease.
	Option 3: Describe the nature of the evidence.
What should be the organizational process to set all DRIs?	Option 1: Continue to use a single DRI development process.
	Option 2: Create 2 separate processes for developing DRIs.
What should be the starting point of chronic disease DRIs?	Option 1: Establish DRIs for individual or small groups of interrelated food substances
	Option 2: Establish DRIs for multiple food substances on the basis of a chronic disease endpoint.

SOURCE: Yetley et al., 2017.

key scientific challenges to develop chronic disease DRIs were identified, related to three questions: (1) What are acceptable levels of confidence that the relationship between an NOFS and a chronic disease is causal?, (2) If a causal relationship exists, what are acceptable levels of confidence in the intake-response relationship data and what are the approaches for identifying and characterizing the intake-response relationship and, if appropriate, to recommend DRIs?, and (3) What should be the organizational process for recommending chronic disease DRIs? Although DRI committees will be sensitive to the uses of the DRIs in nutrition policies, this committee recognized that recommendations for tasks that relate to risk management and policy (see Figure S-1), such as formulating the statement of task for a future DRI committee, are outside the scope of this study. In making its recommendations, the committee assumed that the work of future DRI committees will be done within the context of the current DRI process (see Figure S-1), in which a DRI committee will receive a thorough and rigorously implemented systematic review,[2] will review the totality of the evidence about questions 1 and 2 above, and will recommend chronic disease DRIs, if appropriate. The committee also assumes that a mechanism will be established for communication between those conducting the systematic review and the DRI committee, which will ensure a moderate level of communication while also protecting against inappropriate influence on the systematic review methods.

Broadly, a DRI committee's task is to review the evidence and recommend all DRIs for specific nutrients. With respect to chronic disease DRIs, the subject of this study, if enough evidence exists with respect to both the causality and the intake-response relationships questions, the committee's decisions will include developing chronic disease DRIs along with the adequacy and toxicity DRIs. Making such decisions entails many challenges. Some challenges are universal—measuring the chronic disease clinical outcome directly, or through surrogate markers[3] for clinical outcomes, or both, and extrapolating to populations other than those directly studied—and other challenges are unique to nutrition research. An important

[2] A systematic review "is a scientific investigation that focuses on a specific question and that uses explicit, planned scientific methods to identify, select, assess, and summarize the findings of similar but separate studies. It may or may not include a quantitative synthesis (meta-analysis) of the results from separate studies" (IOM, 2011b, p. 21). A systematic review is typically conducted by a group of experts in the process itself that includes subject matter experts (e.g., a systematic review team) and in consultation with specific subject matter experts (e.g., a technical expert panel).

[3] A surrogate marker is a type of biomarker that "predicts clinical benefit (or harm or lack of benefit or harm) based on epidemiologic, therapeutic, pathophysiologic, or other scientific evidence. A surrogate disease marker is qualified for its intended purposes" (Yetley et al., 2017, p. 263S).

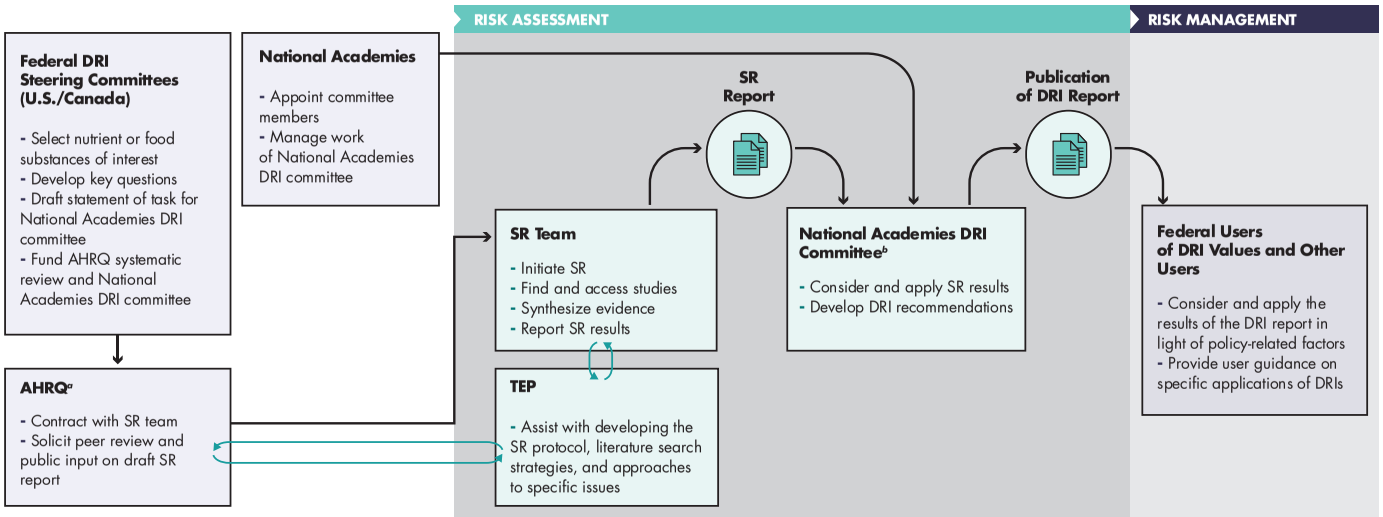


FIGURE S-1 Overall process for developing DRIs with actions, actors, and their relationships. The risk assessment section indicates elements included in the National Academies DRI committee process and their relationships to the systematic review and prior and subsequent risk management elements.

NOTES: AHRQ = Agency for Healthcare Research and Quality; DRI = Dietary Reference Intake; SR = systematic review; TEP = Technical Expert Panel.

[a] The Agency for Healthcare Research and Quality is listed here as the agency that, in the current DRI process, has been responsible for the systematic review aspects of DRI development.

[b] National Academies DRI committees are convened by and positioned with the Food and Nutrition Board of the National Academies of Sciences, Engineering, and Medicine and operate under the National Academies study committee guidelines, which include an external peer review process of the draft DRI report.

SOURCE: Adapted from IOM, 2011a,b; Taylor, 2008.

feature of nutrition studies that explore associations between NOFSs and chronic diseases is their frequent reliance on prospective cohorts, rather than randomized controlled trials (RCTs), as study designs. Reasons for the predominance of observational data include the long-term exposure usually needed to observe an association with the chronic disease, the advantage of following study participants under their usual diet intakes and other living conditions, the ability to observe associations with varying levels of intake of the NOFS of interest, and the considerable cost difference compared to conducting a large RCT. On the other hand, making firm inferences about causal relationships is challenging when observational studies predominate. Two main reasons are the inaccuracy of current methods to measure intake of the NOFS and the potential effect of other factors that, if not accounted for, can lead to erroneous conclusions (e.g., age or other characteristics of the individual, other nutrients in the diet, exposure to other chronic disease factors). Recognizing that causal relationships are inferred mainly from RCTs, observational studies are still critical to inform conclusions about causal relationships and to support evidence on the intake-response relationships.

The committee's recommendations in the following sections consider the options offered in Table S-1, which addresses the pertinent conceptual and methodological challenges, and are organized based on the three questions above. Two of the challenges, evaluating the dietary intake measure and identifying health outcomes, are central to developing chronic disease DRIs and therefore precede the rest of the recommendations. In addition to providing recommendations, the committee developed guiding principles to support DRI committees as they make decisions about chronic disease DRIs. In implementing the recommendations, communication with users and development of guidance for appropriate application of the recommendations are essential.

MEASURING DIETARY INTAKE AND SELECTING CHRONIC DISEASE OUTCOMES

How Should Dietary Intake Measures Be Evaluated?

The nature and quality of nutrient intake ascertainment are diverse. Self-reported measures may be essential for some purposes but intrinsically suffer from both random errors and systematic biases. Biomarkers of intake, which are more objective, can replace self-report for some purposes but have only been developed for a few nutrients. Currently, no one single approach accurately measures dietary intake in a comprehensive manner for all nutrients. Therefore, each study methodology needs to be assessed on the basis of its own merits.

Recommendation 1. Until better intake assessment methodologies are developed and applied widely, Dietary Reference Intake committees should strive to ensure that random and systematic errors and biases of nutrient or other food substance (NOFS) exposure assessment methodologies are considered in their evidence review. In the long term, research agendas should include accelerated efforts to improve NOFS exposure assessments for application in studies of chronic disease risk.

How Should Chronic Disease Outcomes Be Selected?

Studies exploring relationships between NOFSs and chronic disease outcomes vary in the nature and quality of health outcome measurement. High-quality measures of chronic disease outcomes are ideal when developing chronic disease DRIs. However, outcome measures may be flawed and sufficient data on associations between NOFSs and these outcomes do not always exist. The committee supports a variant of option 1 (see Table S-1), where studies that measure qualified surrogate markers—following the criteria adopted by the committee[4]—are considered in evaluating the evidence about causal relationships. The committee does not support option 2, using nonqualified intermediate markers, because they could lead to serious misinterpretation of DRIs by users.

Recommendation 2. The ideal outcome used to establish chronic disease Dietary Reference Intakes should be the chronic disease of interest, as defined by accepted diagnostic criteria, including composite endpoints, when applicable. Surrogate markers could be considered with the goal of using the findings as supporting information of results based on the chronic disease of interest. To be considered, surrogate markers should meet the qualification criteria for their purpose. Qualification of surrogate markers must be specific to each nutrient or other food substance, although some surrogates will be applicable to more than one causal pathway.

[4] Qualification criteria for a surrogate marker are (1) analytical validation exists, (2) surrogate marker is on causal pathway in disease pathogenesis, (3) surrogate marker is significantly associated with the disease in target population, (4) surrogate marker consistently changes with health outcome in response to a nutritional intervention, (5) change in the surrogate marker explains a substantial proportion of the change in the disease response to a nutritional intervention, and (6) context of use is defined (Calder et al., 2017; Clarke, 2017; IOM, 2010).

EVALUATING ACCEPTABLE LEVELS OF CONFIDENCE THAT THE RELATION IS CAUSAL

What Are Acceptable Levels of Confidence That the Relation Is Causal?

Well-established systems for judging evidence for causal relationships related to a variety of exposures and health outcomes are now available. The committee considered the four options in Table S-1 as well as other evidence review systems related to selecting the levels of confidence about the causality of an association, including the Grading of Recommendations Assessment, Development and Evaluation (GRADE) system. GRADE is an approach to rating the certainty of a body of evidence (i.e., high, moderate, low and/or very low) by using five domains (risk of bias,[5] imprecision,[6] inconsistency,[7] indirectness,[8] and publication bias[9]). Although all the systems reviewed involve similar concepts, GRADE is recommended as the basis for DRI committee judgments about causal relationships between NOFS exposures and chronic diseases for the following reasons: (1) it meets the criteria for an appropriate evidence review tool (e.g., sufficient experience; well-structured protocols, clarity, and transparency; sufficiently detailed; ability to address questions about causal relationships; methods applicable over time), (2) it uses appropriate taxonomy for rating the evidence, and (3) it is in wide use for a variety of health matters. The committee recognizes that any discipline (e.g., public health, environmental health) faces specific challenges that need to be considered when embarking on an evidence review evaluation.

Recommendation 3. The committee recommends that Dietary Reference Intake (DRI) committees use Grading of Recommendations Assessment, Development and Evaluation (GRADE) in assessing the certainty of the evidence related to the causal association between nutrient or other food substances and chronic diseases. Using GRADE, the committee recommends that a decision to pro-

[5] Risk of bias is systematic error due to limitations in the study design or execution (Schunemann et al., 2013).

[6] Imprecision is random error that occurs when studies have a small sample size and the number of events is also small; resulting in a wide 95 percent confidence interval around the estimate of the effect (Schunemann et al., 2013).

[7] Inconsistency refers to unexplained heterogeneity or variability of study results (Schunemann et al., 2013).

[8] Indirectness occurs when a study does not compare the interventions of interest, apply the intervention to the population of interest, or measure the outcomes that are important to patients (Schunemann et al., 2013).

[9] Publication bias is a systematic under-estimation or over-estimation of the underlying beneficial or harmful effect due to the selective publication of studies (Schunemann et al., 2013).

ceed with development of chronic disease DRIs be based on at least moderate certainty that a causal relationship exists and on the existence of an intake-response relationship.

APPROACHES TO IDENTIFY AND CHARACTERIZE THE QUANTITATIVE RELATIONSHIP AND DEVELOP DIETARY REFERENCE INTAKES

What Is the Approach to Selecting Indicators and Specifying Intake-Response Relations?

As recommended above, DRI committees should consider only direct measures of disease or surrogate markers deemed appropriate as highly predictive of disease. Characterizing intake-response relationships includes many methodological challenges, including the selection of appropriate models and consideration of confounding factors and interactions. Setting DRIs based on "multiple indicators of a chronic disease" and "multiple indicators for multiple diseases" would require development of multivariate, multi-pathway, intake-response models and, therefore, the committee does not generally recommend those options. The simplest approach of choosing "a single outcome indicator" (option 1) is supported.

Recommendation 4. The committee recommends the use of a single outcome indicator on the causal pathway. However, when a single food substance reduces the risk of more than one chronic disease, reference values could be developed for each chronic disease. The committee, however, does not recommend the use of "multiple indicators of a chronic disease" or "multiple indicators for multiple diseases" unless there is sufficient experience with the use of algorithms or other strong evidence suggesting that multiple indicators point to risk of a chronic disease, due to potential lack of reliability or consistency in the results.

When Should Intake-Response Data Be Extrapolated?

The many factors that influence chronic disease risk in different populations are not well characterized quantitatively, and the likelihood of error is significant. Therefore, the committee supports option 1 in Table S-1. The evidence supporting any departure from this approach should be fully described and should reveal minimal uncertainty.

Recommendation 5. The committee recommends extrapolation of intake-response data for chronic disease Dietary Reference Intakes

only to populations that are similar to studied populations in the underlying factors related to the chronic disease of interest.

What Should Be the Different Types of DRIs Associated with Benefit?

Once intake-response relationships have been identified with acceptable levels of confidence and minimal bias, the challenge is characterizing the relevant intake-response relationships—in terms of their shape and the range of intakes. Several issues complicate translation of intake-response relationships into a DRI, such as the continuous nature of the relationship between nutrients and chronic disease, the multifactorial nature of chronic disease risk, and the diversity of individual baseline risk. Based on this, the committee recommends option 2, which is to recommend DRI ranges at an intake associated with a specified degree of risk reduction or specified benefit but not as single risk-reduction intake value (option 1). The committee also does not recommend developing a family of DRIs for different risk reduction targets for the same chronic disease due to potential difficulties in communicating the uncertainties.

Recommendation 6. The committee recommends that Dietary Reference Intakes (DRIs) for chronic disease risk take the form of a range, rather than a single number. Intake-response relationships should be defined as different ranges of the intake-response relationship where risk is at minimum, is decreasing, and/or is increasing (i.e., slope = 0, negative, or positive). When a nutrient or other food substance reduces the risk of more than one chronic disease, DRIs could be developed for each chronic disease, even if the confidence levels for each chronic disease are different. The magnitude of risk slope considered necessary to set a DRI should be decided based on clearly articulated public health goals, such as those previously identified by other authorities (e.g., Healthy People 2020). The committee does not recommend, however, developing a family of DRIs for any one NOFS for different risk reduction targets for the same chronic disease.

What Should Be the Different Types of DRIs Associated with Reduction in Chronic Disease Risk?

The committee supports a variant of options 1 and 2 and notes that the traditional Tolerable Upper Intake Level (UL)[10] based on toxicity should

[10] The highest level of daily nutrient intake that is likely to pose no risk of adverse health effects to almost all individuals in the general population.

be retained whether a putative chronic disease DRI is above (option 1) or below the traditional UL value for that substance. If the increase in risk occurs only at intakes greater than the traditional UL, then no chronic disease DRI would be required, because avoiding intakes greater than the UL will also avoid increases in chronic disease risk. In cases in which increased intake is associated with increased chronic disease risks at intakes less than the traditional UL, both the traditional UL and a chronic disease DRI could be retained (a variation of option 2). In this case, the intake-response relationship should be characterized as to whether the range over which increased risk occurs overlaps with the traditional UL (option 2). The rationale for retaining both would be that their meanings are both different and valuable: the UL connotes an intake limit that should not be exceeded and the DRI for chronic disease would imply that chronic disease risks will be increased with increasing and long-term exposure over the DRI range.

Recommendation 7. The committee recommends retaining Tolerable Upper Intake Levels (ULs) based on traditional toxicity endpoints. In addition, if increased intake of a substance has been shown to increase the risk of a chronic disease, such a relationship should be characterized as the range where a decreased intake is beneficial. If the increase in risk only occurs at intakes greater than the traditional UL, no chronic disease Dietary Reference Intake would be required, because avoiding intakes greater than the UL will avoid the chronic disease risk.

What Are Acceptable Levels of Confidence in the Intake-Response Data?

Specifying a DRI involves a continuum of options related to specifying a range; the certainty in the evidence might vary at different nutrient levels. Recommending the chronic disease DRIs involves decisions related to the type of DRIs, acceptable level of confidence in the intake-response data, and balancing health risks and benefits. The committee concluded that factors considered in rating the certainty of evidence delineated by GRADE are also appropriate when evaluating the certainty of the intake-response relationship. The committee supports a variant of option 2, with a number of additional considerations that need to be evaluated on a case-by-case basis. Reliable and accurate intake data are particularly important to develop intake-response relationships.

Recommendation 8. The committee recommends that to develop a chronic disease Dietary Reference Intake, the level of certainty

in the intake-response relationship should generally be the same as the level of certainty for a determination of causality, that is, at least "moderate," using the Grading of Recommendations, Assessment, Development and Evaluation (GRADE). However, in some cases, for example when a food substance increases chronic disease risk, the level of certainty considered acceptable might be lower. In all cases, a thorough description of the scientific uncertainties is essential in describing quantitative intake-response relationships. Requiring at least "moderate" certainty extends to cases where relationships between intake and a surrogate marker and between the same surrogate marker and the chronic disease are characterized separately, in a piecemeal (i.e., two-stage) approach.

What Approaches Can Be Taken to Make Decisions When Benefits and Harms Overlap?

DRI committees will explore the decisions related to all DRIs (i.e., based on adequacy, toxicity, and chronic diseases) and, in some cases, harms and benefits could overlap. For example, an NOFS that increases the risk of one chronic disease may decrease the risk of another. It might not always be possible to avoid the overlap between benefits and harms (option 1). It also might not be appropriate for DRI committees to establish criteria related to severity and risk (option 2), which is typically the role of risk managers (e.g., federal agencies). Ideally, DRI committees would conduct analyses with existing models for dealing with competing health risks and benefits. If this is not possible, option 3 is recommended.

Recommendation 9. The committee recommends that, if possible, health risk/benefit analyses be conducted and the method to characterize and decide on the balance be made explicit and transparent. Such a decision needs to consider the certainty of evidence for harms and benefits of changing intake and be based on clearly articulated public health goals. If Dietary Reference Intake committees do not perform such risk/benefit analyses, it is still necessary to describe the disease outcomes and their severities, the magnitudes of risk increases and decreases over various ranges of intakes, and other factors that would allow users to make informed decisions.

INTEGRATING CHRONIC DISEASE DRIs IN THE CURRENT PROCESS

What Should Be the Organizational Process to Develop All DRIs?

Current practice is that DRIs for one or more related nutrients are develop by one committee under the auspices of the National Academies. When chronic disease considerations are integrated in the process, all committee members need to exchange ideas and coordinate their recommendations, particularly when harms and benefits overlap. The committee recommends a variation of option 1 such that for each set of NOFSs under review, a single DRI parent committee would be formed. This organization would allow, if necessary, for the formation of two subcommittees—one subcommittee would address DRIs for the prevention of nutrient deficiencies and minimizing toxicities and the other subcommittee would address DRIs for reducing the risk of chronic disease. Creating two separate processes (option 2) would not allow for sufficient exchange of ideas between the two committees.

Recommendation 10. Because of the need for close coordination and exchange of ideas when setting Dietary Reference Intakes (DRIs) based on indicators of adequacy, toxicity, and chronic disease, one single National Academies of Sciences, Engineering, and Medicine parent committee should develop DRIs for the prevention of nutrient deficiencies and toxicities and for reducing the risk of chronic disease. Due to the need for different expertise and different methodological considerations, two subcommittees could be established at the discretion of the parent committee, for reviewing evidence on (1) adequacy and toxicity and (2) chronic disease, respectively.

What Should Be the Starting Point of Chronic Disease DRIs?

The committee concluded that continuing with the current approach of establishing DRIs for individual or small groups of related NOFSs (option 1) has advantages compared to establishing DRIs for multiple food substances that are related to a single chronic disease endpoint (option 2). First, it may be premature to change the current process before additional experience is gained. Second, the scientific literature and study designs tend to explore relationships substance by substance rather than on all substances related to one disease. Finally, balancing of harms and benefits will be more challenging when NOFSs contribute to more than one chronic

disease because there would be more than one DRI committee addressing the same nutrient.

Recommendation 11. When sufficient evidence exists to develop chronic disease Dietary Reference Intakes for one or more nutrient or other food substances (NOFSs) that are interrelated in their causal relationships with one or more chronic diseases, a committee should be convened to review the evidence of their association with all selected diseases. Using a chronic disease as the starting point for the review is not recommended because balancing health risks and benefits for multiple NOFSs that are related to a single chronic disease endpoint will be a challenge in cases where the same NOFSs might be associated with more than one chronic disease.

GUIDING PRINCIPLES FOR THE PROCESS OF ESTABLISHING CHRONIC DISEASE DRIs

The committee developed the following guiding principles as a foundation for a scientifically credible chronic disease DRI process. Although recommendations about integrating chronic disease as a consideration in setting DRIs should be revisited in the future as more practice and knowledge are acquired, these guiding principles are meant to withstand scientific and methodological advances that will occur in the future.

With respect to systematic reviews:

1. Well-structured and established protocols that include the question of interest and analytical frameworks are necessary to address multiple major and ancillary scientific issues related to the degree of confidence in evidence for causal associations.
2. Protocols should be developed with guidance from a technical expert panel that includes relevant content experts in nutrition science, toxicology, scientific study design and analysis, public health, biostatistics, nutrition epidemiology and chronic disease epidemiology, and disease pathogenesis.
3. In consultation with the technical expert panel, systematic reviews should be sufficiently inclusive of all study designs that potentially contribute to evaluation of the causal NOFS-chronic disease relationship of interest and identification of associated intake-response relationships.
4. Protocols should include studies that use various dietary assessment approaches, including self-report and biomarkers of intake, while

taking the quality of exposure assessment into account when rating study quality.

5. Protocols should include studies that document outcomes or surrogates of outcomes of potential importance for assessing benefits and harms, while taking the quality of outcome assessments into account in rating study quality.
6. Instruments and analytical methods applied to systematic reviews should be thoughtfully chosen and defensible. Instruments to assess the internal validity of the studies should include considerations that apply to nutrition research and various study designs (observational and intervention studies).
7. Results from the systematic review should be clearly presented in study-by-study evidence tables and summary tables of the total evidence for each outcome and study type.

With respect to DRI committee reviews of the totality of the evidence:

8. The DRI committees should include content experts and methodologists relevant to the primary scientific issues and to evidence review. DRI committees should be free of significant financial, intellectual, and professional conflicts of interest. In some cases, the required expertise might not be found without some conflicts of interest. In such cases, it is necessary to identify, disclose, and manage any potential conflicts of interest. Mechanisms to allow for interactions between the DRI committee and members representing both the technical expert panel and systematic review team, while also protecting against inappropriate influence on the systematic review methods, are strongly encouraged.
9. Particular elements of needed expertise will be guided by the general scientific question(s) and specific questions and will generally include nutrition science, scientific study design and analysis, public health, biostatistics, nutrition and chronic disease epidemiology, disease pathogenesis, and evidence review conduct.
10. The evidence review should be sufficiently comprehensive to anticipate the major scientific issues and methods that will likely be a part of the ensuing guideline development process.
11. Sufficient documentation, clarity, and transparency in the evidence review process is needed so that others can comprehend and evaluate this process and its activities, methodological considerations, final decisions, and the rationale for decisions about each outcome.
12. The review of the evidence and other aspects of the systematic review should be replicable and subject to expert peer review.
13. When apparent discrepancies in the evidence exist, DRI commit-

tees should attempt to determine whether they can be explained by differences in methodology or conceptualization of diet–disease relationships and, where possible, incorporate such explanations into the process of rating the evidence.

14. Where they exist, quantitative intake-response relationships should include a thorough description of the scientific uncertainties associated with them.

REFERENCES

Calder, P. C., A. Boobis, D. Braun, C. L. Champ, L. Dye, S. Einother, A. Greyling, C. Matthys, P. Putz, S. Wopereis, J. V. Woodside, and J. M. Antoine. 2017. Improving selection of markers in nutrition research: Evaluation of the criteria proposed by the ILSI Europe Marker Validation Initiative. *Nutr Res Rev* 30(1):73-81.

Clarke, R. 2017. *Pattern of lipid biomarkers and risk of cardiovascular disease*. Presented at the Workshop of the Committee on Development of Guiding Principles for the Inclusion of Chronic Disease Endpoints in Future Dietary Reference Intakes, January 9, 2017, Washington, DC.

IOM (Institute of Medicine). 2010. *Evaluation of biomarkers and surrogate endpoints in chronic disease*. Washington, DC: The National Academies Press.

IOM. 2011a. *Clinical practice guidelines we can trust*. Washington, DC: The National Academies Press.

IOM. 2011b. *Finding what works in health care: Standards for systematic reviews*. Washington, DC: The National Academies Press.

Schunemann, H., J. Brozek, G. Guyatt, and A. Oxman. 2013. *Introduction to GRADE Handbook*. http://gdt.guidelinedevelopment.org/app/handbook/handbook.html (accessed May 16, 2017).

Taylor, C. L. 2008. *Framework for DRI development: Components "known" and components "to be explored."* Washington, DC.

Yetley, E. A., A. J. MacFarlane, L. S. Greene-Finestone, C. Garza, J. D. Ard, S. A. Atkinson, D. M. Bier, A. L. Carriquiry, W. R. Harlan, D. Hattis, J. C. King, D. Krewski, D. L. O'Connor, R. L. Prentice, J. V. Rodricks, and G. A. Wells. 2017. Options for basing Dietary Reference Intakes (DRIs) on chronic disease endpoints: Report from a joint US-/Canadian-sponsored working group. *Am J Clin Nutr* 105(1):249S-285S.

1

Introduction

Food and nutrition policy in the United States and Canada emanates from a variety of activities, including conducting dietary intake surveillance and monitoring, funding and conducting research and education, and reviewing the science. These activities constitute the backbone of developing population-based dietary intake recommendations as well as designing food and nutrition assistance programs for vulnerable populations, with the ultimate goal of improving the nutrition and health of the population. In the United States, the first efforts to establish the science of nutrition date from 1893 when Congress authorized the U.S. Department of Agriculture (USDA) to conduct research on agriculture and human nutrition, and this coincided with a period when nutrition deficiency diseases, such as scurvy and beriberi, were being described (HHS, 1988).

Murphy et al. recently summarized the "long road leading to the Dietary Reference Intakes for the United States and Canada" (Murphy et al., 2016). Briefly, the first Recommended Dietary Allowances (RDAs) were developed by the National Academy of Sciences' Food and Nutrition Board at the request of the National Defense Advisory Commission and adopted at the National Nutrition Conference in 1941. The goal was to assist with World War II food relief efforts where needed. In Canada, a similar process to establish nutrition standards had started in 1938. For the next 40 years, RDAs were periodically reviewed and updated. As knowledge about nutrition and health steadily expanded, research about the contribution of nutrition to noncommunicable chronic diseases began to emerge and was summarized in a landmark report, *Diet and Health: Implications for Reducing Chronic Disease*, which was published by the

National Academy of Sciences in 1989 (NRC, 1989). In addition, the RDAs were being reconsidered as their uses expanded into new arenas. The uses and misuses of RDAs in fortification programs or supplemental food packages for targeted subgroups, for which estimation of levels of inadequacy is needed, prompted exploration of new reference values and a new framework. An approach to identify potential adverse effects of excessive nutrient intake was also important for the regulation of food fortification by federal agencies. In 1983, the Recommended Nutrient Intakes replaced the RDAs in Canada, and in the mid-1990s, the Dietary Reference Intakes (DRIs) replaced the RDAs in the United States.

The publication of the DRIs signified a change in paradigm that was intended partly to allow for better understanding the dietary intakes of populations. As such, DRIs were defined as a set of six reference values related to both adequate intakes and upper levels of intakes.[1] DRIs are specified on the basis of age, sex, and life-stage and cover more than 40 nutrients and food substances. DRIs have been established by different committees of experts (DRI committees) under the auspices of the Food and Nutrition Board of the National Academies of Sciences, Engineering, and Medicine. The most recent report is an update of the 1997 DRIs for calcium and vitamin D (IOM, 2011b).

Since 1997, DRIs have been intended to serve as a roadmap for assessing intake and planning diets for individuals and groups, and to provide the basis for food guidelines in both the United States and Canada. For example, health professionals use DRIs to guide individual nutrition decision-making (Murphy et al., 2016). DRIs also are used in the development of nutrition policy, notably for the *Dietary Guidelines for Americans*. The 2015 Dietary Guidelines for Americans Advisory Committee (DGAC) used the DRI values as evidence to estimate which nutrients were over- and under-consumed, based on National Health and Nutrition Examination Survey (NHANES) data (USDA and HHS, 2015). Data on over- and under-consumption, combined with nutritional biomarker status data and information about the relationship with chronic disease, were used to identify nutrients of public health concern (e.g., sodium and calcium). The DGAC also used the DRIs to determine whether USDA food patterns are adequate to meet nutrient requirements. Similarly, the Canadian Food Guide promulgates food patterns based on the DRIs. Food patterns[2] are significant in the

[1] See definitions in Chapter 2 (see Box 2-1) and all the values and related reports in http://www.nationalacademies.org/hmd/Activities/Nutrition/SummaryDRIs/DRI-Tables.aspx.

[2] Developing food patterns encompasses an iterative process with the following steps: (1) identifying appropriate energy levels for the patterns, (2) identifying nutritional goals for the patterns, (3) establishing food groupings, (4) determining the amounts of nutrients that would be obtained by consuming various foods within each group, and (5) evaluating nutrient levels in each pattern against nutritional goals (USDA and HHS, 2015).

sense that they identify patterns of eating that would meet known nutrient needs, balance intake from various food groups, and can be used as the basis for nutrition communication to the general population (e.g., the Food Guide Pyramid; MyPlate).

Another important use of DRIs is the Nutrition Facts Panel that appears on packaged food products, which uses the term "Daily Value." The U.S. Food and Drug Administration (FDA) developed Daily Values based on DRIs to inform consumers about the percentage that a serving of a particular food contributes to the requirement for that nutrient. Health Canada took a similar approach to informing consumers with nutrition facts based on DRIs. DRIs also are used by food companies in making nutrient claims (e.g., "free," "high," and "low") about a product's nutritive value (IOM, 2003). Under the Nutrition Labeling and Education Act (NLEA) of 1990 (21 U.S.C. § 343), the U.S. government sets strict rules and definitions that a product must meet to make a nutrient claim. For example, "low in sodium" means a product must contain 140 milligrams of sodium or less per serving (IOM, 2003). Other important uses for DRIs include dietary intake surveillance of populations, government food assistance programs, nationwide health programs (e.g., Healthy People 2020), and food planning for military personnel. For example, DRIs are the basis for the Military RDAs and Military DRIs (MDRIs) through Army Regulation (AR) 40-25. These MDRIs have been used to not only plan meals at military installations but also to plan the Meals Ready-to-Eat (MREs), Survival Rations, and other special rations used in the field (IOM, 1994, 2006a,b).

RATIONALE FOR SETTING CHRONIC DISEASE DIETARY REFERENCE INTAKES

Although nutrient deficiencies are still widespread at a global level, they are less common in the western world and the United States in particular. Even so, in the United States, some essential nutrients are consumed at levels below adequacy (e.g., vitamins A, E, B6, and C), and such nutrition deficiencies have persisted from the 1970s to the 1990s for most age and racial/ethnic groups (CDC, 2012). Figure 1-1 shows the percentage of the U.S. population with usual intakes below the Estimated Average Requirement[3] (EAR). These percentages, derived from NHANES 2007-2010 data, estimate dietary nutrient inadequacy. Prevalence estimates of nutrition deficiencies (based on biomarker concentrations that indicate a risk of disease

[3] Estimated Average Requirement is the average daily nutrient intake observed in an apparently healthy life-stage group. It is based on experimentally derived intake levels or observations of mean nutrient intakes by a group of apparently healthy people who are maintaining a defined criterion of adequacy.

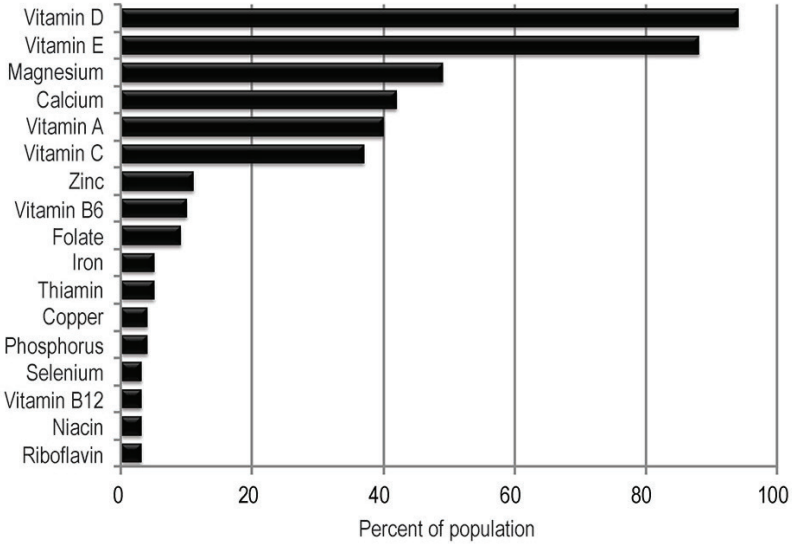


FIGURE 1-1 Percentage of U.S. population with usual intakes below the Estimated Average Requirement. NOTE: Mean intake is estimated directly from the day 1 dietary recall. It does not include nondietary sources, such as skin synthesis of vitamin D. Based on data from What We Eat in America, NHANES 2007-2010.
SOURCE: USDA and HHS, 2015 (Figure D1.1).

rather than on estimates of dietary nutrient inadequacy) among people who live in the United States have not been collected recently, but data from NHANES 2003-2006 have found low levels of deficiency (from 10.5 percent to <1.0 percent). Likewise, data from a 2004 survey show that the majority of Canadians consumed adequate amounts of micronutrients and that the highest prevalence of inadequacy were for magnesium, calcium, vitamin A, and vitamin D (Health Canada, 2012). Possible reasons for this positive outlook are ongoing efforts in enrichment, fortification, and other improvements in the food supply. For example, thanks to fortification policies in the United States, the prevalence of folate deficiency has dropped among women of childbearing age from 10 to 12 percent to less than 1 percent (CDC, 2012).

Today, an even more challenging health problem than nutrient deficiency diseases is chronic diseases.[4] It is estimated that chronic diseases are responsible for 70 percent of all deaths globally. Of these deaths, 82 per-

[4] A chronic disease is "a culmination of a series of pathogenic processes in response to internal or external stimuli over time that results in a clinical diagnosis/ailment and health outcomes" (e.g., diabetes) (IOM, 2010, p. 23).

cent are due to cardiovascular diseases, cancers, respiratory diseases, and diabetes (WHO, 2014). In the United States, half of all adults have at least one chronic health condition such as hypertension, coronary heart disease, stroke, diabetes, cancer and others (Ward et al., 2014). Five of the top 10 causes of death in 2010 were chronic diseases (CDC, 2014). The extent to which specific nutrients or other food substances (NOFSs) contribute to the development of chronic diseases is uncertain, mostly because the etiology of chronic disease is complex (i.e., the result of environment and genetic factors and their interactions) and they develop over long periods of time. However, a robust scientific literature suggests that the contribution of nutrients and diet to the prevention of chronic diseases is important and that many chronic diseases could be prevented, delayed, or alleviated through lifestyle factors, such as changes in diet and exercise (CDC, 2017) and that poor dietary patterns, overconsumption of calories, and physical inactivity are contributors to preventable chronic diseases (USDA and HHS, 2015).

Given the growing understanding of the role of nutrition in chronic disease, a new concept evolved that NOFSs could serve not only to prevent deficiency diseases and toxicity but also to help ameliorate chronic diseases. For example, the evidence suggests that excess intake of some nutrients contributes to some chronic diseases (e.g., sodium and cardiovascular disease), whereas intake of other nutrients that greatly exceed current reference values may help prevent chronic diseases (e.g., omega-3 fatty acids and cardiovascular disease). Past DRI committees had begun to deliberate about including chronic disease in setting DRIs—and did so for a few nutrients (see Chapter 2)—even though the definitions of DRIs were conceived solely with the goals of reaching nutrient adequacy and avoiding toxicity. As more experience was accumulated and public discussions took place, the nutrition community (e.g., researchers and policymakers) recognized that considering chronic diseases raised new challenges and questions for DRI committees, requiring a special focus and specifically oriented guidance. From there, a consensus emerged that a common understanding of the challenges was needed, as were recommendations and guiding principles to drive the activities and approaches of future DRI committees.

In response, in 2015, a multidisciplinary working group sponsored by the Canadian and U.S. government DRI steering committees convened to identify key scientific challenges encountered in the use of chronic disease endpoints to establish DRI values. Their report, *Options for Basing Dietary Reference Intakes (DRIs) on Chronic Disease: Report from a Joint US-/Canadian-Sponsored Working Group* (i.e., the Options Report) (Yetley et al., 2017), outlined and proposed ways to address conceptual and methodological challenges related to three aspects of the work of future DRI committees: (1) What are acceptable levels of confidence that the relationship

between an NOFS and a chronic disease is causal?, (2) If a causal relationship exists, what are acceptable levels of confidence in the data to establish an intake-response relationship and what are approaches for identifying and characterizing the intake-response relationship and, if appropriate, to recommend DRIs?, and (3) What should be the organizational process for recommending chronic disease DRIs?

STATEMENT OF TASK

The statement of task for the current study (see Box 1-1) requests that the committee assess the options presented in the Options Report and determine guiding principles for including chronic disease endpoints for food substances[5] that will be used by future National Academies committees in establishing Dietary Reference Intakes (DRIs). The report uses the term "nutrient or food substance" (NOFS) throughout to reflect the totality of this definition.

The Options Report presented several issues and approaches that this committee was asked to consider in making its recommendations. One issue was whether to continue incorporating chronic disease endpoints into the existing DRI development process or to develop a separate, complementary process. Another issue related to whether the DRI process for considering chronic disease endpoints would continue to consider, separately, food substances that may be interrelated in their apparent causal relationship to a chronic disease or be clustered according to their apparent causal relationships to a single chronic disease. The Options Report also addressed the circumstances under which surrogate markers of disease could be used in place of endpoints based on actual disease incidence, and the level of certainty required for a judgment that an observed association is a causal relationship. *When a causal relationship has been identified with sufficient confidence* specific intake-response relationships need to be defined, and the Options Report laid out six issues where recommendations and guiding principles were needed: (1) the acceptable level of confidence (i.e., low, moderate, or high) in the intake-response relationship that would be used to set reference values, (2) the types of reference values to be established to indicate benefit (i.e., whether similar to the current DRI approaches or based on ranges), (3) types of reference values used to indicate harm (e.g., whether using the current approach to identifying Tolerable Upper Intake Levels [ULs] or developing a new approach), (4) how to address potential overlap between benefits and harms (i.e., whether to avoid any overlap, establish criteria based on degree of risk reduction, or describe

[5] Food substances are nutrients that are essential or conditionally essential, energy nutrients, or other naturally occurring bioactive food components. (Yetley et al., 2017, p. 253S).

BOX 1-1
Statement of Task

An ad hoc committee of the National Academies of Sciences, Engineering, and Medicine (the National Academies) will undertake a study to assess options presented in the document *Options for Consideration of Chronic Disease Endpoints for Dietary Reference Intakes (DRIs): Summary Report from a Joint US-/Canadian-sponsored Expert Panel* (i.e., the Options Report) and determine guiding principles for the inclusion of chronic disease endpoints for food substances that will be used by future National Academies committees in establishing Dietary Reference Intakes (DRIs). The committee shall provide justification for the selection (and non-selection) of options that served as the basis for the guiding principles, including additions not considered in the Options Report. In carrying out its work, the committee shall:

1. Consider that the term food substances for the purposes of this study refers to nutrients that have been established as essential or conditionally essential, energy nutrients, and other naturally occurring bioactive components in foods. Revisions to this definition need to be justified.

2. For the purposes of this study, interpret the terms chronic disease and apparently healthy population (also referred to as general population and healthy population in DRI reports) in a manner consistent with the use of those terms in the Options Report.

3. Organize the guiding principles in a manner consistent with the components of the risk assessment framework (hereafter organizing framework) currently used in developing DRIs. In particular, the guiding principles are to address the initial two organizing framework components that relate to the evidentiary and intake-response key questions addressed in the Options Report:
   a. indicator review and selection, and
   b. intake-response assessment and specification of reference values.

4. Specify, on the basis of its review of the Options Report, guiding principles relevant to the development of chronic disease-based reference values for benefit (disease risk reduction) for the population and for risk (disease risk increase or tolerable upper intake level) for the population or any sub-population that may be susceptible to the food substance. While many principles may be similar for the development of benefit and risk reference values, it is anticipated that there are likely to be differences, often subtle in nature, which will require specification.

5. Ensure that the guiding principles that articulate (a) the evaluation of evidence to assess causal relationships, and (b) development of intake-response relationships and values include, but not be limited to rationale and criteria for:

continued

BOX 1-1 Continued

a. Evaluating the usefulness of different types of studies and data to establish causality and to derive intake-response relationships;
b. Selecting appropriate outcome measures for chronic diseases—for determinations of causality and for determinations of intake-response relationships;
c. Weighing the strength of evidence (degree of confidence) for establishing both causality and intake-response relationships;
d. Evaluating the accuracy and usefulness of intake evidence for assessing causality, intake-response relationships, and population status;
e. Selecting reference values when a food substance is related to more than one chronic disease, possibly with different intake-risk relationships, or for selecting reference values when a chronic disease is related to more than one food substance;
f. Addressing situations where the intake-response curve for benefit overlaps with the potential risk associated with higher intakes; and
g. Types of reference values based on chronic disease endpoints taking into account the wide array of DRI uses.

6. Recognize that recommendations for risk management and policy are outside the scope of this study.

7. Prepare a report written in clear language for use by future DRI committees and other stakeholders. The rationale for principles chosen, and those rejected, should be clear.

the evidence to support user-based decisions), (5) selecting indicators for specifying intake-response relationships (i.e., types of indicators, whether single or multiple, and for single or multiple diseases), and (6) under what circumstances to allow extrapolation for population groups other than those studied (i.e., to similar populations for which the availability of evidence is limited).

APPROACH OF THE COMMITTEE

An ad hoc committee of 12 experts and 1 consultant was selected to respond to the statement of task. Experts were drawn from a broad range of disciplines, including human nutrition, toxicology, biostatistics, major diet-related chronic diseases, preventive medicine, study quality assessment, research methodology, epidemiology, and use of DRIs. Three of the members were among the authors of the Options Report.

During the course of the study, one public meeting and one public

workshop were held to gather data and information in areas requested by the committee, including clarifications from the sponsors related to the statement of task (see Appendix A). Based partly on those clarifications, the committee offers the following key considerations related to interpreting the statement of task.

DRI Framework and Process

The most recent DRI report for vitamin D and calcium (IOM, 2011b) adopted the concept of risk assessment and its components (hazard identification, hazard characterization, intake assessment, and risk characterization) as the organizing framework applicable to the activities in establishing EARs and ULs (see Annex to this chapter for the description of each step as it applies to establishing DRIs). The 2011 report states that adopting risk assessment as a common paradigm for EARs and ULs, and in general for any indicator of interest, will result in a process that is flexible, transparent, and suitable for making decisions related to DRIs. Following this concept and the statement of task (see Box 1-1, item 3), the committee addressed two framework components: (1) indicator review and selection (hazard identification) and (2) intake-response assessment and specification of reference values (hazard characterization).

To fulfill its task, the committee developed recommendations and guiding principles around these two framework components. First, the committee developed recommendations that specifically answer the task of selecting options from the Options Report. Because each NOFS, chronic disease, and their relationships will present idiosyncrasies, the recommendations will be of value only if they are broad and flexible. The guiding principles, in contrast, are meant as a foundation for a scientifically credible chronic disease DRI process. Although recommendations about integrating chronic disease as a consideration in setting DRIs should be revisited in the future as more practice and knowledge are acquired, the guiding principles are meant to withstand scientific and methodological advances that will occur in the future.

The intent of this report is not to change the core DRI process but to provide a degree of consistency and transparency in the approach to making decisions when the process of considering chronic disease endpoints for establishing DRIs is conducted. Therefore, the committee offered its recommendations and guiding principles in the context of the process shown in Figure 1-2, which illustrates actors and selected tasks. Figure 1-2, which is patterned after the general process of guideline review (IOM, 2011a,c), shows that development of DRIs requires the coordinated work of groups of individuals charged with risk assessment and risk management tasks. For example, the U.S. and Canadian DRI steering committees will consider

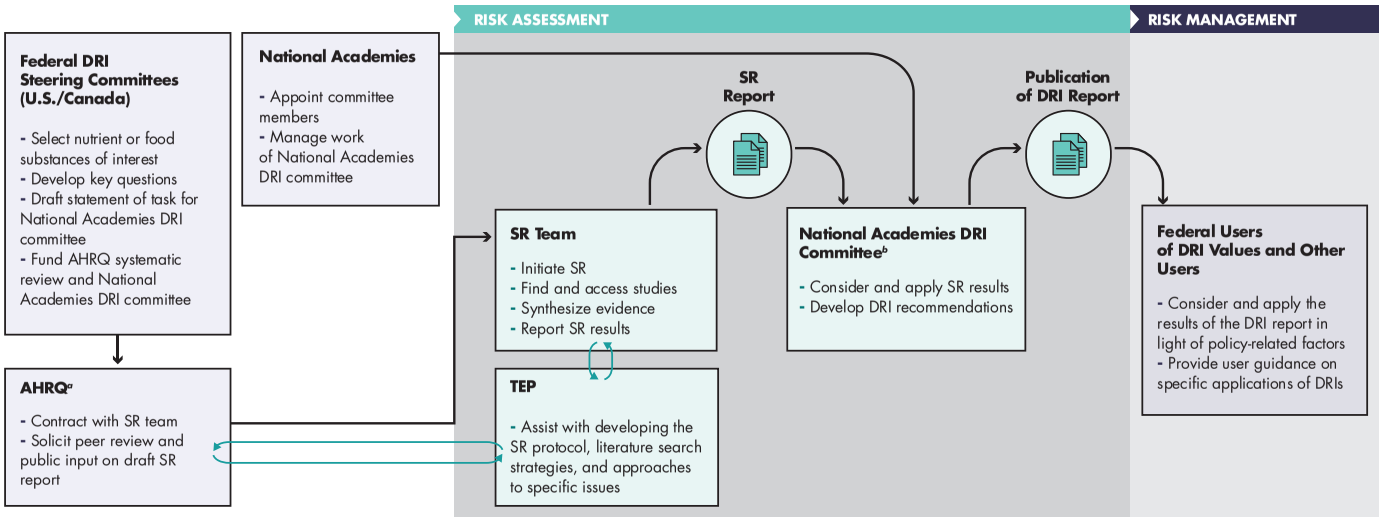


FIGURE 1-2 Overall process for developing DRIs with actions, actors, and their relationships. The risk assessment section indicates elements included in the National Academies DRI committee process and their relationships to the systematic review and prior and subsequent risk management elements.

NOTES: AHRQ = Agency for Healthcare Research and Quality; DRI = Dietary Reference Intake; SR = systematic review; TEP = Technical Expert Panel.

[a] The Agency for Healthcare Research and Quality is listed here as the agency that, in the current DRI process, has been responsible for the systematic review aspects of DRI development.

[b] National Academies DRI committees are convened by and positioned with the Food and Nutrition Board of the National Academies of Sciences, Engineering, and Medicine and operate under the National Academies study committee guidelines, which include an external peer review process of the draft DRI report.

SOURCE: Adapted from IOM, 2011a,c; Taylor, 2008.

many factors when selecting nutrients of interest and formulating statements of task for DRI committees. At the end of the DRI process, federal users of DRI values and other users may consider and apply DRIs in light of policy-related factors in various ways. Although DRI committees should be sensitive to risk management considerations, risk management decisions are not the purview of DRI committees. Also, as shown in Figure 1-2, the committee suggests continuation of the practice started with the 2011 DRIs for calcium and vitamin D, in which formal systematic reviews were conducted by an outside contractor (e.g., an Agency for Healthcare Research and Quality Evidence-based Practice Center) before the formation of the DRI committee. The systematic reviews results were then used by the DRI committee to determine what DRIs were appropriate and to recommend specific levels or ranges. As shown in the figure, systematic reviews designed to inform health-related recommendations, such as DRIs, involve many steps that, for a high-quality result, need to be conducted according to well-established methodologic standards with attention to the most recent updates in these standards (see, e.g., http://handbook.cochrane.org, accessed July 22, 2017, and IOM, 2011c). It includes having guidance from a technical expert advisory committee, an opportunity for peer-review and public comment, and publication of the final systematic review report. The quality of the initial systematic review is critical for an effective DRI committee process. Once initiated, the DRI process may involve additional systematic reviews, as determined by the committee.

In general, guideline development processes require that systematic reviews are formulated, conducted, and interpreted in ways that are maximally responsive to the ultimate objectives (IOM, 2011a) and interactions between systematic review teams and the committee who will eventually use the systematic reviews to develop recommendations are considered critical. Such interactions may occur at the beginning and end of the systematic review process, at various stages during the process, or throughout—in cases where systematic reviews are conducted by members of the guideline review panel. Tradeoffs of different models for such interactions relate to (1) the ability to achieve maximum responsiveness of the systematic review to the guideline panel's questions and needs (which favors more interaction); (2) avoiding undue influence of guideline panel members on the systematic review process (which favors less interaction); and (3) the extent to which the guideline panel can fully understand the systematic review results as they review and interpret them when judging the evidence and making recommendations (which favors at least a minimum level of well-timed interaction). For feasibility or efficiency, the systematic reviews conducted to support DRI committees may be conducted and completed before the DRI committees are appointed. These DRI reviews are conducted by systematic review teams with expertise specific to a nutrient or cluster of

related nutrient. As shown in Figure 1-2, a technical expert panel, which is appointed to assist with developing the systematic review protocol, literature search strategies, and approaches to specific issues, interacts with the systematic review team and also may be represented in the eventual DRI committee membership, providing for overlap. Members of the systematic review team itself also may be represented on the DRI committee as members or consultants. Also of note in Figure 1-2, the final decisions about implementation of DRIs are in the domain of risk management and are, therefore, made by the federal agencies rather than the DRI committees; they draw heavily on the content of the DRI report while also incorporating key policy or programmatic considerations in their deliberations and decisions. In other models, the same committees that use the systematic reviews to evaluate the evidence are also expected to make policy recommendations.

Population of Interest

The phrase "apparently healthy population" (or "general population" or "healthy population") has been used by past DRI committees as a way to define the population covered by the DRIs. It excludes those individuals who (1) have a chronic disease that needs to be managed with medical foods, (2) are malnourished (undernourished), (3) have diseases that result in malabsorption or dialysis treatments, or (4) have increased or decreased energy needs because of disability or decreased mobility. The statement of task requests that this committee use "apparently healthy population" in a manner consistent with the Options Report.

The committee generally agrees that the guiding principles in this report apply to the "apparently healthy population." However, because "apparently healthy population" potentially encompasses a diverse group of individuals with many different health conditions, the committee also highlights the need for DRI committees to characterize the health status of the population in terms of who is included and excluded for each DRI. Specifically, based on the committee's interpretation of the statement of task, the committee recognizes (1) that an "apparently healthy population" includes a substantial proportion of individuals who have obesity and other chronic conditions, such as hypertension or diabetes, (2) that everyone in the population is theoretically at risk of developing such chronic conditions, and (3) that identification of intake levels or ranges related to chronic disease risks has not been the focus of the traditional DRI process. Therefore, this recommendation for a formal approach to adding chronic conditions to the DRI process begins with clarification that the "apparently healthy" population of interest includes people at risk of or with chronic conditions who do not meet the DRI exclusion criteria that exist at that time. Also, the committee recognizes that specific DRIs may be appropriate for certain

subgroups within the apparently healthy population in cases where relevant study design and approaches to risk stratification are considered sufficiently robust to warrant this. With this in mind, DRI committees will address these populations as appropriate.

NUTRIENTS AND OTHER FOOD SUBSTANCES

As the statement of task requests, the committee refers to food substances as "consist[ing] of nutrients that are essential or conditionally essential, energy nutrients, or other naturally occurring bioactive food components." The term "nutrients or other food substances (NOFSs)" is used throughout the report.

NOFSs, however, are not consumed in one single form, but can be consumed in a variety of chemical forms and within matrixes, such as in a dietary supplement or in a fortified food. Reflecting this, the evidence base to evaluate the confidence in the causal relationship between an NOFS and a health outcome will derive from studies where the NOFS is ingested in a variety of chemical forms either as part of a food or as a dietary supplement that might influence health. The chemical form and the matrix might influence its interactions, availability, and bioequivalence. Therefore, consumption of an NOFS as part of a food cannot be assumed to result in equivalent effects as consumption of the same NOFS as a dietary supplement and a scientific evaluation will have to be made about whether the results are comparable.

Out of Scope

In following a risk model paradigm, the process of establishing DRIs includes risk assessment and risk management activities (see Figure 1-2). As mentioned above, the committee's task is limited to providing recommendations and guiding principles in regard to the risk assessment activities within the first two steps of the risk assessment model—risk identification and risk characterization. The committee did not address activities related to formulating the problem or applying the DRIs, which are the purview of the U.S. and Canadian DRI steering committees. Although DRI committees will need to consider health-related risks and benefits and risk reduction goals for the population, other important factors, such as cost, judgments, and values, or implementation factors, are considered the purview of the U.S. and Canadian DRI steering committees and, therefore, also outside of the scope of work of this committee. Because a well-conducted systematic review is essential, this report includes guiding principles related to selected aspects of conducting scientifically rigorous systematic reviews, such as the systematic review protocol.

The committee recognizes that chronic disease DRIs for certain groups might need to be adjusted based on physiological and lifestyle characteristics. Based on the statement of task, which requests that the committee focus on Steps 1 and 2 of the risk assessment framework (see Annex), this report does not address such potential modifications. The committee also recognizes that, in individuals with certain diseases, risk of diseases, or nutrient deficiency diseases, the requirements for nutrients will be different from those for the "apparently healthy population." This report does not address establishing such reference intake levels as those are typically addressed by reference to clinical guidelines for disease management. In addition, this report does not address changes in nutrient requirements in cases where a nutrient may augment the effect of a pharmaceutical. For regulatory purposes, in these cases, the nutrient is considered to be part of the drug (e.g., the possible beneficial effects of higher levels of folate when given in combination with certain pharmaceuticals).

Audience for This Report

The statement of task explains that the guiding principles will be used to include chronic disease endpoints for NOFSs by future National Academies committees as they establish DRIs. The process of establishing DRIs, however, implies the coordinated work of various groups of individuals (e.g., systematic review team, technical expert panel, risk managers, and DRI committees; see Figure 1-2). Therefore, for greatest value and effectiveness of the process, some of the recommendations and guiding principles might need to be considered by the systematic review team, the sponsors or others, as they develop the protocol and tasks for DRI committees.

ORGANIZATION OF THE REPORT

The report is organized into three background chapters (Chapters 1, 2, and 3) and five chapters that directly address the statement of task (Chapters 4, 5, 6, 7, and 8). Chapter 2 describes general aspects of the current DRI process and describes more specifically how chronic diseases have been included in DRIs for some nutrients, and approaches taken by past committees to resolve challenges. Chapter 3 explains the conceptual and methodological challenges when exploring the association of NOFSs with chronic disease. The remainder of the report describes in more detail the conceptual and methodological challenges and justifications for the options selected as well as guiding principles for including chronic disease in the DRI process. In describing the challenges, the reader is frequently directed to the Options Report (see Appendix B) for more extensive explanations. Chapters 4 and 5 describe challenges and approaches regarding the ascertainment of dietary

intake and measurement of health outcomes, respectively. Chapters 6 and 7 aim to answer the two core questions of this report: (1) What are acceptable levels of confidence that the relationship between an NOFS and a chronic disease is causal?, and (2) If a causal relationship exists, what are acceptable levels of confidence in the data to establish an intake-response relationship and what are approaches for identifying and characterizing the intake-response relationship and, if appropriate, to recommend DRIs? (see Chapter 7). Issues related to the DRI organizational process itself are addressed in Chapter 8. Before deliberating the questions in the task, the committee agreed on a set of important definitions (see Appendix D).

Table 1-1 provides a roadmap of the report where the reader can find, by chapter, the questions addressed and the relevant options.

TABLE 1-1 Roadmap of the Report

Chapter	Questions Addressed	Relevant Options (from Yetley et al., 2017)
1: Introduction	Why do we need guiding principles? What is the objective of the guiding principles for establishing DRIs based on chronic disease? What is the task and what subjects are outside of the task? How is the report organized?	N/A
2: Current Process to Establish Dietary Reference Intakes	What are the DRIs and the process to establish them? To whom do they apply? What types of NOFSs and outcomes have DRIs addressed? How are DRIs used in nutrition policy?	N/A

continued

TABLE 1-1 Continued

Chapter	Questions Addressed	Relevant Options (from Yetley et al., 2017)
3: Conceptual and Methodological Challenges in Establishing Chronic Disease Dietary Reference Intakes	What types of conceptual and methodological challenges are associated with assessing NOFS intake and NOFS-chronic diseases research to inform DRIs based on chronic disease outcomes? How might these challenges affect the certainty of judgments about evidence about causal or intake-response relationships between NOFSs and chronic diseases?	N/A
4: Methodological Considerations Related to Assessing Intake of Nutrients of Other Food Substances	How should use of biomarkers of intake and self-report dietary intake methodologies influence ratings of study quality?	N/A
5: Measuring Chronic Disease Outcomes	How should relevant outcomes be selected for inclusion in a systematic review and for inclusion in the review of the evidence? What methodological issues should be considered when assessing the quality of outcome data?	OPTIONS FOR JUDGING THE EVIDENCE • Selecting Chronic Disease Endpoints

TABLE 1-1 Continued

Chapter	Questions Addressed	Relevant Options (from Yetley et al., 2017)
6: Evidence Review: Judging the Evidence for Causal Relationships	What are established approaches for assessing causal factors in relation to chronic diseases generally, and how does one optimally apply these approaches to NOFS-chronic disease questions? How do different study designs potentially contribute to judgments about causal relationships of NOFS intakes or exposures to chronic diseases?	OPTIONS FOR JUDGING THE EVIDENCE • Acceptable Levels of Confidence That the Relation Is Causal
7: Intake-Response Relationships and Dietary Reference Intakes for Chronic Disease	What methods should be used to describe the relationship between NOFS intake and chronic disease? What tools, approaches, or instruments should be used to assess the certainty of the evidence in the data for an intake-response relationship between an NOFS and a chronic disease?	OPTIONS FOR INTAKE-RESPONSE RELATIONSHIPS • Selecting Indicators and Specifying Intake-Response Relations: Qualified Surrogate Disease Markers and Nonqualified Disease Markers • Extrapolating Intake-Response Data • Different Types of Reference Values: Types of Reference Values Associated with Benefit • Different Types of Reference Values: ULs and Reduction in Chronic Disease Risk • Acceptable Level of Confidence in the Intake-Response Data • Overlaps Between Benefits and Harms
8: The Process for Establishing Chronic Disease Dietary Reference Intakes	How should the new DRI process be integrated into the current process of establishing DRIs for adequacy and toxicity? Should the task be to establish DRIs related to individual NOFSs or to NOFSs that are related?	OPTIONS FOR THE NEW DRI PROCESS • Process Components and Options • Starting-Point Issues and Options

REFERENCES

CDC (U.S. Centers for Disease Control and Prevention). 2012. *Second national report on biochemical indicators of diet and nutrition in the U.S. population 2012*. Atlanta, GA: National Center for Environmental Health.

CDC. 2014. *Deaths and mortality*. https://www.cdc.gov/nchs/fastats/deaths.htm (accessed May 16, 2017).

CDC. 2017. *Chronic disease overview*. https://www.cdc.gov/chronicdisease/overview (accessed May 16, 2017).

Health Canada. 2012. *Do Canadian adults meet their nutrient requirements through food intake alone?* http://www.hc-sc.gc.ca/fn-an/alt_formats/pdf/surveill/nutrition/commun/art-nutr-adult-eng.pdf (accessed July 14, 2017).

HHS (U.S. Department of Health and Human Services). 1988. *The Surgeon General's report on nutrition and health*. Washington, DC: Government Printing Office.

IOM (Institute of Medicine). 1994. *Food components to enhance performance: An evaluation of potential performance-enhancing food components for operational rations*. Washington, DC: National Academy Press.

IOM. 2003. *Dietary Reference Intakes: Guiding principles for nutrition labeling and fortification*. Washington, DC: The National Academies Press.

IOM. 2006a. *Mineral requirements for military personnel: Levels needed for cognitive and physical performance during garrison training*. Washington, DC: The National Academies Press.

IOM. 2006b. *Nutrient composition of rations for short-term, high-intensity combat operations*. Washington, DC: The National Academies Press.

IOM. 2010. *Evaluation of biomarkers and surrogate endpoints in chronic disease*. Washington, DC: The National Academies Press.

IOM. 2011a. *Clinical practice guidelines we can trust*. Washington, DC: The National Academies Press.

IOM. 2011b. *Dietary Reference Intakes for calcium and vitamin D*. Washington, DC: The National Academies Press.

IOM. 2011c. *Finding what works in health care: Standards for systematic reviews*. Washington, DC: The National Academies Press.

Murphy, S. P., A. A. Yates, S. A. Atkinson, S. I. Barr, and J. Dwyer. 2016. History of nutrition: The long road leading to the Dietary Reference Intakes for the United States and Canada. *Adv Nutr* 7(1):157-168.

NRC (National Research Council). 1989. *Diet and health: Implications for reducing chronic disease risk*. Washington, DC: National Academy Press.

Taylor, C. L. 2008. *Framework for DRI development: Components "known" and components "to be explored."* Washington, DC.

USDA and HHS. 2015. *Scientific report of the 2015 Dietary Guidelines Advisory Committee*. https://health.gov/dietaryguidelines/2015-scientific-report/PDFs/Scientific-Report-of-the-2015-Dietary-Guidelines-Advisory-Committee.pdf (accessed May 16, 2017).

Ward, B. W., J. S. Schiller, and R. A. Goodman. 2014. Multiple chronic conditions among US adults: A 2012 update. *Prev Chronic Dis* 11:E62.

WHO (World Health Organization). 2014. *Global status report on noncommunicable diseases*. Geneva, Switzerland: World Health Organization.

Yetley, E. A., A. J. MacFarlane, L. S. Greene-Finestone, C. Garza, J. D. Ard, S. A. Atkinson, D. M. Bier, A. L. Carriquiry, W. R. Harlan, D. Hattis, J. C. King, D. Krewski, D. L. O'Connor, R. L. Prentice, J. V. Rodricks, and G. A. Wells. 2017. Options for basing Dietary Reference Intakes (DRIs) on chronic disease endpoints: Report from a joint US-/Canadian-sponsored working group. *Am J Clin Nutr* 105(1):249S-285S.

ANNEX 1-1
STEPS OF THE DRI ORGANIZING FRAMEWORK[6]

Step 1: Indicator Review and Selection

An initial starting point for this report—as for all deliberations based on risk assessment—is the identification and review of the potential indicators to be used. Based on this review, the indicators to be used in developing DRIs are selected. As described within the DRI framework, this step of indicator identification is outlined as follows.

- **Literature reviews and interpretation.** Subject-appropriate and well-done systematic evidence-based reviews as well as other relevant scientific reports and findings serve as a basis for deliberations and development of findings and recommendations for the nutrient under study. De novo literature reviews carried out as part of the study are well documented, including, but not limited to, information on search criteria, inclusion/exclusion criteria, study quality criteria, summary tables, and study relevance to the task at hand consistent with generally accepted methodology used in the systematic review process.
- **Identification of indicators to assess adequacy and excess intake.** Based on results from literature reviews and information gathering activities, the evidence is examined for potential indicators related to adequacy for requirements and the effects of excess intakes of the substance of interest. Chronic disease outcomes are taken into account. The approach includes a full consideration of all relevant indicators, identified for each age, gender, and life stage group for the nutrients under study as data allow.
- **Selection of indicators to assess adequacy and excess intake.** Consistent with the general approach, indicators are selected based on the strength and quality of the evidence and their demonstrated public health significance, taking into consideration sources of uncertainty in the evidence. They are in consideration of the state of the science and public health ramifications within the context of the current science. The strengths and weaknesses of the evidence for the identified indicators of adequacy and adverse effects are documented.

[6] Note: Adapted from the DRI Organizing Framework as described in the 2011 Institute of Medicine report *Dietary Reference Intakes for Calcium and Vitamin D* (IOM, 2011b, pp. 27-29).

Step 2: Intake-Response Assessment and Specification of Reference Values

The intake-response relationships (commonly referred to as dose-response relationships) for the selected indicators of adequacy and excess are specified to the extent the available data allow. If the available information is insufficient, then appropriate statistical modeling techniques or other appropriate approaches that allow for the construction of intake-response curves from a variety of data sources are used. In some instances, most notably for the derivation of UL relative to excess intake, it is necessary to make use of specified levels or thresholds in the absence of the ability to describe a dose-response relationship, specifically a no observed effect level or a lowest observed effect level. Furthermore, the levels of intake determined for adequacy and excess are adjusted as required, appropriate, and feasible by uncertainty factors, variance in requirements, nutrient interactions, bioavailability and bioequivalence, and scaling or extrapolation.

Step 3: Intake Assessment

Consistent with risk assessment approaches, after the reference value is established, based on the information derived from scientific studies, an assessment of the current intake of the nutrient of interest is carried out in preparation for the discussion of implications and special concerns. That is, the known intake is examined in light of the reference value established. Where information is available, an assessment of biochemical and clinical measures of nutritional status for all age, gender, and life-stage groups can be a useful adjunct.

Step 4: Discussion of Implications and Special Concerns

Characterization of the implications and special concerns is a hallmark of the organizing framework. For DRI purposes, it includes an integrated discussion of the public health implications of the DRIs and how the reference values may need to be adjusted for special vulnerable groups within the normal population. As appropriate, discussions on the certainty/uncertainty associated with the reference values are included as well as ramifications of the committee's work that the committee has identified as relevant to its risk assessment tasks.

2

The Current Process to Establish Dietary Reference Intakes

The process that underlies development of Dietary Reference Intakes (DRIs) has evolved over several decades (IOM, 1997, 1998b, 2000b, 2001, 2002/2005, 2005, 2011). In this chapter, the current process is described and an overview of typical applications of the different DRI values is provided. This section is followed by a discussion of the scope of the current DRI process and how chronic disease endpoints have been included as a basis for DRIs to date. This chapter serves as a foundation for understanding the committee's recommendations and guiding principles to establish chronic disease DRIs found in subsequent chapters of this report.

EVOLUTION OF THE DRI PROCESS TO ITS CURRENT STATUS

Brief Overview of the DRI Process

The DRIs are a set of nutrient reference standards that are used for planning and assessing diets of apparently healthy individuals and groups. For each nutrient, the objective is to establish standards for 22 life-stage groups (infants aged 0-6 and 9-12 months; children aged 1-3 and 4-8 years; males and females separately for ages 9-13, 14-18, 19-30, 31-50, 51-70, and ≥70 years; pregnant women aged 14-18, 19-30, and 31-50 years; and lactating women aged 14-18, 19-30, and 31-50 years). DRIs include reference standards for both nutritional adequacy (Estimated Average Requirement [EAR], Recommended Dietary Allowance [RDA], and Adequate Intake [AI]) as well as potential risk of excess nutrient intake (Tolerable Upper Intake Level [UL]). Two additional DRIs have been established—Acceptable

Macronutrient Distribution Range (AMDR) and Estimated Energy Requirement (EER). These six different types of DRI values are defined in Box 2-1. One or more DRIs may be available for a single nutrient. In particular, many nutrients have an EAR and RDA (or AI) and a UL. If adequate data are available, the DRI may have incorporated considerations for

BOX 2-1
Dietary Reference Intakes

The Dietary Reference Intakes (DRIs) were developed to serve as standards for nutrient intake and include the following:

Estimated Average Requirement (EAR): The usual daily intake of a nutrient that is expected to meet the requirement of half of healthy individuals in a group defined by age and sex. The requirement is based on a specific indicator of adequacy.

Recommended Dietary Allowance (RDA): The usual daily intake level that is sufficient to meet the nutrient requirements of 97 to 98 percent of healthy individuals in the specified sex and life-stage group. If the requirements in a specified group are normally distributed, the RDA is equivalent to the EAR plus two standard deviations.

Adequate Intake (AI): When available evidence is not sufficient to determine the EAR for a nutrient, an AI is set. The AI is the average daily nutrient intake observed in an apparently healthy sex and age group. It is based on experimentally derived intake levels or observations of mean nutrient intakes by a group of apparently healthy people who are maintaining a defined criterion of adequacy. It is not certain where an AI level of intake fits relative to an actual nutrient requirement, as no EAR or RDA have been specified for these nutrients. It is generally believed that the AI would be equal to or exceed the RDA (if one existed).

Tolerable Upper Intake Level (UL): The highest usual daily nutrient intake level that is likely to pose no risk of adverse effects to nearly all healthy individuals in the specified sex and life-stage group.

Acceptable Macronutrient Distribution Range (AMDR): A range of usual intakes for a macronutrient that is associated with reduced risk of chronic disease while providing adequate intakes of essential nutrients. An AMDR is expressed as a percentage of total energy intake.

Estimated Energy Requirement (EER): A calculated level of energy intake that is estimated to maintain energy balance (or as appropriate, normal growth), that incorporates weight, height, physiological state (i.e., pregnancy), and level of energy expenditure.

SOURCES: IOM, 2000b, 2002/2005.

reducing the risk of chronic disease (IOM, 2003a). The experience to date with incorporating evidence related to chronic disease in DRIs, which is described later in this chapter, is relevant in the context of developing DRIs for reaching adequacy or preventing toxicity but not fully applicable for development of "chronic disease DRIs" with the goal of preventing chronic disease. This report focuses on the potential for developing DRIs for which the primary focus would be on reducing chronic disease risk rather than ancillary to considerations of adequacy or toxicity.

The setting of quantitative nutrient intake reference values in the United States and Canada began in the 1930s. In Canada, the Dietary Standards/Recommended Nutrient Intakes (RNIs) were published from 1938 until 1990, and in the United States, Recommended Dietary Allowances were published from 1941 until 1989 by the Food and Nutrition Board of the National Academy of Sciences. The current DRI process was initiated in the early 1990s, in recognition of the expanded uses of dietary reference values and newer insights into the role of nutrients. With sponsorship primarily from U.S. and Canadian federal agencies, volunteer expert panels (i.e., DRI committees) of Canadian and U.S. scientists and subcommittees in relevant disciplines (e.g., human nutrition, epidemiology, toxicology) are convened based on a selection process that considers suggestions by stakeholders, and conflicts of interests and biases following the policies of the National Academies of Sciences, Engineering, and Medicine (the National Academies). In general, DRI recommendations result from 1 to 2 years of DRI committee deliberations. Table 2-1 lists the DRI reports that have been published to date as well as other key publications related to the development of the DRI process.

A cornerstone of the current thinking of the role of DRIs in nutrition policy was a 2007 Institute of Medicine workshop called The Development of the DRIs 1994-2004 (IOM, 2008). The workshop explored emerging challenges and issues in the process of establishing DRIs with the goal of gathering ideas for improving the process in the future and as scientific knowledge expands.

Traditionally, a major consideration in the DRI process has been nutritional adequacy. For nutrients deemed nutritionally essential for normal physiological functioning (i.e., the nutrient cannot be synthesized in the body, or cannot be synthesized in sufficient amounts to meet needs and thus must be provided in the diet), the scientific literature is reviewed by the DRI committee to determine the most appropriate indicator of adequacy that will be used to set the requirement for the nutrient. Possible indicators of adequacy could include prevention of signs or symptoms of a nutrient deficiency disease, biomarkers of the nutrient's function (e.g., activity of an enzyme that uses the nutrient as a cofactor), and biomarkers of body stores of the nutrient. Ideally, intake-response data are available for the selected

TABLE 2-1 Chronology of DRI Publications

Year	DRI Publication Title	Reference
1994	How Should the Recommended Dietary Allowances Be Revised?	IOM, 1994
1997	Dietary Reference Intakes for Calcium, Phosphorus, Magnesium, Vitamin D, and Fluoride	IOM, 1997
1998	Dietary Reference Intakes: A Risk Assessment Model for Establishing Upper Intake Levels for Nutrients	IOM, 1998a
1998	Dietary Reference Intakes for Thiamin, Riboflavin, Niacin, Vitamin B6, Folate, Vitamin B12, Pantothenic Acid, Biotin, and Choline	IOM, 1998b
2000	Dietary Reference Intakes: Applications in Dietary Assessment	IOM, 2000a
2000	Dietary Reference Intakes for Vitamin C, Vitamin E, Selenium, and Carotenoids	IOM, 2000b
2001	Dietary Reference Intakes for Vitamin A, Vitamin K, Arsenic, Boron, Chromium, Copper, Iodine, Iron, Manganese, Molybdenum, Nickel, Silicon, Vanadium, and Zinc	IOM, 2001
2002/2005	Dietary Reference Intakes for Energy, Carbohydrate, Fiber, Fat, Fatty Acids, Cholesterol, Protein, and Amino Acids	IOM, 2002/2005
2003	Dietary Reference Intakes: Applications in Dietary Planning	IOM, 2003a
2003	Dietary Reference Intakes: Guiding Principles for Nutrition Labeling and Fortification	IOM, 2003b
2005	Dietary Reference Intakes for Water, Potassium, Sodium, Chloride, and Sulfate	IOM, 2005
2006	Dietary Reference Intakes Research Synthesis: Workshop Summary	IOM, 2006a
2006	Dietary Reference Intakes: The Essential Guide to Nutrient Requirements	IOM, 2006b
2008	The Development of DRIs 1994–2004: Lessons Learned and New Challenges: Workshop Summary	IOM, 2008
2011	Dietary Reference Intakes for Calcium and Vitamin D	IOM, 2011

indicator of adequacy, so that the amount of the nutrient that meets the requirements of half the members of a specified sex and life-stage group can be identified. If such data are available, an EAR is identified and used to establish an RDA (an intake level that meets the needs of nearly all members of a sex and life-stage group, i.e., two standard deviations above the

mean requirement). If intake-response data are not available and an EAR and RDA cannot be established, other types of data (e.g., average intakes of a healthy group of people) are used to set an AI that is used as a recommended intake level.

It is important to recognize that the EAR for a given nutrient could differ depending on the indicator of adequacy that is selected. For example, for iron, possible indicators of adequacy could include prevention of anemia (as identified by a specified hemoglobin level) or maintenance of a certain level of iron stores (as identified by a specified serum ferritin level). If an EAR was established based on maintaining iron stores, it would be higher than an EAR based on prevention of iron-deficiency anemia.

The DRI process also considers risks of adverse effects from excessive intakes. Thus, many nutrients that have an EAR and RDA or AI also have a UL, which represents a maximal daily intake level that is unlikely to lead to adverse health effects when consumed habitually. ULs are set using a risk assessment framework. This involves identifying any adverse effects of high intakes of the nutrient, where "adverse effect" includes impairment of any physiologically important function as well as any detrimental effect of the nutrient on the health benefits of another nutrient (i.e., an adverse nutrient-nutrient interaction). Intake-response data are examined to identify a no-observed-adverse-effect level (NOAEL), which is the highest intake that does not result in adverse effects in any of the individuals studied. If it is not possible to identify a NOAEL, a lowest-observed-adverse-effect level (LOAEL) is identified as the lowest intake at which an adverse effect was observed. An uncertainty factor (UF) reflecting uncertainty associated with extrapolating from the observed data to the general population is then selected. The magnitude of the UF will vary among nutrients, depending on factors such as individual variability in susceptibility to the adverse effect, extrapolation from experimental animals to humans, use of a LOAEL rather than a NOAEL, use of data reflecting subclinical versus chronic exposure, and the severity or irreversibility of the adverse effect. The UL is then set by dividing the NOAEL or LOAEL by the UF. At present, ULs have been identified only for some essential nutrients (i.e., vitamin A, vitamin C, vitamin D, vitamin E, niacin, vitamin B6, folate, choline, calcium, copper, fluoride, iodine, iron, magnesium, manganese, molybdenum, phosphorus, selenium, zinc, sodium, and chloride). ULs also have been set for nickel, vanadium and boron, which are not considered nutritionally essential.

In contrast to the previous method of establishing U.S. RDAs and Canadian RNIs, the current DRI process, which began in the early 1990s also incorporates consideration of chronic diseases. This will be apparent in some of the examples described below.

Derivation of the Six Types of Dietary Reference Intakes

Figure 2-1 illustrates the relationship among the DRIs and the risks of nutrient inadequacy or excess. At extremely low intakes, the risk of inadequacy is 100 percent: Everyone would fail to meet the requirement for the specified indicator of adequacy for the nutrient. As intake increases, the risk of inadequacy decreases: it is 50 percent when intake equals the EAR, and diminishes to near zero (about 2 to 3 percent) when intake equals the RDA. The AI is set when the necessary intake-response data are not available to establish an EAR or an RDA. Although the AI is not shown in Figure 2-1, it is assumed that it would be at or above the RDA if an RDA could be calculated, as it is estimated to meet the needs of almost all healthy individuals. As intakes increase above the RDA (or AI), there is a range of intake that is associated with neither further reductions in risk of inadequacy nor any increase in the potential risk of excess. However, as intake increases above the UL, risk of adverse effects may increase. It should be noted that much less is known about the "shape" of the risk curve for the adverse effects

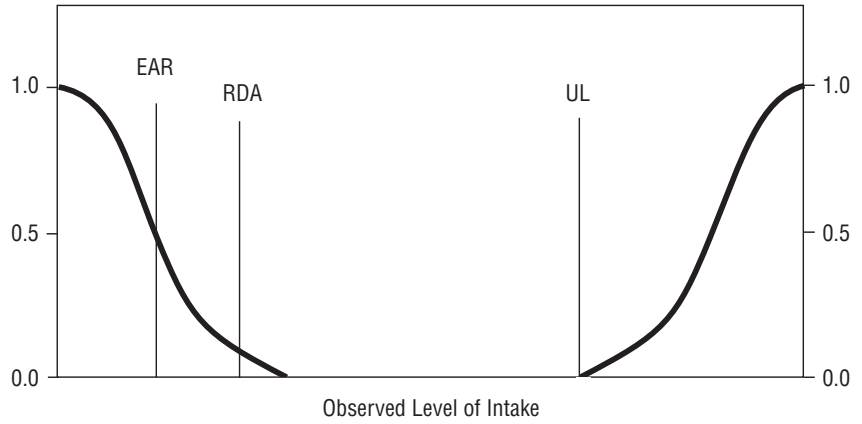


FIGURE 2-1 Relationship between Dietary Reference Intakes (DRIs). This figure shows that the Estimated Average Requirement (EAR) is the intake at which the risk of inadequacy is 0.5 (50 percent) to an individual. The Recommended Dietary Allowance (RDA) is the intake at which the risk of inadequacy is very small—only 0.02 to 0.03 (2 to 3 percent). The Adequate Intake (AI) (not shown in figure) does not bear a consistent relationship to the EAR or the RDA because it is set without the estimate of the requirement. At intakes between the RDA and the Tolerable Upper Intake Level (UL), the risks of inadequacy and of excess are both close to zero. At intakes above the UL, the risk of adverse effects may increase.
SOURCE: IOM, 2006b.

of excessive intake than the risk curve for inadequacy. The point where risk actually does increase likely varies among nutrients, depending on the magnitude of the UF used to set the UL. Furthermore, it is not possible to identify intake levels where a given proportion of a group experiences the adverse effect used to set the UL (whereas it is possible to identify intake levels where a given proportion of a group do not meet the requirement for the indicator of adequacy used to set the EAR).

The AMDR indicates a range of carbohydrate, protein, or fat intake within which essential nutrient needs could be met without increasing the risk of chronic disease. Finally, the EER indicates the level of energy intake that is predicted to maintain energy balance (or, as appropriate, normal growth). The EAR, RDA, AI, UL, AMDR, and EER can be used to assess the probability of intake adequacy or potential risk of excess or to plan for appropriate intakes of individuals or groups.

Steps in the DRI Process

The DRI process generally involves four overarching steps as part of a risk assessment framework (see also Figure 1-2 in Chapter 1). Each step is described briefly in Box 2-2. These steps describe the ideal DRI process, in which the desired evidence would be available. In reality, the process is typically not so straightforward because of gaps in the data and variation in the type and amount of evidence for each nutrient. However, because of their importance to health, establishing reference values for adequacy for essential nutrients has been considered necessary, regardless of the certainty in the evidence. In the past, therefore, the basis for nutrient adequacy has varied for each nutrient depending, in part, on the availability of data that allowed estimation of an EAR. Further examples and considerations are provided later in this chapter.

Applications of the DRIs

The DRIs are used for a variety of nutrition-related objectives, as summarized in Table 2-2. Applications include providing a reference point for assessing the nutrient intake distribution of populations, as is done to develop the *Dietary Guidelines for Americans* (HHS and USDA, 2015), and to provide information for consumer evaluation of food products, such as for food labels. The DRIs also may be used to estimate the effect of altering the food supply on population intakes, such as for food fortification or when a new food product or ingredient is proposed for addition to the food supply.

BOX 2-2
Steps in the DRI Process

Step 1: Indicator Review and Selection

In this report, the term "indicator" broadly refers to clinical endpoints, surrogate endpoints, biomarkers, or risk factors for a chronic disease—all of which are measures that may serve as the basis for estimating nutrient intake requirements or excessive levels of nutrient intake that might result in adverse health effects. Examples of various types of indicators are presented in Figure 2-2. The indicator may reflect a desirable (e.g., tissue saturation) or undesirable (e.g., high blood pressure) outcome and may reflect a level of exposure (another term for intake), a mechanism or functional outcome related to the chronic disease, a physiological effect of nutrient intake that is correlated with a chronic disease, or the actual clinical (or health) outcome that is a marker of the disease (e.g., dental caries are a clinical outcome, where risk is increased by inadequate fluoride intake).

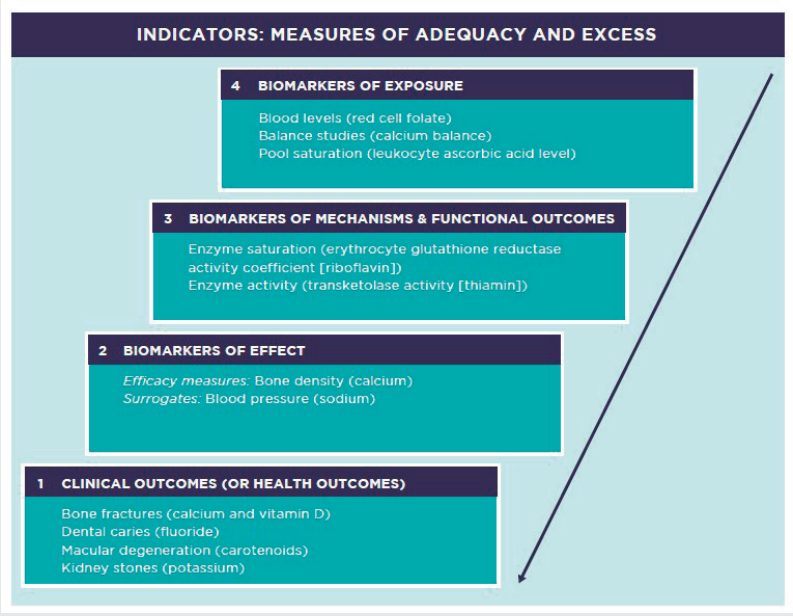


FIGURE 2-2 Measures reflective of indicators for nutrient substances.
NOTES: Numbering and arrows reference hierarchical proximity to the clinical outcome of interest. Blood pressure is a surrogate for cardiovascular disease and bone density is a surrogate for fracture risk.
SOURCE: Adapted from Taylor, 2008.

Selecting an appropriate indicator for a given nutrient and DRI (e.g., EAR versus UL) starts with a review and interpretation of published literature by the DRI committee. Before the 2011 DRI report on calcium and vitamin D (IOM, 2011), this evidence review was done through informal literature searches on the nutrients under consideration. Starting with the 2011 report, formal and exhaustive systematic evidence-based reviews were done for calcium and vitamin D by an outside contractor (e.g., Agency for Healthcare Research and Quality [AHRQ] Evidence-based Practice Center) before the formation of the DRI committee. The results of this systematic review were used by the DRI committee to establish the DRIs. This new, more transparent method set the standard for the DRI process going forward.

Indicators are identified that allow for assessment of nutrient intake adequacy or excess. Chronic disease outcomes also may be taken into account. Finally, indicators are selected for use, with consideration of the state of the science, the strengths and weaknesses of each, and the level of uncertainty. Further discussion of the level of certainty in the indicator is found in Step 2.

Step 2: Intake-Response Assessment and Specification of Reference Values

Once indicators are selected, the intake-response relationships are determined. The responses (outcomes) may include nutrient adequacy, excess, or relationship to a chronic disease outcome. Due to limitations of the available data, characterizing the relationship often requires establishing thresholds at which no effect or a small effect is observed. Adjustments to the available data by accounting for uncertainty, variance in requirements, nutrient interactions, bioavailability and bioequivalences, and scaling or extrapolation may be required. This step culminates in the identification of reference values (DRIs) for all 22 life-stage groups.

Step 3: Intake Assessment

The reference values determined in Step 2 are then compared to current levels of intake. A variety of different national survey data are used for this comparison, including the National Health and Nutrition Examination Survey, the Continuing Survey of Food Intake by Individuals, and the Canadian Community Health Survey. This comparison is useful for discussions related to implications of the determined reference values and special concerns (see Step 4).

Step 4: Discussion of Implications and Special Concerns

Comparing the reference values to intakes allows for a discussion of the public health implications of the reference values and adjustments that may be needed for special groups. In this section, the certainty or uncertainty surrounding the DRI can be described, as well as any other ramifications that the DRI committee has identified as relevant to the overall risk assessment (see later section on "Population groups with special concerns").

SOURCES: Adapted from IOM, 2011, and Taylor, 2008.

TABLE 2-2 Common Applications of Dietary Reference Intake (DRI) Values for Populations and for Individuals

Application	Description	Example	DRI
Population-Level Applications			
Food labeling	Use (in part) to calculate the proportion of recommended intake provided by a food product	The % DV (daily value) (FDA, 2013b)	RDA, AI, UL
Food fortification	Evaluate the effect of adding a nutrient to a staple food, with the intent of improving population-wide nutrient intakes and reducing prevalence of a health risk	Health Canada is examining whether changes to the current vitamin D fortification policy are needed, and modeling has been used to assess the impact of adding vitamin D to various foods on the percent less than the EAR and greater than the UL (CIHR, 2016)	EAR, UL
Federal supplemental food program planning	Assess the nutrient adequacy of specific population groups to determine what supplemental foods should be provided	USDA child nutrition programs (USDA, 2016)	EAR, AI, UL
Research	Design and evaluate data from human studies; analyze dietary intake data	Planning diets for intervention studies (individual-level application); NHANES or CCHS data analysis (population-level application)	EAR, AI, RDA, UL
Product development or modification of existing foods	Evaluate the effect of adding an ingredient to a food for a non-nutrient purpose, such as preservation	Ascorbic acid (vitamin C) (FDA, 2013a)	RDA, AI, UL

TABLE 2-2 Continued

Nutrition surveillance	Assess the prevalence of nutrient inadequacy and potentially excessive intake in a population by sex/life-stage group	NHANES (United States), CCHS (Canada)	EAR, UL, AMDR
Individual-Level Applications			
Development of dietary guidance	Develop recommended food intake patterns to ensure that individuals meet recommendations for nutrient intake, with consideration of typical food intake patterns	*The Dietary Guidelines for Americans* (HHS and USDA, 2015); *Eating Well with Canada's Food Guide*[a] (Health Canada, 2007); Army Regulation 40-25 (U.S. Army, 2017)	EAR, AI, RDA, UL, AMDR
Clinical assessment	Determine the nutrient adequacy of an individual by using a serum marker of nutrient intake	DRI for serum vitamin D (IOM, 2011)	EAR

NOTES: AI = adequate intake; AMDR = acceptable macronutrient distribution range; CCHS = Canadian Community Health Survey; DRI = Dietary Reference Intake; EAR = Estimated Average Requirement; NHANES = National Health and Nutrition Examination Survey; NTD = neural tube defect; RDA = Recommended Dietary Allowance; UL = Tolerable Upper Intake Level; USDA = U.S. Department of Agriculture.

[a] Canada's Food Guide was developed using a modeling process that enables it to be used to plan diets of groups as well as individuals (Katamay et al., 2007).

SCOPE OF THE CURRENT DRI PROCESS

Food Substances Addressed Using the DRI Process

DRIs have been established for all the nutrients that are considered essential (vitamins, minerals, water, electrolytes, carbohydrate, protein, total fat, linoleic acid, and linolenic acid). Essential nutrients cannot be synthesized by the body (or are synthesized in insufficient amounts), but are required for normal human physiological function.

In contrast to earlier efforts to establish dietary reference values, the DRI committees have also explored the possibility of setting DRIs for non-essential food components that are found naturally in foods (also called non-nutrients), but which may have a meaningful impact on human health. For example, dietary fiber is not an essential nutrient; however, it was included and evaluated in the 2002/2005 DRI report. The DRI com-

mittee found sufficient evidence to set an AI for total fiber in foods, based on the intake level shown to protect against coronary heart disease (IOM, 2002/2005) in men and women.

A wide variety of other non-essential nutrients also have been evaluated by the various DRI committees, but they have been unable to set EAR or AI values for any of them. For example, some evidence suggested that arsenic, boron, nickel, silicon, and vanadium play a beneficial role in human health, but no EAR or AI have been set because no clear and consistent evidence of their metabolic role has been available (IOM, 2001). In the case of saturated fatty acids, monounsaturated fatty acids, *trans* fatty acids, cholesterol, beta carotene, and other carotenoids, DRI committees have concluded that no EAR or AI values could be determined because of insufficient evidence that these food components are considered essential to human health (IOM, 2000b, 2002/2005). The DRI committee on macronutrients have considered long-chain omega-3 fatty acids (EPA and DHA) and stated that they could contribute to meeting the AMDR for the essential n-3 and n-6 fatty acids (linoleic and α-linolenic acids) (IOM, 2002/2005); however, no reference value has been set for them.

With some exceptions, DRI evaluations and recommendations have been based on intake of the nutrient or food component from foods. In some instances, DRI committees have evaluated intake from fortified foods and/or dietary supplements. The EAR for folate, for example, is based on the amount of folate from foods and folic acid from fortified foods or dietary supplements, measured as dietary folate equivalents, needed to maintain erythrocyte folate (IOM, 1998b). Another example is the vitamin B12 recommendation for individuals older than age 50 years, who may be unable to absorb naturally occurring vitamin B12. The DRI committee advised that this population group meets its requirement by consuming foods fortified with vitamin B12 or by taking a supplement that contains vitamin B12 (IOM, 1998b).

The UL for a number of nutrients has been based on intake from dietary supplements, such as the UL for vitamin E, which applies to intake of all forms of synthetic vitamin E found in dietary supplements, fortified foods, and pharmacological agents (IOM, 2000b). Similarly, the UL for niacin and for folic acid applies to synthetic forms found in dietary supplements and/or fortified foods, and the UL for magnesium has been based only on intake from pharmacologic agents (IOM, 1997, 1998b). Generally, this is done when adverse effects have been observed only with supplemental or synthetic sources of the nutrient (versus sources that occur naturally in foods).

Incorporating Considerations for Chronic Disease

To date, chronic disease endpoints have been used as the indicator to establish a DRI for six nutrients: calcium, fluoride, potassium, sodium, total fiber, and vitamin D. AIs were set for fluoride, potassium, sodium, and total fiber, and EARs were set for calcium and vitamin D. The specific indicators that were selected as well as those that were considered during the DRI process for these nutrients are shown in Table 2-3. It should be noted that, in most cases, it was not possible to identify an EAR based on chronic disease risk, so AIs were set instead. In 2011, EARs and RDAs were set for calcium and vitamin D. However, for most age groups, the calcium EAR and RDA were based on intake-response data for calcium balance, which while clearly related to bone health, is not a specific chronic disease outcome (IOM, 2011). Chronic disease outcomes or surrogates (e.g., fracture risk, bone mineral density) were used to set the EAR and RDA for older adults, but the committee commented that the absence of good intake-response data made it challenging to clearly identify a requirement in these age groups. The vitamin D EAR and RDA also warrant comment, as they were set based on the amounts of dietary vitamin D (in the absence of sunlight exposure) required to achieve a specified serum level of 25(OH)D. Intake-response data on serum 25(OH)D levels and bone health outcomes were used to identify the "required" serum 25(OH)D levels used for the EAR and RDA (IOM, 2011).

USE OF EVIDENCE TO DEVELOP DRIs: TYPES AND APPLICATIONS

Types of Indicators Used to Develop DRIs

Selecting an appropriate indicator (i.e., a variable for determining adequacy or excess) is the first step in the DRI development process. As previously discussed, indicators may assess nutrient adequacy (EAR or AI) or excess (UL). DRI committees have also taken into account potential reductions or increases in chronic disease outcomes. A wide variety of types of indicators have been used to set the DRIs, including clinical indicators (signs of deficiency, altered body composition, impaired function, or increased morbidity), nutrient balance studies, biochemical measures (blood or urine levels), functional measures (bone health, hormone levels), risk of developmental abnormalities, or risk of chronic disease outcomes (Taylor, 2008). For example, the EAR for vitamin A was based on ensuring adequate liver stores of the vitamin (IOM, 2001). A combination of indicators often has been used. The DRI for copper was based on plasma copper and ceruloplasmin concentrations, erythrocyte superoxide dismutase activity,

TABLE 2-3 Current Dietary Reference Intakes (DRIs) Linked to Chronic Disease and/or Surrogate Endpoints

Nutrient	Reference Value	Indicator Selected as the Basis for Establishment of Chronic Disease-Related DRI (AI, EAR, UL)	Indicators Considered, But Not Selected	
			Biomarkers of Nutrient Adequacy	Surrogate Endpoints
Calcium	EAR	Bone health[c]	Serum calcium Calcium balance[c] Calcium absorption	Parathyroid hormone Bone mineral content/bone mineral density[c] Hypertension

Chronic Disease	Types of Studies[a]	Considerations for Other Populations[b]
Cancer/neoplasms (all cancers, breast cancer, colorectal cancer/colon polyps, prostate cancer) Cardiovascular diseases Diabetes (type 2) and metabolic syndrome Falls Fracture risk Immune responses (asthma, autoimmune disease) Diabetes (type 1) Inflammatory bowel and Crohn's disease Multiple sclerosis Rheumatoid arthritis Systemic lupus erythematosus Infectious diseases Tuberculosis Influenza/upper respiratory infections Neuropsychological functioning (autism, cognitive function, expression) Physical performance Preeclampsia of pregnancy and other non-skeletal reproductive outcomes Rickets/ osteomalacia	Balance studies RCTs Observational studies (ecological, cross-sectional, case-control, cohort)	The committee noted that different ethnic/ racial groups respond to calcium and vitamin D in some biologically different ways, most notably among those of African American ancestry. However, the available data were too limited to permit the committee to assess whether separate, quantitative reference values for such groups are required.

continued

TABLE 2-3 Continued

Nutrient	Reference Value	Indicator Selected as the Basis for Establishment of Chronic Disease-Related DRI (AI, EAR, UL)	Indicators Considered, But Not Selected	
			Biomarkers of Nutrient Adequacy	Surrogate Endpoints
	UL	Incidence of kidney stones		Vascular and soft tissue calcification Interactions with iron and zinc
Vitamin D[d]	EAR	Bone health (operationalized as the intake required, in the absence of sunlight exposure, to achieve serum 25(OH)D levels consistent with desirable changes in bone density and fracture risk)	Serum 25(OH)D level	Serum 25(OH)D level

Chronic Disease	Types of Studies[a]	Considerations for Other Populations[b]
Hypercalcemia Hypercalciuria Prostate cancer Constipation		
Cancer/neoplasms (all cancers, breast cancer, colorectal cancer/colon polyps, prostate cancer) Cardiovascular diseases Diabetes (type 2) and metabolic syndrome Falls Fracture risk[c] Immune responses (asthma, autoimmune disease) Diabetes (type 1) Inflammatory bowel and Crohn's disease Multiple sclerosis Rheumatoid arthritis Systemic lupus erythematosus Infectious diseases Tuberculosis Influenza/upper respiratory infections Neuropsychological functioning (autism, cognitive function, depression) Physical performance Preeclampsia of pregnancy and other non-skeletal reproductive outcomes Rickets/ osteomalacia	RCTs Observational studies (ecological, cross-sectional, case-control, cohort)	The committee noted that different ethnic/ racial groups respond to calcium and vitamin D in some biologically different ways, most notably among those of African American ancestry; however, the available data were too limited to permit the committee to assess whether separate, quantitative reference values for such groups are required.

continued

TABLE 2-3 Continued

Nutrient	Reference Value	Indicator Selected as the Basis for Establishment of Chronic Disease-Related DRI (AI, EAR, UL)	Indicators Considered, But Not Selected	
			Biomarkers of Nutrient Adequacy	Surrogate Endpoints
	UL[d]	Hypercalcemia and related toxicity Emerging evidence for all-cause mortality, some cancers, cardiovascular risk, falls and fractures		
Sodium	AI[e]	Replenish losses of sodium needs of moderately active, apparently healthy individuals Based on ensuring adequate intake of other nutrients	Sodium balance Chloride balance Serum concentration	Blood pressure Plasma renin activity Blood lipids concentration Insulin resistance

Chronic Disease	Types of Studies[a]	Considerations for Other Populations[b]
Hypercalciuria Hypercalcemia (infants) Cancer Cardiovascular risk Falls and fractures All-cause mortality	Observational studies (ecological, cross-sectional, case-control, cohort)	For infants, UL was based on retarded linear growth
	Balance studies RCTs with feeding or behavioral interventions Observational studies	AI for 0-12 months was based on mean sodium intake. For 1-18 years, AI is extrapolated down based on energy intake. For those older than 50 years, AI is extrapolated down based on energy intake. Evidence was insufficient to set levels for: -pregnancy -lactation

continued

TABLE 2-3 Continued

Nutrient	Reference Value	Indicator Selected as the Basis for Establishment of Chronic Disease-Related DRI (AI, EAR, UL)	Indicators Considered, But Not Selected: Biomarkers of Nutrient Adequacy	Indicators Considered, But Not Selected: Surrogate Endpoints
	UL	Continuous and progressive increase in blood pressure (for CVD) with increases in sodium intake. The LOAEL as applied to dietary sodium (2.3 g) is a point on the continuous relationship with blood pressure that corresponds with the next level above the AI that was tested in dose-response trials.	Calcium excretion, Bone mineral density	Left ventricular mass Kidney stones
Total fiber	AI[e]	Intake level found to protect against coronary heart disease. Reduction in risk of diabetes was used as a secondary endpoint to support the AI.	Fiber intake, satiety and weight maintenance	Blood pressure Hyperlipidemia Glucose tolerance insulin response
	No UL set for total fiber			

Chronic Disease	Types of Studies[a]	Considerations for Other Populations[b]
Stroke Coronary heart disease Pulmonary function Gastric cancer		
Colon health (constipation, laxation, fecal weight; fiber fermentation products—energy source for colon; prevention of diverticular disease) Colon cancer Breast cancer Other cancers (endometrial, ovarian) Diabetes	Epidemiological (prospective cohorts), mechanistic, and clinical data, intervention trials	No AI for infants AI for children is extrapolated from adult AI, based on energy intake. Evidence was insufficient to set levels for: -pregnancy -lactation -sex

continued

TABLE 2-3 Continued

Nutrient	Reference Value	Indicator Selected as the Basis for Establishment of Chronic Disease-Related DRI (AI, EAR, UL)	Indicators Considered, But Not Selected	
			Biomarkers of Nutrient Adequacy	Surrogate Endpoints
Fluoride	AI[e]	Prevention of dental caries	Fluoride balance	Bone mineral content
	UL	UL for infants and children ages 0-8 years is based on risk of enamel fluorosis (two studies from 1937 and 1942). UL for all age groups (>8 years), and pregnant or lactating women is based on risk of skeletal fluorosis (clinical studies).	Enamel fluorosis	

Chronic Disease	Types of Studies[a]	Considerations for Other Populations[b]
Dental caries	Epidemiological studies (observational)	AI for 0-12 months was based on mean fluoride intake. Evidence was insufficient to set levels for: -pregnancy -lactation -sex
	Clinical studies	

continued

TABLE 2-3 Continued

Nutrient	Reference Value	Indicator Selected as the Basis for Establishment of Chronic Disease-Related DRI (AI, EAR, UL)	Indicators Considered, But Not Selected	
			Biomarkers of Nutrient Adequacy	Surrogate Endpoints
Potassium	AI[e]	Level of intake from foods that maintains lower blood pressure levels, reduces the adverse effects of sodium chloride intake on blood pressure, reduces risk of kidney stones, and possibly reduces bone loss	Potassium balance Serum potassium concentration Hypokalemia Bone mineralization	Blood pressure Kidney stones Bone loss
	UL	No UL was set for potassium intake from food.		

NOTES: AI = adequate intake; CVD = cardiovascular disease; DRI = Dietary Reference Intake; EAR = Estimated Average Requirement; LOAEL = lowest-observed-adverse-effect level; RCT = randomized controlled trial; UL = Tolerable Upper Intake Level. Systematic reviews, conducted by an outside contractor, were used by the DRI committee to establish DRIs for calcium and vitamin D (IOM, 2011).

[a] General statements were made about studies being adequately powered and about the study quality considerations, but no inclusion/exclusion criteria were provided (e.g., studies seem to be included regardless of the dietary intake assessment tool used) or quality assessment included.

Chronic Disease	Types of Studies[a]	Considerations for Other Populations[b]
CVD Impaired pulmonary function	Epidemiological studies (observational) Metabolic studies Intervention studies	AI for 0-12 months was based on mean potassium intake. AI for children age 1-18 years is extrapolated from adult AI based on median energy intakes. Evidence was insufficient to set levels for: - sex - race/ethnicity - pregnancy - those on a high protein diet -individuals on diuretic therapies -individuals with predisposition to hyperkalemia -individuals older than 50 years of age

[b] Criteria for nutrition adequacy might differ for different ages and justification is described; it is not basal or normative, as in other reports. Levels are based on reducing the risk of developing a negative condition associated with the nutrient for the apparently healthy population, not for those that are malnourished or, in certain disease states, marked by increased nutrient needs. In those instances, qualified medical and nutrition personnel must tailor the recommendations.

[c] Bone health was operationalized differently in the age and life-stage groups. For most age groups, calcium balance (positive or neutral) was a criterion, and for older adults, reduced fracture risk was considered.

[d] This row was revised since prepublication release.

[e] An EAR could not be established for any of these nutrients due to inadequate data.

SOURCES: IOM, 1997, 2002/2005, 2005, 2011.

and platelet copper concentration (IOM, 2001). More than 400 indicators have been considered by the DRI committees (Taylor, 2008).

Various indicators of chronic disease outcomes for all nutrients and non-nutrients, either measured directly or indirectly, have been considered by the various DRI committees. (See Chapter 5 for a more detailed discussion of chronic disease indicators.) Direct measurement of a chronic disease outcome was used to set the DRI in only a few instances (see Table 2-3); the AI for fluoride was based on prevention of dental caries. The AI for total fiber was based on reduced risk of coronary heart disease. DRIs based on indirect measurement of chronic disease outcomes using intermediate outcomes were established for sodium, vitamin D, calcium, and potassium. The UL for sodium was based on risk of increased blood pressure and cardiovascular outcomes, particularly cardiovascular disease and stroke. The EARs and RDAs for vitamin D and calcium were based on reduced risk of bone loss. Lastly, the AI for potassium was based on maintenance of normal blood pressure, reduced risk of bone loss, and possible reduced risk of recurrent kidney stones.

As is clear from Table 2-3, in most cases, although chronic disease endpoints were considered by the nutrient panels as potential indicators of adequacy for an EAR, ultimately they were not used. Setting an EAR requires intake-response data over a range of intakes that span the requirement range, in order to identify the nutrient intake level where half the group meets the requirement for the specified indicator of adequacy and the other half does not. These types of studies are most feasible to conduct and most easily interpreted when the indicator of adequacy (1) responds only (or to a very large extent) to changes in intake of the nutrient of interest; (2) changes over a relatively short period of time (e.g., weeks versus decades); and (3) can be assessed as having been "met" or "not met." For example, the indicator of adequacy used to set the EAR for vitamin C was near-maximal neutrophil ascorbate concentrations with minimal urinary loss, as a marker of antioxidant protection (IOM, 2000b; Levine et al., 1996). This indicator responds to changes in individuals' vitamin C intake, but would not be expected to change if their intake of other nutrients changed. It changes relatively quickly, which allowed investigators to study subjects (adult men) who were first depleted of vitamin C (using a diet with less than 5 mg/d), and then repleted, consecutively, at seven dosage levels, ranging from 30 to 2,500 mg/d. They remained at each repletion dose until steady-state plasma and neutrophil ascorbate concentrations were attained. Urinary ascorbate excretion was also monitored. This design allowed the EAR for adult men to be identified as the average intake at which neutrophils were 80 percent saturated with vitamin C and urine losses were low. EARs for other sex and life-stage groups were extrapolated from the EAR for men, based on differences in body weight, and RDAs were determined

by adding twice the assumed coefficient of variation of 10 percent to the EAR to cover the needs of 97 to 98 percent of the population. Although the DRI panel also considered a number of chronic disease endpoints (e.g., cardiovascular disease, various cancers, cataracts, asthma and chronic obstructive pulmonary disease, and cognitive function) or biomarkers of chronic disease risk (e.g., low density lipoprotein oxidation, cancer biomarkers, DNA damage) as indicators of adequacy, they could not be used to set an EAR because the data were not consistent or specific enough, and in many cases strong intake-response data linking these outcomes to actual disease risk were not available.

Types of Study Designs Used to Develop DRIs

The scientific data used to establish DRIs have been drawn from observational and experimental studies done in humans. For the most part, experimental studies have been used to establish EARs and RDAs, while AIs and ULs have been based more often on evidence from observational studies. The types of observational studies included have been, in descending order of certainty (where "certainty" refers to the ability to make inferences about the possibility of a causal relationship), prospective cohort studies, case-control studies, cross-sectional studies, case series, and case reports. Experimental studies have included randomized and nonrandomized clinical trials, controlled intake-response studies, and balance, turnover, and depletion-repletion physiological studies. Animal studies have been excluded because results are largely not applicable to nutritional deficiencies, chronic diseases, and toxic effects in humans. However, in the absence of human data, animal studies have been considered. The only instances in which animal studies alone were used as the basis for setting a DRI were in the cases of establishing the UL for molybdenum, boron, nickel, and vanadium (IOM, 2001). Study designs that include a determination of the intake-response relationship between intake of the nutrient and the selected indicator are optimal for identifying an EAR. These relationships are necessary to identify intakes that reduce (or increase in some cases) chronic disease risks (see Chapter 7), as well as other outcomes.

Situations in Which Optimal Evidence Is Not Available for All Life-Stage Groups

The need to establish a nutrient recommendation for all 22 life-stage groups often requires extrapolation or interpolation from recommendations based on experimental data that generally come from adults. This circumstance is in contrast to situations when the goal is to establish chronic disease DRIs; preventing chronic disease with diet is highly desirable but

not as necessary as reaching nutrient adequacy; thus, there is no "requirement" to establish a DRI based on chronic disease risk. Experimental data regarding nutrient requirements are limited for infants and children, making it necessary to derive recommendations from adult data. Standards for children are extrapolated from adult data based on a body weight or a metabolic factor and then adjusted for growth or tissue deposition needs. An estimate for tissue deposition is also required for estimating nutrient requirements of pregnant women. Nutrient requirements for lactation are derived from general values for milk nutrient levels that are then adjusted for the bioavailability of each nutrient in a typical maternal diet.

The first step for determining nutrient requirements by extrapolation is to derive reference body weight and height for each age group. Data from a national survey, i.e., the National Health and Nutrition Examination Survey (NHANES), which are based on precise height and weight measurements, are frequently used. Generally, the reference body weights for adults ages 19 to 30 are applied to the older adult age groups. Although body mass index tends to increase with age, this trend is not incorporated into the DRI. If no evidence exists that the metabolic rate influences the nutrient requirement, the nutrient requirement is estimated as being directly proportional to the reference body weight derived from the national surveys. This method is used to determine the nutrient recommendations for children for some, but not all, nutrients. When metabolic rate is thought to influence the requirement, reference values have also been based on metabolic differences related to body weight, estimated as the 0.75 power of body mass. For example, this method was used to extrapolate data from adults to infants and children for all the B vitamins, and to children and adolescents between 1 to 18 years of age for vitamin A. In contrast, sodium and potassium recommendations were adjusted for the combined median energy intakes for adult men and women, i.e., energy intake adults/energy intake younger adolescents and children. Dietary fiber recommendations for children also were derived from the ratio of adult median energy intakes to childhood median energy intakes.

Research on the nutrient requirements of pregnant and lactating women is also limited. To derive recommendations for pregnancy, a factorial approach is used that includes fetal nutrient accretion estimates and additional maternal needs for increased metabolic activity or for fluid or tissue deposition. If known and if it is appropriate, adjustments for insensible fluid losses and physical activity may be added. Although breast milk nutrient composition data are available, the volume produced may vary widely among women. Milk composition also may vary with the mother's nutritional status. To account for those differences in lactation needs, a reference milk volume is used for months 0-6 (780 mL/d) and months 7-12 (600 mL/d), and the milk composition of a well-nourished woman exclu-

TABLE 2-4 Common Considerations for Adjusting DRI Values When Planning Dietary Intakes for Healthy Individuals or Groups

Consideration	Nutrient	Adjustment
Women of childbearing age	Folic acid	Recommended that women capable of becoming pregnant take 400 µg folic acid/d from fortified foods, supplements, or both in addition to meeting the RDA.
Individuals older than 50 years of age	Vitamin B12	Foods fortified with B12 or supplemental B12 should be consumed by those older than 50 years of age due to decreased gastric acid with aging. Persons with any malabsorption syndrome will likely require increased amounts of B12.
Smoking	Vitamin C	The vitamin C requirement for smokers is increased by 35 mg/d due to increased vitamin C turnover. This also may be true for nonsmokers who are regularly exposed to tobacco smoke.
Bioavailability in vegetarian diets	Iron	The iron requirement is 1.8 times higher for vegetarians due to lower iron bioavailability.
	Zinc	The zinc requirement may be as much as 50 percent higher, especially for strict vegetarians who consume grains and legumes as the major food staples.
	Vitamin A	Individuals who do not consume animal-based foods must meet their requirement by consuming sufficient provitamin A carotenoids or fortified foods.

continued

TABLE 2-4 Continued

Consideration	Nutrient	Adjustment
Infants of vegan mothers	Vitamin B12	Supplemental B12 at the AI should be given at birth due to low stores at birth and in mother's milk.
Age of menstruation	Iron	The RDA for women ages 14 to 18 years includes 2.5 mg iron/d to cover menstrual iron losses. If girls start menstruation before age 14 years, 2.5 mg iron should be added to their RDA; if menstruation starts after age 14 years, 2.5 mg could be subtracted from the RDA until menstruation begins.
Women who use oral or patch contraceptives	Iron	Oral or patch contraceptives lower menstrual blood loss, which may lower the requirement.
Age at menopause	Iron	Iron needs decrease if menopause occurs before age 50 years, and would be higher in women older than age 50 years who still menstruate.
Athletes engaged in regular intense exercise	Iron	Average requirements may range from 30 to 70 percent above those of normally active individuals.
Individuals unaccustomed to prolonged physical activity in a hot environment that engage in intense exercise under such environment	Sodium	Due to the excessive loss of sodium, the AI might not apply to individuals in these situations.
Recommendation according to reference body weight	Protein	Recommendation for adults is 0.8 g/kg/day.
Recommendation set per calorie needs	Fiber	Recommendation is 14g/1,000 kcal.

TABLE 2-4 Continued

Consideration	Nutrient	Adjustment
Alcohol consumption	Fatty acids	Significant alcohol intake can depress fatty acid oxidation. Excess may be stored as fat.
	Zinc	Daily requirement may be higher in individuals who exhibit long-term alcohol consumption.
Multiparous pregnancies	Protein	Women carrying twins should increase their protein intake by an additional 50 g/day beginning in the second trimester and ensure sufficient energy intake.
Adolescent mothers, multiparous pregnancies, and mothers nursing multiple infants	Phosphorus, Magnesium	Requirements may be higher due to increased maternal and fetal needs.

SOURCE: IOM, 2006b.

sively breastfeeding a healthy infant born at term is used as a reference (IOM, 2006b).

Population Groups with Special Concerns

The DRIs are standards for apparently healthy people and are inappropriate for those with acute or chronic disease or for the repletion of nutrient levels in previously deficient individuals (IOM, 2006b). The current DRIs cannot be used to estimate the nutrient requirements for populations with these special concerns or needs. The DRI reports emphasize that only qualified medical or nutritional personnel can make appropriate adjustments for individuals with those specific needs. Other factors, such as nutrient bioavailability and physiological and lifestyle characteristics may alter nutrient requirements (IOM, 2006b). When planning dietary intakes for individuals with such altered nutrient requirements, adjustments in the DRI values can be made. Table 2-4 lists the appropriate adjustments for several common considerations.

As the previous DRI committees began to introduce the idea of chronic disease outcomes as a basis for nutrient adequacy or toxicity, they began to

deal with a number of challenges. These challenges, including the selection of outcomes and the limited availability of longitudinal data with appropriate dietary intake assessments, were discussed in various scientific fora and are further elaborated in Chapter 3. The reminder of the report will delve into the unique issues of including chronic disease outcomes when establishing nutrient reference values.

REFERENCES

CIHR (Canadian Institutes of Health Research). 2016. *Best brains exchanges 2015*. http://www.cihr-irsc.gc.ca/e/49563.html (accessed April 30, 2017).

FDA (U.S. Food and Drug Administration). 2013a. *Everything added to food in the United States (EAFUS)*. https://www.accessdata.fda.gov/scripts/fcn/fcnNavigation.cfm?rpt=eafusListing (accessed April 26, 2017).

FDA. 2013b. *Guidance for industry: A food labeling guide (14. Appendix F: Calculate the percent daily value for the appropriate nutrients)*. https://www.fda.gov/Food/GuidanceRegulation/GuidanceDocumentsRegulatoryInformation/LabelingNutrition/ucm064928.htm (accessed April 26, 2017).

Health Canada. 2007. *Eating well with Canada's Food Guide*. http://www.hc-sc.gc.ca/fn-an/food-guide-aliment/index-eng.php (accessed April 30, 2017).

HHS and USDA (U.S. Department of Health and Human Services and U.S. Department of Agriculture). 2015. *2015–2020 Dietary Guidelines for Americans. 8th Edition*. https://health.gov/dietaryguidelines/2015/guidelines (accessed July 24, 2017).

IOM (Institute of Medicine). 1994. *How should the Recommended Dietary Allowances be revised?* Washington, DC: National Academy Press.

IOM. 1997. *Dietary Reference Intakes for calcium, phosphorus, magnesium, vitamin D, and fluoride*. Washington, DC: National Academy Press.

IOM. 1998a. *Dietary Reference Intakes: A risk assessment model for establishing upper intake levels for nutrients*. Washington, DC: National Academy Press.

IOM. 1998b. *Dietary Reference Intakes for thiamin, riboflavin, niacin, vitamin B6, folate, vitamin B12, pantothenic acid, biotin, and choline*. Washington, DC: National Academy Press.

IOM. 2000a. *Dietary Reference Intakes: Applications in dietary assessment*. Washington, DC: National Academy Press.

IOM. 2000b. *Dietary Reference Intakes for vitamin C, vitamin E, selenium, and carotenoids*. Washington, DC: National Academy Press.

IOM. 2001. *Dietary Reference Intakes for vitamin A, vitamin K, arsenic, boron, chromium, copper, iodine, iron, manganese, molybdenum, nickel, silicon, vanadium, and zinc*. Washington, DC: National Academy Press.

IOM. 2002/2005. *Dietary Reference Intakes for energy, carbohydrate, fiber, fat, fatty acids, cholesterol, protein, and amino acids*. Washington, DC: The National Academies Press.

IOM. 2003a. *Dietary Reference Intakes: Applications in dietary planning*. Washington, DC: The National Academies Press.

IOM. 2003b. *Dietary Reference Intakes: Guiding principles for nutrition labeling and fortification*. Washington, DC: The National Academies Press.

IOM. 2005. *Dietary Reference Intakes for water, potassium, sodium, chloride, and sulfate*. Washington, DC: The National Academies Press.

IOM. 2006a. *Dietary Reference Intakes research synthesis: Workshop summary*. Washington, DC: The National Academies Press.

IOM. 2006b. *Dietary Reference Intakes: The essential guide to nutrient requirements*. Washington, DC: The National Academies Press.

IOM. 2008. *The development of DRIs 1994–2000: Lessons learned and new challenges: Workshop summary*. Washington, DC: The National Academies Press.

IOM. 2011. *Dietary Reference Intakes for calcium and vitamin D*. Washington, DC: The National Academies Press.

Katamay, S. W., K. A. Esslinger, M. Vigneault, J. L. Johnston, B. A. Junkins, L. G. Robbins, I. V. Sirois, E. M. Jones-Mclean, A. F. Kennedy, M. A. Bush, D. Brule, and C. Martineau. 2007. Eating well with Canada's Food Guide (2007): Development of the food intake pattern. *Nutr Rev* 65(4):155-166.

Levine, M., C. Conry-Cantilena, Y. Wang, R. W. Welch, P. W. Washko, K. R. Dhariwal, J. B. Park, A. Lazarev, J. F. Graumlich, J. King, and L. R. Cantilena. 1996. Vitamin C pharmacokinetics in healthy volunteers: Evidence for a recommended dietary allowance. *Proc Natl Acad Sci U S A* 93(8):3704-3709.

Taylor, C. L. 2008. *Framework for DRI development: Components "known" and components "to be explored."* Washington, DC.

U.S. Army. 2017. *Army Regulation 40-25: Nutrition and menu standards for human performance optimization*. Washington, DC: Headquarters, Departments of the Army, the Navy, and the Air Force.

USDA (U.S. Department of Agriculture). 2016. *School meals: Child nutrition programs*. https://www.fns.usda.gov/school-meals/child-nutrition-programs (accessed April 26, 2017).

3

Conceptual and Methodological Challenges in Establishing Chronic Disease Dietary Reference Intakes

Chapters 1 and 2 provide context for the committee's task of providing recommendations and guiding principles for chronic disease Dietary Reference Intakes (DRIs). This chapter further highlights several conceptual and methodological considerations that are specific to nutrition science issues or for which the committee felt that explicit guidance was needed for making judgments about causal or intake-response relationships related to nutrients or other food substances (NOFSs) and chronic diseases. As described in more detail in Chapter 6, the general process of judging evidence about causal relationships between exposures or interventions and health outcomes is well established. The use of carefully specified and conducted systematic reviews is at the core of this process.

Systematic reviews to guide health policy and practice require specification of questions of interest and systematic literature reviews, which identify and assess individual studies and summarize and rate the body of evidence for each question (Brannon et al., 2014). The questions themselves relate to judgments about causal attribution of certain benefits and harms to particular health-related exposures or interventions. The challenge in the nutrition field is how this process can best be applied for various food or nutrition policy purposes, and—in this case—specifically for developing NOFS-chronic disease DRIs. The relevant systematic review and evidence rating methods were developed initially to provide guidance for clinical and pharmacological therapeutics. Systematic reviews in nutrition have many special challenges not present in reviews of pharmacologic agents, even without adding the complexity of chronic disease endpoints.

As discussed in this chapter, NOFS-chronic disease questions raise spe-

cial considerations when applying these methods. The evidence base related to these questions is characterized by biological, behavioral, or study design factors that may lower the certainty of judgments associated with a given body of evidence. These factors include

- Characterizations of nutrient intake or exposures of individuals,
- Ways to account for intra- or inter-personal biological variations in effects of nutrient exposure,
- Nutrient interrelationships,
- Subpopulation differences in effects of a given nutrient intake,
- Study designs available for making causal judgments, and
- Intra- and inter-person variability in measures of exposure and outcome.

The nature of the challenges relates in part to the level at which diet or nutrition is being considered. Tapsell et al. (2016) conclude that establishing relationships between nutrients and health outcomes may present the greatest challenges, and should be guided by first establishing relationships for whole dietary patterns and then for specific foods. When DRIs for essential nutrients are the starting point, health effects of nutrients based on adequacy and toxicity are extremely difficult to isolate from the foods and dietary patterns through which nutrient exposures are created. Also, due to these challenges, careful interpretation of results from nutrition research is essential in order to understand the effects of NOFSs. For example, results from observational studies might appear inconsistent with results from large-scale randomized controlled trials (RCTs) unless they are carefully evaluated (Bazzano et al., 2006; Chew et al., 2015; Ye et al., 2013).

This report can be seen as part of a broader effort to articulate the relevant challenges and to provide guidance for how to take nutrition-related evidence issues into account in evidence reviews to support public health nutrition (Brannon et al., 2014; Chung et al., 2009; Dwyer et al., 2016; Lichtenstein et al., 2008; Russell et al., 2009; Tapsell et al., 2016; Tovey, 2014). Many of these challenges were outlined in *Options for Basing Dietary Reference Intakes (DRIs) on Chronic Disease Endpoints: Report from a Joint US-/Canadian-Sponsored Working Group* (i.e., the Options Report) (Yetley et al., 2017) (see Appendix B). This report was the foundation for the committee's statement of task (see Chapter 1, Box 1-1) and offered options that the committee was asked to consider in making its recommendations. The remainder of this chapter explains and illustrates the issues that contribute to the uncertainty of causal judgments about nutrient-chronic disease relationships. It is intended to serve as a reference point when reading Chapters 4 through 8. A case study of vitamin D, adapted from the evidence review for the 2011 update of the DRI for calcium and

vitamin D (IOM, 2011), is included at the end of this chapter to illustrate the general concepts and complexities that may apply, variably, to many nutrients and chronic diseases.

CHARACTERIZING NUTRIENT INTAKES (EXPOSURES)

Documenting what people eat to the greatest possible degree of certainty is fundamental for assessing NOFS-chronic disease relationships because the relationships themselves carry an inherent measure of uncertainty. Chronic diseases are multifactorial, and the role of any one factor will not be known precisely. Controlled feeding studies, which encompass recording of pre- and post-meal weights of known foods, offer fewer opportunities for measurement errors. Some short-duration feeding studies have used surrogate outcomes to estimate effects on chronic disease pathways (e.g., Sacks et al., 1995). However, because of their costs (e.g., they require more time for research staff to monitor and record intakes and conduct quality control procedures), they are mainly used to assess actual intakes of nutrients or foods over short periods of time, primarily in validation studies where reference intake estimates can be used to assess the accuracy of dietary intake reporting or to validate biomarkers. To examine risk associated with long-term exposure, information on nutrient exposures is usually obtained from some type of dietary assessment. These assessments, whether from foods or dietary supplements or botanicals, are always imperfect, in part because there is variation in what people eat from day to day, and also because people's ability to recall or record what they eat involves error. Systematic biases in dietary reporting also may occur, by sex, age, and weight status, and they may affect the validity of the resulting values. Chapter 4 provides an overview of the various methods of food intake assessment, their strengths and limitations, and implications for the tasks of systematic review teams and DRI committees.

For DRI purposes, documenting the amount of a nutrient within foods is of central importance, beyond assessment of the intake of foods themselves. Additional error is introduced due to limitations in nutrient composition databases, which vary in quality for different nutrients. In addition, factors such as the chemical form of the nutrient or its bioavailability have not always been taken into account.

These potential sources of error in exposure assessment can have an impact on the ability to draw clear conclusions from a given study. In observational studies, random error may decrease the ability to see true associations, and unintentional bias in reporting by certain subgroups may affect the distribution of intake, further affecting validity. The impacts of these biases are not always predictable. In RCTs, an intervention to increase or decrease a nutrient assumes that baseline intake has been correctly

characterized and that subsequent reports of food intake will reflect the achieved differences in intake between intervention and control groups. The information on achieved differences between groups, however, is often not available in published RCTs. Furthermore, because relatively few interventions are designed to compare more than one intake level to the control, intake-response relationships are often based on strata of achieved intakes, as reported by trial participants (because respondents will vary in dietary adherence), which may break the random assignment to the treatment groups.

ACCOUNTING FOR BIOLOGICAL FACTORS THAT INFLUENCE NUTRIENT EXPOSURES

Nutrient intakes, even when assessed with minimal error, may not necessarily reflect the biological exposures that influence chronic disease pathways. As indicated above, the biological "dose" associated with a given amount of food is affected by the chemical form of the nutrient and its digestibility and bioavailability in the human gastrointestinal tract. This may or may not be accounted for in nutrient database tables. In addition, digestibility, and therefore availability of certain nutrients, may be affected by other components of a meal. Different chemical forms of nutrients may vary in biological activity, requiring conversion to equivalent units for evaluation of intake. An individual's baseline nutrient status is another potential influence, which may affect absorption and utilization of the dietary source. Some of these issues will affect the interpretation of blood levels or tissue levels that might be used as biomarkers of intake as an alternative to or in conjunction with dietary intake data (see Chapter 4). These biological factors may be most important for DRI considerations when they are known or thought to vary systematically in subpopulations. In this case, such variation can be considered when formulating questions to guide systematic reviews.

ISOLATING EFFECTS OF SINGLE NOFS OR NOFS-CHRONIC DISEASE PATHWAYS

Isolating single NOFS effects is challenging if not theoretically impossible. Many NOFS functions are interrelated and may affect more than one biological pathway, and any one biological pathway may be affected by multiple NOFSs. In addition, there is collinearity in NOFS intake. Thus, both confounding by and interactions among NOFSs must be considered. Mapping these potential relationships in logic models or analytic frameworks helps to identify these considerations when framing questions to guide systematic reviews. Mapping the evidence identified can also be help-

ful for understanding relationships and patterns. For example, in Table 3-1, which is based on a World Health Organization dietary guidance report (WHO/FAO, 2003), several rows indicate associations of nutrients or nutrient sources with more than one of the six disease outcomes, and several of the columns for disease outcomes indicate associations with more than one NOFS variable. The source table for the excerpts in Table 3-1 also included variables based on food or beverage groups that may be of interest because of their nutrient or bioactive food substance (e.g., polyphenol) content with potential chronic disease risk reduction benefits. However, such variables had not been analyzed in terms of these specifics nutrients or food substances.

In individuals who do not use nutrient supplements, the range of intakes may be narrow within a given population with day-to-day variation that makes it difficult to identify group differences. Among those who do use supplements, single nutrient supplements will be associated with a substantially higher range of doses than would be obtained from food sources, facilitating clear comparisons if supplement intake is ascertained. The same would apply to supplements of botanicals (e.g., cucurmin from turmeric). A complication that sometimes remains is that the form of the NOFS in a supplement may be qualitatively different from the form that is in food, with different pathways or potency of effect. Use of multivitamin supplements limits ability to attribute any effect of the supplement to a specific nutrient.

Intervention trials involving supplements can evaluate effects of the supplement dose as an increase over baseline intake in the study population. However, for a variety of behavioral and biological reasons, answers to DRI questions may require studies that vary NOFS intakes based on dietary advice. In this case, the intervention unavoidably involves changes in intake of other NOFSs present in the foods for which consumption is changed, and these NOFSs will vary according to participant food choices as well as the degree of compliance. Changes in targeted and non-targeted NOFSs in comparison groups can be evaluated through dietary reports or biomarkers of intake (where available) to help with the attribution of any observed changes in outcome to the intervention assignment.

ASSESSING BIASES DUE TO STUDY DESIGNS

Basing judgments about causal relationships using the typical biomedical hierarchy of study designs inherently lessens the ability to make definitive judgments about NOFS-chronic disease questions because the relevant evidence is much more likely to rely on observational study designs, rather than RCTs. RCTs are at the top of the hierarchy (i.e., are the strongest study designs) when it comes to internal validity for judging causality (Yetley et

TABLE 3-1 Summary of the Strength of Evidence for Obesity, Type 2 Diabetes, Cardiovascular Disease (CVD), Cancer, Dental Disease, and Osteoporosis[a]

	Obesity	Type 2 Diabetes	CVD	Cancer	Dental Disease	Osteoporosis
Energy and fats						
High intake of energy dense foods	C↑					
Saturated fatty acids		P↑	C↑[b]			
Trans fatty acids			C↑			
Dietary cholesterol			P↑			
Myristic and palmitic acid			C↑			
Linoleic acid			C↓			
Fish and fish oils (EPA and DHA)			C↓			
Plant sterols and stanols			P↓			
α-Linolenic acid			P↓			
Oleic acid			P↓			
Stearic acid			P-NR			
Nuts (unsalted)			P↓			
Carbohydrate						
High intake of NSP (dietary fiber)	C	P↓	P↓			
Free sugars (frequency and amount)					C↑[c]	
Sugar-free chewing gum					P↓[c]	
Starch[d]					C-NR	
Wholegrain cereals			P↓			
Vitamins						
Vitamin C deficiency					C↑[e]	
Vitamin D					C↓[f]	C↓[g]
Vitamin E supplements			C-NR			
Folate			P↓			

Minerals				
High sodium intake	C↑			
Salt-preserved foods and salt		P↑[b]		
Potassium	C↓			
Calcium				C↓[g]
Fluoride, local			C↓[c]	
Fluoride, systemic			C↓[c]	P-NR[g]
Fluoride, excess			C↑[f]	
Hypocalcaemia			P↑[f]	

NOTES: C↑ = convincing increasing risk; C↓ = convincing decreasing risk; C-NR = convincing, no relationship; DHA = docosahexaenoic acid; EPA = eicosapentaenoic acid; P↑ = probable increasing risk; P↓ = probable decreasing risk; P-NR = probable, no relationship.

[a] Only convincing (C) and probable (P) evidence are included in this summary table.

[b] Evidence also summarized for selected specific fatty acids; see myristic and palmitic acid.

[c] For dental caries.

[d] Includes cooked and raw starch foods, such as rice, potatoes and bread. Excludes cakes, biscuits, and snacks with added sugar.

[e] For periodontal disease.

[f] For enamel developmental defects.

[g] In populations with high fracture incidence only; applies to men and women more than 50-60 years old.

[h] For stomach cancer.

SOURCE: Adapted from WHO/FAO (World Health Organization/Food and Agriculture Organization). 2003. *Diet, nutrition and the prevention of chronic diseases: Report of a joint WHO/FAO expert consultation.* Geneva, Switzerland: WHO.

al., 2017). Assigning study participants to treatment or control groups at random balances extraneous factors across treatment groups and permits attribution of differences in study outcomes to the NOFS intervention being tested. Participants in observational studies are assigned to comparison groups based on the estimated exposure to the NOFS of interest, which may be related to other personal characteristics or behaviors. Furthermore, these characteristics or behaviors may also affect study outcomes, which limits the certainty that an apparent causal relationship is due to the NOFS, as opposed to other factors. Residual confounding due to lack of proper statistical adjustments, untestable assumptions, or measurement error limits the causal certainty, even with the best observational study design and execution. However, RCT designs have limitations for answering NOFS-chronic disease questions, and observational data are indispensable for certain aspects of the process of developing DRIs, including chronic disease DRIs.

The availability of NOFS-chronic disease RCTs is limited primarily by cost and feasibility issues associated with conducting very large trials of long duration needed to allow time for disease to become apparent, and in enough people to observe a causal effect if one is present. For example, the dietary modification trial within the U.S. Women's Health Initiative study was designed to follow 48,000 women, to be enrolled at 40 research sites around the United States and followed for an average of 8 years to understand the role of reduction of dietary total fat intake in preventing breast cancer (Prentice et al., 2006). The trial involved only women ages 50 years and older, which increased the likelihood of observing sufficient study endpoints to indicate a disease outcome. Smaller and shorter trials can be conducted if biomarkers of chronic disease risk are acceptable, rather than actual disease outcomes. However, certainty of conclusions from such trials will be limited to the extent that the marker is not completely equivalent to the disease. This issue is discussed in more detail in Chapter 5. Finally, RCTs are limited in their ability to test varying doses or combinations of an intervention, because adding treatment conditions can require a marked increase in trial size.

Other limitations of NOFS-chronic disease RCTs may include incomplete compliance with treatment, which is common in trials that rely on dietary behavior change counseling to achieve the intervention effect. Trials also have the potential for non-targeted changes in dietary behavior among participants, which could confound interpretation of the intervention effect. For example, counseling to reduce total fat intake might incidentally lead to reduction in caloric intake and weight change, which could also affect chronic disease development apart from the fat composition of foods consumed. However, such factors can be partially accounted for with a proper statistical analysis. NOFS interventions that use supplements to increase NOFS intake are easier to implement and sustain than those based on

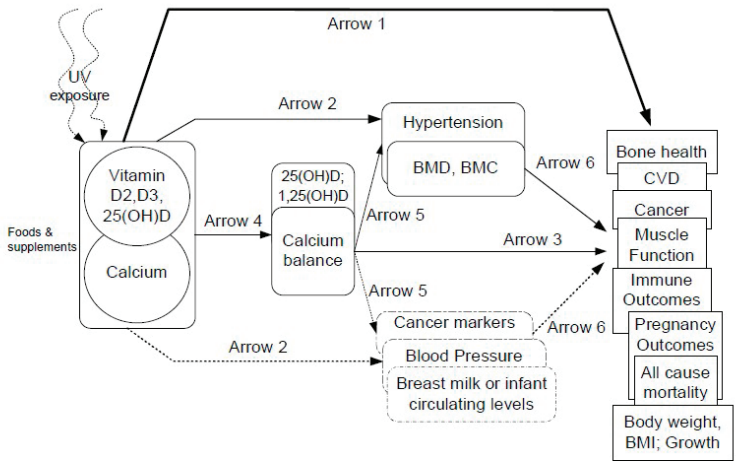
extensive counseling, although control group participants may obtain and take the supplement on their own. Importantly, translation of the effect of the supplement may differ for that from dietary sources, complicating translation to dietary recommendations (as opposed to supplement use recommendations).

In contrast to RCTs, longitudinal cohort studies in which intake of the NOFS is assessed at the time of enrollment and, in many cases, periodically thereafter, lend themselves to long-term follow-up for disease development under natural living conditions. Observational studies are the only source of information about extant eating patterns in populations of interest for DRI development and are, therefore, a reference point for understanding intake-response relationships and interpretation of DRIs by users. They are usually much less costly than large RCTs, especially the multiple RCTs or intervention arms that would be needed to observe effects of varying levels of intake of the NOFS of interest. However, the potential confounding effects of other factors, if not accounted for (e.g., age or other characteristics of the individual, other nutrients in the diet, exposure to other chronic disease factors), can lead to erroneous conclusions, and determining the certainty of conclusions about causal relationships must be carefully considered in each study. The inherent susceptibility of dietary assessments to systematic error also affects the certainty of evidence from observational studies; unlike trials, in which the intervention is the basis for the comparison, the classification on usual dietary intake is the basis for the comparison in observational studies (see Chapter 4). In observational studies, the range of exposure reflects the prevailing intakes. Hence, when optimal intake is outside of the usual range, other studies will be needed. Recognizing that causal relationships are mainly inferred from RCTs, observational studies are still critical to inform conclusions about likely causal relationships, and to support evidence on intake-response relationships.

Chapter 6 reviews considerations for incorporating evidence from observational studies when making causal judgments. In this respect, differences among types of observational studies should be recognized. Prospective studies in which NOFS intake is ascertained before disease develops are stronger designs for inferring causality than are those that obtain dietary and disease data at the same time (e.g., cross-sectional studies or retrospective case-control studies). In the latter, reverse causality—an effect of the disease on the diet—and recall bias—an effect of knowledge of the disease on the accuracy of the self-report—are major threats to study validity, making it unlikely that such studies would be judged other than as low certainty. This issue does not apply to historical or "nested" case-control studies in which pre-existing data on dietary intake or a biomarker of intake are available.

VITAMIN D CASE STUDY

This section uses a case study on the recent process to update vitamin D and calcium DRIs to briefly illustrate some of the nutrition-specific issues discussed above. The overall process for the calcium and vitamin D DRI update was described generally in Chapter 1 of this report (see Figure 1-2 and associated narrative). The reader is also referred to the methods described in detail in Appendix D of the calcium and vitamin D DRI report (IOM, 2011), from which this case study was developed and which includes an explanation of the roles of the various entities involved (e.g., the federal sponsors, technical expert panel, and the Evidence-based Practice Centers [EPCs]) and the conduct of the systematic reviews. The process for gathering and summarizing evidence was based on the Agency for Healthcare Research and Quality Methods Guide for Effectiveness and Comparative Effectiveness Reviews (AHRQ, 2014; IOM, 2011). Figure 3-1



Arrow 1: Association of exposure with clinical outcomes of interest.
Arrow 2: Association of exposure with surrogate or intermediate outcomes (that have good or possible evidence for linkage with clinical outcomes, respectively). (Surrogate outcomes are depicted in boxes with a solid outline, and intermediate outcomes are depicted in boxes with dashed outline.)
Arrow 3: Association of indicators of exposure to clinical outcomes.
Arrow 4: Association between exposure and indicators of exposure.
Arrow 5: Association of indicators of exposure to surrogate or intermediate outcomes.
Arrow 6: Association between surrogate or intermediate outcomes and clinical outcomes.

FIGURE 3-1 Analytic framework for vitamin D and/or calcium generic health outcomes.
NOTES: BMC = bone mineral content; BMD = bone mineral density; BMI = body mass index; CVD = cardiovascular disease.
SOURCE: IOM, 2011.

shows the analytic framework used to identify beneficial effects, developed by the sponsoring agencies, in consultation with the technical expert panel and EPC methodologists. It shows the multiple sources of vitamin D, including sunlight as a non-dietary source, the inseparability of pathways involving vitamin D from those of calcium and the several clinical outcomes or surrogates to be considered when evaluating benefits of a given level of vitamin D intake. A similar figure, developed to identify pathways for potential adverse effects, included several additional outcomes.

The analytic framework guided the development and refinement of key questions (see Box 3-1) used to conduct the systematic review provided to the calcium and vitamin D DRI update committee. For the pathways related to bone health, the systematic review drew on a previously conducted systematic review of evidence on the effectiveness and safety of vitamin D in relation to bone health (Cranney et al., 2007; IOM, 2011).

Using vitamin D as a case example, Table 3-2 lists and explains nutrition-specific issues to be considered in establishing an NOFS-chronic disease DRI. The general challenges in evaluating NOFS-chronic disease relationships for DRI purposes are in the first column, with comments in

BOX 3-1
Key Questions Guiding the Evidence Review for the Calcium/Vitamin D Update

Key Question 1. What is the effect of vitamin D, calcium, or combined vitamin D and calcium intakes on clinical outcomes, including growth, cardiovascular diseases, body weight outcomes, cancer, immune function, pregnancy or birth outcomes, mortality, fracture, renal outcomes, and soft tissue calcification?

Key Question 2. What is the effect of vitamin D, calcium, or combined vitamin D and calcium intakes on surrogate or intermediate outcomes, such as hypertension, blood pressure, and bone mineral density?

Key Question 3. What is the association between serum 25(OH)D concentrations on calcium balance and clinical outcomes?

Key Question 4. What is the effect of vitamin D or combined vitamin D and calcium intakes on serum 25(OH)D concentrations?

Key Question 5. What is the association between serum 25(OH)D concentrations and surrogate or intermediate outcomes?

SOURCE: IOM, 2011.

TABLE 3-2 Challenges in Evaluating NOFS-Chronic Disease Relationships to Inform DRIs, Illustrated Through the Case of Vitamin D

1.	**Characterizing nutrient exposures in comparison groups**	
1.1.	Assessment of nutrient intake from foods	Assessment of intake of any nutrient by self-reports is subject to both random and systematic error; assessment must include vitamin D naturally occurring in foods and in fortified foods; fortified foods may contain different levels of vitamin D.
1.2.	Assessment of intake from supplements	Vitamin D in supplements may be as D_2 or D_3, which may vary in potency, and in various doses. Intakes from supplements may be more reliably ascertained by self-report, at least for regular users who can identify the specific product and dose. However, exposure to vitamin D from multivitamin preparations will be potentially confounded by other nutrients with effects on the same outcomes.
1.3.	Assessing exposure from non-dietary sources	The amount of vitamin D obtained from synthesis in the skin depends on sunlight exposure, for which estimation has a high degree of uncertainty.
1.4.	Assessment of biomarkers of exposure	Serum 25(OH)D concentrations are considered the best measure of total vitamin D exposure, but estimates of levels in serum vary within and across assay type and laboratory.
1.5.	Assessing intake-response relationships	Use of serum 25(OH)D concentrations reduces the uncertainty that would be associated with using self-reported intake data, but DRIs based on serum levels require translation of serum levels into recommendations for dietary intake from food and/or supplements, and consideration of the possibility that the vitamin D in supplements has different levels of biological activity compared to food sources.
2.	**Accounting for biological factors that influence nutrient exposures**	
2.1.	Nutrient bioavailability	Bioavailability and metabolism of vitamin D may differ depending on source, various metabolic factors or other foods present at the time of digestion.
2.2.	Bioequivalence or safety profiles of different chemical forms of a nutrient	The forms of vitamin D (D_2 or D_3) considered as nutritional supplements may not have the same potency at a given dose.
2.3.	Biological stores	Vitamin D stored in body fat tissue is an endogenous source and may influence total exposure differentially, according to need. In addition, the amount of body fat influences the amount of vitamin D that is stored and its release into the circulation. Thus, overall vitamin D metabolism may differ in people with obesity.

TABLE 3-2 Continued

2.4. Endogenous sources	Sunlight (UV radiation) affects the amount of vitamin D synthesized in the skin. Exposure to sunlight is, therefore, of interest in studies of associations of vitamin D with health outcomes.
2.5. Interpretation of biomarkers	Serum 25(OH)D concentrations may be useful as an indicator of exposure. This biomarker reflects effects of intake from foods, supplements, and sun exposure as well as metabolism of vitamin D in the liver and kidneys (activation) that affects vitamin D status.
3. Nutrient interrelationships	
3.1. Multiple potential clinical outcomes and surrogates	Evidence review questions related to health effects of vitamin D address effects of vitamin D intake or serum 25(OH)D concentrations on growth, cardiovascular diseases, body weight outcomes, cancer, immune function, pregnancy or birth outcomes, mortality, fracture, renal outcomes, soft tissue calcification or on surrogate markers such as hypertension, blood pressure, and bone mineral density, and hypercalcemia; potential adverse effects of sunlight (as a "source" of vitamin D) are examined because of the potential for sunlight exposure to increase the risk of risk of non-melanoma or melanoma skin cancers.
3.2. Interrelated biological functions of nutrients	The biological effects of vitamin D include influences of calcium and phosphorous nutriture. Key questions for the calcium/vitamin D systematic reviews attempted to address this issue by evaluating effects of calcium alone, vitamin D alone, and calcium combined with vitamin D.
3.3. Interpreting nutrient effects of food-based interventions	Dietary interventions based on foods will reflect naturally occurring vitamin D as well as fortification with vitamin D and also will potentially be confounded by effects on the same outcomes by other nutrients or food substances in the same foods. For example, depending on the source, foods high in vitamin D may also contain significant amounts of omega-3 fatty acids, B vitamins, and potassium.
3.4. Interpreting nutrient effects based on supplement-based interventions	Randomized controlled trials that provide vitamin D as a supplement potentially allow robust comparisons between intervention and comparison groups. Direct interpretation of such trials with respect to DRIs assumes bioequivalency of food and supplement forms.

continued

TABLE 3-2 Continued

3.5. Potential for non-linear intake-response	As with any nutrient, a linear dose-response across the entire range of possible intakes cannot be assumed. For example, toxic effects are anticipated above the established upper limit for recommended vitamin D intake.
4. Subpopulation differences in effects of a given level of intake	
4.1. Children	Bone-related outcomes include rickets, bone mineral density, bone mineral content, fractures, or parathyroid hormone.
4.2. Women of reproductive age, including those who are pregnant or lactating	Bone-related outcomes include bone mineral density, heel bone fractures, or parathyroid hormone.
4.3. Elderly men and postmenopausal women	Bone-related outcomes are bone mineral density, fractures, and falls.
4.4. Other subpopulations	Effects of supplemental doses of vitamin D on bone outcomes may vary by ethnicity (e.g., due to differences in skin pigmentation and behaviors related to sunlight exposure), body mass index (e.g., due to effects on vitamin D storage and release from stores and possibly to behaviors related to sunlight exposure), or geography (e.g., due to variations in sunlight exposure at different latitudes).
5. Study designs	
5.1. Criteria for inclusion in systematic review	Primary studies eligible for inclusion in the systematic review were "randomized controlled trials (RCTs), non-randomized, prospective comparative studies of interventions; prospective, longitudinal, observational studies (where the measure of exposure occurred before the outcome) and prospective nested case-control studies (case-control study nested in a cohort)." Observational studies with cross-sectional and retrospective case-control designs were excluded (IOM, 2011).

the second column about how that challenge potentially affects evidence reviews and judgments related to benefits or risk of vitamin D exposure.

This chapter concludes the committee's presentation of background and context, which is intended to provide a foundation for a consideration of the conceptual and methodologic issues involved in establishing chronic disease DRIs and the committee's recommendations. The following two chapters take the next step by describing challenges and approaches involved in ascertaining dietary intake and measuring health outcomes. These activities provide the essential data needed to determine whether a causal relationship exists between an NOFS of interest and a chronic disease (Chapter 6) and whether a quantitative relationship between the NOFS and the chronic disease can be described with confidence (Chapter 7). The final chapter of the report addresses questions considered by the committee about the nature of the process to be used when developing chronic disease DRIs in relation to the existing DRI process, as described in Chapter 2.

REFERENCES

AHRQ (Agency for Healthcare Research and Quality). 2014. *Methods guide for effectiveness and comparative effectiveness reviews*. AHRQ Publication No. 10(14)-EHC063-EF. Rockville, MD: AHRQ.

Bazzano, L. A., K. Reynolds, K. N. Holder, and J. He. 2006. Effect of folic acid supplementation on risk of cardiovascular diseases: A meta-analysis of randomized controlled trials. *JAMA* 296(22):2720-2726.

Brannon, P. M., C. L. Taylor, and P. M. Coates. 2014. Use and applications of systematic reviews in public health nutrition. *Annu Rev Nutr* 34:401-419.

Chew, E. Y., T. E. Clemons, E. Agron, L. J. Launer, F. Grodstein, P. S. Bernstein, and Age-Related Eye Disease Study 2 Research Group. 2015. Effect of omega-3 fatty acids, lutein/zeaxanthin, or other nutrient supplementation on cognitive function: The AREDS2 randomized clinical trial. *JAMA* 314(8):791-801.

Chung, M., E. M. Balk, S. Ip, G. Raman, W. W. Yu, T. A. Trikalinos, A. H. Lichtenstein, E. A. Yetley, and J. Lau. 2009. Reporting of systematic reviews of micronutrients and health: A critical appraisal. *Am J Clin Nutr* 89(4):1099-1113.

Cranney A., T. Horsley, S. O'Donnell, H. A. Weiler, L. Puil, D. S. Ooi, S. A. Atkinson, L. M. Ward, D. Moher, D. A. Hanley, M. Fang, F. Yazdi, C. Garritty, M. Sampson, N. Barrowman, A. Tsertsvadze, and V. Mamaladze. 2007. *Effectiveness and safety of vitamin D in relation to bone health. Evidence Report/Technology Assessment No. 158*. (Prepared by the University of Ottawa Evidence-based Practice Center (UO-EPC) under Contract No. 290-02-0021.) AHRQ Publication No. 07-E013. Rockville, MD: Agency for Healthcare Research and Quality.

Dwyer, J. T., K. H. Rubin, K. L. Fritsche, T. L. Psota, D. J. Liska, W. S. Harris, S. J. Montain, and B. J. Lyle. 2016. Creating the future of evidence-based nutrition recommendations: Case studies from lipid research. *Adv Nutr* 7(4):747-755.

IOM (Institute of Medicine). 2011. *Dietary Reference Intakes for calcium and vitamin D*. Washington, DC: The National Academies Press.

Lichtenstein, A. H., E. A. Yetley, and J. Lau. 2008. Application of systematic review methodology to the field of nutrition. *J Nutr* 138(12):2297-2306.

Prentice, R. L., B. Caan, R. T. Chlebowski, R. Patterson, L. H. Kuller, J. K. Ockene, K. L. Margolis, M. C. Limacher, J. E. Manson, L. M. Parker, E. Paskett, L. Phillips, J. Robbins, J. E. Rossouw, G. E. Sarto, J. M. Shikany, M. L. Stefanick, C. A. Thomson, L. Van Horn, M. Z. Vitolins, J. Wactawski-Wende, R. B. Wallace, S. Wassertheil-Smoller, E. Whitlock, K. Yano, L. Adams-Campbell, G. L. Anderson, A. R. Assaf, S. A. Beresford, H. R. Black, R. L. Brunner, R. G. Brzyski, L. Ford, M. Gass, J. Hays, D. Heber, G. Heiss, S. L. Hendrix, J. Hsia, F. A. Hubbell, R. D. Jackson, K. C. Johnson, J. M. Kotchen, A. Z. LaCroix, D. S. Lane, R. D. Langer, N. L. Lasser, and M. M. Henderson. 2006. Low-fat dietary pattern and risk of invasive breast cancer: The Women's Health Initiative Randomized Controlled Dietary Modification Trial. *JAMA* 295(6):629-642.

Russell, R., M. Chung, E. M. Balk, S. Atkinson, E. L. Giovannucci, S. Ip, A. H. Lichtenstein, S. T. Mayne, G. Raman, A. C. Ross, T. A. Trikalinos, K. P. West, Jr., and J. Lau. 2009. Opportunities and challenges in conducting systematic reviews to support the development of nutrient reference values: Vitamin A as an example. *Am J Clin Nutr* 89(3):728-733.

Sacks, F. M., E. Obarzanek, M. M. Windhauser, L. P. Svetkey, W. M. Vollmer, M. McCullough, N. Karanja, P.-H. Lin, P. Steele, M. A. Proschan, M. A. Evans, L. J. Appel, G. A. Bray, T. M. Vogt, T. J. Moore, and DASH Investigators. 1995. Rationale and design of the Dietary Approaches to Stop Hypertension trial (DASH). *Annals of Epidemiology* 5(2):108-118.

Tapsell, L. C., E. P. Neale, A. Satija, and F. B. Hu. 2016. Foods, nutrients, and dietary patterns: Interconnections and implications for dietary guidelines. *Adv Nutr* 7(3):445-454.

Tovey, D. 2014. The role of The Cochrane Collaboration in support of the WHO Nutrition Guidelines. *Adv Nutr* 5(1):35-39.

WHO/FAO (World Health Organization/Food and Agriculture Organization). 2003. *Diet, nutrition and the prevention of chronic diseases: Report of a joint WHO/FAO expert consultation.* Geneva, Switzerland: WHO.

Ye, Y., J. Li, and Z. Yuan. 2013. Effect of antioxidant vitamin supplementation on cardiovascular outcomes: A meta-analysis of randomized controlled trials. *PLOS ONE* 8(2):e56803.

Yetley, E. A., A. J. MacFarlane, L. S. Greene-Finestone, C. Garza, J. D. Ard, S. A. Atkinson, D. M. Bier, A. L. Carriquiry, W. R. Harlan, D. Hattis, J. C. King, D. Krewski, D. L. O'Connor, R. L. Prentice, J. V. Rodricks, and G. A. Wells. 2017. Options for basing Dietary Reference Intakes (DRIs) on chronic disease endpoints: Report from a joint US-/Canadian-sponsored working group. *Am J Clin Nutr* 105(1):249S-285S.

4

Methodological Considerations Related to Assessing Intake of Nutrients or Other Food Substances

One of the most important aspects of evaluating the certainty of the diet-chronic disease evidence is the validity and reliability of the food and nutrient or other food substance (NOFS) intake methodology used in each published study. Accuracy in the NOFS intake methodology is essential because the quantitative relationship between the nutrient and the chronic disease must be characterized with considerable certainty in order to establish a chronic disease Dietary Reference Intake (DRI). As Chapter 3 points out, characterizing dietary exposures is one type of challenge that is unique to nutrition research. Without careful attention, intake assessment may add major uncertainty to the judgments about NOFS-chronic disease relationships. Although the recent *Options for Basing Dietary Reference Intakes (DRIs) on Chronic Disease Endpoints: Report from a Joint US-/Canadian-Sponsored Working Group* (i.e., the Options Report) (Yetley et al., 2017) raised the issue of characterizing NOFS intakes as a challenge, it did not specify options for any particular methodology. Because NOFS intake methodology is a crucial topic, this committee has devoted this chapter to it.

Within the context of the DRI process (see Figure 1-2) and as described in Chapter 6, decisions about dietary intake methodologies are made mainly at two steps. During the first step—developing the protocol for the systematic review—the inclusion criteria for NOFS intake methods are decided. At this step, the inclusion criteria need not be restrictive, but can include relevant and commonly accepted self-reported methods and biomarker measurements. In the second step—conducting the evidence review—the

BOX 4-1
Key Terminology

Accuracy: Closeness of a measured or computed value to its "true" value, where the "true" value is obtained with perfect information. Owing to the natural heterogeneity and stochastic nature of many biologic and environmental systems, the "true" value may be an integrated average over a defined time period.

Calibration of a self-reported dietary intake method: Calibration is the process of using a suitable intake biomarker in an attempt to correct a self-reported intake assessment for measurement error. Calibration equations are typically developed by regressing biomarker intake values on corresponding self-reported values and possibly other study participant characteristics.

Precision: The quality of a measurement that is reproducible in amount or performance. Measurements can be precise in that they are reproducible, but can be inaccurate and differ from "true" values when biases exist. Measurement error can also affect precision. In risk-assessment outcomes and other forms of quantitative information, precision refers specifically to variation among a set of quantitative estimates of outcomes.

Uncertainty: Lack or incompleteness of information. Quantitative uncertainty analysis attempts to analyze and describe the degree to which a calculated value may differ from the true value; it is sometimes expressed as probability distributions. Uncertainty depends on the quality, quantity, and relevance of data, and on the applicability and relevance of models and assumptions.

DRI committee considers and evaluates the internal validity of individual studies (risk of bias), which would include evaluating the dietary intake methodology. This chapter provides information on dietary assessment methodology to help guide committees with evaluating its contribution to the risk of bias. Specifically, the chapter describes methods currently in use to assess NOFSs, describes novel methods for evaluating NOFS intake, provides guidance that will help DRI committees evaluate published literature, and offers suggestions for future research that may help fill gaps in this important area. Key terms used in this chapter are in Box 4-1.

Validation of a biomarker: Validation is the action of checking or proving the accuracy of some measure. Validity can sometimes be established by conducting controlled human feeding studies in a population of interest. Each participant is provided a diet over a defined time period and potential biomarkers in pertinent biofluids (e.g., urine or serum/plasma) are examined for correlation with actual intake of the nutrient or food substance of interest. Biomarkers meeting criteria (e.g., correlation≥0.6[a]) may provide useful objective measures of intake in the population from which feeding study participants were drawn.

Validation of a self-reported dietary intake method: Validation is established by comparing the self-reported measurement with an objective measure of intake (e.g., quantitative recovery biomarkers such as doubly labeled water assessment of (short-term) energy intake, or urinary nitrogen assessment of protein intake). It should be noted that objective intake measures, such as quantitative recovery biomarkers, are not available for all nutrients or food substances.

Variability: True differences in attributes due to heterogeneity or diversity. Variability is usually not reducible by further measurement or study, although it can be more thoroughly characterized. Two important sources of variability are biological variability due to inter-individual differences (i.e., attributable to genetic differences and influenced by environmental factors) and analytical variability (i.e., associated with analysis of dietary component).

[a] The rationale for this criterion was comparison with established intake biomarkers for energy (doubly labeled water) and protein (24-hour urinary nitrogen), which had (following log-transformation) correlations of 0.71 and 0.61, respectively, with estimated actual intakes (Lampe et al., 2017).

METHODS TO ASSESS NUTRIENTS OR OTHER FOOD SUBSTANCES

Self-Reported Measures of Dietary Intake

Self-report of NOFSs has been the primary intake assessment method for most cohort studies and randomized controlled trials with chronic disease endpoints. Much has been written about various types of self-reports and the strengths and weaknesses of each in the research setting (Thompson et al., 2015). Briefly, four major types of self-report dietary assessment have been used in studies of diet and chronic disease outcomes: (1) multiple day dietary records, (2) multiple day 24-hour dietary recalls,

(3) semi-quantitative food frequency questionnaires, and (4) brief instruments focused on specific foods or food groups. Each of these methods assesses short-term intake over a period of days to 1 year. Some strengths and weaknesses of these measures are given in Table 4-1.

Self-report of diet is often used in the research setting because it is relatively low cost and carries a relatively low participant burden. Self-report also may need to be considered because valid biomarkers are available only for certain NOFSs. Moreover, even when biomarkers are available they may sometimes usefully be combined with self-reported intake data in disease association analyses. Self-reported dietary assessment methods also provide data on food sources of NOFSs; such data are needed to translate evidence about NOFS-disease associations into food-based dietary guidance.

Standardized protocols exist for the collection of self-reported dietary data. However, even when collected by the best available measures of short- or long-term intake, dietary self-reports have a high level of uncertainty due to the (1) complexity of many foods, (2) limitations of self-report for accurately describing or recording specific foods consumed, (3) difficulty in accounting for day-to-day variability when estimating usual intakes (variability is highest for micronutrients and most or all other food substances, compared to macronutrients), and (4) limitations of food composition databases. Self-reported dietary intake data are well-known to contain random error, such as day-to-day variability, which adds "noise" in the data and can often lead to imprecise findings with very wide confidence intervals or even null results that obscure the ability to detect an association that is actually present (Beaton et al., 1979). Methods for correcting random error in dietary intake data exist (NCI, 2017; NRC, 1986). For example, in many contemporary cohort studies repeated measures of dietary intake have been collected periodically during the follow-up periods. The use of repeated and updated dietary assessments by validated food frequency questionnaires has been found to reduce measurement errors and represent long-term dietary habits, which are most relevant to chronic disease etiology and prevention. Some studies have analyzed changes in dietary exposures over time and subsequent risk of chronic diseases. This is a stronger observational design than a typical cohort study because it mimics an intervention study. Importantly, dietary self-report data are also subject to systematic error or bias that is dependent upon participant characteristics, such as age, sex, body mass index (BMI), and race/ethnicity (NCI, 2017; Neuhouser et al., 2008; Prentice et al., 2011; Tinker et al., 2011; Zheng et al., 2014). This type of bias is especially problematic because it may influence the size and direction of observed associations with chronic disease and it is also difficult to detect. Although adjustments for variables such as age, sex, and race/ethnicity are frequently done to minimize confounding in nutrition studies,

no amount of adjustment can remove or adjust for this systematic error due to the participants characteristics.

Biomarkers of NOFS Intake

Exposure to NOFSs also can be assessed through biomarkers. Biomarkers are measurements obtained (usually sampled) from a biological system or organism, such as human blood, urine, feces, saliva, hair, skin, nail, and other tissues (i.e., adipose tissue, organ biopsy material). In the context of nutrition, biomarkers have a continuous typology based on the intended purpose, with quality possibly dependent on the intended application (see Figure 4-1).

Biomarkers of nutritional exposure fall into five general categories:

- Compounds or molecules that reflect exposure to nutrients with existing DRIs (i.e., certain vitamins, minerals, macronutrients);
- Bioactive compounds without a DRI (i.e., carotenoids, isoflavones);
- Integrative biomarkers that capture exposure to food substances plus metabolic processing (i.e., metabolomics);
- Functional nutritional biomarkers that may reflect enzyme saturation or functional measures of nutritional status; and
- Food contaminants (i.e., aflatoxin, polycyclic aromatic hydrocarbons [PAHs], nitrosamines, acrylamide, pesticides). These biomarkers are less directly related to nutrients, but may be relevant.

This chapter is concerned *only* with biomarkers of NOFS intake (i.e., biomarkers of exposure or level 4 in Figure 4-1; biomarkers of effect and clinical outcomes are discussed in Chapter 5). Biomarkers of intake may be subject to the same types of random error or bias as self-report, but suitable biomarkers should be less subject to systematic bias and, for this reason, may be viewed as more objective measures of NOFS intake. For an intake biomarker to be suitable, it should measure (perhaps following appropriate transformation) the intake of interest along with possible measurement error that is unrelated to the targeted intake or to other study participant characteristics. A biomarker that substantially meets this classical measurement model criterion will have the greatest utility when the biomarker measurement error variance is small relative to the variance of the targeted nutritional variable in the study population. However, it is important to note that some nutrition-related biomarkers, useful for other purposes, do not reflect intake but, rather, NOFS status. This is an important distinction in establishing quantitative intake-response relationships. For example, an individual can have zero vitamin D intake and have good vitamin D status.

Most nutritional biomarkers of intake (or exposure) are either recovery

TABLE 4-1 Examples of Uses, Strengths, and Limitations of Self-Report Measures of NOFS Intake

Self-Reported Measure	What Is Collected?	How?	Feasibility	Strengths and Limitations
Dietary Records	All food and drink consumed is reported as it occurs; foods served and left over may be weighed to improve quantification.	May be done directly by respondents themselves on paper or into a computer program or may be done by an observer.	For self-report, requires educated and highly compliant respondents; the cost and logistics of observation is prohibitive for most large studies.	Strengths: Weighed intakes with highly compliant participants are accurate, but are not a part of most dietary record protocols. Limitations: High respondent burden usually leads to incomplete records and high drop-out rates. Even when accurate, evidence suggests that eating behavior changes while recording, thus it may not represent usual intake (systematic error).
24-Hour Dietary Recalls	All food and drink consumed is reported for the previous day. A multiple pass system may be used to minimize underestimation due to poor memory or lack of detail.	May be entered by the respondent into a guided computer program or may be conducted by a trained interviewer in person or over the telephone.	Feasible for large studies but requires multiple days to stabilize usual intake estimates of individuals.	Strengths: Most used method for estimating group mean intakes and the usual intake distribution, as used in national surveys (Beaton et al., 1979; Moshfegh et al., 2008). Limitations: High day-to-day variability limits extrapolation to usual intake of individuals. This random error leads to attenuation of associations with individual outcomes like chronic disease. Evidence suggests underreporting in general and bias in reporting, with biases larger among persons having higher body mass (systematic error).

Food Frequency Questionnaire	The usual frequency of intake is reported for a list of foods. Portion sizes may or may not be ascertained.	The questionnaire may be completed by the respondent or it may be interviewer administered.	Highly feasible, as a single application provides an estimate of usual intake over a period of time, usually the past year. Low-cost method.	Strengths: When used appropriately, may allow ranking of wide range of nutrients and food substances, usually after adjusting for energy intake. Provides a moderate-term (e.g., 6 months-1 year) integrated measure of intake, appropriate for use with health outcomes (Willett et al., 1985). Limitations: It is semi-quantitative, based on group foods and recipes, so detail is less than in other methods. Validity depends on the food list, recipe assumptions, and cognitive challenge of integrating frequency of intake over 1 year. May not accurately represent dietary intake for groups with different dietary patterns than for which it was developed. This may introduce systematic bias associated with obesity, race/ethnicity, social class, or unusual dietary lifestyle.
Brief Instruments for specific nutrients or food groups, such as fat screeners	Limited questions about usual intake of major food sources of specific nutrients or food groups.	The questionnaire may be completed by the respondent or it may be interviewer administered.	Highly feasible, low respondent burden.	Strengths: Provides qualitative or general ranking data on specific foods or nutrients. Limitations: Does not allow for energy adjustment or consideration of these items within the full diet. Subject to systematic bias in estimation of NOFSs due to missing food contributors.

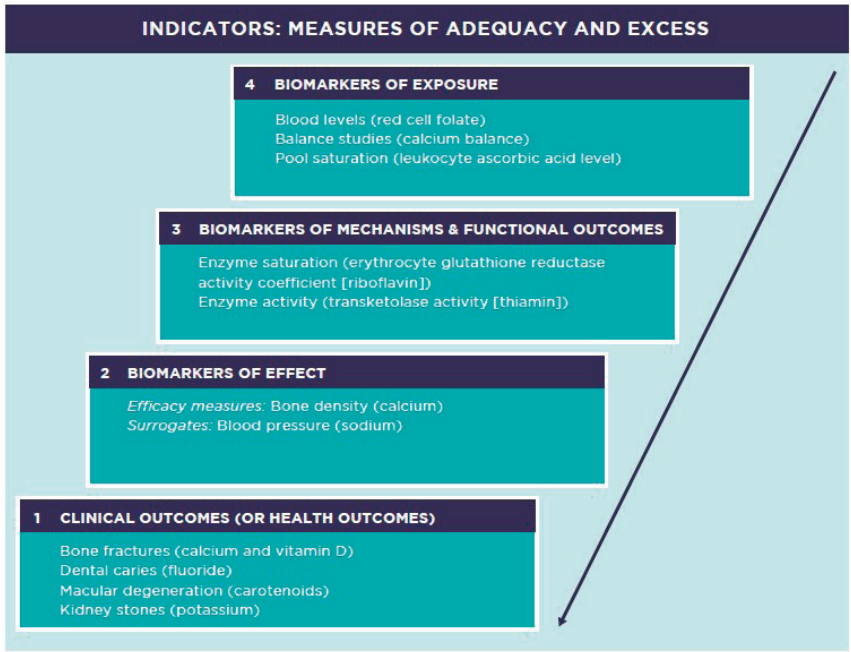


FIGURE 4-1 Types and examples of biomarkers for nutrient or food substances.
NOTE: Numbering and arrows reference hierarchical proximity to the clinical outcome of interest. Blood pressure is a surrogate for cardiovascular disease and bone density is a surrogate for fracture risk.
SOURCE: Adapted from Taylor, 2008.

biomarkers or concentration biomarkers. Recovery biomarkers measure intake and output that can be "recovered" and measured quantitatively (usually in urine). The advantage of recovery biomarkers is that they can be used to assess absolute intake over a defined period of time (typically days or weeks). Current methods for assay are excellent and have high precision. Limitations, however, do exist. Importantly, only a few true recovery biomarkers are available, namely doubly labeled water, which estimates total energy expenditure and is used to approximate total energy intake in weight-stable individuals; urinary nitrogen (from 24-hour urine collections), from which protein intake can be computed; and urinary sodium (from 24-hour urine collections).[1] Although 24-hour urine collections of

[1] However, 24-hour urine collections can substantially underestimate intake in individuals who have heavy sweat losses (e.g., athletes or those working in hot conditions). This is a much greater problem for sodium (an extracellular anion lost in large amounts in sweat) than for potassium (an intracellular anion present in small amounts in sweat).

potassium has been considered a suitable biomarker, it needs to be further studied due to data showing variability based on race and other factors (Turban et al., 2008). Other limitations of current methods for biomarkers are that the protocols for specimen collection are burdensome to study participants and, therefore, may not be collected completely. Also, the collection procedures and assays are expensive. In addition, the measures are short term—reflecting days or weeks of intake—but the time course for the diet-chronic disease risk occurs over a period of many years or decades.

Concentration biomarkers assess concentrations or relative percentages of NOFSs in the blood, urine, or other tissues. Recently, serum phospholipid fatty acids that correlate strongly with intake have been identified as biomarkers of intake of some specific fatty acids, total saturated fatty acids, total *trans* fatty acids, and total carbohydrate in postmenopausal women (Song et al., 2017). Likewise, in the same population, serum biomarkers of certain carotenoids, folate, vitamin B12, and α-tocopherol have been identified (Lampe et al., 2017). Many nutritional biomarkers used in studies of chronic disease risk are these concentration-type biomarkers, such as erythrocyte and serum folate, serum carotenoids, plasma phospholipid fatty acids, and serum vitamin D [as 25(OH)D]. Most of these biomarkers are relatively short term, depending on the half-life of the NOFS as well as on the tissue from which it is collected. For example, erythrocyte folate may represent intake from the past 3 months, whereas serum folate may represent the past few weeks. Furthermore, for some NOFSs, erythrocyte measures could be viewed as markers of status and the serum measures as markers of intake. Concentration biomarkers have typically not been used in the same quantitative manner as recovery biomarkers, but for some (not all) NOFSs they can reflect intake following necessary rescaling or other transformations. Importantly, numerous participant characteristics, such as age, sex, race/ethnicity, and BMI, may strongly influence the serum concentration of a particular NOFS. Concentration biomarkers also may contain bias as a measure of intake due to the effects of other exposures, such as smoking, adiposity, medications, and related NOFS intakes. Furthermore, for many concentration biomarkers, metabolic and physiological factors will influence concentrations. For example, phospholipid fatty acids are influenced by both dietary intake and endogenous fatty acid synthesis. In addition, the time of day of a blood draw (and whether in the fasting state or not) will influence whether the fatty acid profile will reflect a state of beta oxidation, de novo synthesis or participation in various intracellular metabolic processes. More generally, for a concentration biomarker to be suitable for intake assessment, it should (following possible rescaling or other transformation) strongly correlate with the intake of interest along with possible measurement error that is unrelated to the targeted intake or to other study subject characteristics.

IMPROVING THE QUALITY OF THE DIETARY INTAKE DATA IN NUTRITION STUDIES

Like any measurement, estimates of exposure to NOFSs may carry biases and uncertainties related to the accuracy and precision of the assessment. Measuring dietary intake with minimal uncertainties, however, is key to establishing quantitative associations between NOFSs and diseases and is a critical criterion in being able to judge the quality of individual studies (part of the risk-of-bias assessment) and overall certainty in the evidence (see Chapter 6). Identifying potential sources of measurement error, evaluating methods used to correct for such errors, and considering which intake methods may provide estimates that are closest to the true exposure, is one of the most important tasks that future DRI committees will need to carry out when evaluating individual studies.

This section presents the committee's guidance for best approaches related to minimizing uncertainty when measuring long-term NOFS intake with biomarkers of intake (objective measure) and self-reported measures (subjective measure).

Validity and Utility in Assessing Dietary Intake with Biomarkers of Intake

The measurement error associated with dietary self-report may be a significant impediment for DRI committees, whose task will be focused on establishing optimal NOFS intake values for chronic disease risk reduction. Self-report is subject to both random and systematic error, the latter being more troublesome and not resolved through sample size increases or statistical adjustments in chronic disease rate modeling. Although nutritional biomarkers are also subject to random error, they can be objective measures of diet and are, under the classical measurement model mentioned above, not subject to the same types of systematic error commonly found in dietary self-report. However, this does not mean that they are without bias. Biomarkers may not accurately reflect dietary intake due to differential factors affecting absorption, metabolism, and utilization, which must be considered when evaluating their use. For example, ultraviolet (UV) exposure could lead to high serum 25(OH)D even with no vitamin D intake. Calibrating a self-report method with a validated nutritional biomarker, if carefully conducted, may represent an important methodological advance over reliance on self-report alone to improve the accuracy of the NOFS intake data. It may also be possible to directly apply established nutritional biomarkers to a prospective study cohort in a nested case-control study, if they are based on stored specimens (i.e., obtained in both cases and controls before the chronic disease develops).

However, a biomarker of intake must first be validated. Specifically,

the biomarker (typically log-transformed) needs to equal a targeted intake (e.g., log-transformed usual intake over a specified time period), plus random noise that does not depend on the targeted intake, or on other study participant characteristics that are pertinent to the disease under study (e.g., established disease risk factors, other dietary intakes, physical activity patterns). Well-designed and well-conducted human feeding studies can provide evidence of validation for intake biomarkers. A major criterion for biomarker evaluation is the magnitude of the correlation between the known intake and the biomarker values, with corresponding correlations for established biomarkers useful as benchmarks in a specific feeding study context. However, such a context, strictly speaking, can provide direct support for biomarker utility only for the typically short time period of the feeding study in question. Measures of dietary stability over time, along with periodic biomarker measurements over time, are typically needed to establish a biomarker of usual intake over a longer time period, such as the several years or more, that may be pertinent to chronic disease risk. Regression modeling of known intake from a feeding study onto a biomarker that reflects intake, along with participant characteristics that may affect this relationship (e.g., age, sex, BMI), provides a natural framework for biomarker development and evaluation. Such a regression model can provide necessary relocation and rescaling of the biological measures, and the inclusion of relevant characteristics can enhance the resulting biomarker's adherence to the classical measurement model described above. A biomarker that plausibly adheres to this measurement model can "anchor" a chronic disease association analysis, either by providing a framework for correcting dietary self-report data or, if available, through use of biospecimens stored from members of a study cohort, by direct application in a cohort of interest in a case-control mode.

Methodologic advances have been made in biomarker studies of moderate size (e.g., a few hundred persons). As described in the Options Report and elsewhere (Neuhouser et al., 2008; Prentice et al., 2011; Tinker et al., 2011; Zheng et al., 2014), the development of regression calibration equations that use objective measures (biomarkers) of dietary intake to calibrate self-report data has resulted in biologically plausible and clinically meaningful diet-disease associations that are not observed when only self-report is used. Suitable calibration equations may be developed by regressing established biomarker values on corresponding self-report values and other relevant participant characteristics that need to be included in the risk model for the chronic disease of interest, for confounding control. If the resulting equation explains a substantial fraction of the variation in the targeted dietary variable (this may require replicate biomarker measures to account for random error in the biomarker), then the equation may be used to obtain calibrated intake values for all participants with dietary

self-report and associated data in the larger cohort from which the biomarker study derives. For example, the established intake biomarkers for energy (doubly labeled water) and protein (24-hour urinary nitrogen) had correlations of 0.71 and 0.61, respectively, with estimated actual intakes (Lampe et al., 2017). The committee concluded that biomarkers meeting an $R \geq 0.6$ criterion relate to actual intake about as closely as those established biomarkers and may be used to obtain calibrated intake values and provide useful objective measures of intake in the population from which feeding study participants were drawn. These values may be used in disease risk models to obtain disease association analyses in large cohort settings. Such association analyses typically apply to intake over a relatively short time period, and may be limited if the self-report data alone provide only a weak signal for the dietary variable of interest. Correlation with longer term intake, over months or years, is needed for reliable nutritional epidemiology association studies, and may require repeat biomarker application at various times over the cohort follow-up period. The major limitation of the regression calibration approach, however, is the paucity of established biomarkers for the large number of nutritional variables that may be relevant to chronic disease risk.

New Research on Biomarkers of NOFS Intake

Identification and validation of new biomarkers of NOFS intake is urgently needed. In this regard, the use of metabolomics to identify dietary biomarkers is a rapidly growing area of investigation and may provide promise for future nutritional epidemiology research (see Chapter 5). Because metabolomics can reflect both dietary intake as well as the influence of metabolic pathways, it may provide an important approach to the development of additional intake biomarkers (Guertin et al., 2014; Playdon et al., 2017). However, despite its promise, metabolomics has challenges. For example, some analytes generated from a metabolomics platform may link back to a particular NOFS; in other cases, groups of metabolites may reflect the intake of a single NOFS or class of NOFSs (Song et al., 2017).

GUIDANCE FOR FUTURE DRI COMMITTEES

As mentioned in Chapter 1, the committee framed its recommendations and guiding principles in the context of the process shown in Figure 1-2 in which a formal systematic review, including a synthesis of the evidence, of the relevant PICO (population, intervention, comparator, and outcome) questions is conducted by a systematic review team before the formation of the DRI committee and with guidance from a technical expert panel committee. The PICO process is a technique used in evidence-based practice to

frame and answer a clinical or health care–related question and it is also used to develop literature search strategies. The DRI committee would assess such a systematic review, consider any additional evidence, and make decisions about setting chronic DRIs for each outcome.

Assessing NOFS exposures that are valid and relevant for chronic disease outcomes is a challenging task, and the existing systematic reviews that address the intake of NOFS and chronic diseases reveal the diversity in nature and quality of the nutrient intake ascertainment.

Although the committee does not question the important role of dietary self-report data in the field of nutrition (e.g., for nutrition policy purposes where they provide data on food sources of NOFS, allowing interpretation of evidence in the context of dietary guidance for chronic disease prevention), in the context of developing DRIs, random and systematic biases of self-reported methodologies need to be particularly minimized. Considered carefully and within context, studies relying on nutritional biomarkers for intake assessment may present important advantages that should be recognized. However, although methods for addressing some of these challenges are available or emerging, they are not yet reflected in most research studies exploring diet and chronic disease associations. In the near term, and until better dietary methodologies are applied in research, DRI committees will need to identify those studies that provide the maximum level of certainty in exposures data. Such information will then need to be integrated as part of the risk-of-bias evaluation (internal validity of individual studies), an element of the evidence review (see Chapters 6 and 7). In the long term, research agendas should include accelerated efforts to improve dietary exposure assessment for use in chronic disease studies (see Box 4-2).

The committee concludes that no single satisfactory approach exists to accurately measure dietary intake, and that each study and methodology needs to be assessed based on its own merit, taking into consideration potential risk of bias. The committee developed a decision guide (see Figure 4-2) for use in considering dietary intake methodological issues in studies on diet and chronic disease and in incorporating questions in the risk-of-bias assessment. For example, to assess risk of bias[2] related to dietary intake methods, the questions in Figure 4-2 could be used as guidance. In addition, as mentioned in Chapter 6, before conducting a risk of bias assessment, a priori hypotheses could be developed to explain potential heterogeneity in the relevant PICO categories (e.g., intervention). In this way, if an apparent effect modifier is found, for example due to the specific measure of dietary intake used, inferences would be substantially stronger.

[2] As mentioned in Chapter 6, until a validated risk-of-bias tool is developed for the field of nutrition and chronic disease, the existing risk-of-bias tools could be expanded with questions relevant to nutrition, including questions related to dietary intake assessment.

BOX 4-2
Nutrient Intake Assessment Methodologies

Committee's Recommendation 1

Until better intake assessment methodologies are developed and applied widely, Dietary Reference Intake committees should strive to ensure that random and systematic errors and biases of nutrient or other food substance (NOFS) exposure assessment methodologies are considered in their evidence review. In the long term, research agendas should include accelerated efforts to improve NOFS exposure assessments for application in studies of chronic disease risk.

Specifically, based on a systematic search of the scientific literature related to validation of dietary assessment methods, DRI committees will select a priori criteria to define the most accurate dietary intake methods, such as those that use validation and calibration methods to minimize bias. Second, they will carefully review the methodologies in the individual studies to assess the seriousness of the risk of bias (see also Chapter 6), by asking questions about the method and its potential limitations (see Table 4-1), following the decision guide in Figure 4-2. The questions will vary depending on whether a self-report method or a biomarker (or both) have been used. These questions are meant to give committees a sense of the potential biases in the study and whether the biases are serious enough to either exclude or include the studies and the potential limitations. For example, if a food frequency questionnaire was used in a study, questions in Sidebars A and B should provide information about the potential limitations of the methods. Furthermore, questions in Sidebar D will help ascertain whether approaches have been applied to address random error; any systematic error also should be questioned (see also Table 4-1).

In conclusion, accuracy in methods to measure the intake of NOFS is an essential feature in the evidence used as basis for establishing DRIs, particularly when characterizing quantitative relationships. Therefore, consideration of the potential biases in these methods is an essential task of DRI committees.

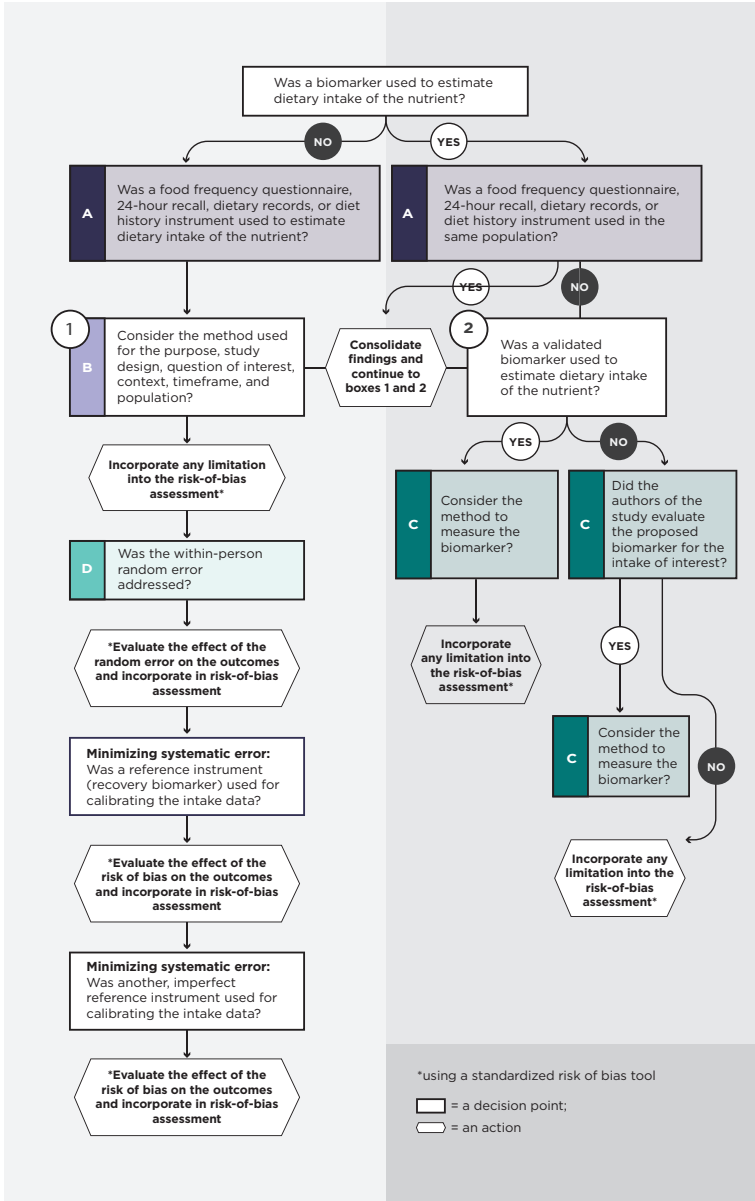


FIGURE 4-2 Decision guide for use in considering dietary intake methodological issues in studies on diet and chronic disease and in incorporating questions in the risk-of-bias assessment.

NOTE: FFQ = food frequency questionnaire; HR = hour.

* Using a standardized risk-of-bias tool.

A

Was the instrument adequately calibrated in the population in which it is being used?

B

FFQ:
- Does the food list include consideration of the dietary patterns of all major subgroups (ethnicity, age group, income group, vegetarianism, or other special diet?)
- Are portion sizes appropriately considered for the included population?
- Are variations in specific food choice or preparation (recipes) appropriately considered in the inclusion and weighting of nutrient database selections?
- Was the FFQ administered by a trained interviewer, if needed (e.g., low literacy, compliance)?
- Was a high-quality nutrient database used?

24-Hour Recall and Diet Records:
- Were recalls administered in person or by telephone by trained interviewers?
- Were individuals trained before keeping diet records?
- Was a high-quality nutrient database used?
- How was likely underreporting addressed?

Diet Records:
- How complete was compliance? Did loss of compliance affect power or generalizability?
- Was a high-quality nutrient database used?

C

- Was it a standardized method, calibrated across laboratories with samples that had been collected, treated, and stored properly and that were not too old?
- Was the method sufficient for the question asked (e.g., ranking)?
- Was the right tissue used to measure the biomarker?
- Was day-to-day variability considered and were sufficient replicates obtained when needed?
- Was the proper time integration and turnover (e.g., recent vs. longer term marker) used for the question asked?
- Was potential confounding by other exposures considered (e.g., use of supplements, body composition, medications, disease states)?

D

FFQ:
- Were repeated measures of intake taken (at least two measures on a sample of the population)?
- Were intake data based on one or more days of the week?
- Was the potential for seasonal variations considered and minimized by having seasonal food questions, if time period of interest is more than a season?

24-Hour Recall:
- Were repeated measures of intake taken (at least two measures on a sample of the population)?
- Were intake data based on one or more days of the week?
- Was the potential for seasonal bias considered and minimized through study design and analysis procedures?
- Were day-of-week and other nuisance effects considered and minimized through study design and analysis procedures?

Diet Records:
- Were repeated measures of intake taken (at least two measures on a sample of the population)?
- Were intake data based on one or more days of the week?
- Was the potential for seasonal bias considered and minimized through study design and analysis procedures?
- Were day-of-week and other nuisance effects, such as season effect, considered and minimized through study design and analysis procedures?

FIGURE 4-2 Continued

REFERENCES

Beaton, G. H., J. Milner, P. Corey, V. McGuire, M. Cousins, E. Stewart, M. de Ramos, D. Hewitt, P. V. Grambsch, N. Kassim, and J. A. Little. 1979. Sources of variance in 24-hour dietary recall data: Implications for nutrition study design and interpretation. *Am J Clin Nutr* 32(12):2546-2559.

Guertin, K. A., S. C. Moore, J. N. Sampson, W. Y. Huang, Q. Xiao, R. Z. Stolzenberg-Solomon, R. Sinha, and A. J. Cross. 2014. Metabolomics in nutritional epidemiology: Identifying metabolites associated with diet and quantifying their potential to uncover diet-disease relations in populations. *Am J Clin Nutr* 100(1):208-217.

Lampe, J. W., Y. Huang, M. L. Neuhouser, L. F. Tinker, X. Song, D. A. Schoeller, S. Kim, D. Raftery, C. Di, C. Zheng, Y. Schwarz, L. Van Horn, C. A. Thomson, Y. Mossavar-Rahmani, S. A. Beresford, and R. L. Prentice. 2017. Dietary biomarker evaluation in a controlled feeding study in women from the Women's Health Initiative cohort. *Am J Clin Nutr* 105(2):466-475.

Moshfegh, A. J., D. G. Rhodes, D. J. Baer, T. Murayi, J. C. Clemens, W. V. Rumpler, D. R. Paul, R. S. Sebastian, K. J. Kuczynski, L. A. Ingwersen, R. C. Staples, and L. E. Cleveland. 2008. The US Department of Agriculture Automated Multiple-Pass Method reduces bias in the collection of energy intakes. *Am J Clin Nutr* 88(2):324-332.

NCI (National Cancer Institute). 2017. *Dietary Assessment Primer.* https://dietassessmentprimer.cancer.gov (accessed May 11, 2017).

Neuhouser, M. L., L. Tinker, P. A. Shaw, D. Schoeller, S. A. Bingham, L. V. Horn, S. A. Beresford, B. Caan, C. Thomson, S. Satterfield, L. Kuller, G. Heiss, E. Smit, G. Sarto, J. Ockene, M. L. Stefanick, A. Assaf, S. Runswick, and R. L. Prentice. 2008. Use of recovery biomarkers to calibrate nutrient consumption self-reports in the Women's Health Initiative. *Am J Epidemiol* 167(10):1247-1259.

NRC (National Research Council). 1986. *Nutrient adequacy: Assessment using food consumption surveys*. Washington, DC: National Academy Press.

Playdon, M. C., S. C. Moore, A. Derkach, J. Reedy, A. F. Subar, J. N. Sampson, D. Albanes, F. Gu, J. Kontto, C. Lassale, L. M. Liao, S. Mannisto, A. M. Mondul, S. J. Weinstein, M. L. Irwin, S. T. Mayne, and R. Stolzenberg-Solomon. 2017. Identifying biomarkers of dietary patterns by using metabolomics. *Am J Clin Nutr* 105(2):450-465.

Prentice, R. L., Y. Mossavar-Rahmani, Y. Huang, L. Van Horn, S. A. Beresford, B. Caan, L. Tinker, D. Schoeller, S. Bingham, C. B. Eaton, C. Thomson, K. C. Johnson, J. Ockene, G. Sarto, G. Heiss, and M. L. Neuhouser. 2011. Evaluation and comparison of food records, recalls, and frequencies for energy and protein assessment by using recovery biomarkers. *Am J Epidemiol* 174(5):591-603.

Song, X., Y. Huang, M. L. Neuhouser, L. F. Tinker, M. Z. Vitolins, R. L. Prentice, and J. W. Lampe. 2017. Dietary long-chain fatty acids and carbohydrate biomarker evaluation in a controlled feeding study in participants from the Women's Health Initiative cohort. *Am J Clin Nutr*.

Taylor, C. L. 2008. *Framework for DRI development: Components "known" and components "to be explored."* Washington, DC.

Thompson, F. E., S. I. Kirkpatrick, A. F. Subar, J. Reedy, T. E. Schap, M. M. Wilson, and S. M. Krebs-Smith. 2015. The National Cancer Institute's Dietary Assessment Primer: A resource for diet research. *J Acad Nutr Diet* 115(12):1986-1995.

Tinker, L. F., G. E. Sarto, B. V. Howard, Y. Huang, M. L. Neuhouser, Y. Mossavar-Rahmani, J. M. Beasley, K. L. Margolis, C. B. Eaton, L. S. Phillips, and R. L. Prentice. 2011. Biomarker-calibrated dietary energy and protein intake associations with diabetes risk among postmenopausal women from the Women's Health Initiative. *Am J Clin Nutr* 94(6):1600-1606.

Turban, S., E. R. Miller, 3rd, B. Ange, and L. J. Appel. 2008. Racial differences in urinary potassium excretion. *J Am Soc Nephrol* 19(7):1396-1402.

Willett, W. C., L. Sampson, M. J. Stampfer, B. Rosner, C. Bain, J. Witschi, C. H. Hennekens, and F. E. Speizer. 1985. Reproducibility and validity of a semiquantitative food frequency questionnaire. *Am J Epidemiol* 122(1):51-65.

Yetley, E. A., A. J. MacFarlane, L. S. Greene-Finestone, C. Garza, J. D. Ard, S. A. Atkinson, D. M. Bier, A. L. Carriquiry, W. R. Harlan, D. Hattis, J. C. King, D. Krewski, D. L. O'Connor, R. L. Prentice, J. V. Rodricks, and G. A. Wells. 2017. Options for basing Dietary Reference Intakes (DRIs) on chronic disease endpoints: Report from a joint US-/Canadian-sponsored working group. *Am J Clin Nutr* 105(1):249S-285S.

Zheng, C., S. A. Beresford, L. Van Horn, L. F. Tinker, C. A. Thomson, M. L. Neuhouser, C. Di, J. E. Manson, Y. Mossavar-Rahmani, R. Seguin, T. Manini, A. Z. LaCroix, and R. L. Prentice. 2014. Simultaneous association of total energy consumption and activity-related energy expenditure with risks of cardiovascular disease, cancer, and diabetes among postmenopausal women. *Am J Epidemiol* 180(5):526-535.

5

Measuring Chronic Disease Outcomes

Studying associations between nutrients or other food substances (NOFSs) and health outcomes starts by formulating specific question(s) about the population of interest, the intervention (e.g., a nutrient level), the comparator (e.g., a different nutrient level), and the health outcome. These elements of the PICO (population, intervention, comparator, and outcome) evidence review framework are described in more detail in Chapter 6. Although these four elements seem simple to address, both the groups preparing and answering the questions face challenges unique to each element. Challenges include using consistent terminology, determining the appropriateness of measurement methods, and being able to identify limitations. This chapter focuses on the component that involves measuring the health outcome. In this case, the outcome, or endpoint, is the chronic disease event (e.g., a diagnosis of heart disease, stroke, diabetes, or a person's death).

The goal of this chapter is to provide Dietary Reference Intakes (DRI) committees with guidance on considerations related to potential biases and limitations in measuring chronic disease outcomes reported in individual research studies and included in systematic reviews and evidence summaries that are used to make judgments about relationships between NOFSs and a chronic disease. Considerations include how well the outcome has been defined and measured, either directly or indirectly with a surrogate marker that is sufficiently specific to reflect the causal pathway of potential interest. The chapter also includes a general section on chronic disease for context and key definitions.

This guidance is provided within the context of the report *Options for Basing Dietary Reference Intakes (DRIs) on Chronic Disease Endpoints:*

Report from a Joint US-/Canadian-sponsored Working Group (i.e., the Options Report) (Yetley et al., 2017) and the committee's statement of task as described in Chapter 1. A critical issue raised in the Options Report relates to the decision to use surrogate markers of disease when information on chronic diseases events is not available or as supplementary information. Research studies often include information on biomarkers along a pathway between exposure and disease, and some can serve as surrogate markers of disease when chronic disease events have not been measured or were observed in numbers too small to permit definitive statistical analyses of causal relationships.

CHRONIC DISEASE OUTCOMES

Defining Chronic Disease

Chronic diseases, including diabetes, cardiovascular and neurodegenerative diseases, and many common cancers, manifest over a lifetime and can begin before birth. As such, aging is widely recognized as the leading risk factor for many human diseases (Niccoli and Partridge, 2012). Chronic diseases are complex traits, and susceptibility to their onset and their progression vary among individuals within a population. Risk of developing a chronic disease is associated with both genetic and non-genetic exposures (Fontana et al., 2010). Family history, including genetic endowment, remain the best predictors of longevity potential (Gavrilov and Gavrilova, 2015), and for many people, their susceptibility to chronic disease. Environmental and lifestyle factors, including diet and exercise, are among the leading modifiable risk factors for chronic disease onset and management (Benziger et al., 2016; GBD 2013 Risk Factors Collaborators, 2015).

Definitions of chronic disease[1] often encompass distinct anatomic, physiological, or behavioral abnormalities; theories and/or evidence of causation; and usually some degree of human suffering. Research on disease detection and diagnosis reveals numerous challenges (IOM, 2015) that are outside of the scope of this study. However, when exploring relationships between NOFS intake and disease occurrence or other outcomes,[2] it is essential to consider the accuracy of the diagnostics and measures as well as

[1] A chronic disease is "a culmination of a series of pathogenic processes in response to internal or external stimuli over time that results in a clinical diagnosis/ailment and health outcomes" (e.g., diabetes) (IOM, 2010, p. 23).

[2] The most common use of the term "outcome," used synonymously with "endpoints," refers to the clinical results of a particular illness(es), often after particular therapeutic interventions. With regard to DRIs, the outcome might be a change in disease incidence (primary prevention of coronary disease) but also can be improvement of the clinical outcome of patients who have already sustained a heart attack (secondary prevention).

the existence of disease category subtypes that may respond to NOFS exposures differentially. As an example, a World Health Organization (WHO) systematic review (WHO, 2012) that explored the effects of potassium on all cardiovascular disease, stroke, and coronary heart disease events (fatal and non-fatal) considered a composite as a measure of cardiovascular disease, which included some or all of the following: fatal and non-fatal stroke, coronary heart disease, myocardial infarction, and/or congestive cardiac failure, or episode of coronary revascularization, bypass grafting, and/or angioplasty. They also considered all-cause mortality and all other outcomes reported by the authors of the original studies.

The role of nutrition in chronic disease risk can begin in utero, as nutrition exposures early in life can influence fetal and neonatal gene expression patterns. As such, nutrient exposures drive the epigenetic programming of physiological cellular networks, and these programs, which may persist throughout the lifetime of mammals, affecting disease susceptibility (Waterland and Michels, 2007). Likewise, the composition of the gut microbiome is associated with chronic diseases (Shreiner et al., 2015). The relative contribution of modifiable (e.g., dietary factors) and non-modifiable (e.g., genetics) contributions to chronic diseases within populations merits consideration when establishing DRIs using chronic disease endpoints, as does the magnitude of population heterogeneity leading to individual variation in the diet-disease relationship (Ohlhorst et al., 2013; Rappaport, 2016).

Although links between nutrition and chronic diseases are frequently identified for common conditions such as diabetes, cardiovascular diseases, and some types of cancer, the committee cannot predict which chronic diseases could be related to NOFSs in the future. Table 2-3 in Chapter 2 indicates that a range of chronic conditions have been considered by prior DRI committees, not all of which are life-threatening. Which potential nutrient-related health outcomes warrant priority for DRIs and population nutrition policy will continue to be an important issue for federal agencies. Therefore, the committee chose to use a broad definition of chronic disease, which includes (1) diseases that last for months or years, or their outcomes (e.g., longevity, disease-specific mortality, and all-cause mortality), even if the cause is unknown, (2) some infectious organism-induced diseases, such as AIDS, chronic hepatitis, cervical cancer, or gastric cancer, that could be potentially mitigated by nutritional interventions, or (3) clinical conditions that lack a formal disease status, such as chronic pain syndromes of unknown cause, or clinical syndromes difficult to characterize, such as Gulf War Syndrome or chronic fatigue syndrome.

Interpreting the effects of NOFSs may be complicated by the nature of chronic diseases. For example, some chronic conditions that surface late in life (e.g., atherosclerotic diseases) may have their origins at a young age. In addition, many chronic disease outcomes are likely to differ in the age at

onset, progression, and severity. Therefore, it might not be clear when an NOFS intervention may have its effect within the disease pathway. Also, study outcomes associated with an NOFS intervention may differ depending on diagnostic procedures and protocols used. For example, diseases that are detected by screening tests when individuals are asymptomatic (e.g., through mammography and Pap smears) may have different clinical outcomes than those detected in response to patients with clinical signs and symptoms (Hillerdal, 2008; Jensen and Vedsted, 2017). Furthermore, variation in access to medical care and in the types of clinical treatments may have important (and confounding) effects on outcome rates, particularly in observational studies of NOFSs.

Regardless of the definition adopted, it is important to delineate the condition of interest with high specificity. Is the disease truly chronic? Is the disease biologically heterogeneous in etiology and in responsiveness to nutrients? Do standard diagnostic methods exist? Does the disease have sub-conditions based on clinical behavior, pathophysiology, molecular markers and biology, and outcomes (prognosis)?

Measuring Chronic Disease Outcomes

Once the outcomes of interest are designated (see Figure 1-2), acceptable methods for measurement (i.e., for inclusion in the systematic review) should be selected a priori based on recommendations from relevant authoritative clinical guidelines.

Each chronic disease can be measured with a different set of methodologies, which have strengths and limitations. When considered a priori, identifying the strengths and limitations of the measurements will allow a DRI committee to make judgments about the potential risk of bias due to the outcome measurements for each study. Assessing the accuracy of occurrence of particular chronic diseases in research studies can be challenging, however. First, diagnosis of a chronic disease itself may be complicated. As an example, coronary heart disease may be suspected based on medical and family history, risk factors, and tests, but no single test for a definitive diagnosis is currently available. Data from studies relying on self-reporting of chronic diseases depend on having had an opportunity for diagnosis as well as accurate recall of the diagnosis; they may, therefore, be regarded as not suitably accurate when establishing DRIs. The accuracy of self-report data depends on the disease and its details, as some diseases are more accurately reported than others (Navin Christina et al., 2016). Self-report data can vary substantially from data based on disease ascertainment from primary clinical records or from population surveys or research studies that involve diagnostic procedures (e.g., the National Health and Nutrition Examination Survey, clinical trials, or observational studies).

Clinical records, including various formats such as electronic records or registries derived from them, are expected to be more accurate than patient self-report, but for various reasons, clinical records of any type may not be fully accurate or complete. Clinical records depend on comprehensive and accurately coded data and the accuracy and completeness of the information stored in electronic (or paper) patient records varies widely. The *International Classification of Diseases* (ICD), which is maintained by WHO, is the international standard diagnostic tool and provides a system of codes for classifying diseases, with a variety of signs and symptoms. In the United States, the ICD-10-CM (*ICD, 10th Revision, Clinical Modification*) is used by hospitals and other health care facilities to better describe the clinical picture of the patient. Cause of death on United States death certificates are also coded in the ICD. One systematic review examined data quality in electronic primary care records and found that the quality of recording in diagnoses of diseases, in particular, varied; completeness was higher when clear diagnostic criteria existed (Thiru et al., 2003). Another systematic review (Jordan et al., 2004) found that the quality of assigning a morbidity code during primary care consultations also varied depending on the condition. Some cited areas where efforts are needed are development of data quality standards and improvements in understanding implementation challenges.

Another critical measure is mortality from the chronic disease. The cause-of-death statistics as reported in the United States by the U.S. Centers for Disease Control and Prevention's National Vital Statistics System[3] are undoubtedly a valuable tool for research and other public health purposes, but accuracy is variable and depends on the disease and other factors, such as (1) potential diagnostic and certification errors, (2) whether autopsies are performed or data are based on medical records, (3) the training of the certifier, and (4) the presence of multiple conditions leading to the death (Lloyd et al., 2017).

Using Surrogate Markers as a Substitute for a Chronic Disease Outcomes

The initial problem formulation (see Figure 1-2) about a general health outcome of interest that may be associated with intake of a given NOFS leads to another fundamental decision point: whether to require evidence based on the chronic disease outcome(s), accept a biomarker(s) of effect that can lead to a valid and reliable test of association between the NOFS and the risk of the chronic disease outcome (i.e., a surrogate marker), or use both types of measures. In regard to establishing DRIs, both biomarkers of functional outcomes (e.g., activity of an enzyme) and biomarkers of effect (e.g., a chronic disease) have been used but they have distinct

[3] See https://www.cdc.gov/nchs/nvss (accessed July 20, 2017).

TABLE 5-1 Characteristics of Biomarkers of Nutrient Function Versus Biomarkers of Effect

Characteristics	Biomarkers of Nutrient Function	Biomarkers of Effect (e.g., cancer, heart disease)
Variables that affect the indicator	Intake of the nutrient is the primary modifiable variable affecting the nutrient function endpoint or development of nutrient deficiency symptoms (unmodifiable variables such as age, sex, genetics, also may play a role, e.g., risk of iron deficiency).	Intake of the nutrient may be one of many modifiable variables (e.g., intake of other nutrients or food substances, physical activity, weight status, environmental exposures) and subgroup characteristics (e.g., age, sex, genetics) that affect development of the chronic disease.
Time course over which the nutrient influences the endpoint	Relatively short (e.g., weeks to months in most cases).	Very long (potentially over the lifespan).
Without adequate intake, proportion of the population who will develop the deficiency or chronic disease	100 percent (e.g., with very low intakes of vitamin C, everyone will develop scurvy, as vitamin C deficiency is the only cause of scurvy).	Always <100 percent (no chronic diseases occur in 100 percent of the population; for example, the lifetime risk of cancer in the United States is 42 percent, not 100 percent) (ACS, 2017).
With adequate intake, proportion of the population who will be protected against the deficiency or the chronic disease	100 percent (e.g., with enough vitamin C, no one will develop scurvy, because vitamin C deficiency is the only cause of scurvy).	Extremely variable, but always <100 percent. In most cases, the reduction in the absolute risk associated with modifying a single nutrient would be expected to be relatively small (as a hypothetical example, perhaps 40 percent of those with adequate intake of the nutrient would develop the chronic disease, versus 44 percent of those without an adequate intake).

TABLE 5-1 Continued

Characteristics	Biomarkers of Nutrient Function	Biomarkers of Effect (e.g., cancer, heart disease)
Relationship between intake and health or functional status with regard to the indicator of adequacy	At intake below the requirement, function would be impaired. No additional increases in function are expected at intakes above the requirement.	In theory, many different types of relationships between nutrient intake and chronic disease risk are possible, and the type of relationship could differ for a single nutrient and different chronic diseases, or for different nutrients and a single chronic disease. Potentially, there could be a broad range of intake over which relative risk of developing a chronic disease could change in a graded manner, but it may not be possible to identify a point at which no further risk reduction occurs. Conversely, it is possible that, beyond a range of intake over which risk is reduced, risk could begin to increase.

characteristics (see Table 5-1). Recognizing the current efforts to modernize and standardize the terminology about biomarkers in the medical field (e.g., Robb et al., 2016), the committee adopted the terms used in the Options Report, "surrogate disease markers" and "nonqualified disease marker" (see Box 5-1) to refer to biomarkers of effect that are of interest for this chapter. Box 5-1 also includes a general definition for biomarkers to provide context. Although biomarkers of nutrient intake (which have been discussed in Chapter 4) are included in the definition for completeness, the focus here is on biomarkers of disease endpoints.

Surrogate Markers in Studies of Nutrient-Chronic Disease Relationships

As noted in Chapter 3, studies that employ randomized controlled trial (RCT) designs have the greatest likelihood of establishing causation compared to observational study designs. However, using disease events as outcome measures may not always be feasible due to study expense, the rarity of the disease in question, time imperatives, or the complexity of diagnosis. In those situations, some studies may resort to *surrogate markers*. These are measures deemed to be on the causal pathways to frank illness,

BOX 5-1
Definitions: Biomarkers, Nonqualified Markers, and Surrogate Markers

Biomarker: A particular measurement sampled from a biological system or organism. It may take many forms, including an anatomic depiction (e.g., brain imaging), a physiological process (e.g., the glomerular filtration rate of the kidney or an electroencephalographic tracing of brain activity), an indicator of dietary intake (e.g., blood vitamin B12 levels), psychological or cognitive functions (e.g., remembering nouns from a recited list), or an indicator of the presence of a disease (e.g., high levels of blood enzymes indicating liver inflammation). All biomarkers have the same general potential problems: measurement error, variation over time and space, and difficulties in biological interpretation. In research and clinical medicine, biomarkers have important uses in understanding biological processes and in predicting the risk, presence, severity, response to, adverse effects of treatment, and outcomes of diseases. More general information on biomarkers is available in the report *Evaluation of Biomarkers and Surrogate Endpoints in Chronic Disease* (IOM, 2010).

Nonqualified disease marker (also known as an intermediate disease outcome marker or intermediate endpoint): A possible biomarker of effect that predicts a chronic disease outcome but lacks sufficient evidence to qualify as an accurate and reliable substitute for that outcome. (Yetley et al., 2017)

Surrogate disease marker (also known as a surrogate marker, surrogate endpoint, or surrogate disease outcome marker): A biomarker of effect that predicts clinical benefit (or harm, or lack of benefit or harm) for an exposure under study based on epidemiologic, therapeutic, pathophysiologic, or other scientific evidence (Yetley et al., 2017). A surrogate disease marker may be qualified by the U.S. Food and Drug Administration (FDA) for its intended purposes. Blood pressure, for example, has been qualified by FDA as a surrogate marker for cardiovascular disease (FDA, 2017a,b).

whose occurrence would often lead to that illness should more time and larger sample sizes be available. Furthermore, the status of the surrogate marker, or more generally the surrogate marker process history, should be able to explain a substantial portion of any relationship between the NOFS and the chronic disease risk (Freedman et al., 1992; Prentice, 1989). A surrogate marker may be any type of biomarker. However, in diagnosing a disease, determining a surrogate marker may require extensive resources, such as a comprehensive clinical evaluation, just as is generally needed for the primary illness. For example, if an adenomatous colon polyp is to be considered a surrogate marker for colon cancer, detecting and diagnosing the polyp usually requires a full clinical examination and colonoscopy.

When evaluating studies relevant to nutrition, identifying true surrogate markers may be problematic for several reasons. Surrogate markers are complex and often relate to the natural history of the specific disease in question. Many categories of conditions, such as many cancers, psychiatric conditions, and neurodegenerative diseases, do not have credible surrogate markers relative to any set of treatments or exposures. Some are not true surrogate markers but may just reflect the activity of existing, active conditions, such as prostate-specific antigen (PSA) for prostate cancer or serological markers of inflammation in various infections or rheumatic conditions. Another challenge is the usual modest ability of candidate surrogate markers to predict full disease outcomes. Disease occurrence may be many years or decades in the future and difficult to evaluate. In addition, other positive or adverse health consequences predicted by that candidate surrogate markers may not be well-evaluated. Perhaps most important, the candidate surrogate marker may only be a risk predictor and not related causally to the condition of interest.

Another complication with selecting candidate surrogate markers is reflected in Figure 5-1. The figure shows how candidate surrogate markers can be the result of multiple exposures and biological pathways, illustrating that they may not be on the causal pathway of the disease. The figure also shows that multiple NOFSs can affect one or more biological pathways. When a chronic disease might be the result of multiple biological pathways, extracting and interpreting specific information about the relationship between one NOFS and the chronic disease outcome in human studies requires a substantial understanding of underlying biological mechanisms, and may not be possible depending on the nature of relevant interrelation-

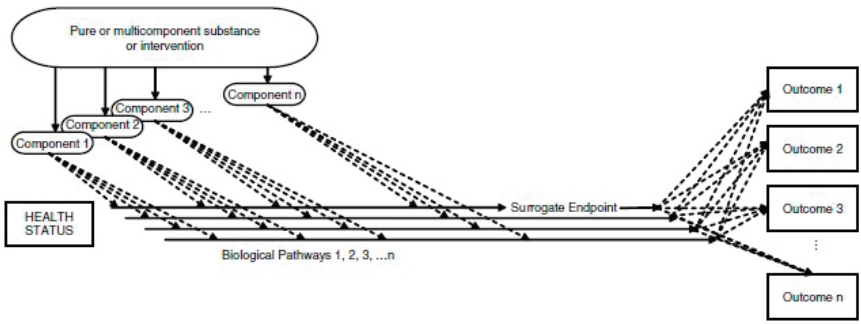


FIGURE 5-1 Illustration of potential complexities of the use of biomarkers and surrogate outcomes (markers) in chronic disease. Solid horizontal arrows indicate biological pathways. Dotted lines indicate possible pathways.
SOURCE: IOM, 2010.

ships. Efforts are ongoing to identify better surrogate markers, which might be important to support causal associations between NOFS and chronic diseases.

One example of the value of surrogate markers is in cancer research where, to be meaningful, studies need to be large and lengthy because specific cancers are not frequent and often take a long time to develop. In exploring the contribution of nutrition to cancer, one area of interest is the use of colorectal adenomas as candidate surrogates because they occur earlier in the disease pathway and relatively frequently compared to the incidence of colorectal cancer. Strong evidence shows a relationship between this marker and colorectal cancer (Fearon and Vogelstein, 1990; Paraskeva et al., 1990; Sugarbaker et al., 1985). The surrogate is still not without challenges, however. Because adenomas occur early in the pathway to disease, if the nutritional intervention has an effect only in the later stages, such an intervention would not have the expected effect on the surrogate marker (false negatives). Another limitation is that only a small proportion of adenomas may ever result in cancer. These considerations need to be accounted for when interpreting the results of interventions. Still, a nutritional intervention reducing the recurrence of adenomas in the large bowel would likely decrease the incidence of colorectal cancer and therefore various nutrition studies have taken advantage of the value of this potential surrogate (Greenberg et al., 1994; Neuhouser et al., 2015; Takata et al., 2014). Another example of the challenges is the ongoing need to find surrogate markers for cardiovascular disease, one of the major causes of death worldwide. The longstanding nutritional guidance of replacing saturated fatty acids (SFAs) in the diet with unsaturated fats has recently been challenged by researchers, who have argued that prospective cohort studies do not show an association between saturated fats, polyunsaturated fatty acids (PUFAs), or monounsaturated fatty acids (MUFAs) with risk of coronary heart disease (Chowdhury et al., 2014). These controversies would be clarified with a reliable surrogate marker. Possible surrogates include blood lipids (low-density lipoprotein [LDL] cholesterol, high-density lipoprotein cholesterol, triglycerides, ApoB, or ApoA1) and fatty acids (PUFA, MUFA, SFA, or omega-3 PUFAs), measured over pertinent subsets of the lifespan. In addition to these more conventional surrogates, the emerging field of metabolomics has recently led to use of plasma metabolites as possible surrogates. Although prospective cohort studies indicate that LDL cholesterol is a good predictor of cardiovascular disease outcome (Lewington et al., 2007), the prediction value of other lipids is questionable (Clarke, 2017). This continued debate shows the difficulties in identifying and interpreting the results when potential surrogate markers are used as predictors of a chronic disease.

New Research on Biomarkers of Chronic Disease

Research on next generation biomarkers of nutrition and chronic disease seeks to (1) identify and classify individuals who are at risk of diet-related chronic disease, the paradigm that currently drives the field of precision medicine (Collins and Varmus, 2015), and (2) quantify the dose-response relationships between individual or groups of nutrients and disease onset and progression (Ohlhorst et al., 2013).

Metabolomics approaches are enabling the identification of comprehensive metabolic signatures in a single assay comprising hundreds of metabolites in serum or other biological fluids (Beger et al., 2016). This snapshot of system biomarkers has the potential to enhance the prediction and identification of disease states at a high level of resolution, as well as inform pharmacological and/or nutrition regimes for chronic disease treatment and prevention. Comprehensive measures of metabolic function can be integrated with other "omics" measures of genetic variation, epigenetic variation, and transcriptional and proteomic measures of gene expression to elucidate the multidimensional biological changes that occur throughout the disease process (Arneson et al., 2017).

Qualified Surrogate Markers

Biomarkers of effect (in this case, a chronic disease) comprise those that are "qualified" (i.e., surrogate marker) and "nonqualified" (i.e., nonqualified disease marker). Therefore, a surrogate marker is, by definition, qualified for its purpose. The concept of "qualified" surrogate markers originated in the medical community, particularly as it applies to pharmaceutical research, but is now being applied to nutrition. The difference between these two types of markers is described in Box 5-1. From a regulatory perspective, qualified surrogate outcomes need to meet certain criteria regarding prediction of clinical outcomes. In this context, the interested parties, such as a pharmaceutical industry, may work collaboratively with FDA's Center for Drug Evaluation and Research (CDER) Biomarker Qualification Program to guide marker development (FDA, 2017a). The qualification process is initiated by submitters, who are typically interested in a drug approval, once they have (1) a clear understanding of the relationship between the candidate surrogate marker and the clinical outcome, (2) a defined use for the candidate surrogate marker in drug development, and (3) an identified candidate surrogate marker measure, preferably analytically validated. After this initial step, a consultation process is initiated with FDA. The process is similar to that used when a biomarker is intended to be used as evidence for submission of a nutrient health claim. Based on the need for scientific rigor in this qualification process, the IOM published a

2010 report *Evaluation of Biomarkers and Surrogate Endpoints in Chronic Disease*. The report recommends a framework for evaluating biomarkers used as qualified surrogate outcomes that includes three elements: analytical validation,[4] evidentiary qualification,[5] and utilization analysis[6] (IOM, 2010). The report also recommends that FDA use the same degree of scientific rigor to evaluate markers across all regulatory areas—drugs, medical devices, biologics, or foods and dietary supplements.

Examples of biomarkers accepted as qualified surrogate endpoints for the purpose of health claims for specific chronic disease endpoints by FDA's Center for Food Safety and Nutrition include (1) serum LDL cholesterol concentration, total serum cholesterol concentration, and blood pressure for cardiovascular disease, (2) bone mineral density for osteoporosis, (3) adenomatous colon polyps for colon cancer, and (4) elevated blood sugar concentrations and insulin resistance for type 2 diabetes (FDA/CFSAN, 2009). Although the surrogate is qualified as a marker of the outcome (independent of the NOFS), the NOFS of interest needs to have an effect on the qualified surrogate marker, not on a different marker. The National Institutes of Health (NIH) and/or FDA's CDER also lists examples of biomarkers of disease risk accepted as "qualified surrogate endpoints" as of December 2015 (FDA, 2017b).

Past DRI committees have attempted to consider chronic disease outcomes when assessing possible indicators of adequacy or excessive intakes of nutrients. Examples of the chronic diseases considered are listed in Table 5-2. Table 5-2 also identifies nutrients for which a DRI was established based on chronic disease risk, as well as whether direct (i.e., disease outcome) or indirect (i.e., biomarker of effect) outcomes were used to set the DRI. Finally, it lists nutrient-disease associations that were considered, but ultimately not used to set DRIs, largely because of insufficient or inconsistent evidence. Although numerous chronic diseases and nutrients were considered, DRIs based on chronic disease risk were set in only a small number of cases (see Chapter 2, Table 2-3). Among those, in most cases an

[4] Analytical validation: Assessment of [an] "assay and its measurement performance characteristics, determining the range of conditions under which the assay will give reproducible and accurate data" (IOM, 2010, p. 3).

[5] Evidentiary qualification: "Assessment of available evidence on associations between the biomarker and disease states, including data showing effects of interventions on both the biomarker and clinical outcomes" (IOM, 2010, p. 2).

[6] Utilization analysis: "Contextual analysis based on the specific use proposed and the applicability of available evidence to this use. This includes a determination of whether the validation and qualification conducted provide sufficient support for the use proposed" (IOM, 2010, p. 2).

TABLE 5-2 Examples of Chronic Disease Endpoints Considered in Setting Previous Dietary Reference Intakes (DRIs)

Chronic Disease	Nutrient for Which a Chronic Disease DRI Was Established (DRI)	Basis for DRI	Nutrients for Which No Chronic Disease DRI Was Established[a]
Age-related Macular Degeneration	—	—	Lutein/zeaxanthin
Cancer	—	—	Vitamin C, beta-carotene, dietary fiber, vitamin E, selenium, calcium, vitamin D, choline, folate, calcium, vitamin D
Cardiovascular Disease	-	—	Magnesium, choline, vitamin B6, calcium, vitamin D
Cataracts	—	—	Vitamin C, Vitamin E, riboflavin
Coronary Heart Disease	Potassium (AI)[b]	Blood pressure[c] (including reduction of salt sensitivity)	
	Sodium (UL)	Blood pressure[c]	
	Dietary fiber (AI)	Decreased disease risk in prospective cohort studies	
Dental Caries	Fluoride (AI)	Decreased disease risk	Calcium
Type 2 Diabetes			Magnesium, vitamin E, chromium, dietary fiber, vitamin D
Immune Response[d]			Vitamin D
Kidney Stones (Recurrent)	Potassium (AI)[b]	Reduced risk of kidney stones	
	Calcium (UL)	Increased risk of kidney stones	

continued

TABLE 5-2 Continued

Chronic Disease	Nutrient for Which a Chronic Disease DRI Was Established (DRI)	Basis for DRI	Nutrients for Which No Chronic Disease DRI Was Established[a]
Neuropsychological Function[e]			Folate, vitamin D, choline
Osteoporosis (Bone Health)	Calcium[f] (EAR/RDA)	Decreased fracture risk, Bone mineral density[c]	
	Vitamin D (EAR/RDA)	Decreased fracture risk, Bone mineral density[c]	
Pulmonary Disease			Vitamin C

NOTES: AI = adequate intake; DRI = Dietary Reference Intake; EAR = Estimated Average Requirement; RDA = Recommended Dietary Allowance; UL = Tolerable Upper Intake Level.

[a] In most cases, chronic disease outcomes were not used to set DRIs because the evidence was insufficient and/or inconsistent.

[b] The AI for potassium was based on multiple indicators related to risk of both cardiovascular disease (through the surrogate outcome of blood pressure) and of recurrent kidney stones.

[c] Indicates that a surrogate outcome qualified by FDA was used to set the DRI.

[d] Chronic diseases associated with the immune response include asthma, type 1 diabetes, inflammatory bowel disease, multiple sclerosis, rheumatoid arthritis, and other conditions.

[e] Chronic diseases associated with neuropsychological function include autism, cognitive decline, depression, Alzheimer's disease, and other conditions.

[f] Indicators of adequacy used to set the EAR and RDA for calcium varied by age, although they were related to bone health in all cases. For growing children, bone accretion and positive calcium balance were primary indicators; for adults up to age 50 years, calcium balance was the primary indicator; for older adults, bone density and fracture risk were used as indicators of adequacy.

SOURCES: IOM, 2011; Taylor, 2008.

Adequate Intake[7] (AI) was set, rather than an Estimated Average Requirement[8] (EAR) because of lack of data for chronic disease indicators.

NIH established the Biomarkers of Nutrition for Development (BOND)

[7] Adequate Intake is the average daily nutrient intake observed in apparently healthy individuals in a specific sex and age group. It is based on experimentally derived intake levels or observations of mean nutrient intakes by a group of apparently healthy people who are maintaining a defined criterion of adequacy.

[8] Estimated Average Requirement is the average daily intake of a nutrient that is expected to meet the requirement of half of healthy individuals in a group defined by age and sex. The requirement is based on a specific indicator of adequacy.

initiative, which was aimed to harmonize the process for identifying biomarkers for nutrition and development (see Box 5-2), but this initiative is no longer funded and, as of today, no efforts are ongoing to identify qualified surrogate outcomes to be used in establishing chronic disease DRIs (NIH, 2017).

As mentioned in Chapter 2, establishing reference values for adequacy for essential nutrients has been considered a critical task because of their importance to health. Chronic disease DRIs, however, are desirable but not essential. Therefore, despite the movement toward using surrogate markers in nutritional applications to chronic disease prevention, the committee urges caution in applying them, unless they meet a high bar for being considered qualified (as discussed in the next section). Some of the caution in this respect comes from experiences with drug studies in which the intervention affected the surrogate in the expected direction but did not have commensurate benefits on the disease outcome (see Lipska and Krumholz, 2017, for controversies around whether hemoglobin A1c can be used as surrogate in studies of diabetes). This suggests that either that the surrogate was not a true surrogate with respect to the causal pathway of

BOX 5-2
Biomarkers of Nutrition for Development

The Biomarkers of Nutrition for Development (BOND) was an initiative of NIH's *Eunice Kennedy Shriver* National Institute of Child Health and Human Development. BOND aimed to harmonize the process for identifying "the best available biomarkers for a given use and provide researchers, clinicians, policy makers, and others with the evidence they need to make informed decisions and diagnoses in the field," including for the development of evidence-based DRIs (NIH, 2017). BOND used the NIH definition of biomarkers as distinct biological or biologically derived molecules from body fluids or tissues that report on a process, event, condition, or disease.

BOND established a process to review the scientific underpinnings supporting the use of nutrient biomarkers to assess functional and health effects of diet and dietary components on health and disease outcomes in individuals and populations (Raiten et al., 2011). The process systematically assessed the relative strengths and weaknesses of nutrition-related biomarkers and the factors and contexts that modify their relationship to a given exposure or outcome. This information was intended to foster consensus among the scientific community for the selection and use of appropriate biomarkers. During the initiative, comprehensive biomarker reviews were conducted for the nutrients folate (Bailey et al., 2015), iodine (Rohner et al., 2014), vitamin A (Tanumihardjo et al., 2016), and zinc (King et al., 2016).

interest (but perhaps only a correlated disease marker) and/or that harms associated with the mechanism of drug action on the surrogate obviated benefits downstream of the intervention. Other examples are homocysteine as a surrogate for cardiovascular disease (Ganguly and Alam, 2015), HDL-cholesterol as surrogate for cardiovascular disease (Mahdy et al., 2012), and biomarkers of Alzheimer's disease (Frisoni and Visser, 2015; Sharma and Singh, 2016). Whether this would be the case with a nutrient intervention is uncertain, although examples of unexpected effects could be cited. For example, unexpected adverse effects on lung cancer and cardiovascular disease were found in a large intervention trial testing high-dose beta-carotene plus vitamin A versus placebo conducted in heavy smokers (ATBC Cancer Prevention Study Group, 1994), when previous observational studies at lower intakes from foods had suggested reduced risk of cancer (NRC, 1982; Peto et al., 1981). One potential explanation is that the dose that is given results in a different effect. However, an alternative explanation could be that because study participants were at high risk of developing lung cancer, they possibly already had preneoplastic lesions. Therefore, the treatment could have acted as a promoter of early stage carcinogenesis (versus intake of carotenoids/vitamin A earlier in life seen in observational studies, which might lower cancer initiation). NOFSs may affect candidate surrogate markers by different mechanisms than will various drugs, requiring specific assessment of effects on both surrogates and outcomes as a criterion for qualification. On the other hand, nutrient-based interventions that fall within the established DRI limits for adequacy and toxicity would be presumed safe, unlike pharmacologic interventions for which safety must be established on a case-by-case basis.

RECOMMENDATION FOR OPTIONS AND JUSTIFICATIONS

The Options Report presents a conceptual framework with three scenarios for assessing whether the relationship between an NOFS and a chronic disease is causal: (1) direct assessment, where both the intake and the chronic disease itself are measured, (2) indirect assessment using a qualified surrogate disease marker as a substitute for the measurement of a chronic disease, or (3) indirect assessment using a nonqualified surrogate disease marker. Two options were offered for consideration (see Box 5-3). This committee was tasked with recommending and justifying one of these two options related to selecting chronic disease outcomes or biomarkers of effect when reviewing the evidence related to establishing DRIs based on the chronic disease outcome of interest. The committee supports a variant of option 1 (see Box 5-3), where studies that measure qualified surrogate markers—following the criteria adopted by the committee in Table 5-3—are considered in evaluating the evidence about causal relationships. The

BOX 5-3
Selecting Chronic Disease Endpoints

Options Report

Option 1: Endpoint (outcome) is the incidence of a chronic disease or a qualified surrogate disease marker

- This option would only accept study endpoints that are assessed by a chronic disease event as defined by accepted diagnostic criteria, including composite endpoints, when applicable, or by a qualified surrogate disease marker. These types of endpoints are associated with higher levels of confidence that the food substance and chronic disease relation is causal than are nonqualified disease markers.

Option 2: Endpoint (outcome) may include nonqualified disease markers

- To implement this option, a DRI committee would also accept studies with outcomes that are possible predictors of the chronic disease of interest but that have not been qualified as surrogate disease markers because they lack sufficient evidence for this purpose.

Committee's Recommendation 2

The ideal outcome used to establish chronic disease Dietary Reference Intakes should be the chronic disease of interest itself, as defined by accepted diagnostic criteria, including composite endpoints, when applicable. Surrogate markers could be considered with the goal of using the findings as supporting information of results based on the chronic disease of interest. To be considered, surrogate markers should meet the qualification criteria for their purpose. Qualification of surrogate markers must be specific to each nutrient or other food substance, although some surrogates will be applicable to more than one causal pathway.

committee does not support option 2, using nonqualified intermediate markers, because they could lead to serious misinterpretation of DRIs by users.

As described in Chapter 3, the evidence base for relationships between NOFSs and chronic disease outcomes includes studies that vary widely in design and health outcome measurement. Observational studies may include direct measurements of chronic disease outcomes in very large populations and over long follow-up periods, but also may include measurements of biomarkers of effect, qualified (i.e., surrogate markers) or nonqualified (i.e., nonqualified intermediate markers). In observational studies with prospective designs (e.g., cohort studies or nested case-control studies),

TABLE 5-3 List of Criteria for Consideration When Making a Decision About Whether a Surrogate Marker Is Qualified to Be Used as an Indicator of a Chronic Disease in the Development of a DRI

Criteria	Description
Analytical Validation	Assessing assays and measurement performance characteristics and determining the range of conditions under which the assays will give reproducible and accurate data.
Evidentiary Qualification	1. The surrogate marker is on a causal pathway in the disease pathogenesis (using Hill's criteria) (Hill, 1965). 2. The surrogate marker is significantly associated with the disease in the target population. 3. The surrogate marker changes consistently with the health outcome in response to the nutrition intervention. If the surrogate marker-clinical endpoint relationship persists over multiple interventions, it is thought to be more generalizable. 4. A change in the surrogate marker explains a clinically significant proportion of the change in response to the nutrition intervention.
Utilization Analysis	Defining the context of use: population and conditions for use to which the assessment applies, such as purpose and when in the development of the intervention the surrogate applies. Because idealized statistical requirements are rarely or never achievable, subjective assessment is necessary to determine when surrogate endpoints can be used. Other variables, such as morbidities and mortalities associated with the disease are important contextual considerations. For more details, see IOM, 2010.

SOURCES: Calder et al., 2017; Clarke, 2017; IOM, 2010.

the surrogate disease marker measures may supplement data on disease outcomes in assessing causality. RCTs are also prospective but often shorter in duration, with smaller populations, and therefore may rely on surrogate disease markers alone. That is, data on the relationship between NOFSs and the clinical outcome do not always exist. Recognizing that chronic disease outcomes are the ideal measurement, the committee recommends that when a surrogate marker is used as proxy for a chronic disease outcome, the evidence related to its qualification is reviewed and surrogate markers

that most faithfully reflect a chronic disease outcome for the purpose are identified. Table 5-3 lists criteria the DRI committees should consider when making a decision about whether a surrogate marker is qualified for its purpose. The criteria are based on the IOM report *Evaluation of Biomarkers and Surrogate Endpoints in Chronic Disease* (IOM, 2010), on the work of the International Life Sciences Institute Europe Marker Validation Initiative (Calder et al., 2017), and on the presentation to the committee by Dr. Robert Clarke (Clarke, 2017). Typically, the most challenging evidentiary criterion to meet is #4 under evidentiary qualification in Table 5-3. This criterion requires the surrogate to substantially explain the relationship between the NOFS and the chronic disease, thereby precluding both biological pathways that bypass the surrogate and NOFS effects on disease following surrogate process ascertainment. This is the major statistical criterion needed to ensure that the chronic disease is associated with the NOFS if and only if the surrogate is also associated with the NOFS (Buyse et al., 2000; Frangakis and Rubin, 2002; Prentice 1989; VanderWeele, 2013). Establishment of a suitable surrogate for a specific chronic disease in relation to a specific NOFS is a complex process. DRI committees will need to carefully evaluate the support for surrogacy claims in studies being reviewed. Using nonqualified disease markers or surrogate disease markers qualified for other purposes (e.g., for drug evaluation) in establishing DRIs could have detrimental effects, including misinterpretation of DRIs by users or unintended diet modifications.

REFERENCES

ACS (American Cancer Society). 2017. *Lifetime risk of developing or dying from cancer*. https://www.cancer.org/cancer/cancer-basics/lifetime-probability-of-developing-or-dying-from-cancer.html (accessed April 27, 2017).

Arneson, D., L. Shu, B. Tsai, R. Barrere-Cain, C. Sun, and X. Yang. 2017. Multidimensional integrative genomics approaches to dissecting cardiovascular disease. *Front Cardiovasc Med* 4:8.

ATBC (Alpha-Tocopherol Beta Carotene) Cancer Prevention Study Group. 1994. The effect of vitamin E and beta carotene on the incidence of lung cancer and other cancers in male smokers. *N Engl J Med* 330(15):1029-1035.

Bailey, L. B., P. J. Stover, H. McNulty, M. F. Fenech, J. F. Gregory, 3rd, J. L. Mills, C. M. Pfeiffer, Z. Fazili, M. Zhang, P. M. Ueland, A. M. Molloy, M. A. Caudill, B. Shane, R. J. Berry, R. L. Bailey, D. B. Hausman, R. Raghavan, and D. J. Raiten. 2015. Biomarkers of nutrition for development-folate review. *J Nutr* 145(7):1636S-1680S.

Beger, R. D., W. Dunn, M. A. Schmidt, S. S. Gross, J. A. Kirwan, M. Cascante, L. Brennan, D. S. Wishart, M. Oresic, T. Hankemeier, D. I. Broadhurst, A. N. Lane, K. Suhre, G. Kastenmuller, S. J. Sumner, I. Thiele, O. Fiehn, R. Kaddurah-Daouk, M., and for "Precision Medicine and Pharmacometabolomics Task Group"-Metabolomics Society Initiative. 2016. Metabolomics enables precision medicine: "A white paper, community perspective." *Metabolomics* 12(10):149.

Benziger, C. P., G. A. Roth, and A. E. Moran. 2016. The Global Burden of Disease Study and the preventable burden of NCD. *Glob Heart* 11(4):393-397.

Buyse, M., G. Molenberghs, T. Burzykowski, D. Renard, and H. Geys. 2000. The validation of surrogate endpoints in meta-analyses of randomized experiments. *Biostatistics* 1(1):49-67.

Calder, P. C., A. Boobis, D. Braun, C. L. Champ, L. Dye, S. Einother, A. Greyling, C. Matthys, P. Putz, S. Wopereis, J. V. Woodside, and J. M. Antoine. 2017. Improving selection of markers in nutrition research: Evaluation of the criteria proposed by the ILSI Europe Marker Validation Initiative. *Nutr Res Rev* 30(1):73-81.

Chowdhury, R., S. Warnakula, S. Kunutsor, F. Crowe, H. A. Ward, L. Johnson, O. H. Franco, A. S. Butterworth, N. G. Forouhi, S. G. Thompson, K. T. Khaw, D. Mozaffarian, J. Danesh, and E. Di Angelantonio. 2014. Association of dietary, circulating, and supplement fatty acids with coronary risk: A systematic review and meta-analysis. *Ann Intern Med* 160(6):398-406.

Clarke, R. 2017. *Pattern of lipid biomarkers to predict disease risk*. Presented at the Workshop of the Committee on Development of Guiding Principles for the Inclusion of Chronic Disease Endpoints in Future Dietary Reference Intakes, January 9, 2017, Washington, DC.

Collins, F. S., and H. Varmus. 2015. A new initiative on precision medicine. *N Engl J Med* 372(9):793-795.

FDA (U.S. Food and Drug Administration). 2017a. *Biomarker qualification program*. https://www.fda.gov/Drugs/DevelopmentApprovalProcess/DrugDevelopmentToolsQualificationProgram/BiomarkerQualificationProgram/default.htm (accessed April 27, 2017).

FDA. 2017b. *Biomarkers used as outcomes*. https://www.fda.gov/Drugs/DevelopmentApprovalProcess/DrugDevelopmentToolsQualificationProgram/BiomarkerQualificationProgram/ucm535926.htm (accessed May 9, 2017).

FDA/CFSAN (FDA/Center for Food Safety and Nutrition). 2009. *Guidance for industry: Evidence-based review system for the scientific evaluation of health claims-final*. https://www.fda.gov/RegulatoryInformation/Guidances/ucm073332.htm#ftn1 (accessed May 9, 2017).

Fearon, E. R., and B. Vogelstein. 1990. A genetic model for colorectal tumorigenesis. *Cell* 61(5):759-767.

Fontana, L., L. Partridge, and V. D. Longo. 2010. Extending healthy life span—from yeast to humans. *Science* 328(5976):321-326.

Frangakis, C. E., and D. B. Rubin. 2002. Principal stratification in causal inference. *Biometrics* 58(1):21-29.

Freedman, L. S., B. I. Graubard, and A. Schatzkin. 1992. Statistical validation of intermediate endpoints for chronic diseases. *Stat Med* 11(2):167-178.

Frisoni, G. B., and P. J. Visser. 2015. Biomarkers for Alzheimer's disease: A controversial topic. *Lancet Neurol* 14(8):781-783.

Ganguly, P., and S. F. Alam. 2015. Role of homocysteine in the development of cardiovascular disease. *Nutr J* 14:6.

Gavrilov, L. A., and N. S. Gavrilova. 2015. Predictors of exceptional longevity: Effects of early-life and midlife conditions, and familial longevity. *N Am Actuar J* 19(3):174-186.

GBD (Global Burden of Disease) 2013 Risk Factors Collaborators. 2015. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks in 188 countries, 1990-2013: A systematic analysis for the Global Burden of Disease Study 2013. *The Lancet* 386(10010):2287-2323.

Greenberg, E. R., J. A. Baron, T. D. Tosteson, D. H. Freeman, Jr., G. J. Beck, J. H. Bond, T. A. Colacchio, J. A. Coller, H. D. Frankl, R. W. Haile, J. S. Mandel, D. W. Nierenberg, R. Rothstein, D. C. Snover, M. M. Stevens, R. W. Summers, R. U. van Stolk, for the Polyp Prevention Study Group. 1994. A clinical trial of antioxidant vitamins to prevent colorectal adenoma. Polyp Prevention Study Group. *N Engl J Med* 331(3):141-147.

Hill, A. B. 1965. The environment and disease: Association or causation? *Proc R Soc Med* 58:295-300.

Hillerdal, G. 2008. Indolent lung cancers—Time for a paradigm shift: A review. *J Thorac Oncol* 3(3):208-211.

IOM (Institute of Medicine). 2010. *Evaluation of biomarkers and surrogate endpoints in chronic disease*. Washington, DC: The National Academies Press.

IOM. 2011. *Dietary Reference Intakes for calcium and vitamin D*. Washington, DC: The National Academies Press.

IOM. 2015. *Improving diagnosis in health care*. Washington, DC: The National Academies Press.

Jensen, H., and P. Vedsted. 2017. Exploration of the possible effect on survival of lead-time associated with implementation of cancer patient pathways among symptomatic first-time cancer patients in Denmark. *Cancer Epidemiol* 49:195-201.

Jordan, K., M. Porcheret, and P. Croft. 2004. Quality of morbidity coding in general practice computerized medical records: A systematic review. *Fam Pract* 21(4):396-412.

King, J. C., K. H. Brown, R. S. Gibson, N. F. Krebs, N. M. Lowe, J. H. Siekmann, and D. J. Raiten. 2016. Biomarkers of Nutrition for Development (BOND)—Zinc review. *J Nutr* 146(4):858S-885S.

Lewington, S., G. Whitlock, R. Clarke, P. Sherliker, J. Emberson, J. Halsey, N. Qizilbash, R. Peto, and R. Collins. 2007. Blood cholesterol and vascular mortality by age, sex, and blood pressure: A meta-analysis of individual data from 61 prospective studies with 55,000 vascular deaths. *The Lancet* 370(9602):1829-1839.

Lipska, K. J., and H. M. Krumholz. 2017. Is hemoglobin A1c the right outcome for studies of diabetes? *JAMA* 317(10):1017-1018.

Lloyd, J., E. Jahanpour, B. Angell, C. Ward, A. Hunter, C. Baysinger, and G. Turabelidze. 2017. Using national inpatient death rates as a benchmark to identify hospitals with inaccurate cause of death reporting—Missouri, 2009-2012. *Morb Mortal Wkly Rep* 66(1):19-22.

Mahdy Ali, K., A. Wonnerth, K. Huber, and J. Wojta. 2012. Cardiovascular disease risk reduction by raising HDL cholesterol—Current therapies and future opportunities. *Br J Pharmacol* 167(6):1177-1194.

Navin Cristina, T. J., J. A. Stewart Williams, L. Parkinson, D. W. Sibbritt, and J. E. Byles. 2016. Identification of diabetes, heart disease, hypertension and stroke in mid- and older-aged women: Comparing self-report and administrative hospital data records. *Geriatr Gerontol Int* 16(1):95-102.

Neuhouser, M. L., T. Y. Cheng, S. A. Beresford, E. Brown, X. Song, J. W. Miller, Y. Zheng, C. A. Thomson, J. M. Shikany, M. Z. Vitolins, T. Rohan, R. Green, and C. M. Ulrich. 2015. Red blood cell folate and plasma folate are not associated with risk of incident colorectal cancer in the Women's Health Initiative observational study. *Int J Cancer* 137(4):930-939.

Niccoli, T., and L. Partridge. 2012. Ageing as a risk factor for disease. *Curr Biol* 22(17): R741-752.

NIH (National Institutes of Health). 2017. *About BOND*. https://www.nichd.nih.gov/global_nutrition/programs/bond/Pages/about.aspx (accessed April 27, 2017).

NRC (National Research Council). 1982. *Diet, nutrition, and cancer.* Washington, DC: National Academy Press.

Ohlhorst, S. D., R. Russell, D. Bier, D. M. Klurfeld, Z. Li, J. R. Mein, J. Milner, A. C. Ross, P. Stover, and E. Konopka. 2013. Nutrition research to affect food and a healthy life span. *Am J Clin Nutr* 98(2):620-625.

Paraskeva, C., A. P. Corfield, S. Harper, A. Hague, K. Audcent, and A. C. Williams. 1990. Colorectal carcinogenesis: Sequential steps in the in vitro immortalization and transformation of human colonic epithelial cells (review). *Anticancer Res* 10(5A):1189-1200.

Peto, R., R. Doll, J. D. Buckley, and M. B. Sporn. 1981. Can dietary beta-carotene materially reduce human cancer rates? *Nature* 290(5803):201-208.

Prentice, R. L. 1989. Surrogate endpoints in clinical trials: Definition and operational criteria. *Stat Med* 8(4):431-440.

Raiten, D. J., S. Namaste, B. Brabin, G. Combs, Jr., M. R. L'Abbe, E. Wasantwisut, and I. Darnton-Hill. 2011. Executive summary—Biomarkers of Nutrition for Development: Building a consensus. *Am J Clin Nutr* 94(2):633S-650S.

Rappaport, S. M. 2016. Genetic factors are not the major causes of chronic diseases. *PLOS ONE* 11(4):e0154387.

Robb, M. A., P. M. McInnes, and R. M. Califf. 2016. Biomarkers and surrogate endpoints: Developing common terminology and definitions. *JAMA* 315(11):1107-1108.

Rohner, F., M. Zimmermann, P. Jooste, C. Pandav, K. Caldwell, R. Raghavan, and D. J. Raiten. 2014. Biomarkers of nutrition for development—Iodine review. *J Nutr* 144(8): 1322S-1342S.

Sharma, N., and A. N. Singh. 2016. Exploring biomarkers for Alzheimer's Disease. *J Clin Diagn Res* 10(7):KE01-KE06.

Shreiner, A. B., J. Y. Kao, and V. B. Young. 2015. The gut microbiome in health and in disease. *Curr Opin Gastroenterol* 31(1):69-75.

Sugarbaker, P. H., L. L. Gunderson, and R. E. Wittes. 1985. Cancer of the anal region. In *Cancer: Principles and Practice of Oncology*, edited by V. T. J. DeVita, S. Hellman and S. A. Rosenberg. Philadelphia, PA: Lippincott, Co.

Takata, Y., M. J. Shrubsole, H. Li, Q. Cai, J. Gao, C. Wagner, J. Wu, W. Zheng, Y. B. Xiang, and X. O. Shu. 2014. Plasma folate concentrations and colorectal cancer risk: A case-control study nested within the Shanghai Men's Health Study. *Int J Cancer* 135(9):2191-2198.

Tanumihardjo, S. A., R. M. Russell, C. B. Stephensen, B. M. Gannon, N. E. Craft, M. J. Haskell, G. Lietz, K. Schulze, and D. J. Raiten. 2016. Biomarkers of Nutrition for Development (BOND)—Vitamin A review. *J Nutr* 146(9):1816S-1848S.

Taylor, C. L. 2008. *Framework for DRI development: Components "known" and components "to be explored."* Washington, DC

Thiru, K., A. Hassey, and F. Sullivan. 2003. Systematic review of scope and quality of electronic patient record data in primary care. *BMJ* 326(7398):1070.

VanderWeele, T. J. 2013. Surrogate measures and consistent surrogates. *Biometrics* 69(3): 561-569.

Waterland, R. A., and K. B. Michels. 2007. Epigenetic epidemiology of the developmental origins hypothesis. *Annu Rev Nutr* 27:363-388.

WHO (World Health Organization). 2012. *Effect of increased potassium intake on cardiovascular disease, coronary heart disease and stroke.* Geneva, Switzerland: World Health Organization.

Yetley, E. A., A. J. MacFarlane, L. S. Greene-Finestone, C. Garza, J. D. Ard, S. A. Atkinson, D. M. Bier, A. L. Carriquiry, W. R. Harlan, D. Hattis, J. C. King, D. Krewski, D. L. O'Connor, R. L. Prentice, J. V. Rodricks, and G. A. Wells. 2017. Options for basing Dietary Reference Intakes (DRIs) on chronic disease endpoints: Report from a joint US-/Canadian-sponsored working group. *Am J Clin Nutr* 105(1):249S-285S.

6

Evidence Review: Judging the Evidence for Causal Relationships

Previous chapters illustrate the current process and framework (see Chapter 1, Figure 1-2) as well as the relevant activities to date (see Chapter 2) for developing Dietary Reference Intakes (DRIs) based on indicators of adequacy, toxicity, or chronic disease. Although the current DRI process considers all three indicators, this committee has been asked to provide recommendations and guiding principles for developing chronic disease DRIs, specifically. As mentioned in Chapter 1, a key activity relevant to chronic disease endpoints within the purview of DRI committees is evaluating the certainty of the evidence with regard to two main questions: (1) What are acceptable levels of confidence that the relationship between an NOFS and a chronic disease is causal? and (2) If a causal relationship exists, what are acceptable levels of confidence in the intake-response relationship data, what are approaches for identifying and characterizing such relationship, and if appropriate, to recommend chronic disease DRIs? This chapter discusses the methodological tools that can be used to answer the first question (does a causal relationship exist?). The chapter starts with an overview of the conduct of a systematic review, as a modern tool to answer scientific questions and to provide context for the reminder of the chapter. It also provides an overview of the systems and tools that are often used to evaluate scientific evidence. Finally, it discusses in more depth the strengths and cautions in applying one of those systems, Grading of Recommendations, Assessment, Development and Evaluation (GRADE), in determining causal associations between nutrients or other food substances (NOFSs) and chronic diseases for DRI development. Resources for more detailed application of the procedures are provided. At the end of the chapter, the

guiding principles as foundations for a scientifically credible chronic disease DRI process are listed. A list of key terms and their definitions is in Box 6-1.

Throughout the report and particularly within this chapter, the committee makes a clear distinction between the task of the systematic review team, which conducts the systematic review, and the task of the DRI committee, which reviews the totality of evidence and recommends chronic

BOX 6-1
Key Terminology

Bias: A systematic error or deviation in results or inferences from the truth. The main types of bias arise from systematic differences in the groups that are compared (**selection bias**), exposure to other factors apart from the intervention of interest (**performance bias**), withdrawals or exclusions of people entered into a study (**attrition bias**), or inaccuracies in the dietary intake or outcome assessment methodologies (**ascertainment bias**). Systematic reviews of studies may also be particularly affected by **reporting bias**, where a biased subset of all the relevant data is available. **Risk of bias** (internal validity) is the evaluation of systematic error due to limitations in the study design or execution. More rigorously designed (better quality) trials are more likely to yield results that are closer to the truth.

Case-control study: An observational study that identifies "cases" based on a diagnosis of a disease or identification of risk factors. "Controls" are those who are without the disease or risk factor. A case-control study compares characteristics of the cases to those of the controls to determine what risk factors may account for who does or does not get the disease being studied. This design is particularly useful where the outcome is rare and past exposure can be validly measured. Measures of past exposure obtained after diagnosis (retrospective case-control studies) are more likely subject to biases that compromise validity than when measures obtained substantially before diagnosis, as in "nested" case-control studies.

Certainty (as it relates to judgments about evidence): The extent to which one can be confident that an estimate of effect is correct.

Cohort study: An observational study in which a defined group of people (the cohort) is without the disease of interest at the time of cohort enrollment and is followed over time, often for many years. The disease outcomes of people in the cohort are compared, to examine people who were exposed or not exposed (or exposed at different levels) to a particular factor (exposure) of interest. A **prospective** cohort study assembles participants and follows them into the future. A **retrospective** (or historical) cohort study identifies subjects from past records and follows them from the time of those records to the present.

disease DRIs, if appropriate. A systematic review is a scientific investigation that focuses on a specific question and that uses explicit, planned scientific methods to identify, select, assess, and summarize the findings of similar but separate studies. It may or may not include a quantitative synthesis (meta-analysis) of the results from separate studies (IOM, 2011b). A systematic review is typically conducted by a group of experts in the process itself

Confidence interval: A measure of the uncertainty around the main finding of a statistical analysis. Estimates of unknown quantities, such as the relative risk comparing an experimental intervention with a control, are usually presented as a point estimate and a 95 percent confidence interval. This means that if someone were to keep repeating a study in other samples from the same population, 95 percent of the calculated confidence intervals from those studies would include the true underlying value. Wider intervals indicate less precision; narrow intervals, greater precision.

Cross-sectional study: An observational study that analyzes data collected from a population, or a representative subset, *at a specific point in time*—that is, cross-sectional data.

Evidence profile: Presentation of detailed information about the quality of evidence assessed and the summary of findings for each of the included outcomes. It presents information about the body of evidence (e.g., number of studies), the judgments about the underlying quality of evidence, key statistical results, and the quality of evidence rating for each outcome. Guideline panels (e.g., Dietary Reference Intake [DRI] committees) are expected to review evidence profiles to ensure that members agree about the judgments underlying the quality assessments.

Guideline panel: A panel of a knowledgeable, multidisciplinary group of experts and representatives from key affected groups that are charged with developing clinical practice guidelines. Standards for panel composition and managing members' conflicts of interests exist and should be followed as closely as possible. In the DRI process, a DRI committee is equivalent to the guideline panel in the Clinical Practice Guideline process.

Heterogeneity: The variation in study outcomes within the body of evidence for a particular outcome. It can be due to variability in participants, outcomes, or interventions, or intake response (clinical heterogeneity) or to variability in methods used, such as blinding, participant recruitment or data collected (methodological heterogeneity).

continued

BOX 6-1 Continued

Imprecision: Measurement of random error, which often occurs when studies within the body of evidence for a particular outcome have a small sample size and the number of events is also small, resulting in a wide 95 percent confidence interval around the estimate of the effect.

Indirectness: A situation when the body of evidence for particular outcome studies does not directly compare the interventions of interest, apply the intervention to the population of interest, or measure the important outcomes.

Inconsistency: Unexplained heterogeneity or variability of in the body of evidence for a particular outcome.

Meta-analysis: A systematic review technique that uses statistical methods to combine quantitatively the results of similar studies in an attempt to allow inferences to be made from the sample of studies and be applied to the population of interest.

Observational study: A study in which the investigators do not intervene, but simply observe a study population. Changes or differences in characteristics or exposures are studied in relation to changes or differences in other characteristic(s) (e.g., whether or not individuals died), without action by the investigator. Observational studies have a greater risk of selection bias and ascertainment bias than do experimental studies. Cross-sectional studies, cohort studies, and case-control studies are types of observational studies.

PICO: A technique used in evidence-based practice to frame and answer a clinical or health care–related question. The PICO framework is also used to develop literature search strategies. The PICO acronym stands for population (P), intervention (I), comparator (C), and outcome (O).

Publication bias: A systematic under-estimation or over-estimation of the underlying beneficial or harmful effect due to the selective publication of studies.

Quasi-experiment: Experimental research designs that test causal hypotheses of an intervention. In contrast to a randomized controlled trial, a quasi-experiment lacks random assignment and assignment to conditions (e.g., treatment versus no treatment or comparison condition) is by means of self-selection or administrator selection. Quasi-experimental designs identify a comparison group that is as

similar as possible to the treatment group in terms of baseline (pre-intervention) characteristics.

Randomized controlled trial (RCT): An experimental study in which two or more interventions are compared by being randomly allocated to participants. In most RCTs, one intervention is assigned to each individual, but sometimes an assignment is to defined groups of individuals (e.g., in a household, worksite, or a community) or interventions are assigned within individuals (e.g., in different orders or to different parts of the body).

Review of the totality of the evidence: In the context of setting chronic disease DRIs, it refers to evaluating the evidence about whether a chronic disease DRI should be recommended. It includes assessing the systematic review evidence profiles, quantitatively characterizing the intake-response, considering relationships with various chronic diseases and potential overlapping benefits and harms, determining the need for, and appropriateness of, extrapolation to other populations, and reviewing other relevant evidence.

Synthesis of evidence: Evaluating the body of evidence collected in a systematic manner and using quantitative and qualitative synthesis strategies. Standards for methods to synthesize the evidence include the use of consistent language to characterize the level of certainty in the estimates of the effect and the use of criteria for evaluating the body of evidence (i.e., risk of bias, consistency, precision, directness, and publication bias), including specific criteria for evaluating bodies of evidence of observational studies (i.e., dose-response association, plausible confounding, and size of the effect).

Systematic review: A scientific investigation that focuses on a specific question and that uses explicit, planned scientific methods to identify, select, assess, and summarize the findings of similar but separate studies. It may or may not include a quantitative synthesis (meta-analysis) of the results from separate studies.

Systematic review team: A group of experts contracted specifically to conduct a systematic review.

Technical Expert Panel: A group of subject matter experts who serve as consultants to the systematic review team in scientific matters related to the questions of interest.

that includes subject matter experts (e.g., a systematic review team) and in consultation with external subject matter experts (e.g., a technical expert panel). In contrast, the DRI committee's review consists of drawing conclusions about the totality of the findings. The guidelines and recommendations in this chapter are envisioned to be followed in the context of the current process, where DRI committees do not conduct systematic reviews themselves but are the recipients of such reviews. Based on the systematic reviews, the DRI committee is tasked with the review of the totality of the evidence, including evidence about quantitative intake-response relationships described in Chapter 7, and, if appropriate, with recommending a DRI. It is further assumed that the initial and any subsequent systematic reviews and protocols are conducted independently of the DRI committee. The first use of this type of systematic review and evidence review process for DRI purposes was reflected in the development of the 2011 update of the DRIs for calcium and vitamin D, which included consideration of chronic disease endpoints (Brannon et al., 2014; IOM, 2011a) (see the vitamin D case example in Chapter 3).

This chapter particularly considers the nutrition-specific issues outlined in Chapter 3. It draws on other sections of the report that discuss methods to measure dietary intake (see Chapter 4), and disease outcomes (see Chapter 5) and leads up to the discussion, in Chapter 7, of approaches to specifying intake-response relationships when causation has been established with sufficient certainty (i.e., the second main question). In addition, although it is not the task of this report to provide direction in moving to guidelines and policies, in various places the committee offers comments on how the conduct of systematic reviews can and should anticipate the guideline decisions that follow them as best as possible, in order to facilitate that process.

OVERVIEW OF THE SYSTEMATIC REVIEW PROCESS

In the context of setting DRIs, the systematic review process is an essential activity before the totality of the evidence is reviewed by the DRI committee. Figure 6-1 is an overview of an ideal systematic review as conceptualized in the IOM report *Finding What Works in Health Care: Standards for Systematic Reviews* (IOM, 2011b). Depending on the context for a systematic review and the group that conducts it, the specific nature of the activities may vary. This section includes a description of some of the activities highlighted as especially important in the context of developing DRIs. These include formulating the questions, registering the systematic reviews and adopting an appropriate format for reporting methods and results, and anticipating additional information needs beyond the systematic review results.

Problem Formulation, Scoping and Developing a Systematic Review Protocol

The initial selection of particular NOFSs and health outcomes is within the purview of the federal agencies (see Figure 1-2), which prioritize NOFSs and health outcomes for study based on public health priorities and other factors. Whatever the choices, each one implies some tentative perspective that a scientific basis and a literature exist to proceed. Therefore, once the general scientific questions that form the basis for the systematic evidence review have been formulated, an important next preliminary step is to sample the available literature, to determine that it is sufficiently robust to make the review useful for each of the central scientific questions identified. Often, a preliminary literature review may be available, including previously conducted systematic reviews and meta-analyses. If not, a preliminary "scoping" review should be conducted (see, e.g., Brannon et al., 2016). If the studies are clearly insufficient, then setting a DRI is not scientifically justifiable. In addition to assessing the literature's breadth, depth, and relevance, the scoping exercise could have other potential dividends, such as determining the need for translation services and identifying additional keywords to aid in searching. This step is different from conducting the comprehensive literature search. Scoping may also identify systematic reviews, including meta-analyses already published on nutrition-disease topics. These previously published systematic reviews can help target relevant literature and also can be incorporated as part of the evidence review. Instruments are available to assess systematic review quality (e.g., AMSTAR [Pieper et al., 2014; Shea et al., 2007, 2009] and risk of bias [ROBIS] [Whiting et al., 2013, 2016]). It may be possible to add recent research reports to existing systematic reviews for reasons of efficiency if the search is comprehensive up to a particular date and all other quality criteria are met (Chung et al., 2012; Garner et al., 2016; Shekelle et al., 2009).

Any comprehensive systematic review can have multiple scientific questions (e.g., multiple dimensions of efficacy and safety of an NOFS-focused intervention, including nutrient-nutrient interactions, varied effects in subpopulations defined by age or other characteristics, and effects on multiple disease outcomes or surrogate markers). Each question should be specific and clearly defined, and narrow enough to be aligned with the evidence base under review. Framing questions as specific scientific hypotheses is essential.

The structure for questions to guide systematic and evidence reviews involves detailed and explicit characterization of the study level and design details, including target population, candidate interventions, and outcomes important to individuals. The PICO (population, intervention, comparator,

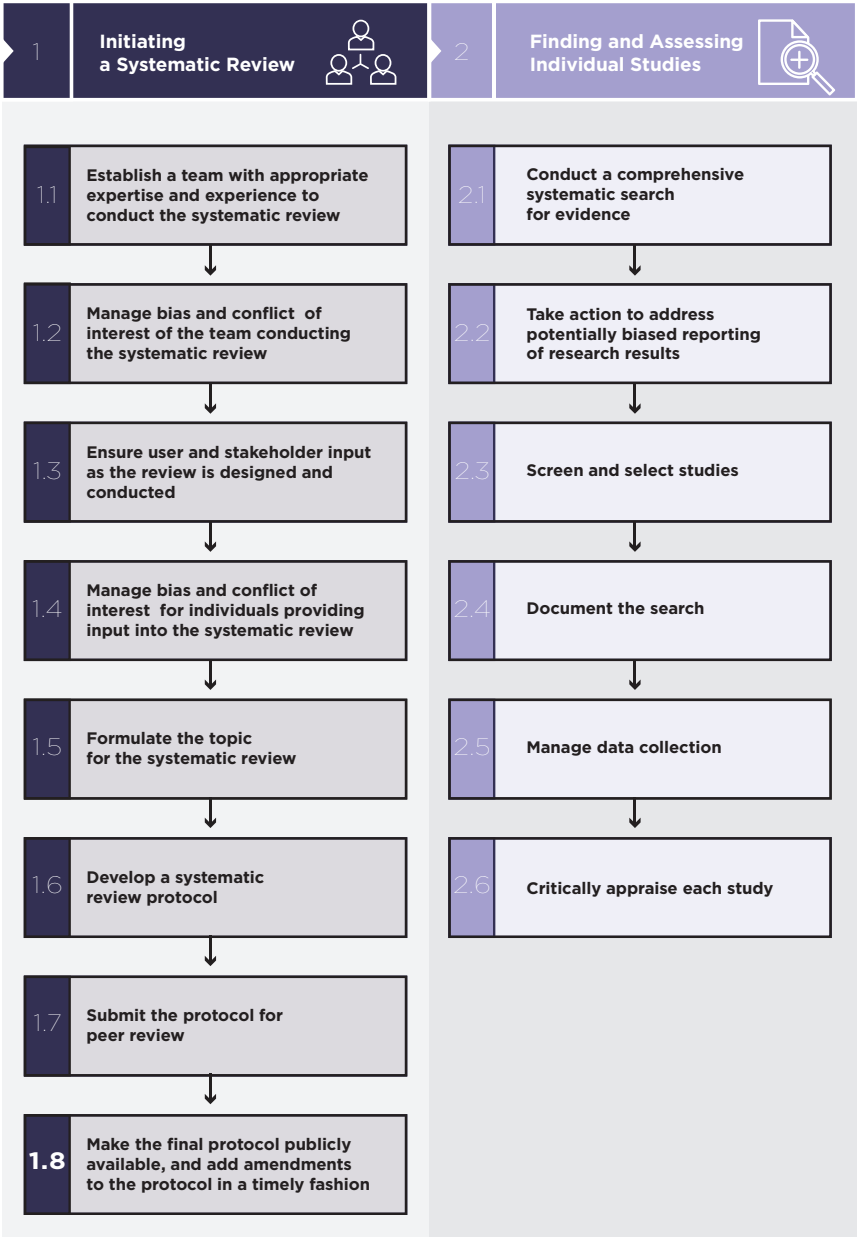


FIGURE 6-1 Standards for systematic reviews.
SOURCE: IOM, 2011b.

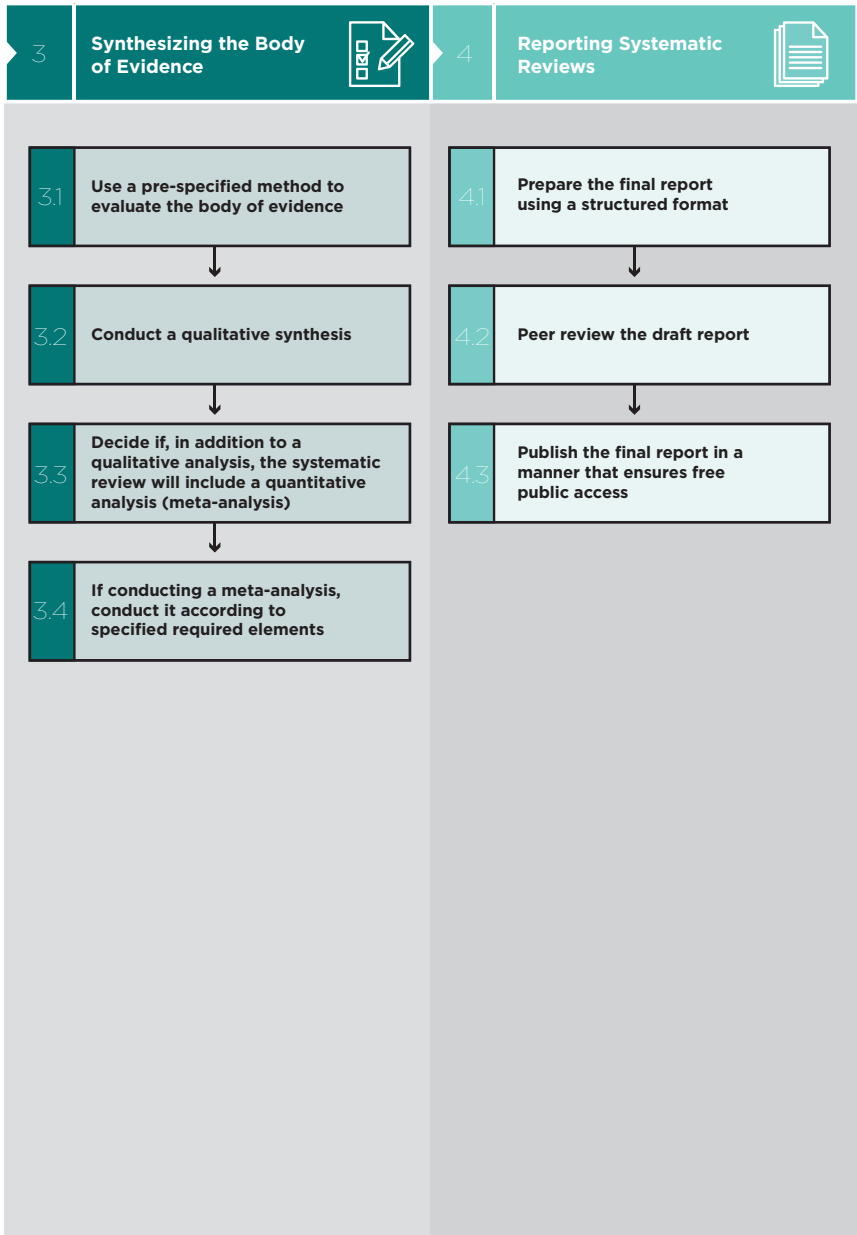


FIGURE 6-1 Continued

and outcome) structure[1] is considered helpful in bringing focus and order to the questions at hand (Guyatt et al., 2011b). In the context of chronic disease DRIs, key questions would be formulated, and the parameters of interest to the sponsoring agency or agencies in consultation with a technical expert panel would be specified using the PICO framework. The vitamin D case example in Chapter 3 demonstrates how a map of the various biological pathways of interest can be created as an analytic framework to guide the process of framing the key questions (see Figure 3-1). The intervention and comparator in these questions are implicit, i.e., effects of interest are not specified with respect to a given level of vitamin D intake. Data abstraction methods that identify the P, I, C, and O elements for each study are included in the systematic review protocol. As noted in Chapter 3, because everyone in the population consumes NOFSs at some level, for the purpose of establishing a causal relationship, some questions may be framed as an increase or decrease relative to existing consumption patterns rather than above or below a certain threshold.

As another example, formulated questions for systematic reviews to support the World Health Organization (WHO) guideline on free sugars[2] were as follows: "What is the effect of a decrease or increase in free sugars in adults and children?" and "What is the effect of restricting intake of free sugars to below 10 percent of total energy intake?"—accompanied by a definition of the term "free sugars" (WHO, 2015). Annex 6-1 shows the full detail of the free sugar questions in a PICO table format that specifies population considerations, exposure definitions, comparisons, confounders, effect modifiers and intermediates, outcomes, and time frames (WHO, 2015). Evidence on these questions about sugar was gathered through several systematic reviews (peer reviewed and published) in order to inform the eventual guideline. The questions in Box 6-2, also drawn from systematic reviews to support WHO guideline development, illustrate more complex NOFS-chronic disease questions using the general PICO approach (Brouwer, 2016; Mensink, 2016). As with the free sugar questions, these questions frame "interventions" about intake level in terms of increases or decreases, sometimes related to specific thresholds in association with particular outcomes—in this case blood lipids as surrogates for noncommunicable disease risk. Other questions address effects of different types

[1] Some groups use PICOTS: population, intervention, comparators, outcomes, timing, and setting.

[2] Free sugars is defined by the United Nations World Health Organization and the Food and Agriculture Organization as "all monosaccharides and disaccharides added to foods by the manufacturer, cook, or consumer, plus sugars naturally present in honey, syrups, and fruit juices" (WHO, 2015, p. 1). In the United States, the term "free sugars" is generally equivalent to "added sugars," the term used by the 2015-2020 *Dietary Guidelines for Americans*.

BOX 6-2
Examples of Questions Guiding Systematic Reviews to Inform WHO Guidelines for Intakes of Saturated and *Trans* Fatty Acids

Saturated Fatty Acid Intake

1. What is the effect in the population of reduced percentage of total energy intake from saturated fatty acids (SFAs) relative to higher intake for reduction in risk of noncommunicable diseases (NCDs)?
2. What is the effect in the population of consuming less than 10 percent of total energy as SFA relative to more than 10 percent total energy as SFA for reduction in risk of NCDs?
3. What is the effect in the population of a reduction in percentage of total energy intake from SFA from 10 percent in gradual increments relative to higher intake for reduction in risk of NCDs?
4. What is the effect in the population of reduced percentage of total energy intake from long-chain SFA, very long-chain SFA, and medium-chain SFA relative to higher intake for reduction in risk of NCDs?
5. What is the effect in the population of reduced percentage of total energy intake from lauric acid, myristic acid, palmitic acid, or stearic acid relative to higher intake for reduction in risk of NCDs?
6. What is the effect in the population of replacing SFA with carbohydrates (refined versus unrefined), cis-monounsaturated fatty acids (cis-MUFA), cis-polyunsaturated fatty acids (cis-PUFA), protein, or *trans* fatty acids (TFAs) relative to no replacement for reduction in risk of NCDs?

SOURCE: Mensink, 2016, Annex 1.

***Trans* Fatty Acid Intake**

1. What is the effect in the population of reduced percentage of total energy intake from *trans* fatty acids (TFAs) relative to higher intake for reduction in risk of noncommunicable diseases (NCDs)?
2. What is the effect in the population of a reduction in percentage of total energy intake from TFA from 1 percent in gradual increments relative to higher intake for reduction in risk of NCDs?
3. What is the effect in the population of reduced percentage of total energy intake from industrial/ruminant TFA relative to higher intake for reduction in risk of NCDs?
4. What is the effect in the population of consuming 0 percent of total energy intake as industrial/ruminant TFA relative to >0 percent of total energy intake as industrial/ruminant TFA intake for reduction in risk of NCDs?
5. What is the effect in the population of reduced percentage of total energy intake from 18:2n-6/18:3n-3 isomers of TFA relative to higher intake for the reduction in risk of NCDs?

continued

BOX 6-2 Continued

6. What is the effect in the population of replacing percentage of total energy intake from TFA with conjugated linoleic acid (CLA) isomers (9-cis, 11-trans and 10-trans, 12-cis)?
7. What is the effect in the population of replacing TFA with monounsaturated fatty acids, polyunsaturated fatty acids, carbohydrates (refined versus unrefined), or saturated fatty acids, relative to no replacement on reduction in risk of NCDs?

SOURCE: Brouwer, 2016, Annex 1.

of saturated or *trans* fat and substitution or replacement effects that are relevant to food manufacturing.

Systematic Review Registration and Report Format

Once the systematic review team is assembled and as the systematic review process begins, the committee strongly suggests registering the review in advance with the international prospective register of systematic reviews (PROSPERO).[3] This registration applies to the initial systematic review and any subsequent reviews needed. Registering the review in advance with the PROSPERO system is mandatory for the Agency for Healthcare Research and Quality (AHRQ) Evidence-based Practice Center (EPC) program, which currently leads the systematic reviews for setting DRIs and has value in several ways, such as avoiding duplication of effort and promoting transparency. In planning the report, the team should follow the Preferred Reporting Items for Systematic Reviews and Meta-Analyses for Protocols (PRISMA-P) approach,[4] which specifies structure and completeness for reporting systematic reviews and meta-analyses.

Anticipating Additional Literature Needs When Planning Systematic Reviews

When planning a systematic review, anticipating downstream needs for literature searches and guideline creation can be important, particularly if the systematic review cannot be repeated or re-opened after completion (e.g., due to financial or other constraints). Examples of such activities

[3] See https://www.crd.york.ac.uk/PROSPERO.
[4] See http://www.prisma-statement.org.

include collecting literature on validation of potential surrogate markers or biomarkers of intake, and the general health characteristics or diets of the target populations of interest, identifying disease risks in those groups of special interest, searching for published computational or other public health decision models that could aid in recommending the DRI, identifying studies of adverse effects of selected NOFS interventions, and cataloging relevant literature on population interventions. These tasks will be accomplished by the systematic review team as it develops its search strategy.

The committee also notes that mechanistic studies, whether in humans, animals, or various in vitro systems, may be important to generating hypotheses regarding the ability of an NOFS to affect the risk of developing a chronic disease. The World Cancer Research Fund Expert Panel on Diet, Nutrition, and Cancer Prevention, for example, used a hierarchy of robustness based on categorization of human and animal experimental findings into three classes of which only one class was thought to be useful for making judgments about diet and cancer risk (WCRF/AICR, 2007). Although it is not generally practical to include mechanistic studies in systematic reviews of the effect of NOFSs on health, mechanistic studies can contribute to addressing the study question by providing additional information. Thus, it would be important to anticipate mechanistic research questions at the start of the systematic review process. Box 6-3 lists some of the important information that can be gathered from mechanistic studies.

Forming the Systematic Review Team, Technical Expert Panel, and DRI Committee

Developing chronic disease DRIs depends on the contributions of many stakeholders, including the federal agencies, groups of experts, and the public. Within the current process, three groups of experts—the systematic review team, the technical expert panel, and the DRI committee—are involved in different capacities and time frames. Briefly, the systematic review team conducts the systematic review in consultation with the technical expert panel, and the DRI committee reviews the totality of the evidence and recommends chronic disease DRIs, if appropriate. Formation of these three groups necessitates a formal process. Although no dictates on group structure and the disciplines represented in any of the three groups have been established, experience suggests that several types of expertise are of great value. While the systematic review team includes mainly methodologists, the technical expert panel would include relevant content experts in nutrition science, toxicology, scientific study design and analysis, public health, biostatistics, nutrition epidemiology and chronic disease epidemiology, and disease pathogenesis. In terms of the DRI committee, disciplines include content experts for the main themes of the review, methodologists

BOX 6-3
Value of Mechanistic Studies as DRIs Are Established

The following are examples of information that can be gathered from mechanistic studies:

- Toxicological studies in various animal test systems can help anticipate adverse effects of interventions.
- Mechanistic studies can assay for food contaminants and provide alerts about potential health effects.
- Test systems can assess the presence of potential nutrient–nutrient interactions.
- Nutrient-drug studies can be performed to help screen for or determine the presence of food-drug interactions in order to anticipate unwanted difficulties with drug therapy.
- Mechanistic studies can identify population subgroups that may respond differently to a particular NOFS.
- Studies of disease mechanisms can help provide evidence for the validity of surrogate markers as predictors of specific diseases.
- Mechanistic study findings can help direct research agendas related to human observational and intervention studies to gather the necessary data to move population policies forward.

experienced with the evidence review approach, biostatisticians, and persons with experience in guideline construction and policy applications. The size of the group should be consonant with the breadth of review topics, and available resources.

To maintain objectivity, scientific rigor, and acceptability of a systematic review, it is important to recognize and manage potential bias and conflict of interests of the individuals conducting or providing input into the systematic review (IOM, 2011b). AHRQ, which is the agency that in the current DRI process has been responsible for administering the systematic review for DRI development, has policies to reduce the risk of bias from financial and non-financial interests. This requirement was also highlighted in a recent report from the National Academies of Sciences, Engineering, and Medicine (the National Academies) in the context of advisory panels (NASEM, 2017). Specifically, the report made recommendations on the selection process, including disclosure of how conflict of interest and biases are identified and managed, of members of the Dietary Guidelines Advisory

Committee.[5] The National Academies, which has convened all DRI committees in the past, has policies for assessing potential conflicts of interest among committee members.[6] To promote transparency and enhance public and stakeholder assurances of committees, the National Academies publishes the biographies of individual members for public review and comment.

Regardless of the subject matter and the disciplines required, it is important to establish a mechanism for communication among the three groups that will ensure a moderate level of communication while also protecting against inappropriate influence on the systematic review methods.

SYSTEMS OF EVIDENCE SYNTHESIS AND THEIR APPLICATIONS

The committee concluded that addressing diet-chronic disease relationships requires a credible, established, structured system for systematic review and review of the totality of the evidence, for all important topics that may be translated into population policy, and that this applies where "diet" refers to specific NOFSs, as is the focus of this DRI-oriented report, as well as when applied to foods or entire dietary patterns. The guiding principles discussed throughout the chapter reflect the fact that systematic reviews and evidence reviews in general are well-developed scientific activities for informing policy and practice. The several different systematic review and evidence review systems that have been applied to nutrition have many features in common. Requirements for such systems include:

- A record of robust experience and precedent to approach important scientific issues, including those involving questions about health effects of NOFS consumption,
- Well-structured protocols that can accommodate multiple major and ancillary scientific issues,
- Sufficient clarity and transparency to be useful for evidence review and policymaking groups,
- Sufficient detail to be replicable and to withstand expert peer review,
- The ability to cogently address causal pathways, which can potentially be used to develop policy recommendations, and
- A core set of methods stable enough to be applied over time but sufficiently flexible to be sensitive to the important details for different content areas. The methods should improve and evolve with thoughtful input from content experts and the user communities

[5] The Dietary Guidelines Advisory Committee independently evaluates the scientific evidence and makes recommendations about how the *Dietary Guidelines for Americans* could be revised.

[6] See www.nationalacademies.org/coi.

and with sufficient documentation so that changes can be understood and carried forth.

The committee acknowledges that no system will meet all needs and contingencies equally, and that all systems have strengths and weaknesses. The committee identified and reviewed key features of a variety of available evidence review and systematic review systems and instruments, contained in Annex 6-2, in sufficient detail to confirm that they have much in common conceptually although they vary in intent and comprehensiveness. The Annex includes systems or instruments intended only for assessing specific elements of primary empirical research reports, such as "risk-of-bias" tools, or instruments to evaluate the quality of reporting and completed systematic reviews. As others have concluded (Dwyer et al., 2016), all of these systems and instruments were deemed to have important value and apply similar principles to the review of the evidence. Some instruments for reviewing evidence are intended only for RCTs while others apply to both trials and observational studies. Some systems are used for a variety of issues related to primary and secondary prevention issues, including nutrition-related issues (AHRQ, 2014; USPSTF, 2015). Another system is used for assessing toxic environmental exposures (NTP, 2015) and others have been adapted for specific nutrition-related uses (NEL, 2015; WCRF/AICR, 2007). Still other systems are designed to evaluate the quality of reporting of completed systematic reviews and meta-analyses (e.g., AMSTAR [Pieper et al., 2014; Shea et al., 2007, 2009]).

THE "GRADE" (GRADING OF RECOMMENDATIONS, ASSESSMENT, DEVELOPMENT AND EVALUATION) SYSTEM: DEFINITION AND JUSTIFICATION

After reviewing the systems in Annex 6-2 and based on the knowledge and experience of committee members about what would meet the needs of DRI committees and users of DRI reports, the committee adopted the GRADE system as the reference point for the evidence reviews relating to NOFS-chronic disease considerations.

The GRADE system is an approach for rating certainty in the evidence in systematic reviews and guidelines. The system is designed for examining alternative management strategies or interventions, which may include no intervention or current best management. Although originally developed in the context of health care (diagnosis, screening, prevention, and therapy), it can be (and has been) applied to public health questions. In the particular case of evaluating the associations of NOFSs with chronic disease outcomes, GRADE will be used to assess the certainty in the association between an NOFS and benefit and harm, which implies causality. In addi-

tion to rating the certainty of a body of evidence, the GRADE approach includes tools for grading the strength of recommendations (see also Chapter 7). The rationale for the selection of GRADE is presented below.

Meets Criteria for an Appropriate Evidence Review Tool

The GRADE system meets the requirements for an established system of guidance for evidence synthesis. It has been used extensively in many scientific domains, and comprehensive documentation and precedent exist to address many of the issues relevant to NOFS-chronic disease questions. A particular advantage of the GRADE system is that it provides a comprehensive and clearly structured taxonomy for rating the certainty of evidence and extensive, specific guidance for making the ratings. In addition, GRADE working subgroups continue to address specific methodological or subject matter issues as needed. Publications for using GRADE include a six-part series of papers in the *British Medical Journal* (Guyatt et al., 2008a,b,c,d; Jaeschke et al., 2008; Schunemann et al., 2008), a large series of papers in the *Journal of Clinical Epidemiology* (Andrews et al., 2013a,b; Balshem et al., 2011; Guyatt et al., 2011a,b,c,d,e,f,g,h), and two additional papers focusing on the framework for translating evidence to guideline decisions (Alsonso-Coello et al., 2016a,b). The online GRADE Handbook (Schunemann et al., 2013) summarizes the guidance in these articles.

As long as the fundamental tenets are followed and a clear rationale for specific decisions stated, GRADE can accommodate particular perspectives. Indeed, the merit of this system is not that all members of a group using it will ultimately agree, but that it provides a structured, explicit, and transparent system for making decisions. Therefore, the GRADE working group strongly discourages modifications to the approach itself and recommends adherence to the currently available methods and applications. However, the GRADE system is neither rigid nor static. GRADE continues to evolve, and the GRADE working group is open to participation and collaboration from those facing challenges in applying the system, such as for informing nutritional guidelines in the context of the many unique conceptual and methodological issues outlined in Chapter 3.

Is in Wide Use and Applicable to Nutrition Policy Questions

The GRADE system provides an approach to rating the certainty of evidence for the outcomes relevant to optimizing health, either in terms of personal habits and lifestyle, or in decisions about health care interventions. It has been adopted by more than 100 organizations including WHO (2014), the Cochrane Collaboration, the American College of Physicians, the National Institute for Health and Care Excellence (NICE), and

UpToDate®), increasing its use and familiarity. GRADE has been applied in more than 100 countries. This facilitates the ability to compare reviews developed by diverse users. GRADE has been applied to nutrition-related systematic reviews (Garcia-Larsen et al., 2016; Mayhew et al., 2016; Naude et al., 2014; Pennant et al., 2015; Santesso et al., 2012) and is the method generally used by WHO to support guideline development, including for nutrition topics.

Has Criteria for Assessing Strength of Recommendations

The GRADE working group has seriously considered the issues of policy guidelines creation after evidence review, though the latter is not the task of this report. GRADE separates questions about confidence in estimates of effect in studies from questions about the strength of recommendations. This is critical for preserving the integrity of the judgments made about causal associations. Strong recommendations are far more likely when evidence is of high or moderate certainty and far less likely when evidence is of low or very low certainty. Nevertheless, when desirable and undesirable consequences and outcomes are closely balanced, it is possible to have a weak recommendation in the context of high certainty of the evidence. Similarly, GRADE has identified five paradigmatic situations in which strong recommendations are warranted in the face of low certainty evidence (Andrews et al., 2013b).

APPLYING THE GRADE APPROACH FOR SYNTHESIZING EVIDENCE ABOUT THE RELATIONSHIP BETWEEN AN NOFS AND CHRONIC DISEASE

This section is not intended to provide specific methods or protocols, which are published in peer-reviewed journals and the online handbook (Schunemann et al., 2013). The section describes how the GRADE system (see Figure 6-2 for an overview) is applied to evidence review in general, with particular suggestions and examples for application to the development of NOFS-chronic disease DRIs. In particular, this section addresses four critical and challenging issues in conducting an evidence review and determining the certainty of the evidence using the GRADE system: (1) selecting and ranking the outcomes, (2) addressing multiple comparators, (3) developing evidence summaries, and (4) determining the certainty of the evidence for each outcome. Published information that provides additional detail is highlighted.

It is envisioned that future DRI committees will review the evidence synthesis provided to them by applying the GRADE approach. Through that process, the committees will make judgments about causality and,

where indicated, intake-response. Thus, although some of the activities shown in Figure 6-2 and described above under the "overview of systematic review" are not conducted by DRI committees, they are essential to ensuring that DRIs are established with the best possible evidence. For example, as Figure 6-2 illustrates and as described earlier in this chapter, in the GRADE approach, a systematic review of the relevant scientific literature by the systematic review team precedes the work of the DRI committees to evaluate the totality of the evidence. The following discussion assumes that the systematic review and subsequent evidence reviews by the DRI committee are largely independent and that these processes precede and are separate from the development of guidelines, as described in Chapter 1 and above.[7]

Selecting and Ranking Health and Disease Outcomes

GRADE emphasizes the importance of identifying all outcomes that are meaningful to the population to whom a recommendation or guideline will be applied. For example, a particular NOFS may have causal relationships with several disease outcomes, beneficial or harmful. For risk assessment purposes, effects on physical and mental health are of primary importance. Effects on individual well-being (quality of life) may be considered during the risk management process, but these general effects are more relevant to consuming foods or following dietary patterns than to the intake of specific NOFSs.

Selecting the chronic disease outcomes to be considered as the basis for a DRI is a two-step process. First, all important chronic disease outcomes are identified during the formulation of the PICO questions. For the WHO free sugars guideline (WHO, 2015), several indicators were identified for body weight or body adiposity outcomes, and the outcome of dental caries was specifically defined as *not* including dental erosion (see Annex 6-1 for free sugar guidelines PICO specifications). In this step, designating and recording other secondary outcomes from relevant scientific reports is likely to be valuable, even if not originally part of the review's goal. Defining longer-term health outcomes from a nutritional intervention will be central for developing chronic disease DRIs. If available, general health outcomes, such as all-cause mortality or disability rates, may inform summative population effects of the NOFS under consideration, both positive and adverse. The inclusion of such broad outcomes is of particular importance for NOFS

[7] In other contexts, these activities may be performed as a continuum in which the same committee acts as or liaises to the technical advisory group before the conduct of systematic reviews, receives the systematic review, proceeds with evidence review, and then develops policy recommendations.

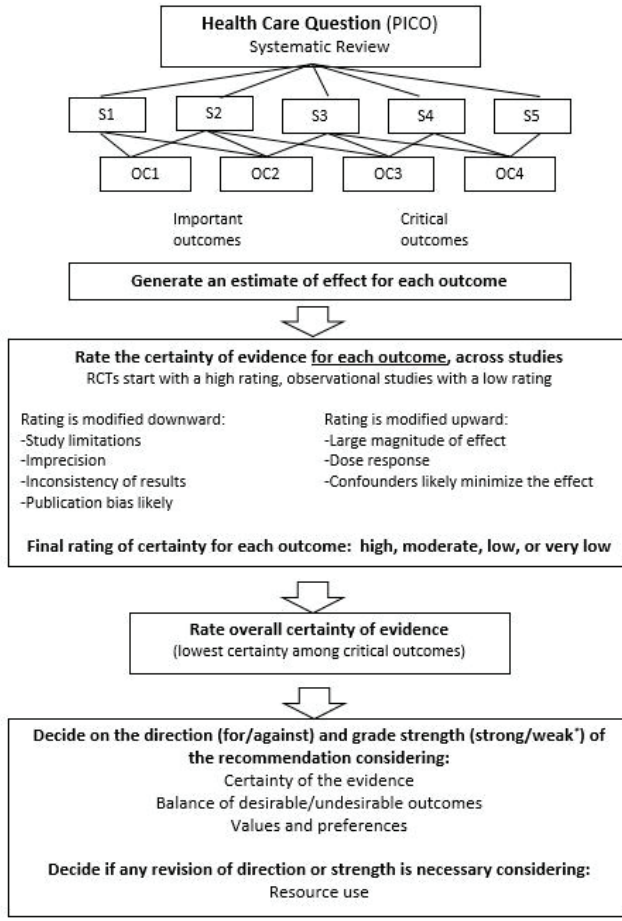


FIGURE 6-2 Schematic view of GRADE process for developing recommendations. In contrast to the GRADE approach, DRI committees will make DRI recommendations based on the balance of risk and benefits of health outcomes alone and will not consider values and preferences or resource use, which are within the purview of policy considerations.

NOTE: OC = outcomes; PICO = population, intervention, comparator, and outcome; RCT = randomized controlled trial; S = studies.

* Also labeled "conditional" or "discretionary."

SOURCE: Reprinted from Guyatt, G., A. D. Oxman, E. A. Akl, R. Kunz, G. Vist, J. Brozek, S. Norris, Y. Falck-Ytter, P. Glasziou, H. DeBeer, R. Jaeschke, D. Rind, J. Meerpohl, P. Dahm, and H. J. Schunemann. 2011. GRADE guidelines: 1. Introduction—GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol* 64(4):383-394, with permission from Elsevier.

questions because so many different outcomes may be affected, directly or indirectly. However, because of the many other NOFSs and other factors (e.g., other behaviors, medication use) that may be in the same pathway, attributing total mortality to a specific NOFS is challenging. The systematic review protocols also may include relevant surrogate markers (see Chapter 5 and the discussion of indirectness in this chapter). As mentioned in Chapter 5, conceptually, surrogate markers have no importance in themselves unless they are evaluated as "qualified" to serve as surrogates for the important policy relevant outcome. The protocol also should specify the time frame to which the outcomes apply (e.g., the interval between intervention and expected effects, such as 1 year, 5 years, or a lifetime).

Dealing with Multiple Comparators

The structure of the PICO questions, i.e., specifying one comparator for a given intervention, requires modification when more than two alternative courses of action exist. Under these circumstances, all potential courses of action can be characterized as alternative interventions in a series of questions in which the intervention is compared with each relevant alternative, or as a single intervention with multiple comparators by group (possible alternatives within a single question). For example, one could ask in one question "what is the effect of replacing saturated fat with monounsaturated fatty acids or various types of polyunsaturated fatty acids on cardiovascular disease outcomes" but extract data to support examination of each type of replacement separately. However, when multiple alternatives exist, it is unlikely that the primary literature will have compared each alternative to every other alternative. In this case, inferences regarding the relative merits of the alternatives require indirect comparisons. A novel statistical methodology for simultaneously considering all direct and indirect comparisons, network meta-analysis, is now available for dealing with such situations. GRADE has developed guidance for rating certainty of evidence in the context of network meta-analysis (Puhan et al., 2014).

Developing Evidence Summaries

Identifying, collecting, and evaluating published scientific studies (i.e., the evidence) is described in detail in the GRADE handbook (Schunemann et al., 2013). Once the studies have been identified, and the evidence evaluated, GRADE suggests aggregating the findings into two types of evidence summary tables: evidence profiles and summary-of-findings tables (Guyatt et al., 2011a). Although this task would be conducted by the systematic review team, not the DRI committee, it is an essential task that allows the committee to assess the totality of the evidence. The DRI committee would

review the evidence tables in order to make their own decisions regarding the certainty of the evidence for each outcome. Any discrepancies should be clearly documented and justified. As mentioned above, sufficient interaction between the DRI committee and the systematic review team and technical expert panel will ensure a more efficient process because any differences in judgments would have been discussed before the DRI committee receiving the final systematic review.

For binary or other categorical outcomes, evidence profiles include, at a minimum, a list of all outcomes, the numbers of participants and studies addressing these outcomes, the fundamental study designs (RCTs or observational studies), and measures of both the relative and absolute effect of the intervention compared to control. For continuous outcomes, evidence profiles differ only in the presentation of the effects. Possible presentation of continuous outcomes includes natural units and a number of transformations—standardized mean differences, minimal important difference units, and ratio of means, and binary outcomes—which will then be presented as relative and absolute effects.

RCTs start with a high certainty of evidence. Following the GRADE approach (see below under Factors Determining the Certainty of Evidence), the systematic review team will specify judgments (i.e., not serious, serious, or very serious) about specific factors (i.e., risk of bias, imprecision, inconsistency, indirectness, and publication bias) that may warrant rating down the certainty of evidence about specific outcomes from high to moderate (or low) and influence the judgment about the overall certainty of evidence. These factors, which also apply to observational studies, are discussed in the following section.

Ideally, for observational studies, the evidence profiles from the systematic review team will document each reason for rating down, and may also document factors for rating up, such as the presence of large effect sizes, dose-response gradients, and coverage of all plausible biases (see the following section for more detail). These can reinforce rather than undermine inferences regarding the effects of NOFS interventions. For each outcome and factor, the specific judgment would vary. Furthermore, the overall certainty rating would vary for each outcome.

Summary-of-findings tables are identical to evidence profiles, with one exception: they do not document judgments for each factor (rating up or down) but only the final overall certainty rating and specification of the domains that were rated down or up.

In practice, full documentation of the evidence reviewed can require multiple tables in order to characterize all important details, usually published as supplemental material in conjunction with the systematic review itself. For example, the systematic review of prospective studies of associations of saturated fat and *trans* fat with noncommunicable diseases as part

of a WHO guideline development process (de Souza et al., 2015) required six appendixes in order to provide details of the search strategy, supplemental tables to describe characteristics of studies reviewed and included in the GRADE summary (20 tables), characteristics of studies reviewed but not included in the GRADE summary, supplemental figures (68 figures), and two final GRADE summary tables, one each for saturated fat and *trans* fat and including all outcomes.

Figure 6-3 lists the 20 summary tables from the de Souza et al. review (de Souza et al., 2015). It shows that tables may address methodological issues such as the accuracy of measurements (eTable 1) and reported dietary intakes (eTable 20) in addition to findings grouped by study design or analytic approach, exposure, and outcomes. These tables are typically documented with footnotes to clarify definitions, analytic issues, and reasons for rating up or down, as appropriate (Santesso et al., 2016). The GRADE summary tables for saturated fat and *trans* fat were annotated with 40 and 86 footnotes, respectively. The format for GRADE table entries is shown in Figure 6-4, using the example for *trans* fat.

Once the systematic review team submits to the DRI committee its systematic review with evidence profiles and summary-of-findings tables for all outcomes of interest, positive and adverse, the committee evaluates the evidence presented, notes and justifies agreements and disagreements, and evaluates the importance of the health outcomes. The committee first makes a decision on the certainty of the evidence with regard to a causal relationship for each outcome they view as important. If such a relationship exists, then the DRI committee characterizes the certainty of the intake-response relationships and considers benefits and harms and their relative importance, as explained in Chapter 7.

Factors Determining the Certainty of Evidence

Initial Rating Based on Study Design: RCTs Versus Observational Studies

GRADE initially rates the certainty of the evidence as high, moderate, low, or very low with the possibility of changing the rating—up or down—based on the factors described below (i.e., risk of bias, imprecision, inconsistency, indirectness, and publication bias). Because randomization—when concealed, practiced with an intention-to-treat approach to the analysis, and with a sufficiently large sample size—deals with the problem of prognostic balance between intervention and control groups, RCTs begin with a high certainty of evidence. However, as discussed in Chapter 3, many NOFS-chronic disease questions will not have been addressed through RCTs for reasons of feasibility or cost. Also, the trials that do exist might have limited value for several reasons related to the form or dose of the

FIGURE 6-3 Example of detailed supplemental information accompanying a published systematic review to support evaluation of the evidence in GRADE summary tables.

NOTE: RR = relative risk; TFA = *trans* fatty acid.

SOURCE: de Souza et al., 2015, Appendix 2. Reprinted with permission under the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license: http://creativecommons.org/licenses/by-nc/4.0 (accessed July 15, 2017).

NOFS, study duration, study population, or other factors. In the GRADE approach, RCTs may be rated down for these and other reasons as the following sections describe.

Chapter 3 also discusses the fact that the relevant evidence on NOFSs and chronic disease risk is mostly observational, with large cohort studies playing a central role. In the GRADE system, conclusions based solely on observational studies are judged as having low certainty that a causal relationship is present because of concerns about residual confounding (i.e., confounding that remains after statistical adjustment), bias, and the accuracy of dietary exposure assessment. For observational studies, the certainty that the relationship is causal may be rated up to moderate under specific circumstances. Observational studies also provide information that is used in assessing intake-response relationship and in risk management. However, as noted in Chapter 3, observational studies that have a retrospective case-control or cross-sectional design may be excluded from eligibility for inclusion in systematic reviews because of their lower suitability for supporting judgments about causal relationships, in comparison to prospective studies. If included, this type of evidence would almost always be rated as low or very low quality, for reasons discussed earlier.

Addressing potential confounding typically requires appropriate use of individual-level data capable of characterizing both chronic disease risk and exposure propensity in the study population. Also, when dietary intake measurement depends on self-report, which is affected by individual characteristics, the study exposure assessment would need to have been shown to correlate strongly with a suitable objective measure of exposure (e.g., exposure biomarker) (see also Chapter 4). In the above-cited systematic review of *trans* fatty acids and noncommunicable disease risk, the authors used a three-tiered rating system to classify the quality of exposure assessment and applied this during their rating of study quality (de Souza et al., 2015, Appendix 2, eTable1).

Observational studies that address both sources of bias (confounding and exposure assessment) in a convincing way, and that report clear intake and chronic disease relationships, can potentially contribute usefully to establish chronic disease DRIs. However, these studies remain as low certainty evidence unless (1) the effect on the outcome is large (relative effect of two or more[8]), (2) a dose-response gradient is present, or (3) all important biases would either diminish an effect that is present or create an effect where none is observed. In Figure 6-4, the observation of a dose response led to rating up, by one level (from low to moderate) for the body of evidence on associations of both coronary heart disease (CHD) mortality and total

[8] For continuous variables, a statistical method of conversion to binary variables should be used.

FIGURE 6-4 Example of GRADE summary evidence table entries (prospective cohort studies of *trans* fatty acids and health outcomes).
NOTES: Table entries for associations with ruminant *trans* fatty acids are not shown. The 86 footnotes annotating this table are also not shown.
SOURCE: de Souza et al., 2015. Excerpt from Appendix 6. Reprinted with permission under the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license: http://creativecommons.org/licenses/by-nc/4.0 (accessed July 15, 2017).

CHD events with intake of total *trans* fatty acids (based on table footnotes 15 and 22 in the source; not shown in the figure). In contrast, all evidence relating total *trans* fatty acids to other outcomes and industrial *trans* fatty acids to CHD outcomes was rated down from low to very low due to one or more of the following certainty factors: serious risk of bias, serious inconsistency, serious imprecision (footnotes 29, 36, 43, 50, and 57 in the source; not shown in the figure). The nature of these factors, which are especially critical with respect to observational studies, is discussed in the next sections. Bodies of evidence comprising observational studies that are subject to being rated down on the basis of serious limitations would raise substantial uncertainty with respect to supporting causal relationships.

Risk of Bias of Individual Studies (Internal Validity and Study Limitations)

One of the factors considered in assessing the certainty of a body of evidence is the risk of bias of the individual studies. Systematic reviews are expected to critically appraise individual studies and evaluate the totality of evidence for each outcome in terms of the risk of bias (i.e., not serious, serious, or very serious). Key features to consider in risk of bias assessment for RCTs are concealment of randomization, blinding, co-intervention,[9] loss to follow-up, sufficient experimental contrast, and conduct of an intention-to-treat analysis. Although GRADE does not mandate the use of particular instruments to assess risk of bias in RCTs, guidance is offered. One suggestion is to use a modified version of the Cochrane Risk of Bias instrument (Guyatt and Busse, 2017).

For observational studies, important features to consider in risk of bias assessment are selection of exposed and unexposed individuals from different cohorts, flawed measurement of exposure and outcome (see Chapters 4

[9] Interventions other than the treatment under study that are applied differently to the treatment and control groups.

Exposure	Outcome	Participants (# studies)	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rate (%)	Absolute – adjusted (per 10,000)[1]	Dose-Response	Most-adjusted MV RR	Least-adjusted MV RR	Importance
Total TFA	Total mortality	20,346 (2 studies; 2 comparisons)[2]	No serious risk of bias[3]	No serious inconsistency[4]	No serious indirectness	No serious imprecision[5]	Not assessed[6]	⊕⊕◯◯ LOW[7]	2,141/20,346(8.5%)	388 more (from 182 more to 639 more)	Yes[8]	1.34(1.16 to 1.56)[9i]	1.80(1.57 to 2.07)	CRITICAL
	CHD mortality	70,864 (5 studies; 6 comparisons)[10]	No serious risk of bias[11]	No serious inconsistency[12]	No serious indirectness	No serious imprecision[13]	Not assessed[14]	⊕⊕⊕◯ MODERATE[15]	1,234/70,864 (1.7%)	56 more (from 18 more to 101 more)	Yes[16]	1.28 (1.09 to 1.50)	1.34 (1.00 to 1.79)	CRITICAL
	CHD total	145,922 (6 studies; 7 comparisons)[17]	No serious risk of bias[18]	No serious inconsistency[19]	No serious indirectness	No serious imprecision[20]	Not assessed[21]	⊕⊕⊕◯ MODERATE[22]	4,579/145,922 (3.1%)	88 more (from 42 more to 139 more)	Yes[23]	1.21(1.10 to 1.33)	1.31(1.15 to 1.48)	CRITICAL
Total TFA	Ischemic Stroke	190,284 (3 studies; 4 comparisons)[24]	No serious risk of bias[25]	Serious inconsistency[26]	No serious indirectness	Serious imprecision[27]	Not assessed[28]	⊕◯◯◯ VERY LOW[29]	1,905/190,284 (1.0%)	5 more (from 8 fewer to 20 more)	No[30]	1.07(0.88 to 1.28)	1.13(0.94 to 1.37)	CRITICAL
	Type 2 diabetes	230,135(6 studies; 6 comparisons)[31]	Serious risk of bias[32]	Serious inconsistency[33]	No serious indirectness	Serious imprecision[34]	Not assessed[35]	⊕◯◯◯ VERY LOW[36]	8,690/230,135 (3.8%)	56 more (from 28 fewer to 151 more)	Unclear[37]	1.10 (0.95 to 1.27)	1.28 (1.05 to 1.55)	CRITICAL
	Total mortality	71,464 (1 study; 2 comparisons)[38]	Serious risk of bias[39]	No serious inconsistency[40]	No serious indirectness	Serious imprecision[41]	Not assessed[42]	⊕◯◯◯VERY LOW[43]	11,890/71,464 (16.7%)	23 fewer (from 91 fewer to 46 more)	No[44]	0.98(0.92 to 1.04)	1.09(1.05 to 1.14)	CRITICAL
Industrial TFA	CHD mortality	93,394(2 studies; 2 comparisons)[45]	Serious risk of bias[46]	No serious inconsistency[47]	No serious indirectness	No serious imprecision[48]	Not assessed[49]	⊕◯◯◯ VERY LOW[50]	3,018/93,394 (3.2%)	36 more (from 8 more to 67 more)	Unclear[51]	1.18(1.04 to 1.33)	1.16(1.02 to 1.31)	CRITICAL
	CHD total	69,848 (2 studies; 2 comparisons)[52]	Serious risk of bias[53]	No serious inconsistency[54]	No serious indirectness	No serious imprecision[55]	Not assessed[56]	⊕◯◯◯ VERY LOW[57]	454/69,848 (0.7%)	176 more (from 21 more to 387 more)	Yes[58]	1.42 (1.05 to 1.92)	1.42 (1.05 to 1.92)	CRITICAL

and 5), failure to accurately measure all important prognostic factors and to adjust for these factors in the analysis, and loss to follow-up. Because of the diversity in study designs, the availability of a universal tool to assess risk of bias for all study designs is unlikely. It is also not within the study scope or feasible for this committee to delineate an exhaustive list of questions related to risk of bias assessment. The CLARITY Group at McMaster University offers candidate instruments for measuring risk of bias within observational studies (Busse and Guyatt, 2017a,b; Tikkinen et al., 2017a,b). Also, it is expected that DRI committees will include sufficient expertise in epidemiological study designs and their potential limitations, to be able to appropriately assess the risk of bias of the individual studies. As discussed below, the field of nutrition has no well-accepted risk-of-bias tools. There are no tools that prompt for evaluation of certain methodological features that are encountered with respect to assessing causal associations of NOFSs or other aspects of diet with chronic diseases (Chung, 2017). Current efforts are being directed at formally modifying risk-of-bias tools with the addition of questions that are relevant for nutrition (see also Chapter 3 for unique characteristics of nutrition research). Some nutrition questions are topic dependent, such as those related to 25(OH)D assay methods, while others are more generic, such as compliance issue for RCTs and measurement errors/biases in dietary assessment methods (see also Chapter 4). Because intake-response relationships are very important for ultimately developing DRIs, any biases in dietary assessment methods need to be considered in a risk-of-bias assessment by the addition of questions relevant to nutrition (see Chapter 4).

For RCTs involving NOFS exposures from foods, blinding of the intervention exposure is often not possible. The exception is in some controlled feeding studies, but these are usually of short duration and dependent on intermediates as outcome variables. Also, unlike trials of pharmaceuticals, the control group is always exposed to some level or form of the NOFS through intake of foods even in most trials of NOFSs given as a pill supplements. In such cases, unless the baseline exposure of subjects is considered in the trial, the real question being examined in nutrition RCTs involves the health benefits or risks of higher versus lower levels of intake, rather than specific exposure levels of the NOFS. For RCTs, therefore, a question related to whether baseline exposure has been measured and considered would be necessary in order to assess the certainty of the evidence.

In the current absence of a validated, nutrition-specific tool, the Cochrane risk-of-bias tool for RCTs[10] and Newcastle-Ottawa Scale[11] for cohort studies, supplemented with items specific to nutrition have been used

[10] See http://methods.cochrane.org/bias/assessing-risk-bias-included-studies.

[11] See http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.

in the field of nutrition. Another risk-of-bias tool used in nutrition is the Nutrition Evidence Library Bias Assessment Tool (NEL BAT).[12] This tool is based on Cochrane risk-of-bias domains (Selection Bias, Performance Bias, Detection Bias, Attrition Bias) and it is tailored by study design, with different sets of questions applying to RCTs (14 questions), non-RCTs (14 questions), and observational studies (12 questions). However, the tool has no nutrition-specific items.

Inconsistency of Results

Inconsistency is defined as heterogeneity or variability of study results. Issues to consider regarding consistency of results are the extent to which point estimates are similar, the extent to which confidence intervals overlap, and statistical analyses, including I^2, a measure of variability, and tests for heterogeneity (Guyatt et al., 2011f). Consistency of outcomes may depend on matters related to the design and conduct of the systematic review itself, such as the sharpness of study question focus, and the inclusion/exclusion criteria for studies included in the analysis, such as the precision of the outcome assessment.

When results have large inconsistencies, effect modifiers can be explored. Subgroup analyses may be misleading, and methodologists have developed criteria to assess the credibility of such analyses. These criteria include whether evidence comes from comparison within or between studies; whether chance can explain the apparent subgroup differences; whether the study questions were developed a priori, deriving from among a small number of hypotheses; and whether the expected direction of the apparent effect was accurately specified in advance (Sun et al., 2010).

Indirectness of Evidence

The available evidence may not provide an ideal match with the question of interest. This occurs, for example, when studies are not available for certain age groups or particular ethnic groups with respect to exposures or outcomes of interest. For example, the WHO systematic review on free sugars and dental caries identified only studies of children, suggesting that conclusions about adults would be indirect and therefore subject to being rated down on this basis. However, in this particular case, a judgment of serious indirectness did not occur, with the following justification:

> Data were not downgraded for indirectness though all 8 cohort studies were conducted in children. Aetiology of dental caries is the same in

[12] See http://www.nel.gov/topic.cfm?cat=3384.

> children and adults (though enamel of the primary dentition is softer and more vulnerable to demineralization by plaque acid). There were no cohort studies in adults; however, data from 5 out of 5 studies of other study design in adults included in the systematic review detected a statistically significant positive relationshzip between dietary free sugars and levels of dental caries. Dental caries is progressive and tracks from childhood (permanent dentition) to adulthood. 4/8 cohort studies include permanent dentition. Studies were conducted in worldwide populations—Argentina, Brazil, S. Africa, Finland, Sweden, England, USA. (Moynihan and Kelly, 2014, Supplemental data 4, Table S-3, Footnote 4)

The evidence reviewed by Te Morenga et al. relating intake of free sugars to body weight outcomes raised the potential for another type of indirectness based on short study duration (Te Morenga et al., 2012). The RCTs that addressed the impact of free sugar consumption on weight were all relatively short term. This could be considered only indirect evidence for what would happen over a more extended period. However, the evidence was not rated down for serious indirectness on the basis that, "These short-term studies were of sufficient duration to detect a change on the outcome of interest and, thus, the data provide relevant evidence for the association between increased free sugars intakes and weight gain" (Te Morenga et al., 2012, supplement 2, footnote 4).

Other types of indirectness are situations in which available studies of individual NOFSs do not reflect dietary patterns of a population of interest or when only surrogate markers rather than the disease outcomes of interest have been studied. The comment in the above quotation from Moynihan and Kelly about the range of populations studied addresses the breadth of population coverage. In the case of fatty acids, evidence from blood lipids, a possible surrogate marker, was used to facilitate judgments about certain questions, but the overall evidence base included data on mortality and other events.

Studies also may provide only indirect comparisons in other ways. The effect of Diet A versus Diet B may be of interest, but no direct comparisons are available. Yet, if both Diet A and Diet B have been compared to Diet C, this allows the possibility of an indirect comparison between A and B. Such indirect comparisons are at greater risk of producing misleading results than are direct comparisons. However, this linking approach may be useful in complementing information not available in direct comparisons.

Imprecision

In the context of nutritional studies, examination of the boundaries of the confidence interval for a given overall effect informs the decision

regarding precision. Imprecision is the random error in studies that results in wide 95 percent confidence intervals around the estimate of the effect. This differs from inconsistency, which refers to the consistency across individual studies. For each outcome, the authors of each study, or ideally the DRI committee, evaluate the decision regarding whether a proposed intervention or comparator is preferable by selecting estimates at both ends of the confidence interval, to see whether these alternatives would change the interpretation. If it would, then one rates down the evidence for imprecision. For instance, in Figure 6-4, evidence from cohort studies of total *trans* fatty acid on ischemic stroke was rated down for imprecision because of 95 percent confidence intervals that suggested benefit at the lower bound and exceeded the threshold for harm at the upper bound (de Souza et al., 2015, appendix 6, footnote 27).

This guidance has one exception. If a study or studies report a large or very large effect with a relatively small number of events, experience has shown that the results do not stand the test of time. Therefore, under these circumstances, even if the confidence interval standard suggests satisfactory precision, the study certainty may be rated down on the grounds of imprecision.

If a DRI committee wished to be quantitative about the decisions to rate down a body of evidence under these circumstances, it could specify a magnitude of effect (typically a modest relative risk reduction (RRR), for example ≤30 percent) and calculate the sample size required, in a single study, to have adequate power (80 or 90 percent) with the usual Type 1 error threshold of 0.05 and assume the baseline risk in the studies available. The sample size that results from this exercise is labeled the optimal information size (OIS). If the total number of individuals in a systematic review is less than the number of individuals generated by a conventional sample size calculation for a single adequately powered trial, or the OIS, the committee should consider rating down for imprecision (Schunemann et al., 2013). Many online calculators are available for sample size calculation.[13] Alternatively, Table 6-1 provides a guide for what happens, given the number of events observed in the studies, when this exercise is conducted with RRRs of 20 to 30 percent. For example, if an RRR of ≤30 is chosen when the total number of events is 100 or less, the thresholds for precision will most likely not be met (Guyatt et al., 2011e; Pogue and Yusef, 1997).

Publication Bias

Publication bias is defined as the systematic under-estimation or an over-estimation of the underlying beneficial or harmful effect due to

[13] See http://www.stat.ubc.ca/~rollin/stats/ssize/b2.html.

TABLE 6-1 Optimal Information Size Implications When an Adequate Power and Magnitude of Effect Is Desired

Total Number of Events	Relative Risk Reduction (%)	Implications for Meeting Optimal Information Size Threshold
≤100	≤30	Will almost never meet threshold whatever control event rate
200	30	Will meet threshold for control group risks of ~25 percent or greater
200	25	Will meet threshold for control group risks of ~50 percent or greater
200	20	Will meet threshold only for control group risks of ~80 percent or greater
300	≥30	Will meet threshold
300	25	Will meet threshold for control group risks of ~25 percent or greater
≥400	20	Will meet threshold for control group risks of ~40 percent or greater

SOURCE: Adapted from Guyatt et al., 2011e.

the selective publication of studies. In the GRADE approach, options include rating down when the evidence comes from a number of small industry-funded trials, creating and evaluating a funnel plot, and using one of a number of statistical tests that evaluate the asymmetry of the data (small studies showing larger effects than the larger studies) (Guyatt et al., 2011d). The methods guide on grading the strength of a body of evidence by AHRQ describes a conceptual framework and recommends both qualitative and quantitative methods, such as the funnel plot, to assess publication bias. Although both of these groups offer important guidance, they also recognize that measuring publications bias is challenging (AHRQ, 2014; Guyatt et al., 2011d) and that there are no completely satisfactory strategies. This is particularly the case when the body of evidence is from observational studies because assessors might not know if the published observational studies are representative of the studies conducted due to the lack of published or registered protocols. With this in mind, the terms GRADE suggests using are "undetected" and "strongly suspected" publication bias. In the examples in Figure 6-4, publication bias was not assessed formally due to too few studies.

Factors That Can Increase the Certainty of the Evidence

Large Magnitude of an Effect

Rating up for large effect size is typically applied in the context of observational studies. GRADE's rule of thumb is that evidence may be rated up one level if the odds ratio or relative risk is >2.0 or <0.5 or two levels if the relative risk is >10 or <0.1. The WHO free sugars review (Te Morenga et al., 2012) rated up certainty in evidence based on a large effect of sugar consumption on dental caries. Although large effects are rarely seen with any nutrition intervention unless individuals are deficient in a particular nutrient, the committee concluded that uprating the confidence in the body of evidence of observational studies requires a large effect in the health outcome. This requirement provides higher assurance that the association could not be due to residual confounding.

Intake-Response

Another reason for rating up evidence certainty is observing an intake-response gradient, in this case between the NOFS and the health endpoint of interest. An intake-response gradient implies that the magnitude of the outcome increases in a linear or non-linear fashion with increasing magnitude of exposure and the question being answered would be: "What is the level of confidence that an intake-response relationship gradient exists?" In contrast, Chapter 7 addresses not only whether an intake-response relationship exists but also how to characterize it quantitatively.

The type of evidence needed for rating up the certainty of the evidence based on an intake-response relationship is less stringent than that needed to characterize the relationship quantitatively. For instance, intake levels may be categorical (low, medium, high), so that while evidence of an increase in response with intake exists, a quantitative relationship cannot be determined. Alternatively, one might have high confidence that the "slope" of a linear relationship is greater than zero, but because of potential bias toward the null, confidence is lower as to what the value of the slope might be. In attempting to make this distinction, this report reserves the term quantitative intake-response for a relationship that can be characterized quantitatively.

Effect of Plausible Residual Confounding

Very rarely, evidence certainty may be rated up when all plausible confounding factors would minimize an observed effect, or create an effect when an effect has not been observed. A systematic review that reported

higher death rates in private for-profit versus private not-for-profit hospitals provides an example of the phenomenon of plausible confounding (Guyatt et al., 2011h). Plausible biases included patients in the not-for-profit hospitals being sicker and higher numbers of patients with excellent private insurance coverage in the for-profit hospitals, both of which would be expected to result in better outcomes for private for-profit hospital patients. Because these biases would diminish the observed effect, even if they were not accounted for in the analysis, the circumstances surrounding this confounder might lead the evidence panel to consider the evidence from observational studies as moderate rather than low certainty (Guyatt et al., 2011h).

Rating the Overall Certainty of Evidence

Within the GRADE approach, the evidence is rated by outcome and study design. That is, observational studies are rated separately from RCTs for each outcome. Within evidence from observational studies, those with different designs may also be rated separately. This is illustrated in the list of supplemental tables from a systematic review of observational studies of fatty acids and noncommunicable diseases in Figure 6-3. Typically, certainty of evidence differs across outcomes. For each outcome, when evidence from both RCTs and observational studies is available, the focus should be on the higher certainty evidence, and the overall rating of certainty should be that of the higher certainty. With regard to the overall certainty of evidence, GRADE specifies that it should be designated as the lowest detected among the critical outcomes.

ACCEPTABLE LEVEL OF EVIDENCE THAT THE RELATIONSHIP IS CAUSAL

As already mentioned, once it receives the systematic review with evidence profiles and summary-of-findings tables for all outcomes of interest, positive and adverse, the DRI committee will evaluate the evidence presented, note and justify agreements and disagreements, and evaluate the importance of the health outcomes. The DRI committee first makes a decision on the certainty of the evidence with regard to a causal relationship for each outcome it views as important. If such a relationship exists, then the DRI committee will characterize the certainty of the intake-response relationships, consider benefits and harms and their relative importance, and recommend chronic disease DRIs levels or ranges when appropriate, as explained in Chapter 7.

Determining causation between exposures and outcomes has been a longstanding philosophical as well as practical challenge in population

research as well as for science in general. As discussed above, RCTs, as true experiments, provide the best overall evidence of causation, but the actual certainty in the evidence depends on the certainty of those trials. Other criteria and systems for approaching causation exist, such as the Bradford Hill criteria (Hill, 1965). Parascandola and Weed reviewed the epidemiological literature on causation and found multiple definitions (Parascandola and Weed, 2001).

The Options Report (Yetley et al., 2017) had one set of options that directly addressed causation. Four options were presented, shown in Box 6-4, relating to the levels of confidence that an association between an NOFS exposure and an outcome is causal. All the options portrayed are feasible approaches to defining acceptable levels. The GRADE approach addresses the issue of whether the association between an exposure and an outcome is causal (Schunemann et al., 2013). After considering these approaches in detail and based on the committee's evaluation of the GRADE approach, the recommendation is that to accept the likelihood of a causal relationship

BOX 6-4
Acceptable Levels of Confidence That the Relation Is Causal

Options Report

Option 1: Require a high level of confidence
- This option requires a high level of confidence that a proposed relation is causal. This level of confidence likely requires at least some evidence from high-quality RCTs in which the measured outcome is a chronic disease event or qualified surrogate disease marker.

Option 2: Use level B evidence
- This option also includes level B evidence as a basis for DRI decisions about causation. This level of evidence suggests a moderate degree of confidence that the relation of interest is causal, but new findings could change the DRI decision.

Option 3: Use actual level of certainty
- This option identifies the actual level of certainty for each DRI reference value based on a chronic disease endpoint.

Option 4: Make decisions on a case-by-case basis
- This option makes decisions about the strength of evidence appropriate to support a conclusion about the relation between a given food substance and a chronic disease endpoint on a case-by-case basis.

continued

BOX 6-4 Continued

TABLE 6-2 Level of Confidence in DRI Decisions

	Overall Evidence Rating Based on Evidence Review		
Chronic Disease Endpoint	High	Medium	Low
Chronic disease event	Level A	Level B	Levels C or D
Qualified surrogate disease marker	Levels A or B	Levels B or C	Levels C or D
Nonqualified outcome	Level C	Levels C or D	Level D

NOTES: Level A: highest degree of confidence that results are valid (e.g., "high"); level B: some uncertainty about validity of results (e.g., "moderate"); level C: considerable uncertainty about validity of results (e.g., "low"); level D: substantial uncertainty about validity of results (e.g., "insufficient"). DRI = Dietary Reference Intake.
SOURCE: Yetley et al., 2017. Reprinted with permission under this article's license: http://nutrition.org/publications/guidelines-and-policies/license (accessed July 18, 2017).

Committee's Recommendation 3

The committee recommends that Dietary Reference Intake (DRI) committees use Grading of Recommendations, Assessment, Development and Evaluation (GRADE) in assessing the certainty of the evidence related to the causal association between nutrient or other food substances and chronic diseases. Using GRADE, the committee recommends that a decision to proceed with development of chronic disease DRIs be based on at least moderate certainty that a causal relationship exists and on the existence of an intake-response relationship.

between an NOFS and a chronic disease outcome, a moderate rating of the relevant overall evidence base derived using GRADE criteria is needed. Lesser levels of confidence in causation should not lead to actionable recommendations or, in the case here, the development of chronic disease DRIs. Presumably, to move forward with developing chronic disease DRIs, determination of an intake-response in data from observational studies as part of the evidence rating also would be needed.

The GRADE system does not specify either a minimum number of studies or participants (although it does provide guidance for minimum number of participants to avoid rating down certainty for imprecision) or the characteristics of the study design (e.g., observational studies have established the causal relationship of smoking and lung cancer) necessary for determining that a relationship has at least moderate certainty, and

therefore is likely to be causal. In the same manner, the DRI committees will apply their judgment relative to this matter.

GUIDING PRINCIPLES AS FOUNDATION FOR A CHRONIC DISEASE DRI PROCESS

Guiding principles suitable for the evidence review process are summarized in Box 6-5. These principles reflect the fact that evidence review in general is well developed as a scientific activity for informing policy and practice. The guiding principles are the foundation of a rigorous scientific process for setting chronic disease DRIs.

BOX 6-5
Guiding Principles for Setting Chronic Disease DRIs

With respect to systematic reviews:

1. Well-structured and established protocols that include the question of interest and analytical frameworks are necessary to address multiple major and ancillary scientific issues related to the degree of confidence in evidence for causal associations.
2. Protocols should be developed with guidance from a technical expert panel that includes relevant content experts in nutrition science, toxicology, scientific study design and analysis, public health, biostatistics, nutrition epidemiology and chronic disease epidemiology, and disease pathogenesis.
3. In consultation with the technical expert panel, systematic reviews should be sufficiently inclusive of all study designs that potentially contribute to evaluation of the causal NOFS-chronic disease relationship of interest and identification of associated intake-response relationships.
4. Protocols should include studies that use various dietary assessment approaches, including self-report and biomarkers of intake, while taking the quality of exposure assessment into account when rating study quality.
5. Protocols should include studies that document outcomes or surrogates of outcomes of potential importance for assessing benefits and harms, while taking the quality of outcome assessments into account in rating study quality.
6. Instruments and analytical methods applied to systematic reviews should be thoughtfully chosen and defensible. Instruments to assess

continued

BOX 6-5 Continued

the internal validity of the studies should include considerations that apply to nutrition research and various study designs (observational and intervention studies).

7. Results from the systematic review should be clearly presented in study-by-study evidence tables and summary tables of the total evidence for each outcome and study type.

With respect to DRI committee reviews of the totality of the evidence:

8. The DRI committees should include content experts and methodologists relevant to the primary scientific issues and to evidence review. DRI committees should be free of significant financial, intellectual, and professional conflicts of interest. In some cases, the required expertise might not be found without some conflicts of interest. In such cases, it is necessary to identify, disclose, and manage any potential conflicts of interest. Mechanisms to allow for interactions between the DRI committee and members representing both the technical expert panel and systematic review team, while also protecting against inappropriate influence on the systematic review methods, are strongly encouraged.
9. Particular elements of needed expertise will be guided by the general scientific question(s) and specific questions and will generally include nutrition science, scientific study design and analysis, public health, biostatistics, nutrition and chronic disease epidemiology, disease pathogenesis, and evidence review conduct.
10. The evidence review should be sufficiently comprehensive to anticipate the major scientific issues and methods that will likely be a part of the ensuing guideline development process.
11. Sufficient documentation, clarity, and transparency in the evidence review process is needed so that others can comprehend and evaluate this process and its activities, methodological considerations, final decisions, and the rationale for decisions about each outcome.
12. The review of the evidence and other aspects of the systematic review should be replicable and subject to expert peer review.
13. When apparent discrepancies in the evidence exist, DRI committees should attempt to determine whether they can be explained by differences in methodology or conceptualization of diet-disease relationships and, where possible, incorporate such explanations into the process of rating the evidence.
14. Where they exist, quantitative intake-response relationships should include a thorough description of the scientific uncertainties associated with them.

REFERENCES

AHRQ (Agency for Healthcare Research and Quality). 2014. *Methods guide for effectiveness and comparative effectiveness reviews. AHRQ Publication No. 10(14)-EHC063-EF.* Rockville, MD: AHRQ.

Alonso-Coello, P., H. J. Schunemann, J. Moberg, R. Brignardello-Petersen, E. A. Akl, M. Davoli, S. Treweek, R. A. Mustafa, G. Rada, S. Rosenbaum, A. Morelli, G. H. Guyatt, A. D. Oxman, and Grade Working Group. 2016a. GRADE Evidence to Decision (EtD) frameworks: A systematic and transparent approach to making well informed healthcare choices. 1: Introduction. *BMJ* 353:i2016.

Alonso-Coello, P., A. D. Oxman, J. Moberg, R. Brignardello-Petersen, E. A. Akl, M. Davoli, S. Treweek, R. A. Mustafa, P. O. Vandvik, J. Meerpohl, G. H. Guyatt, H. J. Schunemann, and Grade Working Group. 2016b. GRADE Evidence to Decision (EtD) frameworks: A systematic and transparent approach to making well informed healthcare choices. 2: Clinical practice guidelines. *BMJ* 353:i2089.

Andrews, J., G. Guyatt, A. D. Oxman, P. Alderson, P. Dahm, Y. Falck-Ytter, M. Nasser, J. Meerpohl, P. N. Post, R. Kunz, J. Brozek, G. Vist, D. Rind, E. A. Akl, and H. J. Schunemann. 2013a. GRADE guidelines: 14. Going from evidence to recommendations: The significance and presentation of recommendations. *J Clin Epidemiol* 66(7):719-725.

Andrews, J. C., H. J. Schunemann, A. D. Oxman, K. Pottie, J. J. Meerpohl, P. A. Coello, D. Rind, V. M. Montori, J. P. Brito, S. Norris, M. Elbarbary, P. Post, M. Nasser, V. Shukla, R. Jaeschke, J. Brozek, B. Djulbegovic, and G. Guyatt. 2013b. GRADE guidelines: 15. Going from evidence to recommendation—Determinants of a recommendation's direction and strength. *J Clin Epidemiol* 66(7):726-735.

Balshem, H., M. Helfand, H. J. Schunemann, A. D. Oxman, R. Kunz, J. Brozek, G. E. Vist, Y. Falck-Ytter, J. Meerpohl, S. Norris, and G. H. Guyatt. 2011. GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol* 64(4):401-406.

Brannon, P. M., C. L. Taylor, and P. M. Coates. 2014. Use and applications of systematic reviews in public health nutrition. *Annu Rev Nutr* 34:401-419.

Brannon, P. M., C. M. Weaver, C. A. Anderson, S. M. Donovan, S. P. Murphy, and A. L. Yaktine. 2016. Scanning for new evidence to prioritize updates to the Dietary Reference Intakes: Case studies for thiamin and phosphorus. *Am J Clin Nutr* 104(5):1366-1377.

Brouwer, I. A. 2016. *Effect of trans-fatty acid intake on blood lipids and lipoproteins: A systematic review and meta-regression analysis*. Geneva, Switzerland: World Health Organization.

Busse, J. W., and G. Guyatt. 2017a. *Tool to assess risk of bias in case control studies*. https://www.evidencepartners.com/resources/methodological-resources (accessed May 16, 2017).

Busse, J. W., and G. Guyatt. 2017b. *Tool to assess risk of bias in cohort studies*. https://www.evidencepartners.com/resources/methodological-resources (accessed May 16, 2017).

Chung, M. 2017. *Overview and comparisons of risk of bias and strength of evidence assessment tools: Opportunities and challenges of application in developing DRIs*. Presented at the Workshop of the Committee on Development of Guiding Principles for the Inclusion of Chronic Disease Endpoints in Future Dietary Reference Intakes, January 9, 2017, Washington, DC.

Chung, M., S. J. Newberry, M. T. Ansari, W. W. Yu, H. Wu, J. Lee, M. Suttorp, J. M. Gaylor, A. Motala, D. Moher, E. M. Balk, and P. G. Shekelle. 2012. Two methods provide similar signals for the need to update systematic reviews. *J Clin Epidemiol* 65(6):660-668.

de Souza, R. J., A. Mente, A. Maroleanu, A. I. Cozma, V. Ha, T. Kishibe, E. Uleryk, P. Budylowski, H. Schunemann, J. Beyene, and S. S. Anand. 2015. Intake of saturated and trans unsaturated fatty acids and risk of all cause mortality, cardiovascular disease, and type 2 diabetes: Systematic review and meta-analysis of observational studies. *BMJ* 351:h3978.

Dwyer, J. T., K. H. Rubin, K. L. Fritsche, T. L. Psota, D. J. Liska, W. S. Harris, S. J. Montain, and B. J. Lyle. 2016. Creating the future of evidence-based nutrition recommendations: Case studies from lipid research. *Adv Nutr* 7(4):747-755.

Garcia-Larsen, V., S. R. Del Giacco, A. Moreira, M. Bonini, D. Charles, T. Reeves, K. H. Carlsen, T. Haahtela, S. Bonini, J. Fonseca, I. Agache, N. G. Papadopoulos, and L. Delgado. 2016. Asthma and dietary intake: An overview of systematic reviews. *Allergy* 71(4):433-442.

Garner, P., S. Hopewell, J. Chandler, H. MacLehose, H. J. Schunemann, E. A. Akl, J. Beyene, S. Chang, R. Churchill, K. Dearness, G. Guyatt, C. Lefebvre, B. Liles, R. Marshall, L. Martinez Garcia, C. Mavergames, M. Nasser, A. Qaseem, M. Sampson, K. Soares-Weiser, Y. Takwoingi, L. Thabane, M. Trivella, P. Tugwell, E. Welsh, E. C. Wilson, H. J. Schunemann, and Panel for Updating Guidance for Systematic Reviews. 2016. When and how to update systematic reviews: Consensus and checklist. *BMJ* 354:i3507.

Guyatt, G., and J. W. Busse. 2017. *Modification of Cochrane tool to assess risk of bias in randomized trials*. https://www.evidencepartners.com/resources/methodological-resources (accessed May 12, 2017).

Guyatt, G. H., A. D. Oxman, R. Kunz, Y. Falck-Ytter, G. E. Vist, A. Liberati, H. J. Schunemann, and Grade Working Group. 2008a. Going from evidence to recommendations. *BMJ* 336(7652):1049-1051.

Guyatt, G. H., A. D. Oxman, G. E. Vist, R. Kunz, Y. Falck-Ytter, P. Alonso-Coello, H. J. Schunemann, and Grade Working Group. 2008b. GRADE: An emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 336(7650):924-926.

Guyatt, G. H., A. D. Oxman, R. Kunz, R. Jaeschke, M. Helfand, A. Liberati, G. E. Vist, H. J. Schunemann, and Grade Working Group. 2008c. Incorporating considerations of resources use into grading recommendations. *BMJ* 336(7654):1170-1173.

Guyatt, G. H., A. D. Oxman, R. Kunz, G. E. Vist, Y. Falck-Ytter, H. J. Schunemann, and Grade Working Group. 2008d. What is "quality of evidence" and why is it important to clinicians? *BMJ* 336(7651):995-998.

Guyatt, G., A. D. Oxman, E. A. Akl, R. Kunz, G. Vist, J. Brozek, S. Norris, Y. Falck-Ytter, P. Glasziou, H. DeBeer, R. Jaeschke, D. Rind, J. Meerpohl, P. Dahm, and H. J. Schunemann. 2011a. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol* 64(4):383-394.

Guyatt, G. H., A. D. Oxman, R. Kunz, D. Atkins, J. Brozek, G. Vist, P. Alderson, P. Glasziou, Y. Falck-Ytter, and H. J. Schunemann. 2011b. GRADE guidelines: 2. Framing the question and deciding on important outcomes. *J Clin Epidemiol* 64(4):395-400.

Guyatt, G. H., A. D. Oxman, G. Vist, R. Kunz, J. Brozek, P. Alonso-Coello, V. Montori, E. A. Akl, B. Djulbegovic, Y. Falck-Ytter, S. L. Norris, J. W. Williams, Jr., D. Atkins, J. Meerpohl, and H. J. Schunemann. 2011c. GRADE guidelines: 4. Rating the quality of evidence—Study limitations (risk of bias). *J Clin Epidemiol* 64(4):407-415.

Guyatt, G. H., A. D. Oxman, V. Montori, G. Vist, R. Kunz, J. Brozek, P. Alonso-Coello, B. Djulbegovic, D. Atkins, Y. Falck-Ytter, J. W. Williams, Jr., J. Meerpohl, S. L. Norris, E. A. Akl, and H. J. Schunemann. 2011d. GRADE guidelines: 5. Rating the quality of evidence—Publication bias. *J Clin Epidemiol* 64(12):1277-1282.

Guyatt, G. H., A. D. Oxman, R. Kunz, J. Brozek, P. Alonso-Coello, D. Rind, P. J. Devereaux, V. M. Montori, B. Freyschuss, G. Vist, R. Jaeschke, J. W. Williams, Jr., M. H. Murad, D. Sinclair, Y. Falck-Ytter, J. Meerpohl, C. Whittington, K. Thorlund, J. Andrews, and H. J. Schunemann. 2011e. GRADE guidelines 6. Rating the quality of evidence—Imprecision. *J Clin Epidemiol* 64(12):1283-1293.

Guyatt, G. H., A. D. Oxman, R. Kunz, J. Woodcock, J. Brozek, M. Helfand, P. Alonso-Coello, P. Glasziou, R. Jaeschke, E. A. Akl, S. Norris, G. Vist, P. Dahm, V. K. Shukla, J. Higgins, Y. Falck-Ytter, H. J. Schunemann, and Grade Working Group. 2011f. GRADE guidelines: 7. Rating the quality of evidence—Inconsistency. *J Clin Epidemiol* 64(12):1294-1302.

Guyatt, G. H., A. D. Oxman, R. Kunz, J. Woodcock, J. Brozek, M. Helfand, P. Alonso-Coello, Y. Falck-Ytter, R. Jaeschke, G. Vist, E. A. Akl, P. N. Post, S. Norris, J. Meerpohl, V. K. Shukla, M. Nasser, H. J. Schunemann, and Grade Working Group. 2011g. GRADE guidelines: 8. Rating the quality of evidence—Indirectness. *J Clin Epidemiol* 64(12):1303-1310.

Guyatt, G. H., A. D. Oxman, S. Sultan, P. Glasziou, E. A. Akl, P. Alonso-Coello, D. Atkins, R. Kunz, J. Brozek, V. Montori, R. Jaeschke, D. Rind, P. Dahm, J. Meerpohl, G. Vist, E. Berliner, S. Norris, Y. Falck-Ytter, M. H. Murad, H. J. Schunemann, and Grade Working Group. 2011h. GRADE guidelines: 9. Rating up the quality of evidence. *J Clin Epidemiol* 64(12):1311-1316.

Hill, A. B. 1965. The environment and disease: Association or causation? *Proc R Soc Med* 58:295-300.

IOM (Institute of Medicine). 2011a. *Dietary Reference Intakes for calcium and vitamin D*. Washington, DC: The National Academies Press.

IOM. 2011b. *Finding what works in health care: Standards for systematic reviews*. Washington, DC: The National Academies Press.

Jaeschke, R., G. H. Guyatt, P. Dellinger, H. Schunemann, M. M. Levy, R. Kunz, S. Norris, J. Bion, and Grade Working Group. 2008. Use of GRADE grid to reach decisions on clinical practice guidelines when consensus is elusive. *BMJ* 337:a744.

Mayhew, A. J., R. J. de Souza, D. Meyre, S. S. Anand, and A. Mente. 2016. A systematic review and meta-analysis of nut consumption and incident risk of CVD and all-cause mortality. *Br J Nutr* 115(2):212-225.

Mensink, R. P. 2016. *Effects of saturated fatty acids on serum lipids and lipoproteins: A systematic review and regression analysis*. Geneva, Switzerland: World Health Organization.

Moynihan, P. J., and S. A. Kelly. 2014. Effect on caries of restricting sugars intake: Systematic review to inform WHO guidelines. *J Dent Res* 93(1):8-18.

NASEM (National Academies of Sciences, Engineering, and Medicine). 2017. *Optimizing the process for establishing the* Dietary Guidelines for Americans*: The selection process*. Washington, DC: The National Academies Press.

Naude, C. E., A. Schoonees, M. Senekal, T. Young, P. Garner, and J. Volmink. 2014. Low carbohydrate versus isoenergetic balanced diets for reducing weight and cardiovascular risk: A systematic review and meta-analysis. *PLOS ONE* 9(7):e100652.

NEL (Nutrition Evidence Library). 2015. *Nutrition Evidence Library—2015 Dietary Guidelines Advisory Committee systematic reviews*. https://www.cnpp.usda.gov/nutrition-evidence-library-2015-dietary-guidelines-advisory-committee-systematic-reviews (accessed April 28, 2017).

NTP (National Toxicology Program). 2015. *Handbook for conducting a literature-based health assessment using OHAT approach for systematic review and evidence integration*. https://ntp.niehs.nih.gov/pubhealth/hat/noms/index-2.html (accessed July 16, 2017).

Parascandola, M., and D. L. Weed. 2001. Causation in epidemiology. *J Epidemiol Community Health* 55(12):905-912.

Pennant, M., M. Steur, C. Moore, A. Butterworth, and L. Johnson. 2015. Comparative validity of vitamin C and carotenoids as indicators of fruit and vegetable intake: A systematic review and meta-analysis of randomised controlled trials. *Br J Nutr* 114(9):1331-1340.

Pieper, D., T. Mathes, and M. Eikermann. 2014. Can AMSTAR also be applied to systematic reviews of non-randomized studies? *BMC Res Notes* 7:609.

Pogue, J. M., and S. Yusuf. 1997. Cumulating evidence from randomized trials: Utilizing sequential monitoring boundaries for cumulative meta-analysis. *Control Clin Trials* 18(6):580-593; discussion 661-666.

Puhan, M. A., H. J. Schunemann, M. H. Murad, T. Li, R. Brignardello-Petersen, J. A. Singh, A. G. Kessels, G. H. Guyatt, and Grade Working Group. 2014. A GRADE Working Group approach for rating the quality of treatment effect estimates from network meta-analysis. *BMJ* 349:g5630.

Santesso, N., E. A. Akl, M. Bianchi, A. Mente, R. Mustafa, D. Heels-Ansdell, and H. J. Schunemann. 2012. Effects of higher-versus lower-protein diets on health outcomes: A systematic review and meta-analysis. *Eur J Clin Nutr* 66(7):780-788.

Santesso, N., A. Carrasco-Labra, M. Langendam, R. Brignardello-Petersen, R. A. Mustafa, P. Heus, T. Lasserson, N. Opiyo, I. Kunnamo, D. Sinclair, P. Garner, S. Treweek, D. Tovey, E. A. Akl, P. Tugwell, J. L. Brozek, G. Guyatt, and H. J. Schunemann. 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.

Schunemann, H. J., A. D. Oxman, J. Brozek, P. Glasziou, R. Jaeschke, G. E. Vist, J. W. Williams, Jr., R. Kunz, J. Craig, V. M. Montori, P. Bossuyt, G. H. Guyatt, and Grade Working Group. 2008. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ* 336(7653):1106-1110.

Schunemann, H., J. Brozek, G. Guyatt, and A. Oxman. 2013. *Introduction to GRADE Handbook*. http://gdt.guidelinedevelopment.org/app/handbook/handbook.html (accessed May 16, 2017).

Shea, B. J., J. M. Grimshaw, G. A. Wells, M. Boers, N. Andersson, C. Hamel, A. C. Porter, P. Tugwell, D. Moher, and L. M. Bouter. 2007. Development of AMSTAR: A measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol* 7:10.

Shea, B. J., C. Hamel, G. A. Wells, L. M. Bouter, E. Kristjansson, J. Grimshaw, D. A. Henry, and M. Boers. 2009. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. *J Clin Epidemiol* 62(10):1013-1020.

Shekelle, P., S. Newberry, M. Maglione, R. Shanman, B. Johnsen, J. Carter, A. Motala, B. Hulley, Z. Wang, D. Bravata, M. Chen, and J. Grossman. 2009. *Assessment of the need to update comparative effectiveness reviews: Report of an initial rapid program assessment (2005-2009)*, Rockville, MD: Agency for Healthcare Research and Quality. https://www.ncbi.nlm.nih.gov/pubmed/21204320 (accessed July 16, 2017).

Sun, X., M. Briel, S. D. Walter, and G. H. Guyatt. 2010. Is a subgroup effect believable? Updating criteria to evaluate the credibility of subgroup analyses. *BMJ* 340:c117.

Te Morenga, L., S. Mallard, and J. Mann. 2012. Dietary sugars and body weight: Systematic review and meta-analyses of randomised controlled trials and cohort studies. *BMJ* 346:e7492.

Tikkinen, K. A., J. W. Busse, and G. Guyatt. 2017a. *Tool to assess risk of bias in observational studies of natural history of medical symptoms/conditions in general populations*. https://www.evidencepartners.com/resources/methodological-resources (accessed May 16, 2017).

Tikkinen, K. A., R. Cartwright, and G. Guyatt. 2017b. *Tool to assess risk of bias in studies of prognosis*. https://www.evidencepartners.com/resources/methodological-resources (accessed May 16, 2017).

USPSTF (U.S. Preventive Services Task Force). 2015. *Procedure manual*. https://www.uspreventiveservicestaskforce.org/Page/Name/methods-and-processes (accessed July 16, 2017).

WCRF/AICR (World Cancer Research Fund/American Institute for Cancer Research). 2007. *Food, nutrition, physical activity, and the prevention of cancer: A global perspective*. Washington, DC: AICR.

Whiting, P., P. Davies, J. Savovic, D. M. Caldwell, and R. Churchill. 2013. *Evidence to inform the development of ROBIS, a new tool to assess the risk of bias in systematic reviews*. http://www.bristol.ac.uk/media-library/sites/social-community-medicine/robis/ROBIS%20Report%204_9.pdf (accessed July 16, 2017).

Whiting, P., J. Savovic, J. P. Higgins, D. M. Caldwell, B. C. Reeves, B. Shea, P. Davies, J. Kleijnen, R. Churchill, and ROBIS group. 2016. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol* 69:225-234.

WHO (World Health Organization). 2014. *WHO handbook for guideline development*. Geneva, Switzerland: World Health Organization.

WHO. 2015. *Sugars intake for adults and children. Guideline*. Geneva, Switzerland: World Health Organization.

Yetley, E. A., A. J. MacFarlane, L. S. Greene-Finestone, C. Garza, J. D. Ard, S. A. Atkinson, D. M. Bier, A. L. Carriquiry, W. R. Harlan, D. Hattis, J. C. King, D. Krewski, D. L. O'Connor, R. L. Prentice, J. V. Rodricks, and G. A. Wells. 2017. Options for basing Dietary Reference Intakes (DRIs) on chronic disease endpoints: Report from a joint US-/Canadian-sponsored working group. *Am J Clin Nutr* 105(1):249S-285S.

ANNEX 6-1
PRIORITY QUESTIONS IN THE FORMAT OF POPULATION, INTERVENTION, CONTROL, AND OUTCOMES (PICO)[14] FOR GUIDING THE SYSTEMATIC REVIEWS IN SUPPORT OF THE WORLD HEALTH ORGANIZATION GUIDELINE ON FREE SUGARS INTAKE FOR ADULTS AND CHILDREN

What is the effect of a decrease or increase in free sugars intake in adults and children?
What is the effect of restricting intake of free sugars to below 10 percent of total energy intake?[15]

Adults

Population	Apparently healthy adults in low-, middle-, and high-income countries • In each, consider population characteristics, such as age, gender, ethnicity, country/region (urban/rural), socioeconomic status/demographic factors/sanitation, health background, and health status
Intervention/exposure	Definitions • Total sugars • Free sugars • Added sugars (sucrose; table sugar; sugars in processed foods) • Percent of total energy intake from sugars • Consumption of sugar-sweetened beverages • Fruit juices
Control	Comparison of levels Continuous or categorical Adherence to recommendations

[14] Reprinted with permission from WHO (World Health Organization). 2015. *Sugars intake for adults and children. Guideline.* Geneva, Switzerland: World Health Organization.

[15] Less than 10 percent of total energy intake is the existing population nutrient intake goal for free sugars (WHO/FAO, 2003).

Confounders/effect modifiers/intermediates	• Baseline level of all categories of sugars intake • Energy intake • Energy expenditure, fitness, and physical activity • Consider other interventions in design, dietary and non-dietary (protocol to specify) • Consider influence of other aspects of diet/ dietary patterns In cohort studies: unadjusted and adjusted estimates; what adjusted for, how (protocol to specify), and impact Consider whether artificial sweeteners/milk/ other foods are used as control Intermediates • Take into account effect of energy density
Outcome	• Body weight or fatness gain measured by: - weight change, BMI - incidence of obesity and overweight - body fatness[16] and distribution assessed in a variety of ways • Dental caries (not erosion)
Time frame	• For controlled feeding studies where a high proportion of food is directly provided and there is no caloric restriction, outcomes are change in weight or body fatness within a minimum study duration of 8 weeks • For studies where the intervention is advisory or shopping type, without caloric restriction, outcomes are obesity incidence, change in weight or body fatness with a minimum study duration of 6 months (26 weeks)

[16] The percentage of fat (i.e., adipose tissue) that a person's body contains.

Children

Population	Apparently healthy children in low-, middle-, and high-income countries • In each, consider population characteristics, such as age, gender, ethnicity, country/region (urban/rural), socioeconomic status/demographic factors/sanitation, health background, and health status
Intervention/exposure	Definitions • Total sugars • Free sugars • Added sugars (sucrose; table sugar; sugars in processed foods) • Percent of total energy intake from sugars • Consumption of sugar-sweetened beverages • Fruit juices
Control	Comparison of levels Continuous or categorical Adherence to recommendations
Confounders/effect modifiers/intermediates	• Baseline level of all categories of sugars intake • Energy intake • Energy expenditure, fitness, and physical activity • Consider other interventions in design, dietary and non-dietary (protocol to specify) • Consider influence of other aspects of diet/dietary patterns In cohort studies: unadjusted and adjusted estimates; what adjusted for, how (protocol to specify), and impact Consider if artificial sweeteners/milk/other foods are used as control Intermediates • Take into account effect of energy density

Outcome	• Body weight or fatness gain measured by: - weight change, BMI - incidence of obesity and overweight - body fatness and distribution assessed in a variety of ways • Dental caries (not erosion)
Time frame	• For controlled feeding studies where a high proportion of food is directly provided and there is no caloric restriction, outcomes are change in weight or body fatness, with a minimum study duration of 8 weeks • For studies where the intervention is advisory or shopping type, without caloric restriction, outcomes are obesity incidence, change in weight or body fatness, with a minimum study duration of 6 months (26 weeks)

ANNEX 6-1 REFERENCES

WHO (World Health Organization). 2015. *Sugars intake for adults and children. Guideline.* Geneva, Switzerland: World Health Organization.

WHO/FAO (Food and Agriculture Organization). 2003. *Diet, nutrition and the prevention of chronic diseases: Report of a joint WHO/FAO expert consultation.* Geneva, Switzerland: World Health Organization.

ANNEX 6-2
ASSESSMENT TOOLS USED IN SYSTEMATIC REVIEWS

TABLE ANNEX 6-2-1 Examples of Tools for Risk of Bias (or Quality) Assessment for Individual Studies Included in the Systematic Reviews

Tool (and Modifications)	Purpose/Who Is Using It	Risk of Bias Domains	Grading Method	Notes
Nutrition Evidence Library Bias Assessment Tool (NEL BAT)[a]	To assess the risk of bias of each individual study (any design) included in a systematic review Used by the 2015 Dietary Guidelines Advisory Committee	• Selection bias • Performance bias • Detection bias • Attrition bias	The NEL BAT is tailored by study design, with different sets of questions applying to RCTs, non-RCTs, and observational studies. There are four response options, which are given a score: • Yes (0 points) • No (2 points) • Cannot Determine (1 point) • N/A (0 points) The lower the number of points received, the lower the risk of bias.	Adapted from the Cochrane Bias Methods Group No nutrition-specific items

continued

TABLE ANNEX 6-2-1 Continued

Tool (and Modifications)	Purpose/Who Is Using It	Risk of Bias Domains	Grading Method	Notes
Cochrane Risk of Bias Tool[b]	To assess the quality of intervention studies for use in the context of a systematic review and elsewhere It is a domain-based evaluation, in which critical assessments are made separately for different domains. Used in Cochrane systematic reviews	• Selection bias • Performance bias • Detection bias • Attrition bias • Reporting bias • Other bias	This is a two-part tool. Each part contains criteria that reviewers use to formulate their descriptions and judgments. Part 1: Within each domain entry, reviewers describe what was reported to have happened in the study in sufficient detail to support a judgment about the risk of bias. Part 2: Reviewers assign a judgement relating to the risk of bias for that entry. This is achieved by assigning a judgement of Low risk of bias, High risk of bias, or Unclear risk of bias.	Used for randomized studies Supplemented with nutrition-specific items

Risk of Bias in Non-randomized Studies- of Interventions (ROBINS-I)[c]	To evaluate the risk of bias in the results of non-randomized intervention studies that compare the health effects of two or more interventions Developed by members of the Cochrane Bias Methods Group and the Cochrane Non-randomized Studies Methods Group	• Bias due to confounding • Bias in selection of participants into the study • Bias in classification of interventions • Bias due to deviations from intended interventions • Bias due to missing data • Bias in measurement of outcomes • Bias in selection of the reported result	Domain-level risk of bias judgements are made, followed by overall judgements about risk of bias. Response options are: 1. Low risk of bias 2. Moderate risk of bias 3. Serious risk of bias 4. Critical risk of bias 5. No information on which to base a judgement about risk of bias	Used for non-randomized studies Not inclusive of study characteristics that are relevant for nutrition Can be outcome specific Provides a detailed assessment of bias domains
The Newcastle-Ottawa Scale (NOS)[d]	To assess the quality of non-randomized studies (cohort or case-control studies) in meta-analyses and systematic reviews Used by AHRQ	1. **Selection** of the study groups 2. **Comparability** of the groups 3. Ascertainment of either the **exposure** or **outcome** of interest for case-control or cohort studies, respectively	A star system in which a study is judged on the three domains. Each domain has one to four questions to answer. A different set of questions is used for cohort studies and case-control studies.	Designed for case-control and cohort studies Supplemented with nutrition-specific items The face/content validity of the NOS has been established. The NOS has been refined based on experience using it in several projects.

continued

TABLE ANNEX 6-2-1 Continued

Tool (and Modifications)	Purpose/Who Is Using It	Risk of Bias Domains	Grading Method	Notes
National Toxicology Program (NTP) Office of Health Assessment and Translation (OHAT) Risk of Bias Tool[e]	To assess internal validity of individual studies (human and non-human animal studies) typically considered in environmental health reviews	Selection bias • Confounding bias • Performance bias • Attrition/Exclusion bias • Detection bias • Selective reporting bias • Other sources of bias	Each domain has one to three questions, depending on study type. Response options are: **Definitely Low** risk of bias: There is direct evidence of low risk of bias practices. **Probably Low** risk of bias: There is indirect evidence of low risk of bias practices, or it is deemed that deviations from low risk of bias practices for these criteria during the study would not appreciably bias results, including consideration of direction and magnitude of bias. **Probably High** risk of bias: There is indirect evidence of high risk of bias practices, or there is insufficient information provided about relevant risk of bias practices.	This tool was developed using established guidelines for experimental human studies (randomized clinical trials).

	Definitely High risk of bias: There is direct evidence of high risk of bias practices.

[a] See http://www.nel.gov/topic.cfm?cat=3384 (accessed July 24, 2017).

[b] See http://methods.cochrane.org/bias/assessing-risk-bias-included-studies (accessed July 24, 2017); http://handbook.cochrane.org (accessed July 24, 2017); Higgins et al., 2011.

[c] See https://sites.google.com/site/riskofbiastool/welcome/home (accessed July 24, 2017); Sterne et al., 2016.

[d] See http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp (accessed July 24, 2017).

[e] See https://ntp.niehs.nih.gov/pubhealth/hat/noms/index-2.html (accessed July 24, 2017).

TABLE ANNEX 6-2-2 Tools for Evaluating the Strength (or Quality) of the Body of Evidence*

Tool (and modifications)	Purpose/ Who Is Using It	Elements (or Quality Criteria)	Grading Method	Notes
Grading of Recommendations, Assessment, Development and Evaluation (GRADE)[a] The working group discourages the use of "modified GRADE approaches" that differ from the approach described by the GRADE Working Group.	GRADE is an approach to evaluate the quality of the evidence and the strength of recommendations Used for clinical practice guideline development Used to evaluate the evidence in systematic reviews Recommended approach for use in Cochrane Reviews	An initial confidence is determined based on study design: RCTs start as High quality evidence. Observational studies start with a Low quality rating. Studies may be downgraded based on: • Study limitations (risk of bias) • Inconsistency of results • Indirectness of evidence • Imprecision • Reporting bias Studies may be upgraded based on: • Magnitude of treatment effect • Evidence of dose-response relationship • If all plausible biases would decrease the magnitude of an apparent treatment effect	The quality of the body of evidence is classified in one of four levels: ***High quality*:** We are very confident that the true effect lies close to that of the estimate of the effect ***Moderate quality*:** We are moderately confident in the effect estimate. The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different ***Low quality*:** Our confidence in the effect estimate is limited. The true effect may be substantially different from the estimate of the effect ***Very low quality*:** We have very little confidence in the effect estimate. The true effect is likely to be substantially different from the estimate of effect The strength of the recommendations are graded as strong or weak based on: • Quality of evidence • Uncertainty about the balance between desirable and undesirable effects • Uncertainty or variability in values and preferences • Uncertainty about whether the intervention represents a wise use of resources	No nutrition-specific items

Nutri-Grade[b]	Assess the meta-evidence (from RCTs and cohort studies) of an association or effect between different nutrition factors and outcomes, taking into account nutrition research-specific requirements not considered by other tools	Seven items for meta-analyses of RCTs and eight items for meta-analyses of cohort studies: • Risk of bias, study quality, and study limitations • Precision • Heterogeneity • Directness • Publication bias • Funding bias • Study design (only for meta-analyses of RCTs) • Effect size (only for meta-analyses of cohort studies) • Dose-response (only for meta-analyses of cohort studies)	Each item is scored from 0-3 points, depending on the item. A total score is calculated: High (≥8 total points): There is high confidence in the effect estimate, and further research probably will not change the confidence in the effect estimate. Moderate (6-7.99 total points): There is moderate confidence in the effect estimate; further research could add evidence on the confidence and may change the effect estimate. Low (4-5.99 total points): There is low confidence in the effect estimate; further research will provide important evidence on the confidence and likely change the effect estimate. Very low (0-3.99 total points): There is very low confidence in the effect estimate; meta-evidence is very limited and uncertain.	Takes into account nutrition research-specific requirements (e.g., dietary assessment methods and their validation or funding bias) for assessing the meta-evidence Based on GRADE, but has modified the classification for meta-analyses of RCTs and cohort studies compared with GRADE
Nutrition Evidence Library (NEL) Grading Rubric[c]	To judge the strength of the body of evidence for each conclusion statement in the Dietary Guidelines Advisory Committee (DGAC) report Used by the 2010 and 2015 DGACs	For each study, researchers carefully evaluated five elements: • Quality/Risk of bias (based on NEL BAT) • Quantity • Consistency • Impact • Generalizability	A grade for the strength of the body of evidence supporting each conclusion statement was determined for each of the five elements: Grade I: Strong Grade II: Moderate Grade III: Limited Grade IV: Expert opinion only** Grade V: Grade not assignable **Only included in the 2010 DGAC	The grading rubric was developed for the 2010 DGAC, and was modified for the 2015 DGAC No nutrition-specific items

continued

TABLE ANNEX 6-2-2 Continued

Tool (and modifications)	Purpose/ Who Is Using It	Elements (or Quality Criteria)	Grading Method	Notes
AHRQ Methods Guide for Evidence-based Practice Centers (EPC)[d]	To grade the strength of the evidence for each outcome in a systematic review	Required domains: • Study limitations (risk of bias) • Consistency • Directness • Precision • Reporting bias Additional domains: • Dose-response association • Strength of association (magnitude of effect) • Plausible confounding that would decrease observed effect	Summary tables that rate the strength of the evidence Overall grade for strength of evidence for an outcome is based on domain scores and a holistic summary recognition of the possible complex interaction among domains: • High • Moderate • Low • Insufficient Strength of evidence grades are first done separately for RCTs and observational studies and then combined into one overall grade.	Conceptually similar to the GRADE approach, but tailored for AHRQ's Evidence Practice Center reviews No nutrition-specific items No explicit instruction on "hierarchy" of the evidence, and not a grading for "strength of the recommendations" because AHRQ systematic reviews do not make clinical recommendations. The clinical recommendation process is a totally independent process thus not covered in ARHQ methods guide.
U.S. Preventive Services Task Force (USPSTF)[e]	USPSTF provides evidence-based recommendations about clinical preventive services and health promotion.	1. Assess individual studies: • Internal validity of individual studies • External validity (applicability) of individual studies (subjects, setting, provider)	Each recommendation is assigned a letter grade (A, B, C, or D or an I statement) based on the strength of the evidence and the balance of benefits and harms of a preventive service:	USPSTF conducts its evidence reviews with assistance from AHRQ/EPC.

Used by health care professionals and the American public	2. Assess the level of evidence at the key question level: • Quantitative or qualitative synthesis, depending on number of studies and heterogeneity 3. Overall summary of evidence considers: • Key question • Number of studies • Summary of findings • Consistency/precision • Reporting bias • Overall study quality • Body of evidence limitations • Applicability • Overall strength of evidence 4. Make a recommendation: • Assess the adequacy of evidence at the key question level • Assess the adequacy of evidence at the linkage level • Estimate the magnitude of benefit and harm of the preventive service • Evaluate the certainty of the evidence for net benefit for the preventive service (High, Moderate, Low) • Estimate the magnitude of the net benefit of the preventive service • Develop a recommendation grade for the preventive service in the relevant population, based on the above parameters	A: The USPSTF recommends the service. There is high certainty that the net benefit is substantial. B: The USPSTF recommends the service. There is high certainty that the net benefit is moderate or there is moderate certainty that the net benefit is moderate to substantial. C: The USPSTF recommends selectively offering or providing this service to individual patients based on professional judgment and patient preferences. There is at least moderate certainty that the net benefit is small. D: The USPSTF recommends against the service. There is moderate or high certainty that the service has no net benefit or that the harms outweigh the benefits. I statement: The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of the service. Evidence is lacking, of poor quality, or conflicting, and the balance of benefits and harms cannot be determined.

continued

TABLE ANNEX 6-2-2 Continued

Tool (and modifications)	Purpose/ Who Is Using It	Elements (or Quality Criteria)	Grading Method	Notes
National Toxicology Program (NTP) Office of Health Assessment and Translation (OHAT) Handbook[f]	Framework for systematic review and evidence integration for literature-based health assessments of environmental substances	Seven-step process 1. Formulate problem and develop protocol 2. Search for and select studies for inclusion 3. Extract data from studies 4. Assess quality of individual studies using OHAT risk of bias tool (see above in Table 1) 5. Synthesize evidence and rate confidence in the body of evidence 6. Translate confidence ratings into level of evidence of health effects 7. Integrate evidence to develop hazard identification conclusions **Step 5** An initial confidence is determined based on four criteria related to study design: 1. Controlled exposure 2. Exposure occurs before outcome 3. Outcome assessed at individual level 4. Comparison group used	Level of confidence in the body of evidence: • **High Confidence (++++)** in the association between exposure to the substance and the outcome. The true effect is highly likely to be reflected in the apparent relationship. • **Moderate Confidence (+++)** in the association between exposure to the substance and the outcome. The true effect may be reflected in the apparent relationship. • **Low Confidence (++)** in the association between exposure to the substance and the outcome. The true effect may be different from the apparent relationship. • **Very Low Confidence (+)** in the association between exposure to the substance and the outcome. The true effect is highly likely to be different from the apparent relationship. In Step 6, determine the conclusions for the level of evidence of health effect (assess human and animal data separately).	Confidence rating is based primarily on guidance from the GRADE working group Confidence rating accommodates the need to integrate data from multiple evidence streams (human, animal, in vitro) and focus on observational human studies.

Then, the following factors decrease confidence:
Risk of bias, unexplained inconsistency, imprecision, publication bias

Or increase confidence:
Large magnitude of effect, dose response, residual confounding, consistency

- **High** level of evidence: There is high confidence in the body of evidence for an association between exposure to the substance and the health outcome(s).
- **Moderate** level of evidence: There is moderate confidence in the body of evidence for an association between exposure to the substance and the health outcome(s).
- **Low** level of evidence: There is low confidence in the body of evidence for an association between exposure to the substance and the health outcome(s), or no data are available.
- **Evidence of No Health Effect:** There is high confidence in the body of evidence that exposure to the substance is not associated with the health outcome(s).
- **Inadequate evidence:** There is insufficient evidence available to assess whether the exposure to the substance is associated with the health outcome(s).

continued

TABLE ANNEX 6-2-2 Continued

Tool (and modifications)	Purpose/ Who Is Using It	Elements (or Quality Criteria)	Grading Method	Notes
World Cancer Research Fund/ American Institute for Cancer Research (WCRF/AICR)[g]	Methodology used by the WCRF/ AICR to ascertain causal relationships between food, nutrition, physical activity, and cancer	• Study type • Heterogeneity within or between study types or in different populations • Random or systematic error (including confounding, measurement error, and selection bias) • Plausible biological gradient (dose-response) • Strong and plausible experimental evidence (human studies or animal models)	**Criteria for grading evidence:** ***Convincing*** = Evidence is strong enough to support a judgment of a convincing causal relationship, which justifies goals and recommendations designed to reduce the incidence of cancer. ***Probable*** = Evidence is strong enough to support a judgment of a probable causal relationship, which would generally justify goals and recommendations designed to reduce the incidence of cancer. ***Limited-suggestive*** = Evidence that is too limited to permit a probable or convincing causal judgment, but where evidence suggests a direction of effect. The evidence may have methodological flaws, or be limited in amount, but shows a generally consistent direction of effect. This almost always does not justify recommendations designed to reduce the incidence of cancer. ***Limited-no conclusion*** = Evidence is so limited that no firm conclusion can be made. ***Substantial effect on risk unlikely*** = Evidence is strong enough to support a judgment that a particular food, nutrition, or physical activity exposure is unlikely to have a substantial causal relation to a cancer outcome.	Judgments were based on an assessment of the evidence available in the scientific literature, with due consideration given to the advantages and disadvantages of each type of study design and to the quality of individual studies. An inclusive approach was taken that recognized the relative strengths and weaknesses of different types of study, but in which no single type of study design is given pre-eminence.

Special upgrading factors

- Presence of a plausible biological gradient (dose-response) in the association. Such a gradient need not be linear or even in the same direction across the different levels of exposure, so long as this can be explained plausibly.
- A particularly large summary effect size (an odds ratio or relative risk of 2.0 or more, depending on the unit of exposure) after appropriate control for confounders.
- Evidence from RCTs in humans.
- Evidence from appropriately controlled experiments demonstrating one or more plausible and specific mechanisms actually operating in humans.
- Robust and reproducible evidence from experimental studies in appropriate animal models showing that typical human exposures can lead to relevant cancer outcomes.

---

* Any grading tool for strength of the body of evidence can be applied to grade the body of evidence of any study design that is defined by the inclusion criteria of the systematic reviews. If a systematic review included both observational studies and randomized controlled trials (RCTs), the strength of the body of evidence is rated as such. If a systematic review only included RCTs, the strength of the body of evidence can only rate the body of RCT evidence.

[a] See www.gradeworkinggroup.org (accessed July 24, 2017).

[b] See http://www.nutrigrade.net (accessed July 24, 2017); Schwingshackl et al., 2016.

[c] See http://www.nel.gov/topic.cfm?cat=3368 (accessed July 24, 2017); https://health.gov/dietaryguidelines/2015-scientific-report/05-methodology.asp (accessed July 24, 2017).

[d] See http://effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productid=318 (accessed July 24, 2017); Berkman et al., 2015.

[e] See https://www.uspreventiveservicestaskforce.org/Page/Name/methods-and-processes (accessed July 24, 2017).

[f] See https://ntp.niehs.nih.gov/pubhealth/hat/noms/index-2.html (accessed July 24, 2017).

[g] WCRF/AICR, 2007.

TABLE ANNEX 6-2-3 Tools for Evaluating the "Quality" (or Reporting Quality*) of Systematic Reviews

Tool (and Modifications)	Purpose/Who Is Using It	Checklist	Grading Method	Notes
A Measurement Tool to Assess Systematic Reviews (AMSTAR)[a]	To develop and evaluate the methodological quality of systematic reviews To guide the conduct of systematic reviews To aid in teaching about systematic reviews Used by professional health care associations and other policy institutions Used by the 2015 Dietary Guidelines Advisory Committee	AMSTAR Checklist (Answer yes/no/can't answer/not applicable): 1. Was an a priori design provided? 2. Was there duplicate study selection and data extraction? 3. Was a comprehensive literature search performed? 4. Was the status of publication (i.e., grey literature) used as an inclusion criterion? 5. Was a list of studies (included and excluded) provided? 6. Were the characteristics of the included studies provided? 7. Was the scientific quality of the included studies assessed and documented? 8. Was the scientific quality of the included studies used appropriately in formulating conclusions? 9. Were the methods used to combine the findings of studies appropriate?	A well-done systematic review is one that has addressed all the items on the checklist. (Sometimes this checklist is quantified by assigning number values to each response.)	AMSTAR can be used for systematic reviews of non-RCTs, but revisions and extensions are needed. Has good inter-rater agreement, test-retest reliability, and face and construct validity.

		10. Was the likelihood of publication bias assessed? 11. Was the conflict of interest included?		
Revised Assessment of Multiple Systematic Reviews (R-AMSTAR)[b]	To quantify the quality of systematic reviews	Added quantification to the AMSTAR checklist	For each of the 11 questions in the AMSTAR checklist, a set of criteria is given and a number value is assigned based on how many criteria are met. Each question's score ranges between 1 and 4, and the R-AMSTAR total score has a range of 11 to 44. A total score of 11 signifies that none of the AMSTAR criteria was satisfied. A score of 44 shows that all of the criteria of systematic review excellence were verified in every domain. Low R-AMSTAR total scores should lead to prudence on the part of the clinician. High R-AMSTAR total scores should impart a certain degree of confidence about the clinical relevance and implications of the findings discussed in the high scoring systematic review.	Not developed by the AMSTAR group

continued

TABLE ANNEX 6-2-3 Continued

Tool (and Modifications)	Purpose/Who Is Using It	Checklist	Grading Method	Notes
Risk of Bias in Systematic Reviews Tool (ROBIS)[c]	To assess risk of bias in systematic reviews Target audience is guideline developers, authors of overviews of systematic reviews, and review authors who might want to assess or avoid risk of bias in their reviews.	3-phase tool Phase 1: Assess relevance (optional) Phase 2: Identify concerns with the review process (study eligibility criteria; identification and selection of studies; data collection and study appraisal; synthesis and findings) Phase 3: Judge risk of bias Signaling questions are used in phases 2 and 3. Responses are: yes, probably yes, probably no, no, or no information. Yes = low concern.	Phase 2: Depending on the responses to the signaling questions, the level of concern about bias associated with each domain is judged as "low," "high," or "unclear." If the answers to all signaling questions for a domain are "yes" or "probably yes," then level of concern can be judged as low. If any signaling question is answered "no" or "probably no," potential for concern about bias exists. Phase 3: Overall bias is judged based on phase 2 analysis and further signaling questions. Results can be summarized in text or depicted in a table or graph.	Used for assessing risk of bias in systematic reviews of RCTs systematic reviews of non-RCTs

* Only ROBIS was developed to assess the risk of bias of a systematic review. AMSTAR is primarily a rating of reporting quality and cannot assess the other potential biases in a systematic review. AMSTAR is generally used as a quality assessment tool for complex systematic reviews that want to include existing systematic reviews to determine whether a systematic review has sufficient quality to be included.

[a] See https://amstar.ca/index.php (accessed July 24, 2017); Pieper et al., 2014; Shea et al., 2007, 2009.

[b] Kung et al., 2010; Pieper et al., 2015.

[c] See http://www.bristol.ac.uk/social-community-medicine/projects/robis (accessed July 24, 2017); Whiting et al., 2013, 2016.

ANNEX 6-2 REFERENCES

Berkman, N. D., K. N. Lohr, M. T. Ansari, E. M. Balk, R. Kane, M. McDonagh, S. C. Morton, M. Viswanathan, E. B. Bass, M. Butler, G. Gartlehner, L. Hartling, M. McPheeters, L. C. Morgan, J. Reston, P. Sista, E. Whitlock, and S. Chang. 2015. Grading the strength of a body of evidence when assessing health care interventions: An EPC update. *J Clin Epidemiol* 68(11):1312-1324.

Higgins, J. P., D. G. Altman, P. C. Gotzsche, P. Juni, D. Moher, A. D. Oxman, J. Savovic, K. F. Schulz, L. Weeks, J. A. Sterne, G. Cochrane Bias Methods, and G. Cochrane Statistical Methods. 2011. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 343:d5928.

Kung, J., F. Chiappelli, O. O. Cajulis, R. Avezova, G. Kossan, L. Chew, and C. A. Maida. 2010. From systematic reviews to clinical recommendations for evidence-based health care: Validation of revised assessment of multiple systematic reviews (R-AMSTAR) for grading of clinical relevance. *Open Dent J* 4:84-91.

Pieper, D., T. Mathes, and M. Eikermann. 2014. Can AMSTAR also be applied to systematic reviews of non-randomized studies? *BMC Res Notes* 7:609.

Pieper, D., R. B. Buechter, L. Li, B. Prediger, and M. Eikermann. 2015. Systematic review found AMSTAR, but not R(evised)-AMSTAR, to have good measurement properties. *J Clin Epidemiol* 68(5):574-583.

Schwingshackl, L., S. Knüppel, C. Schwedhelm, G. Hoffmann, B. Missbach, M. Stelmach-Mardas, S. Dietrich, F. Eichelmann, E. Kontopanteils, K. Iqbal, K. Aleksandrova, S. Lorkowski, M. F. Leitzmann, A. Kroke, and H. Boeing. 2016. Perspective: NutriGrade: A scoring system to assess and judge the meta-evidence of randomized controlled trials and cohort studies in nutrition research. *Advances in Nutrition: An International Review Journal* 7(6):994-1004.

Shea, B. J., J. M. Grimshaw, G. A. Wells, M. Boers, N. Andersson, C. Hamel, A. C. Porter, P. Tugwell, D. Moher, and L. M. Bouter. 2007. Development of AMSTAR: A measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol* 7:10.

Shea, B. J., C. Hamel, G. A. Wells, L. M. Bouter, E. Kristjansson, J. Grimshaw, D. A. Henry, and M. Boers. 2009. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. *J Clin Epidemiol* 62(10):1013-1020.

Sterne, J. A., M. A. Hernan, B. C. Reeves, J. Savovic, N. D. Berkman, M. Viswanathan, D. Henry, D. G. Altman, M. T. Ansari, I. Boutron, J. R. Carpenter, A. W. Chan, R. Churchill, J. J. Deeks, A. Hrobjartsson, J. Kirkham, P. Juni, Y. K. Loke, T. D. Pigott, C. R. Ramsay, D. Regidor, H. R. Rothstein, L. Sandhu, P. L. Santaguida, H. J. Schunemann, B. Shea, I. Shrier, P. Tugwell, L. Turner, J. C. Valentine, H. Waddington, E. Waters, G. A. Wells, P. F. Whiting, and J. P. Higgins. 2016. ROBINS-I: A tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* 355:i4919.

WCRF/AICR (World Cancer Research Fund/American Institute for Cancer Research). 2007. *Food, nutrition, physical activity, and the prevention of cancer: A global perspective*. Washington, DC: AICR.

Whiting, P., P. Davies, J. Savovic, D. M. Caldwell, and R. Churchill. 2013. *Evidence to inform the development of ROBIS, a new tool to assess the risk of bias in systematic reviews*. http://www.bristol.ac.uk/media-library/sites/social-community-medicine/robis/ROBIS%20Report%204_9.pdf (accessed July 16, 2017).

Whiting, P., J. Savovic, J. P. Higgins, D. M. Caldwell, B. C. Reeves, B. Shea, P. Davies, J. Kleijnen, R. Churchill, and ROBIS group. 2016. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol* 69:225-234

7

Intake-Response Relationships and Dietary Reference Intakes for Chronic Disease

The existence of a causal relationship between intake of a nutrient or food substance (NOFS) and risk of a chronic disease (see Chapter 6) is necessary, but not sufficient to identify intakes that reduce (or increase) chronic disease risks. Whether Dietary Reference Intakes (DRIs) can be recommended depends upon the availability of identifiable and quantitative intake-response relationships. Risks of chronic diseases will typically be expressed as relative risks (RRs), and intakes may be expressed as daily exposures to a food substance, cumulative exposures over specified periods of time, or as a measured biomarker of intake exposure. RRs may be shown to increase or decrease in relation to intake. In some cases, it is possible that, for the same substance, RRs may decrease over a given range of intakes and then increase over another range. This is only one of the many challenges when characterizing intake-response relationships.

Options for Basing Dietary Reference Intakes (DRIs) on Chronic Disease Endpoints: Report from a Joint US-/Canadian-Sponsored Working Group (i.e., the Options Report) provides a comprehensive review of issues that arise in identifying and applying intake-response relationships (Yetley et al., 2017, Appendix B) and offers a series of options for resolving them so that if certainty in the relationship exists, chronic disease DRIs can be recommended. Complementing Chapter 6—which addresses whether a causal relationship exists—this chapter provides recommendations on scientific and methodological considerations regarding the quantitative characterization of the intake-response relationship and implications for risk assessment. In addition to all guiding principles in Chapter 6 being applicable

to characterizing intake-response relationships, one guiding principle that applies specifically to quantitative relationships is added in this chapter.

Consistent with the approach for evaluating evidence about causal relationships, the GRADE (Grading of Recommendations, Assessment, Development and Evaluation) system provides guidance for evaluating evidence about the intake-response relationships. However, the GRADE working group has not explicitly addressed or published any documents addressing the specific issue of the credibility of intake-response relationships. The committee also offers commentary and guidance on conceptual challenges raised in the Options Report, such as establishing risk-reductions goals for DRIs based on RR, rather than on absolute risk (the measure traditionally used for DRIs based on deficiency diseases and high intake toxicity) and developing DRIs that take into account various interactions among food substances that may influence risk of chronic disease.

Experience with the application of various methods described in this chapter to characterize the NOFS-chronic disease relationships is limited. Therefore, the recommendations in this chapter should be considered preliminary, subject to revision as experience with their application accumulates and as new scientific evidence emerges relevant to intake-response relationships for chronic diseases. For example, several options were suggested for developing multiple DRIs for a single nutrient, based on different degrees of risk reduction (sometimes referred to in the Options Report as "families of DRI values"). The committee does not recommend this approach because of challenges communicating the DRIs' interpretation and possible confusion as they are applied by users. Such an option could be reconsidered after experience with chronic disease DRIs accumulates or when there is sufficient experience with the use of risk algorithms that integrate results from multiple indicators. The committee also notes that without significant justification, such as compelling evidence that the benefit of chronic disease reduction greatly exceeds any risk of nutrient deficiency or excess, chronic diseases DRIs should not be recommended at intakes less than the Recommended Dietary Allowance (RDA) or greater than the Tolerable Upper Intake Level (UL). For nutrients with only an Adequate Intake (AI), DRI committees may need to grapple with what quantitative parameters to set.

IDENTIFYING AND DESCRIBING INTAKE-RESPONSE RELATIONSHIPS

This section describes methodologic considerations for modeling and quantifying the intake-response relationship that DRI committees should attend to as they review the analyses conducted (see also the later section on "Examples of Approaches in Nutrition"). Modeling the intake-response

relationship for the purpose of recommending DRIs is a complex and novel situation. Choices made when modeling the relationship (e.g., modeling approach, dietary intake transformations) can lead to bias, and, as DRI committees review the evidence, such choices should be assessed as part of the risk of bias assessment of individual studies or when assessing the appropriateness of the analysis itself (i.e., pooled analysis or meta-analysis). DRI committees should not be constrained by one particular modeling approach, as new methods may emerge that could be optimal for a particular nutrient or chronic disease endpoint.

The Options Report (see Appendix B) recognized that the shape of the intake-response relationship can be quite diverse for various NOFS-chronic disease relationships (see also Chapter 4, where intake may be characterized by a biomarker of intake or by self-report that may have been calibrated using biomarker data). The relationship may be linear or non-linear. For example, the Committee to Review Dietary Reference Intakes for Vitamin D and Calcium (IOM, 2011) reported on several studies showing a U-shaped or reverse J-shaped association of serum 25(OH)D, a biomarker of vitamin D status, and total or all-cause mortality. No firm biological rationale was provided to explain these relationships. However, mortality from a toxic or adverse event is not thought to be the underlying cause. DRI committees will need to determine whether any non-linear relationships, such as U-shaped or J-shaped, are due to the intake-response relationship with the actual disease under investigation or due to adverse events or toxicities. In that sense, although mechanistic data are not directly considered in assessing the certainty of the evidence (see Chapter 6), and it is not, of itself, sufficient for DRI development, having an understanding of the kinetics and dynamics of an NOFS in the body can be of great value to several considerations that are central to DRI recommendations. For example, mechanistic data can be helpful in providing biological plausibility of observations, determining the reliability of nutritional exposure surrogate markers, and understanding plausible shapes to intake-response relationships, including both increases and decreases in chronic disease risks (e.g., Tan et al., 2016).

Characterization of the intake-response relationships (e.g., with more than one chronic disease) will be one of the goals of DRI committees. The ability to quantify how much reduction in chronic disease risk would result from a given change in the intake of an NOFS will depend on the nature of the available data from randomized controlled trials (RCTs) or from observational studies. RCTs test investigator-assigned doses of a treatment (for the purposes of the DRI, a nutrient-based RCT would have one or a few test doses of an NOFS) and test the treatment's effect on the outcome (ideally, a chronic disease endpoint but measuring a qualified surrogate is useful, see Chapter 5). A successful RCT can draw a cause-effect inference

that the quantifiable level of "NOFS x" tested in the trial reduced "chronic disease (or qualified surrogate) y" by the observed effect size. It is important to note that many RCTs are typically powered for a 10 to 20 percent difference in disease outcomes by treatment arm. The true quantifiable relationship may be at lower or higher values of the nutrient than were tested, and the true effect size in the general population also may be lower or higher, but inferences can be drawn only on what was tested in the trial so this may limit usable information for DRI committees. Nonetheless, RCTs can offer quantifiable, cause-effect data on the intake-response relationship.

Observational studies will have less quantifiable data on the intake-response relationship due to the nature of their design and analysis and residual confounding, but they may have a wider range of intakes that is more similar to the population as a whole (see Chapter 3). Most available observational data will come from prospective cohorts or from case-control studies that are nested within cohorts. Data from these studies are typically presented as the RR or hazard ratio (HR), which compares risk of one intake group relative to another (or across a range of intakes). Box 7-1 provides more information on RR and absolute risk estimates and how they might inform the intake-response relationship. In describing the intake-response relationship, it is particularly important to rely on epidemiological studies where the nutrient intake methodology minimizes the potential for systematic biases (see Chapter 4).

Methodological Considerations

Was the Appropriate Model of the NOFS-Chronic Disease Relationship Selected?

One of the first things to consider in modeling the intake-response relationship is the pattern or nature of the intake distribution. Intake of any one nutrient is almost never zero due to the diverse food supply, which includes an array of micronutrient-enriched and micronutrient-fortified foods. The shape of the nutrient intake distribution infrequently resembles a normal distribution (see Figure 7-1). Rather, the intake distribution is usually skewed and asymmetrical, often with a long tail at the upper end of the intake range. Skewed intake distributions tend to exist for both self-reported dietary exposures and for nutrient biomarker exposures. The manner in which skewed intake data are mathematically transformed may depend on the nature of the analysis. For example, if the intake-response relationship models the intake in quantiles, then transformations may not be necessary because the effect of extreme values or outliers is minimized in such modeling. Multiple methods exist for appropriate data transformation, including natural logarithmic, square root, cubic, and quadratic

BOX 7-1
Absolute Risk Versus Relative Risk

Because chronic diseases tend to have complex etiologies, of which history of exposure to a particular food substance may be but one component, and because multiple chronic diseases may share risk factors, particularly in such areas as diet and physical activity patterns, it is important to consider some form of overall health benefit versus risk when considering a chronic disease DRI. Although no universal means of summarizing health benefits versus risks exists, some food substances may have a relatively short list of chronic disease outcomes that are plausibly related to the dietary variable under study, and are considered as serious or important health outcomes. For example, in the Women's Health Initiative RCT of a low-fat eating pattern, breast, colorectal, ovary, and endometrial cancer, coronary heart disease, and mortality from all other causes were combined to form a "global index" that played a role in trial monitoring and reporting (e.g., WHI Study Group,1998).

Consideration of some form of overall health benefits versus risks, or even consideration of the motivation for a DRI relative to a single chronic disease, naturally involves some focus on absolute risks in addition to that for RR. In fact, from some perspectives, absolute risks are more directly relevant when considering potential DRIs for a specific food substance. However, absolute risks tend to be much more sensitive to study population characteristics than are RRs. This means that RRs, but not absolute risks, can often be transported from one population to another, or can usefully be combined across studies in multiple populations. (Note, however, that the possibility of interactions of RRs with study population characteristics must still be kept in mind.)

In recent years, investigators have strongly emphasized absolute risk assessment for a number of chronic diseases. Risk calculators, which typically project chronic disease occurrence over a period of a few years while taking into account competing risks (e.g., Gail et al., 1989) are widely used for disease screening and health monitoring, as well as for other purposes. These calculators mostly rely on a model for RR (or more accurately for the HR), which is combined with estimated "baseline" absolute risks from a study cohort or, in some instances, is combined with baseline risk estimates from a larger population (e.g., Surveillance, Epidemiology, and End Results [SEER] population in the United States) to project absolute disease risks. To date, such calculators have had little dependence on or consideration of dietary exposures, though the inclusion of diet-related body composition data are commonplace. Hence, even though absolute risk projections may not be available that incorporate the food substance under consideration by a DRI committee, it will be useful for such committees to have available data on the overall (absolute) risks (e.g., from national disease registers) for the chronic diseases that are thought to be plausibly related to the food substance. This information can be used in potential DRI considerations as informal supplementary material to RRs estimates for each such disease.

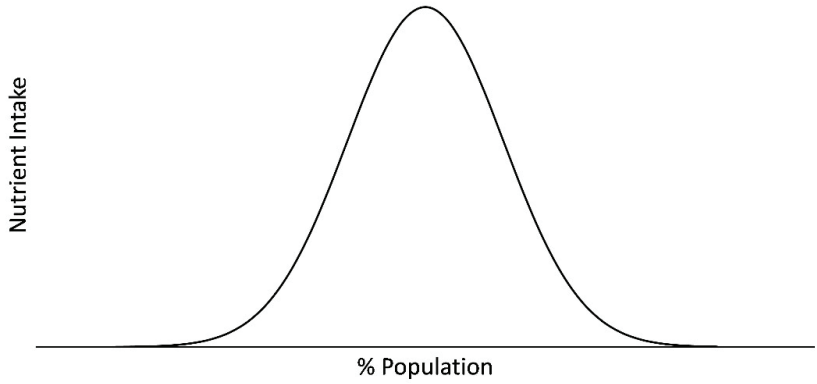


FIGURE 7-1 Example of a normal nutrient intake distribution curve.

transformations (Rosner, 2016). These transformations are often necessary to meet with assumptions of a particular modeling procedure. Intake-response can be estimated by treating the quantiles as ordered, continuous variables and applying an appropriate test for trend across the categories. Alternatively, if the data are suitable for a linear model, and the intake data are modeled as continuous, then interpretation is that for every unit change in X then the expected value of the response Y (outcome) changes by β_i, where β_i may be either positive or negative. For example, Yu et al. reported that, in the Prevención con Dieta Mediterránea (PREDIMED) trial, every per standard deviation (SD) increase in plasma tryptophan was associated with a 21 percent reduced risk (HR per SD = 0.79; 95% confidence interval [CI]: 0.63-0.98) of incident cardiovascular disease (Yu et al., 2017). This type of linear model may on the surface appear to be a better quantification of the intake-response relationship than are categorical models, but caution should still be applied regarding causal inference (see Chapter 6), as all models have some degree of uncertainty and error.

Were the Measurements of Intake and Chronic Disease Outcome Appropriate?

Important factors that will influence, and potentially bias, the intake-response relationship include the nutrient intake and the chronic disease measures. Reliability and accuracy of those measures is essential when identifying chronic disease DRIs. Many of these measurement issues and approaches to assess the risk of bias of individual studies are covered both in the Options Report as well as in Chapters 4 and 5.

Were Confounding Factors Considered?

DRI committees will need to consider a number of challenges regarding modeling and interpreting intake-response data, regardless of whether the intake exposure data come from self-report, biomarker-calibrated self-report, or biomarkers of intake (see Chapter 4). One such important issue is whether studies have appropriately measured (or measured with reasonable precision) potential confounding variables and whether these variables are included in the modeling. Control for confounding is critical so that spurious associations are not made between an NOFS exposure and a disease outcome. The committee assumes that the systematic review team would have evaluated whether individual studies adjusted for the appropriate confounders during the risk-of-bias assessment. DRI committees could come up with a priori lists of likely confounders of the food substance/outcome relationship.

Then, DRI committees will need to determine whether studies included in systematic reviews, for example, had sufficient control for confounding to be able to evaluate the certainty of a causal relationship DRI (see Chapter 6). As mentioned in Chapter 3, residual confounding due to lack of proper statistical adjustments, untestable assumptions, or measurement errors limits the ability to assess causality in observational studies. Control of confounding is critical because chronic diseases have multiple causes and risk factors and dietary intake is only one of them. Age, sex, family history of a particular chronic disease, physical activity, body weight, body adiposity, socio-economic status, health disparities, intake of other NOFSs (including kilocalories), co-morbid conditions, medication use, and other lifestyle habits (e.g., smoking, use of alcohol) can confound the relationship between the intake and the response, act as effect modifiers, be important mediators of the association under study, or play multiple such roles. The ability to control confounding depends in part on the ability to measure these demographic, health, and lifestyle-related variables with precision, and on the appropriate use of these variables in data analysis, given their multiple possible roles. Sensitivity analysis can rule out uncontrolled confounding. For example, in an analysis of the relationship between a nutrient and a health outcome, such as carbohydrate and risk of type 2 diabetes, a sensitivity analysis could be conducted excluding those with "prediabetes" or evidence of glucose intolerance. If the HRs or RRs do not differ for the models with the entire cohort, then the sensitivity analysis suggests that those excluded populations do not have characteristics that confound the overall analysis. If the disease risk estimates differ substantially, then those populations with confounding characteristics are often excluded from the final models.

Were Other Potential Effects Considered?

Other variables that may influence the intake-response relationship are the gut microbiome and polymorphisms in various metabolizing, transport, or degradation enzymes. However, such factors as genetic characteristics and the microbiome are not commonly measured (or they are measured on a subset of participants only) in most population-based observational studies or RCTs, so they likely contribute to understanding unexplained variation in the intake-response relationship.

Another critical analytic consideration in the intake-response relationship is the time course of the intake exposure and its relationship to the chronic disease. Available data often do not identify the relevant time course or exposure window for most food substances in terms of their relationship to the response. More comprehensive data may be needed to guide DRI committees to better evaluate the likely long latency period between intake and chronic disease endpoints and to help determine whether they should restrict data evaluation to a defined time frame before the response or rely on data with various time exposures of intakes.

Were Interactions with Other Nutrients Considered?

An understanding of nutrient interactions is critical for correct interpretation of the nutrient intake response and risk assessment. The term "nutrient interactions" covers four possible scenarios. First, it is possible that the effect of a nutrient on health differs based on the intake of another nutrient, as in the example of the effects of high dietary sodium intake being mitigated by high potassium intake (Crillo et al., 1994; Sellmeyer et al., 2002; Whelton et al., 1997). Second, two nutrients may compete for absorption or transport, as in the case of zinc and iron, where high iron concentrations can negatively affect zinc absorption or when gamma and alpha tocopherol compete for the same binding protein (Fung et al., 1997; Gutierrez et al., 2009; Meadows et al., 1983; O'Brien et al., 2000; Solomons and Jacob, 1981; Solomons et al., 1983). Third, the composition of the diet may change unavoidably when a macronutrient has been modified. For example, if the amount of total fat in the diet is lowered, the amount of protein or carbohydrate in the diet increases if energy intake is kept constant (Hall et al., 2012). Lastly, the consumption pattern of a nutrient may depend heavily on kilocalorie intake and whether the recommended intake for that nutrient is indexed to kilocalories, as in the case of fiber (i.e., 14 grams per 1,000 calories) (IOM, 2002/2005). Although some interactions are already well-known, such as vitamin D and calcium, the interactions and their effects on health may either involve other nutrients and/or be far more complex.

Rationale for considering interactions between nutrients Nutrient interactions have several implications for the intake response and risk assessment considerations. In the case where the intake requirements for a nutrient may differ depending on another nutrient, the joint effects of those nutrients must be characterized and considered in determining risk for relevant health outcomes. For example, it has been observed that dietary potassium modulates both the pressor and hypercalciuric effects of excessive sodium, and a higher intake of potassium attenuates the adverse effects of sodium on blood pressure, thus suggesting that sodium and potassium guidelines should be considered simultaneously (Crillo et al., 1994; Sellmeyer et al., 2002; Whelton et al., 1997).

In the scenario where nutrient absorption or transport in the circulation depends on the presence of another nutrient, their metabolic relationships must be examined to understand the consequences of any dietary imbalances of the two nutrients and their relationship to chronic disease risk. For example, some prior investigations of vitamin E and chronic disease risk may have overlooked the biological interaction of vitamin E and selenium in the design and interpretation of the data (Lippman et al., 2009).

Similarly, the consequences of interactions must be considered in the context of dietary macronutrient intake. In meal planning and consumption, individuals eat a mix and foods and therefore it is relatively common to encounter interactions between the macronutrients (i.e., fat, protein, and carbohydrates) as well as micronutrients. Consideration of these interactions in intake-response and risk assessment may minimize unintentional consequences of nutritional guidelines for macronutrients. For example, a guideline to lower dietary fat may have an unintentional consequence of leading to an increase in intake of refined carbohydrates if the interactions are not well-described and addressed.

The interactions between nutrients and kilocalories may necessitate the development of guidelines where nutrients are indexed to kilocalories. For example, evidence suggests that dietary fiber from whole foods reduces risk of cardiovascular disease, obesity, and type 2 diabetes (Astrup et al., 1990; Meyer et al., 2000; Rimm et al., 1996; Wolk et al., 1999). Fiber also is important for good digestive health (Aldoori et al., 1994, 1995; Roberts and Veidenheimer, 1990; Watters and Smith, 1990). However, most Americans under-consume dietary fiber and the major sources in the diet are foods that are relatively low in fiber, but are widely consumed, such as white flour and other refined grains (Slavin, 2008). The epidemiological studies that were the basis for the current AI expressed intakes of fiber per 1,000 kilocalories to correct for under-reporting. For this reason, current fiber guidelines explicitly acknowledge the interaction between fiber and calories by indexing fiber to calories (HHS and USDA, 2015). As such,

adherence to the fiber guideline requires consumption of concentrated dietary fiber sources per 1,000 kilocalories of intake.

Considerations when recommending a chronic disease DRI Considerations about nutrient interactions are necessary in every step of the process for recommending chronic disease DRIs. For example, if a nutrient interaction is identified, a key question(s) about nutrient interactions would be included, a systematic review of nutrient interactions would be conducted, the GRADE framework would be applied to the evidence on nutrient interactions, and, if appropriate, a DRI would be established for a chronic disease endpoint that explicitly acknowledges the nutrient interaction and provides appropriate recommendations as needed.

Did Other Factors Contribute to the Chronic Disease?

As noted here and throughout the Options Report, the intake-chronic disease relationship is complicated because intake is rarely the sole risk factor for the outcome of interest. Furthermore, the signal from some nutrients in relation to a particular chronic disease may be strong whereas for others, the signal may be weak but still carry public health importance. Analytic approaches are needed that can reliably parse out the independent, quantitative contribution of nutrient intake and its relationship to the response (e.g., in silico mathematical modeling, systems modeling). DRI committees will need to take into consideration novel analytic procedures that may develop over time.

Examples of Approaches Used in Nutrition

Pooled Studies

Over the past 10 to 15 years, many investigator groups around the globe have worked together on sets of pooled data projects (Key et al., 2010, 2015; Roddam et al., 2008). These are not systematic reviews or meta-analyses. Rather, original data (intake from either self-report or biomarkers and response data) are harmonized and pooled using random effects models. The end result is a larger sample size with a greater number of disease endpoints, thus increasing statistical power. In addition, pooled relative risks can be estimated both for main effects and, importantly for effect modification by age, race/ethnicity, sex, body mass index, and other characteristics where individual studies would not have the statistical power for interaction tests. Other advantages are that less common subtypes of a chronic disease (i.e., triple negative breast cancer) can be examined in a larger number of individuals from harmonized and pooled outcomes

data. This may be particularly important for DRI committees, as intake-response relationships may differ across chronic disease phenotypes. Typically, investigators will present both the individual study estimates as well as the pooled estimates, the former being very useful for examining the overall pattern of the intake-response relationships across multiple populations. Linear models can be used when appropriate. Notably, studies used in pooled analyses can provide heterogeneous data, and appropriate tests for heterogeneity should be performed. Limitations of pooled studies are that measures of both exposures and outcomes often vary across the studies and while data harmonization can be performed, it is often not precise and can add noise to the estimates. These pooled data studies are subject to the same limitations, for example, concerning dietary exposure assessment and confounding, as are the individual studies being pooled.

Systematic Reviews and Meta-Analyses

A second approach to modeling the intake-response relationship for consideration by DRI committees is meta-analyses. A meta-analysis is a systematic review that uses statistical methods to combine quantitatively the results of similar studies in an attempt to allow inferences to be made from the sample of studies and be applied to the population of interest. Meta-analyses began to appear along with systematic reviews more frequently in the 1970s in the medical research field and are now commonly conducted in nutrition research, initially for interventions trials and subsequently for observational studies. Meta-analyses follow very specific data gathering and analysis guidelines, commonly known as PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) (Moher at al., 2009). Tools for assessing systematic reviews are covered in Chapter 6.

Strengths of meta-analyses include their ability to increase the power and precision for chronic disease outcomes (or qualified surrogates), reduce problems of over-interpreting individual studies, and provide a summary of strength and consistency in evidence, which is a key component of the Bradford-Hill criteria for causal inference (Hill, 1965). Other strengths of meta-analyses are that they can help answer questions not posed by any one individual study. In the nutrition area, for example, individual studies will typically explore the effect of limited nutrient intake or exposure levels, making it difficult to estimate the intake-response relationship with a single study. However, when a meta-regression of several individual studies is conducted, other analyses become possible, such as examinations of the quantitative intake response-response relationship between the NOFS and a health outcome. Potential limitations of meta-analyses are that the quality of the results greatly depends on the availability of quality studies with low heterogeneity—both clinical heterogeneity, due to variability in

participants, outcomes, or interventions, or intake response, and methodological heterogeneity, due to variability in methods used, such as blinding. Another potential limitation is bias. Bias must be assessed and reported, and PRISMA provides appropriate guidance. Two general types of models have been used in conducting meta-analyses: the fixed-effects model, which assumes that the sole source of variation is within the study, and the more conservative random-effects model, which assumes that study populations and other factors add to the variability. More recently a mixed model, which addresses both types of effects, has been developed (McCullough et al., 2008).

For constructing intake-response curves for nutrients, integrating observational studies may have more utility than integrating RCTs because observational studies typically cover the range of intakes that the general public typically consumes. However, results from these studies are sometimes not as valuable as they could be because of insufficient information comparisons on exposure levels in the different groups and also potential systematic error (see Chapter 4). Observational studies often report odds ratios (ORs) or RR by quantile categories, and the standard meta-analysis cannot accommodate effect estimates from a single study but only high versus low level types of comparisons. As a result, most meta-analyses provide results as "high-low" meta-analyses (comparisons of extreme categories). A key element in conducting the kind of meta-analysis needed to construct intake-response curves has been the availability of models and statistical packages that are suitable for observational nutrition studies. One of the most popular multivariate analysis methodologies to adjust the correlation between effect estimates from a single study was first published by Greenland and Longnecker (1992) and it has been used in nutrition to compare health effects at different levels of nutrients (Kim and Je, 2016) and foods (Aune et al., 2016). Other methods have been used, such as meta-regression analysis, which allows the modeling of various levels of intake and incorporates the fact that reference values are not zero in nutrition studies (Chung et al., 2016; del Gobbo, 2013; Jiang et al., 2016). A summary of the meta-regression analysis by del Gobbo et al. is in Annex 7-1, as an example of modeling the intake response curves for nutrients. New models and statistical packages continue to be created that could further improve the methodologies in the future.

EVALUATING INTAKE-RESPONSE RELATIONSHIPS: APPLYING PRINCIPLES FROM THE GRADE APPROACH

The certainty of evidence for causality, described in Chapter 6, is necessary, but not sufficient to support development of a DRI. For instance, evidence of an intake-response gradient that can support a causal relation-

ship can be based on categorical intakes (e.g., low, medium, high), whereas evidence for intake-response to support DRI is necessarily quantitative (e.g., mg per day). Moreover, the variability around an intake-response line of best fit may be sufficiently low to conclude that an intake-response relationship exists, but not sufficiently low to justify a specific threshold. Therefore, although in practice questions related to causality and intake-response relationships will be included in the same systematic reviews, a separate evaluation of certainty of evidence is needed for intake-response relationships.

Drawing upon the conceptual and methodologic issues described above and considering the set of options relevant to this chapter described in the Options Report, the committee has developed recommendations (and their rationale) for evaluating the certainty of evidence for intake-response relationships and one guiding principle that is specifically relevant to identify intake-response relationships and set chronic disease DRIs. Although the GRADE working groups have not developed guidance specific to evaluating intake-response relationships, the committee concluded that the underlying principles and process embodied in GRADE are applicable.

The section follows the order of steps and approach to assess the evidence within the GRADE system, and boxes with the options and recommendations of the committee are provided as they apply. However, because the GRADE approach was discussed in detail in Chapter 6, it is not entirely reiterated here. Issues are highlighted below because they are distinct from evaluating causality, such as framing the questions, considering factors that determine certainty in the evidence, and rating the overall certainty in the body of evidence.

Framing the Question

The PICO (population, intervention, comparator, outcome) structure for intake-response has some important differences from the PICO structure for causality. These differences are illustrated graphically in the Venn diagram in Figure 7-2. First, in terms of "intervention" and "comparator," the PICO for characterizing intake-response relationships will likely be more restrictive in terms of requiring quantitative measures of intake. Second, in cases in which a "piecemeal" (two-stage) approach is being considered, the evidence for causality may require only evaluation of a surrogate marker with outcome, whereas for intake-response, a separate PICO statement(s) may be needed to evaluate the relationship between intake and the surrogate marker. Third, in terms of "outcomes," only those with "high" or "moderate" (and in some exceptional circumstances, "low") certainty of a causal relationship would likely be included (see Chapter 6). An additional issue with outcomes involves the use of surrogate markers rather than incidence of disease (see Chapter 5). This issue can be thought of as relating

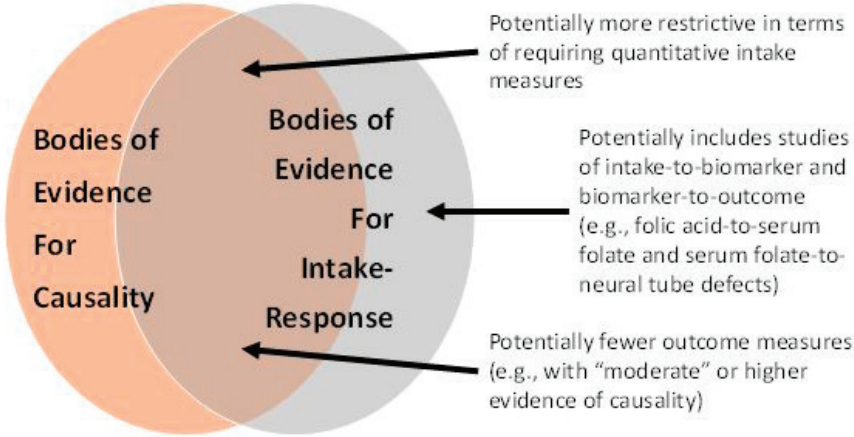


FIGURE 7-2 Venn diagram illustrating overlaps between bodies of evidence for evaluating causality and intake-response.

"indirectness" in the GRADE framework, which is discussed separately below.

Considering Factors That Determine Certainty in the Evidence

The following factors within the GRADE approach determine whether the certainty of the evidence can be upgraded. Application to intake-response relationships involves unique questions to consider that are highlighted in these sections.

Risk of Bias of Individual Studies

Risk of bias is defined as the systematic error due to limitations in the study design or execution, or inappropriate analysis. The features discussed under risk of bias for causality in Chapter 6 are applicable to assessing the evidence for the quantitative relationship between an NOFS and a chronic disease. Additionally, the modeling of the intake-response relationship (choice of mathematical form, method for accounting for covariates/confounders) and method for estimating model parameters and their precision or uncertainty also can lead to bias (see the previous section on "Identifying and Describing Intake-Response Relationships"). Furthermore, measurement error in intake estimates can create bias in the magnitude of effect and

also distort the shape of the intake-response relationship (see Chapter 4). This can create differential bias across the range of intakes. Therefore, methodologic considerations described under "Identifying and Describing Intake-Response Relationships" should be considered when evaluating risk of bias for intake-response. Considerations related to uncertainty are discussed further later under "Imprecision."

Inconsistency of Results

The degree of inconsistency (i.e., the unexplained heterogeneity or variability of study results) in intake-response may differ as a function of intake, among other reasons. For instance, studies may be consistent over one range of intakes but inconsistent over a different range of intake or when assessed in different populations. Because of this added dimension of possible inconsistency, evaluating heterogeneity statistically may be more difficult for intake-response relationships. Circumstances where this might be possible include those when all studies use the same intake-response function (e.g., reporting beta coefficients) or when data from multiple studies can be pooled for re-analysis (e.g., using mixed effects meta-regression) providing that the original measures were assessed with precision and a low degree of bias.

Imprecision

Imprecision is the random error that occurs in studies that results in wide 95 percent CIs around the estimate of the effect. When addressing imprecision in the overall evidence (or other factors for rating down in the context of the GRADE approach), "uncertainty" needs to be addressed. Characterizing uncertainties is ultimately aimed at providing information on the range of possibilities that are consistent with the available data regarding the nature of the intake-response relationship(s). However, as discussed at length in two reports (IOM, 2013; NRC, 2009), while evaluating, assessing, and communicating uncertainties is always necessary, overly complex uncertainty analyses that provide little or no value-added in terms of the ultimate decision should not be pursued.

The precision of the quantitative relationship between nutrient intake and the risk of developing a chronic disease is always limited by the available data. The uncertainties in this relationship are usually disaggregated into two parts: model uncertainty and parameter uncertainty. Model uncertainty refers generally to uncertainty in the mathematical formulation describing the relationship—for instance, a linear model versus an exponential model. Given a specific model's mathematical formulation, parameter uncertainty refers to uncertainty in the terms that determine the quantita-

tive relationship—for instance, the intercept and slope of a linear model. Most statistical model fitting approaches, including both frequentist and Bayesian methods, address parameter uncertainty through CIs on the model parameters (although covariances are not always fully addressed). Model uncertainty is an area of active research, and approaches include model averaging (where different model forms are given weights that depend on the model fit) (e.g., Fang et al., 2016; Liu et al., 1998), model expansion (where different models are combined or embedded into a larger model with additional parameters), and use of non- or semi-parametric models that are less constrained in terms of their model shape (e.g., Guha et al., 2013).

When sufficient data do not exist about an intake-response relationship directly, a piecemeal (two-stage) approach can be used to recommend chronic diseases DRIs, where information about the level of a surrogate marker that is associated with a chronic disease can be supplemented with data on the dietary intake associated with that level of surrogate marker (see Figure 7-3). In addition to the conventional requirements for an accurate surrogate marker and intake methodology, using this method requires the application of models that can integrate diverse data sets and their associated uncertainties. This approach has been used to infer levels of folic acid intake necessary to avoid the risk of neural tube defects in women of childbearing age (Crider et al., 2014; Marchetta et al., 2015). Briefly, Bayesian statistical methods were used to estimate the relationship between intake of food folate and serum red blood cell (RBC) folate (Marchetta et al., 2015), and separate models were used to estimate the association of serum RBC folate with the risk of neural tube defects using various data sets (Crider et al., 2014). Interpreting the findings together, 450 μg DFEs (dietary folate equivalents) per day for women of childbearing age were estimated to result in optimal RBC folate concentration to prevent neural tube defects in children. Investigators are already attempting to study the relationship between folic acid intake and chronic diseases (e.g., see relationship between folic acid intake and plasma folate in Yang et al., 2010, and plasma folate and

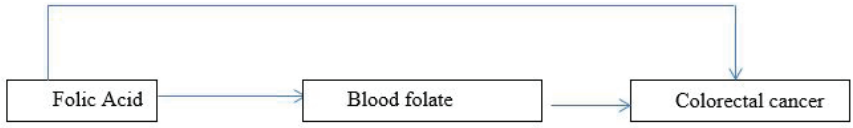


FIGURE 7-3 Example of potential relationships between intake, a surrogate marker, and a disease that would allow indirect determinations of intake-response relationships.

colorectal cancer in Takata et al., 2014, and Neuhouser et al., 2015), and the committee envisions that the piecemeal (two-stage) approach could be pursued in the future for setting chronic disease DRIs. Note, however, that evidence that the relationship between the dietary variable and the disease outcome is completely mediated through the intermediate variable needs to be compelling in using this approach.

However, integration of multiple sources of imprecision can be an additional challenge, particularly if some of them are not quantified. For quantified uncertainties (i.e., with a probability distribution reflecting the likelihood of different values), Monte Carlo simulation is often applied to propagate and combine different types and sources of uncertainties (EPA, 1997). For unquantified uncertainties, sensitivity analyses can be used to understand the impact of different discrete options that reflect a plausible range.

Indirectness of Evidence

Two major sources of indirectness are likely to be at issue for establishing intake-response relationships. The first concerns the use of disease markers of the chronic disease outcome (e.g., surrogate markers); the second concerns the extrapolation of intake-response relationships from studied to unstudied populations. These two issues are discussed in more detail as follows.

Using surrogate markers As explained in Chapter 5, although ideally, DRIs would be determined based on data from chronic disease outcomes, such data on clinical outcomes often are more limited than data on (qualified) surrogate markers. As a result, qualified surrogate markers will necessarily have a significant role in the development of chronic disease DRIs.

The approaches to describing intake-response relationships for qualified surrogate markers and factors such as confounding and inter-individual variability would proceed as outlined earlier in the section titled "Was the Appropriate Model of the Nutrient-Chronic Disease Relationship Selected?" It should be noted, however, that the confounders and covariates for such markers might differ from those for clinical outcomes. These considerations, along with mechanistic studies and analysis (discussed in Chapters 5 and 6), will all contribute to the decision as to whether a down-rating for "indirectness" is warranted. For example, mechanistic data have been used to study the effect of different vitamin A interventions in the kinetics of tissue retinol, a biomarker of vitamin A (Tan et al., 2016). As another example, animal models have been used to study the vascular pathology of Alzheimer's disease leading to hypothesis about potential surrogate markers (Klohs et al., 2014).

With respect to the options put forth in the Options Report (see Box 7-2), the committee notes that the simplest approach of choosing "a single outcome indicator" is likely to be the most feasible (option 1). Many potential complexities and challenges are involved in considering "multiple indicators of a chronic disease" and "multiple indicators for multiple diseases" (option 2), so the committee is reluctant to recommend these options as a general matter, as they require development of multivariate,

BOX 7-2
Selecting Indicators and Specifying Intake-Response Relations: Qualified Surrogate Disease Markers and Nonqualified Disease Markers

Options Report

Option 1: Choose a single outcome indicator on the causal pathway

This option selects a single outcome indicator that is on the causal pathway, provided that it is sufficiently sensitive to quantify the relation between a food substance and a chronic disease.

Option 2: Use multiple indicators of a chronic disease

This option integrates information from multiple indicators of a given chronic disease that add substantially to the accuracy of the intake-response relation and the development of a reference value.

Option 3: Use multiple indicators for multiple diseases

This option may be necessary when a single food substance has different intake-response relations with multiple chronic diseases. In this situation, the DRI committee might need to develop criteria for selecting appropriate disease indicators to establish multiple intake-response relations, methods to integrate multiple endpoints, and approaches to account for the inevitable inter-individual variability in the relations of interest. A committee might develop different reference values for each disease endpoint.

Committee's Recommendation 4

The committee recommends the use of a single outcome indicator on the causal pathway. However, when a single food substance reduces the risk of more than one chronic disease, reference values could be developed for each chronic disease. The committee, however, does not recommend the use of "multiple indicators of a chronic disease" or "multiple indicators for multiple diseases" unless there is sufficient experience with the use of algorithms or other strong evidence suggesting that multiple indicators point to risk of a chronic disease, due to potential lack of reliability or consistency in the results.

multi-pathway, intake-response models. If there is sufficient experience with the use of risk algorithms or other strong evidence to suggest that multiple indicators point to risk of a chronic disease, then a DRI committee could attempt to integrate the findings into its own work.

The Options Report also raises the possibility that a single food substance may reduce the risk of more than one chronic disease (option 3). The Options Report suggests that this approach might be used only when the level of confidence in the evidence for all of the chronic diseases is similar. When such a situation arises, and the available intake-response data are adequate, there is no special difficulty with recommending reference values for each chronic disease, even when the level of confidence in the evidence for all of the chronic diseases is different; DRI committees should adhere to the approaches spelled out in the following sections for each endpoint. However, an overall DRI across multiple chronic diseases would need to consider the risks and benefits of each chronic disease across the range of intakes, which is discussed later in the section on "Balance Between Desirable and Undesirable Outcomes."

Extrapolating intake-response data across populations The committee generally holds that a high degree of caution should guide extrapolation in the intake-response data across populations that differ greatly in important underlying risk factors related to the chronic disease (see Box 7-3). The number of factors that influence chronic disease risk can be large and not well characterized quantitatively, and the likelihood of error is substantial. Unless a significant scientific justification can be made, the committee recommends extrapolation of chronic disease DRIs only to populations similar to studied populations in respect to the underlying factors related to the chronic disease of interest. The evidence supporting any departure from this approach should be fully described and should reveal minimal uncertainty, or otherwise be down-rated for "indirectness."

Publication Bias

As discussed in Chapter 6, publication bias is challenging to assess. Additionally, little or no methodological work has been conducted to evaluate this factor in the context of an intake-response relationship.

Large Magnitude of an Effect

In the GRADE approach, uprating the certainty of the evidence of a causal relationship (e.g., from low to moderate certainty of evidence) when a large effect exists between a nutrient and a chronic disease outcome is typically applied when the body of evidence derives from observational

BOX 7-3
Extrapolation of Intake-Response Data

Options Report

Option 1: Establish reference intake values only for similar populations

This option establishes DRI values on the basis of chronic disease endpoints only for populations that are similar to studied groups. This differs from setting traditional DRI values for essential nutrients for which a value was set for all groups.

Option 2: Allow extrapolation when sufficient evidence is available

This option allows extrapolation when sufficient evidence shows that specific intakes of a food substance can increase or decrease the risk of a chronic disease.

Committee's Recommendation 5

The committee recommends extrapolation of intake-response data for chronic disease Dietary Reference Intakes only to populations that are similar to studied populations in the underlying factors related to the chronic disease of interest.

studies (Guyatt et al., 2011). In the context of an intake-response relationship, the magnitude could also be defined as a "slope" rather than a pairwise comparison between groups. Moreover, for surrogate markers, which tend to be continuous rather than binary outcomes, a "large magnitude" may be expressed in terms of a percentage increase or decrease. These will involve expert judgment on the part of each DRI committee, as "rules of thumb" have not been established for what constitutes a "large magnitude" for either "slopes" or "percentage changes." Alternatively, such continuous outcomes could be re-expressed as a binary outcome in terms of the risk (OR or RR) of being above versus below a particular cut-point (IOM, 2011). As with rating the evidence for causality, committees will apply their judgements about rating in special cases, for example when a large effect size is seen but the evidence has a serious risk of bias or imprecision.

Intake-Response Gradient

Bodies of evidence that are used to evaluate intake-response relationships would necessarily be adequate to rate up for an intake-response gradient. Without such a gradient, it would be very difficult to establish a chronic disease DRI.

Effect of Plausible Residual Confounding

As was discussed in Chapter 6, when assessing causality, it is possible to change the rating of the certainty of the evidence when it is determined that all plausible confounding of the studies would lead to attenuation of the estimate. In the context of intake-response relationships, this also may be the case, though here the "attenuation" would be related to the intake-response gradient. This factor may differ across the range of intakes, which complicates interpretation.

Rating the Overall Certainty in the Body of Evidence for Intake-Response

In the GRADE approach to assess causality, the certainty in the evidence for a particular outcome is generally designated as the highest level among the bodies of evidence for that outcome (e.g., if both RCTs and observational studies examined blood pressure, then the body of evidence with the highest certainty should be used) (Guyatt et al., 2013). However, this certainty may differ across outcomes (e.g., blood pressure versus stroke), and if several outcomes are deemed "critical," then the overall evidence would correspond to the outcome with the lowest certainty. In the case of the "piecemeal" (two-stage) approach—when intake-to-biomarker and biomarker-to-clinical outcome relationships are combined—the overall certainty of evidence from intake-to-clinical outcome would correspond to the lowest certainty among the two steps (i.e., weakest link in the chain). Moreover, rating the certainty in intake-response relationships has an additional dimension in that the level of certainty may differ across the range of intakes due to different reasons. For example, the precision of the intake-response estimate might differ across the range of intakes or by differing population characteristics. Differential bias across the range of intakes also can create differences in certainty across intakes. Thus, for each outcome, ranges of intake where the overall certainty differs should be described.

USING GRADE TO MOVE FROM EVIDENCE FOR INTAKE-RESPONSE RELATIONSHIP TO DRI RECOMMENDATIONS

As described in Chapter 6, GRADE provides a framework for moving from evidence to decisions. Existing GRADE guidance is focused on either "binary" (e.g., intervene or not) or a small set of "discrete" recommendations (e.g., use intervention A, B, or C). Specifically, for each recommendation, a number of factors are considered in grading the strength of recommendation (Andrews et al., 2013a,b).

By contrast, recommending a chronic disease DRI involves a continuum of options related to specifying a number or range; the certainty in the evi-

TABLE 7-1 Possible DRIs for Chronic Disease

Possible DRI for Single Chronic Disease	Description	Region of Intake-Response Relationship[a]	Comments
Acceptable Range of Intakes (ARI)	Range of usual intakes of a food substance without increased risk of chronic disease.	Region where slope is flat, outside of which there is increased risk of chronic disease, deficiency, or toxicity.	See green shaded region in Figure 7-4. Analogous to AMDR for macronutrients. Implies the intake should ideally be in this range.
Range of Beneficial Increased Intakes (RBII) [Alt: Range where Increased Intake is Beneficial (RIIB)]	Range of usual intakes of a food substance where increasing intake can reduce risk of chronic disease.	Region where slope is negative, outside of which slope is non-negative, or there is increased risk of deficiency or toxicity.	See orange shaded region in Figure 7-4.
Range of Beneficial Decreased Intakes (RBDI) [Alt: Range where Decreased Intake is Beneficial (RDIB)]	Range of usual intakes of a food substance where decreasing intake can reduce risk of chronic disease.	Region where slope is positive, outside of which slope is non-negative, or there is increased risk of deficiency or toxicity.	See yellow shaded region in Figure 7-4.

NOTE: AMDR = Acceptable Macronutrient Distribution Range.

[a] In each case, defining the region of the intake-response relationship corresponding to the DRI requires judgment required as to what "slope" is small or large enough, and at what confidence level, to consider flat, negative, or positive.

dence might not only vary by nutrient and by disease outcome but by the different nutrient levels. Additionally, it may not become clear what kind of DRI (see possible types of chronic disease DRIs in Table 7-1) would be supported until after the intake-response relationship is evaluated. Therefore, selecting chronic diseases DRIs involves additional decisions related to the type of DRIs, acceptable level of confidence in the intake-response data, and balancing health risks and benefits. The factors considered in rating the certainty of evidence delineated by GRADE are still appropriate, but the committee suggests that they be used more iteratively to determine the most appropriate DRI.

Options for Specifying a DRI for Chronic Disease

The Options Report raises two key issues related to specifying a DRI for chronic disease. The first is a general question as to what types of refer-

ence values (i.e., DRIs) might be established (see Box 7-4). The second is a more focused question as to whether a UL can be established based on increased risks of chronic disease (see Box 7-5). The Options Report contains a lengthy discussion of reference values that would, if implemented, serve to decrease chronic disease risks (i.e., provide health benefits), and offers several options for establishing such DRIs. In general, DRIs might be established at an intake associated with a specified degree of risk reduction, or be described as a range of beneficial intakes.

The Options Report also considers the use of ULs in past DRI development efforts. The use of a UL to deal with situations in which a food substance has been shown to increase chronic disease risks is offered as an option.

Reference Values Associated with Benefit

This section addresses situations in which a substance has been shown to decrease chronic disease risk. The following section addresses substances that have been shown to increase chronic disease risk. The Options Report introduces the difficult problem of interactions among food substances that affect chronic disease risk. Approaches to address these questions are in the previous section on "Was the Appropriate Modeling of the NOFS-Chronic Disease Relationship Selected?" This section will first comment on the reference value for the simplest situation where one substance is associated with one chronic disease (or marker of chronic disease). Although the possible value of using absolute risk measures, described earlier, should not be ignored, the committee's views on this matter rest for the time being on intake-response relationships in which RR, OR, or HR is the risk measured.

As mentioned in the introduction to this chapter, the committee does not recommend the use of a "family" of reference values due to the lack of experience with developing or implementing DRIs for chronic diseases as well as significant difficulties in describing and communicating uncertainties. Therefore, the committee focuses on the simplest case, one NOFS that is associated with one or various diseases. The committee emphasizes that selecting a DRI for a chronic disease(s), as either a point estimate or as a range of beneficial values, depends upon the availability of intake-response relationships that have been generated from studied populations and then extrapolated to unstudied populations (see in the previous section on "Identifying and Describing Intake-Response Relationships" and Box 7-3). Only when intake-response relationships have been identified with acceptable levels of confidence (as this concept has been described in Box 7-6) can DRIs be recommended.

Once intake-response relationships have been identified with acceptable levels of confidence, the truly difficult tasks involve characterizing

BOX 7-4
Different Types of Reference Values: Types of Reference Values Associated with Benefit

Options Report

Option 1: Establish chronic disease risk-reduction intake values (e.g., CD_{CVD})

DRI committees could modify the traditional Estimated Average Requirement (EAR)/Recommended Daily Allowance (RDA) approach to estimate the mean intakes of individuals and the inter-individual variability associated with specified disease risk reductions. This option is conceptually very similar to the traditional EAR/RDA approach, but the definitions and interpretations of reference values based on chronic disease endpoints are different from those based on classical deficiency endpoints. This option uses relative risks and requires knowledge of baseline disease prevalence, whereas the traditional approach is based on absolute risks and is independent of baseline prevalence. The mean intake values and associated variances for given magnitudes of risk reduction give valuable information on the "typical" person and population variability. These values could, therefore, be useful for assessing population and group prevalence. Several adaptations of this option are possible, depending on the nature of the available data.

Option 2: Identify ranges of beneficial intakes

In some cases, available data might be adequate only for deriving an intake range that can reduce the relative risk of a chronic disease to a specified extent. One end of this intake range is close to the point at which risk begins to decline or increase, depending on the relation, and the other end extends as far as the available evidence permits. The DRI committee could establish the range so that it does not increase the risk of adverse health effects.

Committee's Recommendation 6

The committee recommends that Dietary Reference Intakes (DRIs) for chronic disease risk take the form of a range, rather than a single number. Intake-response relationships should be defined as different ranges of the intake-response relationship where risk is at minimum, is decreasing, and/or is increasing (i.e., slope = 0, negative, or positive). When a nutrient or other food substance reduces the risk of more than one chronic disease, DRIs could be developed for each chronic disease, even if the confidence levels for each chronic disease are different. The magnitude of risk slope considered necessary to set a DRI should be decided based on clearly articulated public health goals, such as those previously identified by other authorities (e.g., Healthy People 2020). The committee does not recommend, however, developing a family of DRIs for any one NOFS for different risk reduction targets for the same chronic disease.

BOX 7-5
Different Types of Reference Values:
ULs and Reduction in Chronic Disease Risk

Options Report

Option 1: Base ULs on the traditional threshold model

This option continues to base ULs on the traditional threshold model when UL values based on chronic disease endpoints are higher than those based on traditional adverse effects.

Option 2: Base UL_{CD} on intakes associated with chronic disease risk

When the risk of a chronic disease increases at an intake below the traditional or current UL, a DRI committee could base a UL on chronic disease endpoints by using approaches analogous to the derivation of chronic disease values (e.g., the development of 1 or multiple values for specified levels of relative risk reduction) or a threshold approach (e.g., identifying the inflection point at which absolute or relative risk increases). These values could be denoted as a chronic disease UL (UL_{CD}) to distinguish them from a traditional UL. The UL_{CD} would be set at a level below which lower intakes are unlikely to achieve additional risk reduction for a specified disease. The traditional UL definition would have to be expanded to include intakes associated with changes in relative risk (in contrast to absolute risk) of an adverse effect. Because the UL_{CD} is based on changes in the relative risk of the chronic disease, intakes below the UL_{CD} might reduce but not necessarily eliminate disease risk, reflecting the multifactorial nature of chronic diseases.

Committee's Recommendation 7

The committee recommends retaining Tolerable Upper Intake Levels (ULs) based on traditional toxicity endpoints. In addition, if increased intake of a substance has been shown to increase the risk of a chronic disease, such a relationship should be characterized as the range where a decreased intake is beneficial. If the increase in risk only occurs at intakes greater than the traditional UL, no chronic disease Dietary Reference Intake would be required, because avoiding intakes greater than the UL will avoid the chronic disease risk.

the relevant intake-response relationships—in terms of their shape and the range of intakes over which they apply. In the simplest case, when the relationship appears linear, this characterization could include the slope of the relationship (amount of change in risk for a given change in intake), the range over which this relationship is supported, and the CIs for each of these (see Figure 7-4, panels A and A'). For more complicated intake-response relationships, such as sigmoidal, U-, or J-shaped curves (includ-

BOX 7-6
Acceptable Level of Confidence in the Intake-Response Data

Options Report

Option 1: Require a high confidence level

This option requires a high level of confidence by, for example, using RCTs with a chronic disease event or a qualified surrogate disease marker as the outcome measure.

Option 2: Accept a moderate confidence level

This option accepts a moderate level of confidence in the data for decisions about intake-response relations.

Option 3: Piecemeal approach

This option pieces together different relations in which the biomarker of interest is a common factor when direct evidence of the biomarker's presence on the causal pathway between the food substance and a chronic disease is lacking.

Committee's Recommendation 8

The committee recommends that to develop a chronic disease Dietary Reference Intake, the level of certainty in the intake-response relationship should generally be the same as the level of certainty for a determination of causality, that is, at least "moderate," using the Grading of Recommendations, Assessment, Development and Evaluation (GRADE). However, in some cases, for example when a food substance increases chronic disease risk, the level of certainty considered acceptable might be lower. In all cases, a thorough description of the scientific uncertainties is essential in describing quantitative intake-response relationships. Requiring at least "moderate" certainty extends to cases where relationships between intake and a surrogate marker and between the same surrogate marker and the chronic disease are characterized separately, in a piecemeal (i.e., two-stage) approach.

ing inverted versions), this characterization could include the range over which the slope is positive, the range over which the slope is negative, the point (if any) at which the slope is zero, any inflection points (maximum or minimum slope), and each respective CI (see Figure 7-4, panels B, B', B", B"', and C). These examples encompass the Options Report suggestions of a point estimate at which maximum risk reduction or a range of beneficial intakes is achieved, but additional information could be valuable for derivation and application of DRIs.

Several issues complicate translation of intake-response relationships into a DRI:

- The relationship between intake and a chronic disease is often continuous over a range of intakes, with incremental changes in intake resulting in incremental changes in risk. In such a case, a DRI consisting of a single intake level could only be specified if there is a point of minimum risk with a high degree of certainty.
- The multifactorial nature of chronic disease implies that individuals across the population would have different "baseline" risks, even if their food or nutrient intakes were comparable. Therefore, setting a DRI based on an absolute level of "acceptable" risk would be complicated because the same level of intake may represent a range of different risks across individuals due to factors other than dietary intake, such as family history, adiposity, smoking, genetic characteristics, physical activity, and other characteristics.
- Setting a DRI for chronic disease also may need to consider potential harm from deficiency at the lower end of intake and toxicity at the upper end of intake. Risks and benefits related to multiple chronic diseases introduce further complexity (see the later section on "Balance Between Desirable and Undesirable Consequences of a DRI").

Based on these issues, the committee envisions that DRIs for chronic disease risk should take the form of a *range* (option 2, Box 7-4), rather than a single number that estimates mean intakes of individuals and their inter-individual variability (option 1) because of the multifactorial nature of chronic diseases and range of different risks across individuals. Suggestions for how to define such ranges are illustrated in Figure 7-4 and described in Table 7-1. Conceptually, such ranges correspond to regions of the intake-response relationship where risk is at minimum, is decreasing, or is increasing (i.e., slope = 0, negative, or positive). The mathematical approach to such a determination has been discussed under "Identifying and Describing Intake-Response Relationships." In making decisions about which ranges to establish as DRIs, the magnitude of the risk slope considered necessary to support a DRI recommendation must be considered.

Both reductions and increases in intake may sometimes have unintended consequences (e.g., changes in dietary patterns that alter the intakes of essential nutrients in possibly harmful ways). For this reason, efforts to quantify the risk reductions achieved as a result of specified levels of intake reduction would be valuable. Describing them as absolute risk would be even more valuable. The committee recommends that steps be taken to describe RR (or OR or HR) reductions as a function of intake reduction

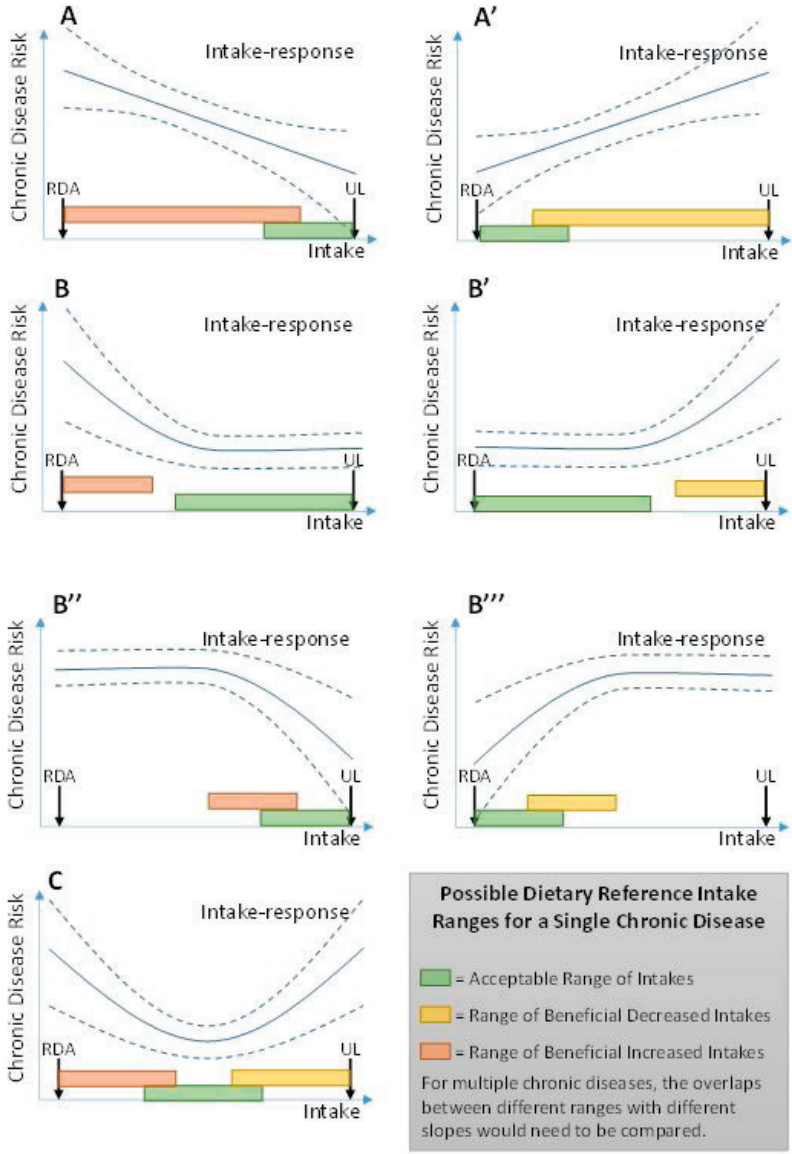

A
Intake-response
Chronic Disease Risk
RDA
UL
Intake
A'
Intake-response
Chronic Disease Risk
RDA
UL
Intake
B
Intake-response
Chronic Disease Risk
RDA
UL
Intake
B'
Intake-response
Chronic Disease Risk
RDA
UL
Intake
B''
Intake-response
Chronic Disease Risk
RDA
UL
Intake
B'''
Intake-response
Chronic Disease Risk
RDA
UL
Intake
C
Intake-response
Chronic Disease Risk
RDA
UL
Intake
Possible Dietary Reference Intake Ranges for a Single Chronic Disease
= Acceptable Range of Intakes
= Range of Beneficial Decreased Intakes
= Range of Beneficial Increased Intakes
For multiple chronic diseases, the overlaps between different ranges with different slopes would need to be compared.

FIGURE 7-4

(i.e., the slope), to be accompanied by efforts to communicate the public health benefits associated with such reductions. Furthermore, efforts should be made to convert RR to absolute risks, which will give users of these reference values increased understanding of health benefits achieved with a given degree of intake reduction.

ULs and Reduction of Chronic Disease Risk

ULs were introduced in 1997 as a new type of DRI (see Figure 4 in the Options Report, Appendix B) (IOM, 1997). The UL is not a recommended level of intake, but rather a level that, if exceeded, may cause adverse health effects or toxicities. Existing ULs for nutrients are based on various types of studies, including animal toxicity studies, in which adverse health effects have been observed. These effects are generally not chronic diseases, as defined in this report. The methodology used to establish ULs is similar to that used in the safety assessment of chemical toxicity. It is based on the identification of no-observed adverse effect levels (NOAELs) from either observational studies in humans or from animal studies. Various uncertainty factors (largely a matter of expert judgment) are introduced to allow extrapolation from studied populations to the general population, to account for inter-individual variation, and to deal with limitations in the available data. The lack of consistency in defining uncertainty factors might relate to the lack of data available for high-dose nutrient toxicity or to a need to avoid establishing any UL at intakes less than an EAR or an RDA

FIGURE 7-4 Possible DRI ranges for a single chronic disease, depending on the shape of the intake-response relationship. These relationships, and their confidence intervals, are "idealized" and meant for illustration, and are likely to be more complicated (e.g., less smoothly changing) in practice. The different scenarios are qualitatively the same whether absolute or relative risk is considered. However, to estimate the significance of the impact on the population of the different choices of ranges, absolute risks are also needed. Panels A and A' represent strictly monotonically changing intake-response relationships; Panels B, B', B", and B"' represent different "J-shaped" relationships, where there is a plateau at one end of the intake range. Panel C represents a "U-shaped" relationship, where there is an intake level that minimizes risk.

NOTES: RDA = Recommended Daily Allowance; UL = Tolerable Upper Intake Level; solid line = best estimate of intake-response; dashed lines = confidence intervals of intake-response. See Table 7-1 for descriptions of the possible DRI ranges. In contrast with an EAR, an ARI would not allow calculation of a prevalence of "inadequacy" as is possible with EAR. Users could calculate, however, the prevalence of intakes below (or above) the ARI range.

(see Figure 4 in the Options Report, Appendix B). See the later section on "Balance Between Desirable and Undesirable Consequences of a DRI."

The traditional approach to developing ULs assumes the existence of a threshold in the observed intake-response curve for toxicity (approximated by the NOAEL), which is said to apply, after the inclusion of uncertainty factors, to the general population. Absolute risk of toxicity is said to increase in some manner, not described quantitatively, as intakes increase above the UL. Although ULs are currently applied as if they represent high-certainty cutpoints between safe and unsafe intakes, the current risk model allows no statement to be made about the magnitude of any risk incurred at intakes greater or less than the UL.

Irrespective of the approach taken to develop DRIs for NOFSs that are demonstrated to increase the risks of chronic diseases, it is important to retain ULs based on traditional toxicity endpoints. Intakes that exceed ULs may well cause harm, and should not be recommended, without significant justification, such as in a hypothetical case where an NOFS has been demonstrated to reduce a chronic disease risk at intakes greater than a UL, and the benefit (perhaps for certain subpopulations) greatly exceeds any risk of toxicity.

The committee supports a variant of options 1 and 2 described in the Options Report (see Box 7-5) and notes that traditional ULs should be retained whether a putative chronic disease DRI is below or above a traditional UL value for that substance. If increased intake of a substance has been shown to increase the risk of a chronic disease, the intake-response relationship should be characterized as to whether the range over which increased risk occurs overlaps with the traditional UL. If the increase in risk occurs only at intakes greater than the traditional UL, then no chronic disease DRI would be required, because avoiding intakes greater than the UL will also avoid increases in chronic disease risk. Therefore, the committee also recommends retaining the traditional UL when DRI values based on chronic disease endpoints are higher than those based on traditional adverse effects (option 1). In cases in which increased intake is associated with increased chronic disease risks at intakes less than the traditional UL (see Figure 7-4 A', B', and C and Figure 5b in the Options Report, Appendix B), both the traditional UL and a chronic disease DRI could be retained (a variation of option 2). The rational for retaining both would be that the UL connotes a "bright line" intake limit that should not be exceeded (although the actual definition[1] acknowledges the possibility of "residual" risk). On the other hand, the DRI for chronic disease in this case would imply that

[1] "The highest level of daily nutrient intake that is *likely* to pose no risk of adverse health effects to almost *all individuals* in the general population [emphasis added] (IOM, 1997, p.4)." See Annex 7-2 for additional discussion of this point.

chronic disease risks will be increased with increasing exposure over the DRI range (see Figure 5b in the Options Report, Appendix B). Imposing a "bright line" (as implied by calling it a UL) in the case of a chronic disease could be misleading and suggest to some that crossing the bright line will lead directly to disease. Such a suggestion is scientifically incorrect, because for chronic diseases, it is the likelihood of developing the disease that is modified by intake. Even if a bright line for a chronic disease could be justified based on a specific "acceptable" level of increased risk, this would not be consistent with how the term "UL" is defined, which implies virtually no risk if the intake is not exceeded.

Possible New Approaches to UL Development

The committee was not asked to offer opinions on the data and methods used to derive ULs, but notes that significant developments have occurred in these areas in the world of chemical risk assessment. Relevant information is included in an Annex 7-2 to this chapter, titled "Possible New Approaches to UL Development." Recognizing the challenge described in the Options Report—that is, the lack of data and implications for setting chronic disease DRIs higher that the traditional UL—the committee notes the need for better information on the adverse health effects of high intake levels for many NOFSs that may become the subjects of DRIs.

Certainty in Evidence for Intake-Response Relationship

As already mentioned, Chapter 6 describes the use of GRADE in evaluating the existence of a causal relationship between intake of an NOFS and risk of a chronic disease. GRADE also has been applied for assessing intake-response relationships in this chapter. Collecting the body of evidence for evaluating the intake-response relationships would presumably need its own separate PICO (see Figure 7-2). Even when studies are used to evaluate both causality and intake-response relationships, the application of the criteria (e.g., risk of bias) may be different (e.g., a bias may be "not serious" for causality but "serious" for intake-response). Therefore, although the evidence for causality will be presented to the DRI in the form of evidence tables, the evidence for intake-response relationships may be best discussed in a narrative manner. In addition, although the factors considered in rating the certainty of evidence delineated by GRADE are still appropriate, they may be used more iteratively to determine the most appropriate DRI, as mentioned above. Also, as with evaluating causality, DRI committees will apply their judgment as to the number of studies or participants necessary for determining that a quantitative relationship can be established.

The certainty in the evidence is a key factor in moving from evidence

to recommendations in the GRADE framework (Andrews et al., 2013a,b). Box 7-6 shows the three potential approaches in the Options Report for accepting the necessary confidence level in the intake-response relationships for developing DRIs. Any decision regarding the acceptable level of confidence in the underlying scientific data necessary to support DRI development is at least in part a policy decision because it implies selecting a risk reduction goal. With this recognition in mind, the committee offers the following comments. As described previously, moving from the evidence on causation to evidence of quantitative intake-response relationships may introduce additional uncertainties.[2] Particularly, reliable and accurate intake data (see Chapter 4) are required to develop useful intake-response relationships. Reliance on RCTs alone (option 1) to establish causal relationships may be problematic for developing intake-response relationships because RCTs often involve only a single intervention dose. Under such circumstances, DRI development may require use of both observational and experimental data, perhaps in combination.

The committee concludes that in describing quantitative intake-response relationships, a thorough description of the scientific uncertainties associated with them is essential, as discussed previously under "Imprecision." Once uncertainties in the analyses are described, decisions can be made regarding the level of uncertainty that is tolerable in specific cases.

Although in general there should be at least "moderate" certainty in range(s) selected as chronic disease DRIs (option 2), the level of uncertainty considered tolerable in DRI decisions may be different in situations in which intake of an NOFS increases chronic disease risk than in those in which risk is decreased. Opinions on whether it is more appropriate to be cautious (accept greater uncertainty) when disease risk is increased than when it is decreased may differ, and, therefore the committee will not offer a specific opinion on this question. What is essential is that in making DRI recommendations, the scientific uncertainties associated with the recommendation and the reasoning behind their acceptance must be made completely transparent. Recognizing the challenges in integrating of multiple sources of imprecision (see the previous section "Considering Factors That Determine Certainty in the Evidence"), a piecemeal (two-stage) approach is potentially feasible (option 3).

Finally, as already mentioned and depicted in Figure 7-4, relationships between nutrient intake and disease outcomes are not always linear and can be J-shaped, U-shaped, or inverted U-shaped. For example, the fact that very low levels of intake may lead to insufficiency and very high levels of intake may lead to toxicities is typically shown in an inverted U curve.

[2] The uncertainties referred to here do not include those introduced when extrapolating from study populations to other populations.

Regarding evidence related to a specific nutrient and non-linear relationships, DRI committees will need to carefully evaluate the data presented in individual studies and in the systematic reviews. In some cases, non-linear relationships will be biologically plausible, and well-conducted studies that meet GRADE criteria for at least moderate evidence will support the shape of the relationship between the nutrient and the disease or the surrogate endpoint. In other studies, however, the non-linear relationship will seem paradoxical to the biological understanding of both the nutrient and the disease. Still other cases may have limitations in study designs, and the strength of the evidence using GRADE (see especially Chapter 6) could limit the plausibility of potential non-linear relationships. DRI committees will need to use the totality of the evidence available to them to critically evaluate these issues.

Balance Between Desirable and Undesirable Consequences of a DRI

The consequences of a particular choice of DRI need to be evaluated in terms of potential for both desirable and undesirable consequences, consistent with the GRADE evidence to recommendations framework. It is essential that the balance be considered on the basis of absolute risk, as the RR for different endpoints may be very different in terms of actual number of cases. Several issues specific to DRIs for chronic disease need to be considered, including overlaps between benefits and harm.

Overlaps Between Benefits and Harm

Deficiency, toxicity, and multiple chronic diseases need to be considered when balancing benefits and harms. Several scenarios can be anticipated. The simplest would be when chronic disease risk increases with intake, and it would be possible, albeit remote, that recommendations to reduce intake will result in intakes that result in deficiency. A chronic disease DRI, therefore, should never go below the RDA.

In situations where one might reach toxicity levels, the solutions may be more complicated. Even what seems to be a relatively simple choice—avoiding proposing any DRI range that exceeds the UL—becomes unclear when a situation similar to that depicted in Figure 7-4, panels A and B", occurs for chronic disease. Note that the risk of disease would continue to decline at intakes well above the UL. Thus, cutting off the range of beneficial intakes at the UL would result in the loss of significant benefits, perhaps for one subpopulation; if the risk of toxicity at and above the UL is poorly documented or not well described (as is sometimes the case), it would seem counterproductive not to extend the DRI (range of beneficial intakes) to levels greater than the UL (assuming the risk reductions achieved in that range

are adequately documented and have minimal bias). Of course, if serious toxicity at levels greater than the UL can be adequately documented, cutting off the range of beneficial intakes at the UL would probably be necessary. As mentioned previously and described in Annex 7-2, this situation can be ameliorated by using more sophisticated methods for deriving the UL that provide an intake-response relationship for toxicity rather than a single "bright line" point estimate.

A similar situation exists when available evidence relates to more than one chronic disease, but the slopes differ for the intake-response relationships. In cases in which a substance has been shown to increase the risk of one chronic disease and to decrease the risk of another, the first step should be the identification of those sections of the two ranges that do not overlap (taking into account the variance and other uncertainties in these sections). These two "sections" could conceivably become DRIs, one for benefit (disease A) and one for risk (disease B). The UL would, as a first step, become another limit on the ranges.

Among the three options in the options report (see Box 7-7), option 1 does not provide a satisfactory response because cases may occur where *avoid an overlap between beneficial intakes and intakes associated with adverse events* is simply not possible. Option 2 requires committees to determine a minimum level of severity and risk reduction targets, which is generally considered to be a policy decision. Although models for dealing with competing health risks and benefits have been developed (see Box 7-8), which attempt to achieve a type of "balancing" of the two by assigning various factors to account for differing severities and other disease characteristics (including costs of treatment) in order to develop some common measure of impact, conducting such an analysis might be beyond what is expected of a DRI committee. For example, even if sources of information exist (e.g., primary studies examining population values and preferences or numerical population impact, focus groups organized by the guideline producers, or other) ranking relative importance of health outcomes may be challenging. In addition, other factors included in certain models (e.g., resources) are within the purview of policy decisions.

If the above cannot be achieved, then the best approach is simply descriptive (option 3), that is, description of the diseases and their severities, the magnitudes of risk increases and decreases over various ranges of intakes, and other factors that would allow maximum utility and flexibility for users of the information. In making their conclusions, however, DRI committees should explicitly specify the certainty in the evidence used to develop the DRI and the populations and other circumstances to which it applies. In the case of chronic disease DRIs, committees will likely find that certainty of evidence for intake-response relationships and the balance between desirable and undesirable outcomes are the most influential fac-

BOX 7-7
Overlaps Between Benefits and Harms

Options Report

Option 1: Avoid overlap between beneficial intakes and intakes associated with adverse events

This option ensures that no point estimate or range of beneficial intakes for chronic disease risk reduction extends beyond the intake at which the risk of adverse events, including chronic diseases, increases.

Option 2: Establish criteria related to severity and risk of chronic disease

This option establishes criteria for ULs on the basis of the minimum level of severity and prevalence of targeted chronic diseases and the degree of risk reduction associated with specified intakes. The DRI committee would apply analogous information on the nature of candidate adverse outcomes when establishing ULs.

Option 3: Describe the nature of the evidence

This option describes the nature of the evidence (e.g., type of evidence, quality, strength) and the public health implications of benefits and risks for the full range of intakes for which inferences are reasonably possible, along with remaining uncertainties. Ultimately, users would choose an appropriate balance between benefits and harms for their population of interest.

Committee's Recommendation 9

The committee recommends that, if possible, health risk/benefit analyses be conducted and the method to characterize and decide on the balance be made explicit and transparent. Such a decision needs to consider the certainty of evidence for harms and benefits of changing intake and be based on clearly articulated public health goals. If Dietary Reference Intake committees do not perform such risk/benefit analyses, it is still necessary to describe the disease outcomes and their severities, the magnitudes of risk increases and decreases over various ranges of intakes, and other factors that would allow users to make informed decisions.

tors and transparency in the thinking process and rationale is required. For example, in addition to differences in populations, DRI committees should consider the possibility that certainty in the evidence may differ depending on the ranges of intake. Guiding Principle 14 was developed to emphasize the importance of describing all uncertainties when characterizing intake-response relationships (see Box 7-9).

BOX 7-8
Approaches to Characterize the Balance Between Risk and Benefits

The following is a list of some techniques that have been applied to balance risks and benefits, and while they often do not provide complete answers, they can potentially provide useful information on summary health choices for population interventions, or at least to weigh the various outcomes discovered in evidence review:

- Full analysis of RCTs that have carefully monitored all available positive and adverse health outcomes. A main problem is that many secondary outcomes are uncommon and they may have inadequate statistical power for useful evaluation.
- Apply cost-benefit and cost-utility approaches to the various positive and negative outcomes to weigh their relative monetary costs if credible attributions can be made (IOM, 2006; Russell, 2015).
- Use health utility weight assignments (e.g., disability-adjusted life years [DALYs]) to assess individual and population health-state preferences (Bansback et al., 2014; Mulhern et al., 2016). Basically, this process allows relevant community populations, patients, and other stakeholders to assign the importance and priorities of diverse negative and positive outcomes based on their personal views and clinical experience.
- Use decision-analytic techniques (Elkin et al., 2006; Vickers and Elkin, 2006) to predict the net effects (endpoints) of disparate outcomes in available clinical trials literature. Computational modeling of the interventions and their outcomes can assist in decision making.
- Use methods for public health-relevant studies that assess the net benefit of trial outcomes (Sawaya et al., 2007). This technique has been used by the U.S. Preventive Services Task Force for evidence review and guideline promulgation.

BOX 7-9
Guiding Principle Related to Describing Uncertainties

14. Where they exist, quantitative intake-response relationships should include a thorough description of the scientific uncertainties associated with them.

POLICY CONSIDERATIONS AND IMPLEMENTATION OF DRI RECOMMENDATIONS

Under the GRADE framework, all recommendations are accompanied by a strength of recommendation classification of "strong" or "weak" based on consideration of four factors established in earlier GRADE writings (magnitude of effects, uncertainty, values and preferences, cost) (Guyatt et al., 2013). Subsequently, GRADE has presented evidence to decision frameworks that include equity, feasibility, and acceptability. However, DRI committees' charges relate only to health outcomes, and, as explained in Chapter 1, policy considerations, and values and preferences related to quality of life (e.g., taste, convenience), equity considerations, and considerations of cost and resources that might ensue from chronic disease DRIs are outside of the scope of DRI committees. Instead, such considerations will be taken into account in the subsequent development of nutrition policy.

REFERENCES

Aldoori, W. H., E. L. Giovannucci, E. B. Rimm, A. L. Wing, D. V. Trichopoulos, and W. C. Willett. 1994. A prospective study of diet and the risk of symptomatic diverticular disease in men. *Am J Clin Nutr* 60(5):757-764.

Aldoori, W. H., E. L. Giovannucci, E. B. Rimm, A. Ascherio, M. J. Stampfer, G. A. Colditz, A. L. Wing, D. V. Trichopoulos, and W. C. Willett. 1995. Prospective study of physical activity and the risk of symptomatic diverticular disease in men. *Gut* 36(2):276-282.

Andrews, J., G. Guyatt, A. D. Oxman, P. Alderson, P. Dahm, Y. Falck-Ytter, M. Nasser, J. Meerpohl, P. N. Post, R. Kunz, J. Brozek, G. Vist, D. Rind, E. A. Akl, and H. J. Schunemann. 2013a. GRADE guidelines: 14. Going from evidence to recommendations: The significance and presentation of recommendations. *J Clin Epidemiol* 66(7):719-725.

Andrews, J. C., H. J. Schunemann, A. D. Oxman, K. Pottie, J. J. Meerpohl, P. A. Coello, D. Rind, V. M. Montori, J. P. Brito, S. Norris, M. Elbarbary, P. Post, M. Nasser, V. Shukla, R. Jaeschke, J. Brozek, B. Djulbegovic, and G. Guyatt. 2013b. GRADE guidelines: 15. Going from evidence to recommendation-determinants of a recommendation's direction and strength. *J Clin Epidemiol* 66(7):726-735.

Astrup, A., E. Vrist, and F. Quaade. 1990. Dietary fibre added to very low calorie diet reduces hunger and alleviates constipation. *Int J Obes* 14(2):105-112.

Aune, D., N. Keum, E. Giovannucci, L. T. Fadnes, P. Boffetta, D. C. Greenwood, S. Tonstad, L. J. Vatten, E. Riboli, and T. Norat. 2016. Whole grain consumption and risk of cardiovascular disease, cancer, and all cause and cause specific mortality: Systematic review and dose-response meta-analysis of prospective studies. *BMJ* 353:i2716.

Bansback, N., A. R. Hole, B. Mulhern, and A. Tsuchiya. 2014. Testing a discrete choice experiment including duration to value health states for large descriptive systems: addressing design and sampling issues. *Soc Sci Med* 114:38-48.

Chung, M., A. M. Tang, Z. Fu, D. D. Wang, and S. J. Newberry. 2016. Calcium intake and cardiovascular disease risk: An updated systematic review and meta-analysis. *Ann Intern Med* 165(12):856-866.

Cirillo, M., M. Laurenzi, W. Panarelli, and J. Stamler. 1994. Urinary sodium to potassium ratio and urinary stone disease. The Gubbio Population Study Research Group. *Kidney Int* 46(4):1133-1139.

Crider, K. S., O. Devine, L. Hao, N. F. Dowling, S. Li, A. M. Molloy, Z. Li, J. Zhu, and R. J. Berry. 2014. Population red blood cell folate concentrations for prevention of neural tube defects: Bayesian model. *BMJ* 349:g4554.

Del Gobbo, L. C., F. Imamura, J. H. Wu, M. C. de Oliveira Otto, S. E. Chiuve, and D. Mozaffarian. 2013. Circulating and dietary magnesium and risk of cardiovascular disease: A systematic review and meta-analysis of prospective studies. *Am J Clin Nutr* 98(1):160-173.

Elkin, E. B., A. J. Vickers, and M. W. Kattan. 2006. Primer: Using decision analysis to improve clinical decision making in urology. *Nat Clin Pract Urol* 3(8):439-448.

EPA (U.S. Environmental Protection Agency). 1997. *Guiding principles for Monte Carlo analysis*. EPA/630/R-97/001. Washington, DC: Risk Assessment Forum, U.S. Environmental Protection Agency.

Fang, X., R. Li, H. Kan, M. Bottai, F. Fang, and Y. Cao. 2016. Bayesian model averaging method for evaluating associations between air pollution and respiratory mortality: A time-series study. *BMJ Open* 6(8):e011487.

Fung, E. B., L. D. Ritchie, L. R. Woodhouse, R. Roehl, and J. C. King. 1997. Zinc absorption in women during pregnancy and lactation: A longitudinal study. *Am J Clin Nutr* 66(1):80-88.

Gail, M. H., L. A. Brinton, D. P. Byar, D. K. Corle, S. B. Green, C. Schairer, and J. J. Mulvihill. 1989. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *J Natl Cancer Inst* 81(24):1879-1886.

Greenland, S., and M. P. Longnecker. 1992. Methods for trend estimation from summarized dose-response data, with applications to meta-analysis. *Am J Epidemiol* 135(11): 1301-1309.

Guha, N., A. Roy, L. Kopylev, J. Fox, M. Spassova, and P. White. 2013. Nonparametric Bayesian methods for benchmark dose estimation. *Risk Anal* 33(9):1608-1619.

Gutierrez, A. D., D. G. de Serna, I. Robinson, and D. S. Schade. 2009. The response of gamma vitamin E to varying dosages of alpha vitamin E plus vitamin C. *Metabolism* 58(4):469-478.

Guyatt, G. H., A. D. Oxman, S. Sultan, P. Glasziou, E. A. Akl, P. Alonso-Coello, D. Atkins, R. Kunz, J. Brozek, V. Montori, R. Jaeschke, D. Rind, P. Dahm, J. Meerpohl, G. Vist, E. Berliner, S. Norris, Y. Falck-Ytter, M. H. Murad, H. J. Schunemann, and Grade Working Group. 2011. GRADE guidelines: 9. Rating up the quality of evidence. *J Clin Epidemiol* 64(12):1311-1316.

Guyatt, G., A. D. Oxman, S. Sultan, J. Brozek, P. Glasziou, P. Alonso-Coello, D. Atkins, R. Kunz, V. Montori, R. Jaeschke, D. Rind, P. Dahm, E. A. Akl, J. Meerpohl, G. Vist, E. Berliner, S. Norris, Y. Falck-Ytter, and H. J. Schunemann. 2013. GRADE guidelines: 11. Making an overall rating of confidence in effect estimates for a single outcome and for all outcomes. *J Clin Epidemiol* 66(2):151-157.

Hall, K. D., S. B. Heymsfield, J. W. Kemnitz, S. Klein, D. A. Schoeller, and J. R. Speakman. 2012. Energy balance and its components: Implications for body weight regulation. *Am J Clin Nutr* 95(4):989-994.

HHS and USDA (U.S. Department of Health and Human Services and U.S. Department of Agriculture). 2015. *2015-2020 Dietary Guidelines for Americans. 8th Edition*. http://health.gov/dietaryguidelines/2015/guidelines (accessed October 3, 2017).

Hill, A. B. 1965. The environment and disease: Association or causation? *Proc R Soc Med* 58:295-300.

IOM (Institute of Medicine). 1997. *Dietary Reference Intakes for calcium, phosphorus, magnesium, vitamin D, and fluoride*. Washington, DC: National Academy Press.

IOM. 2002/2005. *Dietary Reference Intakes for energy, carbohydrate, fiber, fat, fatty acids, cholesterol, protein, and amino acids*. Washington, DC: The National Academies Press.

IOM. 2006. *Valuing health for regulatory cost-effectiveness analysis*. Washington, DC: The National Academies Press.

IOM. 2011. *Dietary Reference Intakes for calcium and vitamin D*. Washington, DC: The National Academies Press.

IOM. 2013. *Environmental decisions in the face of uncertainty*. Washington, DC: The National Academies Press.

Jiang, R., A. Botma, A. Rudolph, A. Husing, and J. Chang-Claude. 2016. Phyto-oestrogens and colorectal cancer risk: A systematic review and dose-response meta-analysis of observational studies. *Br J Nutr* 116(12):2115-2128.

Key, T. J., P. N. Appleby, G. K. Reeves, A. W. Roddam, and Endogenous Hormones Breast Cancer Collaborative Group. 2010. Insulin-like growth factor 1 (IGF1), IGF binding protein 3 (IGFBP3), and breast cancer risk: Pooled individual data analysis of 17 prospective studies. *The Lancet Oncol* 11(6):530-542.

Key, T. J., P. N. Appleby, R. C. Travis, D. Albanes, A. J. Alberg, A. Barricarte, A. Black, H. Boeing, H. B. Bueno-de-Mesquita, J. M. Chan, C. Chen, M. B. Cook, J. L. Donovan, P. Galan, R. Gilbert, G. G. Giles, E. Giovannucci, G. E. Goodman, P. J. Goodman, M. J. Gunter, F. C. Hamdy, M. Heliovaara, K. J. Helzlsouer, B. E. Henderson, S. Hercberg, J. Hoffman-Bolton, R. N. Hoover, M. Johansson, K. T. Khaw, I. B. King, P. Knekt, L. N. Kolonel, L. Le Marchand, S. Mannisto, R. M. Martin, H. E. Meyer, A. M. Mondul, K. A. Moy, D. E. Neal, M. L. Neuhouser, D. Palli, E. A. Platz, C. Pouchieu, H. Rissanen, J. M. Schenk, G. Severi, M. J. Stampfer, A. Tjonneland, M. Touvier, A. Trichopoulou, S. J. Weinstein, R. G. Ziegler, C. K. Zhou, N. E. Allen, and Endogenous Hormones Nutritional Biomarkers Prostate Cancer Collaborative Group. 2015. Carotenoids, retinol, tocopherols, and prostate cancer risk: Pooled analysis of 15 studies. *Am J Clin Nutr* 102(5):1142-1157.

Kim, Y., and Y. Je. 2016. Prospective association of sugar-sweetened and artificially sweetened beverage intake with risk of hypertension. *Arch Cardiovasc Dis* 109(4):242-253.

Klohs, J., M. Rudin, D. R. Shimshek, and N. Beckmann. 2014. Imaging of cerebrovascular pathology in animal models of Alzheimer's disease. *Front Aging Neurosci* 6:32.

Lippman, S. M., E. A. Klein, P. J. Goodman, M. S. Lucia, I. M. Thompson, L. G. Ford, H. L. Parnes, L. M. Minasian, J. M. Gaziano, J. A. Hartline, J. K. Parsons, J. D. Bearden, 3rd, E. D. Crawford, G. E. Goodman, J. Claudio, E. Winquist, E. D. Cook, D. D. Karp, P. Walther, M. M. Lieber, A. R. Kristal, A. K. Darke, K. B. Arnold, P. A. Ganz, R. M. Santella, D. Albanes, P. R. Taylor, J. L. Probstfield, T. J. Jagpal, J. J. Crowley, F. L. Meyskens, Jr., L. H. Baker, and C. A. Coltman, Jr. 2009. Effect of selenium and vitamin E on risk of prostate cancer and other cancers: The Selenium and Vitamin E Cancer Prevention Trial (SELECT). *JAMA* 301(1):39-51.

Liu, C. H., D. B. Rubin, and Y. N. Wu. 1998. Parameter expansion to accelerate EM: The PX-EM algorithm. *Biometrika* 85(4):755-770.

Marchetta, C. M., O. J. Devine, K. S. Crider, B. L. Tsang, A. M. Cordero, Y. P. Qi, J. Guo, R. J. Berry, J. Rosenthal, J. Mulinare, P. Mersereau, and H. C. Hamner. 2015. Assessing the association between natural food folate intake and blood folate concentrations: A systematic review and Bayesian meta-analysis of trials and observational studies. *Nutrients* 7(4):2663-2686.

McCullough, C. E., S. R. Searle, and J. M. Neuhaus. 2008. *Generalized, linear, and mixed models, 2nd Edition*. New York: John Wiley & Sons.

Meadows, N. J., S. L. Grainger, W. Ruse, P. W. Keeling, and R. P. Thompson. 1983. Oral iron and the bioavailability of zinc. *Br Med J (Clin Res Ed)* 287(6398):1013-1014.

Meyer, K. A., L. H. Kushi, D. R. Jacobs, Jr., J. Slavin, T. A. Sellers, and A. R. Folsom. 2000. Carbohydrates, dietary fiber, and incident type 2 diabetes in older women. *Am J Clin Nutr* 71(4):921-930.

Moher, D., A. Liberati, J. Tetzlaff, D. G. Altman, and P. Group. 2009. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Ann Intern Med* 151(4):264-269, W264.

Mulhern, B., K. Shah, M. F. Janssen, L. Longworth, and R. Ibbotson. 2016. Valuing health using time trade-off and discrete choice experiment methods: Does dimension order impact on health state values? *Value Health* 19(2):210-217.

Neuhouser, M. L., T. Y. Cheng, S. A. Beresford, E. Brown, X. Song, J. W. Miller, Y. Zheng, C. A. Thomson, J. M. Shikany, M. Z. Vitolins, T. Rohan, R. Green, and C. M. Ulrich. 2015. Red blood cell folate and plasma folate are not associated with risk of incident colorectal cancer in the Women's Health Initiative observational study. *Int J Cancer* 137(4):930-939.

NRC (National Research Council). 2009. *Science and decisions: Advancing risk assessment*. Washington, DC: The National Academies Press.

O'Brien, K. O., N. Zavaleta, L. E. Caulfield, J. Wen, and S. A. Abrams. 2000. Prenatal iron supplements impair zinc absorption in pregnant Peruvian women. *J Nutr* 130(9):2251-2255.

Rimm, E. B., A. Ascherio, E. Giovannucci, D. Spiegelman, M. J. Stampfer, and W. C. Willett. 1996. Vegetable, fruit, and cereal fiber intake and risk of coronary heart disease among men. *JAMA* 275(6):447-451.

Roberts P. L., Veidenheimer, M. C. 1990. Diverticular disease of the colon. In: Bayless, T. M., ed. *Current therapy in gastroenterology and liver disease—3*. Toronto: Decker Mosby. Pp. 416-419.

Roddam, A. W., N. E. Allen, P. Appleby, T. J. Key, L. Ferrucci, H. B. Carter, E. J. Metter, C. Chen, N. S. Weiss, A. Fitzpatrick, A. W. Hsing, J. V. Lacey, Jr., K. Helzlsouer, S. Rinaldi, E. Riboli, R. Kaaks, J. A. Janssen, M. F. Wildhagen, F. H. Schroder, E. A. Platz, M. Pollak, E. Giovannucci, C. Schaefer, C. P. Quesenberry, Jr., J. H. Vogelman, G. Severi, D. R. English, G. G. Giles, P. Stattin, G. Hallmans, M. Johansson, J. M. Chan, P. Gann, S. E. Oliver, J. M. Holly, J. Donovan, F. Meyer, I. Bairati, and P. Galan. 2008. Insulin-like growth factors, their binding proteins, and prostate cancer risk: Analysis of individual patient data from 12 prospective studies. *Ann Intern Med* 149(7):461-471, W83-W88.

Rosner, B. 2016. *Fundamentals of biostatistics, 8th edition*. Boston, MA: Cenage Learning.

Russell, L. B. 2015. The science of making better decisions about health. In *Population health: Behavioral and social science insights*. AHRQ Publication No. 15-0002, edited by R. Kaplan, M. Spittel, and D. David. Rockville, MD: Agency for Healthcare Research and Quality and Office of Behavioral and Social Sciences Research, National Institutes of Health. Pp. 251-270.

Sawaya, G. F., J. Guirguis-Blake, M. LeFevre, R. Harris, D. Petitti, and U.S. Preventive Services Task Force. 2007. Update on the methods of the U.S. Preventive Services Task Force: Estimating certainty and magnitude of net benefit. *Ann Intern Med* 147(12):871-875.

Sellmeyer, D. E., M. Schloetter, and A. Sebastian. 2002. Potassium citrate prevents increased urine calcium excretion and bone resorption induced by a high sodium chloride diet. *J Clin Endocrinol Metab* 87(5):2008-2012.

Slavin, J. L. 2008. Position of the American Dietetic Association: Health implications of dietary fiber. *J Am Diet Assoc* 108(10):1716-1731.

Solomons, N. W., and R. A. Jacob. 1981. Studies on the bioavailability of zinc in humans: Effects of heme and nonheme iron on the absorption of zinc. *Am J Clin Nutr* 34(4):475-482.

Solomons, N. W., O. Pineda, F. Viteri, and H. H. Sandstead. 1983. Studies on the bioavailability of zinc in humans: Mechanism of the intestinal interaction of nonheme iron and zinc. *J Nutr* 113(2):337-349.

Takata, Y., M. J. Shrubsole, H. Li, Q. Cai, J. Gao, C. Wagner, J. Wu, W. Zheng, Y. B. Xiang, and X. O. Shu. 2014. Plasma folate concentrations and colorectal cancer risk: A case-control study nested within the Shanghai Men's Health Study. *Int J Cancer* 135(9):2191-2198.

Tan, L., A. E. Babbs, M. H. Green, and A. C. Ross. 2016. Direct and indirect vitamin A supplementation strategies result in different plasma and tissue retinol kinetics in neonatal rats. *J Lipid Res* 57(8):1423-1434.

Vickers, A. J., and E. B. Elkin. 2006. Decision curve analysis: A novel method for evaluating prediction models. *Med Decis Making* 26(6):565-574.

Watters, D. A., and A. N. Smith. 1990. Strength of the colon wall in diverticular disease. *Br J Surg* 77(3):257-259.

Whelton, P. K., J. He, J. A. Cutler, F. L. Brancati, L. J. Appel, D. Follmann, and M. J. Klag. 1997. Effects of oral potassium on blood pressure. Meta-analysis of randomized controlled clinical trials. *JAMA* 277(20):1624-1632.

WHI (Women's Health Initiative) Study Group. 1998. Design of the Women's Health Initiative clinical trial and observational study. The Women's Health Initiative Study Group. *Control Clin Trials* 19(1):61-109.

Wolk, A., J. E. Manson, M. J. Stampfer, G. A. Colditz, F. B. Hu, F. E. Speizer, C. H. Hennekens, and W. C. Willett. 1999. Long-term intake of dietary fiber and decreased risk of coronary heart disease among women. *JAMA* 281(21):1998-2004.

Yang, Q., M. E. Cogswell, H. C. Hamner, A. Carriquiry, L. B. Bailey, C. M. Pfeiffer, and R. J. Berry. 2010. Folic acid source, usual intake, and folate and vitamin B-12 status in US adults: National Health and Nutrition Examination Survey (NHANES) 2003-2006. *Am J Clin Nutr* 91(1):64-72.

Yetley, E. A., A. J. MacFarlane, L. S. Greene-Finestone, C. Garza, J. D. Ard, S. A. Atkinson, D. M. Bier, A. L. Carriquiry, W. R. Harlan, D. Hattis, J. C. King, D. Krewski, D. L. O'Connor, R. L. Prentice, J. V. Rodricks, and G. A. Wells. 2017. Options for basing Dietary Reference Intakes (DRIs) on chronic disease endpoints: Report from a joint US-/Canadian-sponsored working group. *Am J Clin Nutr* 105(1):249S-285S.

Yu, E., M. Ruiz-Canela, M. Guasch-Ferre, Y. Zheng, E. Toledo, C. B. Clish, J. Salas-Salvado, L. Liang, D. D. Wang, D. Corella, M. Fito, E. Gomez-Gracia, J. Lapetra, R. Estruch, E. Ros, M. Cofan, F. Aros, D. Romaguera, L. Serra-Majem, J. V. Sorli, F. B. Hu, and M. A. Martinez-Gonzalez. 2017. Increases in plasma tryptophan are inversely associated with incident cardiovascular disease in the Prevencion con Dieta Mediterranea (PREDIMED) Study. *J Nutr* 147(3):314-322.

ANNEX 7-1
EXAMPLE OF INTAKE-RESPONSE ANALYSIS

Example of Meta-Analysis

TABLE 7-1-1 Summary of Meta-Analysis of Circulating and Dietary Magnesium and Risk of Cardiovascular Disease (Del Gobbo et al., 2013)

Author, Year	del Gobbo et al., 2013
Aims/Key Questions	To investigate prospective associations of circulating and dietary magnesium with incidence of cardiovascular disease (CVD) ischemic heart disease (IHD), and fatal IHD
Study Eligibility Criteria	Inclusion criteria: • *Study design:* All prospective studies (cohort and nested case–control) • *Exposure:* Circulating or dietary magnesium • *Outcome measures*: CVD, defined as any CVD, including cardiovascular or IHD incidence or death and stroke or angina as part of a broader composite CVD outcome. IHD, defined as IHD incidence or death. IHD death, defined as any fatal IHD, including sudden cardiac death (SCD) • *Population:* Adults • All eligible studies included a multivariate-adjusted effect estimate with a measure of uncertainty for circulating or dietary magnesium and incident CVD, IHD, or fatal IHD including SCD Exclusion criteria: • Studies reporting stroke as a distinct outcome • Studies focused on children • Studies that only evaluated drinking water magnesium or water hardness, dietary patterns/food groups, intracellular free magnesium, or extracellular ionized magnesium • Studies focused on populations with disturbed mineral homeostasis • Studies with only crude risk estimates, ecologic studies, case reports, cross-sectional studies, retrospective case-control studies, editorials/commentaries, letters, and reviews
Literature Search Dates or Year Range	Earliest available online indexing year to May 2012
Number of Studies Included	16 prospective studies, mostly cohort (11 with estimates of dietary magnesium; 9 with estimates of circulating magnesium)

TABLE 7-1-1 Continued

Author, Year	del Gobbo et al., 2013
Synthesis Methods (Summary Tables, Narrative Text, and/or Meta-Analysis)	Summary tables, narrative text, and meta-analysis
Key Findings	**Circulating Mg and CVD** (N=9 studies) RR: 0.70 (95% CI: 0.56-0.88) I^2=49.5% (moderate heterogeneity) Meta-regression: study location, percentage baseline CVD, and event type (incidence compared with death) significantly modified the association between circulating magnesium and CVD (P-heterogeneity=0.04, 0.02, and 0.02, respectively) **Circulating Mg and IHD** (N=5 studies) RR: 0.83 (95% CI: 0.65-1.05) I^2=49.5% (moderate heterogeneity) Fixed effects model: RR: 0.88 (95% CI: 0.76-1.02) No significant sources of between-study heterogeneity were identified **Circulating Mg and fatal IHD** (N=4 studies) RR: 0.61 (95% CI: 0.37-1.00) I^2=80.2% Fixed effects model: RR: 0.77 (95% CI: 0.64-0.93) Meta-regression did not identify any statistically significant sources of heterogeneity **Dietary magnesium and total CVD** (N=11 studies) RR: 0.89 (95% CI: 0.75-1.05) I^2=67.7% Fixed effects model: RR: 0.87 (95% CI: 0.72-0.89) No statistically significant sources of between-study heterogeneity were identified, but trends were seen toward stronger associations with lower risk among studies with lower median BMI (P-heterogeneity=0.09) or evaluating IHD rather than CVD (P-heterogeneity=0.07) **Dietary magnesium and CHD** (N=9 studies) RR: 0.78 (95% CI: 0.67-0.92) I^2=44.1 Fixed-effects model: RR: 0.80 (95% CI: 0.72-0.89) Trends toward stronger associations in cohorts with more men (P-heterogeneity=0.06) and studies evaluating fatal IHD death rather than total IHD (P-heterogeneity=0.07) were observed

continued

TABLE 7-1-1 Continued

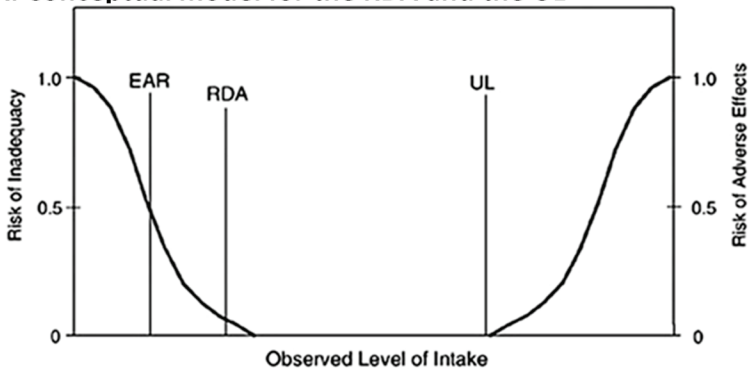
Author, Year	del Gobbo et al., 2013
Key Findings	**Dietary magnesium intake and fatal IHD** (N=4 studies) RR: 0.73 (95% CI: 0.52-1.03) I^2=43.2% Fixed-effects model: RR: 0.77 (95% CI: 0.60-0.98) Meta-regression did not identify any statistically significant pre-specified sources of heterogeneity, although the power to identify heterogeneity was limited given the number of studies No evidence of nonlinear associations between circulating magnesium and CVD (P=0.64), IHD (P=0.42), or fatal IHD (P=0.67) or between dietary magnesium and CVD (P=0.56) or IHD (P=0.26) Significant nonlinear association between dietary magnesium and fatal IHD (P-nonlinearity, 0.001). Compared with lower intakes, a 27 percent lower risk of fatal IHD was seen up to a threshold of ~250 mg/d (RR: 0.73 [95% CI: 0.62-0.86])
Limitations	Findings were constrained by the availability of published or unpublished data on magnesium-CVD associations Most of the included studies did not report on potential contribution of multivitamins or supplements to magnesium intake Possibility of residual confounding by dietary potassium in the dietary magnesium analyses

ANNEX 7-2
POSSIBLE NEW APPROACHES TO UL DEVELOPMENT

The UL is a DRI intended to address the potential risk associated with nutritional excess. As described in Chapter 2, it is defined as the "highest level of daily nutrient intake that is likely to pose no risk of adverse health effects to almost all individuals in the general population" (IOM, 1997, p. 4). As noted in Chapter 7, the committee recommends against using chronic disease endpoints as the basis of a UL, in part because in practice the UL is treated as a "bright line" between safe and unsafe intakes. Due to the multifactorial nature of chronic disease, no such bright line exists, as chronic diseases do not exhibit such "threshold"-like behavior and typically exhibit a continuous change in risk with changing intakes. However, this "threshold" interpretation of the UL is not entirely satisfactory either. In fact, although the conceptual model for a UL, illustrated in Figure 7-2-1, Panel A, depicts this level of intake as a point of "zero" risk, in practice such a level cannot be estimated because a small residual risk can never be ruled out.

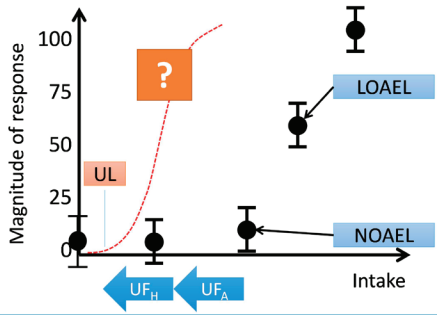
However, the way in which ULs have traditionally been derived does not characterize the degree of residual risk that may be present. This is because, as summarized in Chapter 7, the UL is usually obtained by identifying a NOAEL of intake from an experimental study (predominantly animal studies), and then dividing this intake level by a number of "uncertainty factors" to account for limitations in the data (see Figure 7-2-1, Panel B). The most commonly applied uncertainty factors are a factor of 10 to address differences between experimental animals (UF_A) and a second factor of 10 to address variability among humans (UF_H). This "NOAEL divided by 100" concept dates back to the 1950s in the context of FDA regulation of food additives (Lehman and Fitzhugh, 1954). Each of these components—the NOAEL, UF_A, and UF_H—is assumed to be "conservative" in the sense of erring on the side of protecting public health, but without much specificity as to "how conservative" they actually are (WHO/IPCS, 2014). For instance, with respect to the NOAEL, it is assumed that the severity of effects at this exposure level is negligible, but the extent to which this is true depends on the endpoint examined and the statistical power of the study (Crump, 1984; EPA, 2012). For UF_A, it is assumed that humans are generally no more than 10-fold more sensitive than the experimental animal species, but it is unclear at what confidence level this 10-fold factor is supposed to be (i.e. whether it should be 90 percent, 95 percent, 99 percent or other). Similarly, for UF_H, it is assumed that individuals who are more susceptible to toxicity are no more than 10-fold more sensitive than more typical individuals. Two ambiguities are relevant here: first, like UF_A, the confidence level of this 10-fold factor is unclear; and second, it is

A. Conceptual model for the RDA and the UL



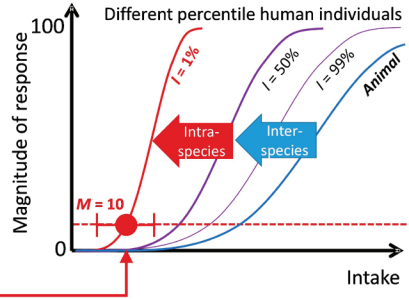
B. Traditional Approach to UL Derivation

$$UL = \frac{NOAEL}{UF_A \times UF_H}$$



C. Conceptual Illustration of how a "Target Human Dose" HD_M^I Can Represent the UL

Target Human Dose (HD_M^I): human dose (or intake) that at which a ***fraction I*** of the population shows an effect of ***magnitude (or severity) M*** or greater (for the critical effect considered).



unclear what "susceptible" means in terms of the more sensitive tail of the population distribution (i.e., 5 percent, 1 percent, 1 in a million or other).

This same approach of "NOAEL divided by 100" has been applied in chemical risk assessment, where exposure limits are set for contaminants in food or the environment. However, in the chemical risk assessment field, a number of refinements to the approach have been developed that may be equally applicable to the UL. Most recently, the World Health Organization/International Program on Chemical Safety (WHO/IPCS, 2014) developed a guidance document describing a "probabilistic" framework that results in substantially better characterization of the intake-response for adverse effects. The key concept underlying the WHO/IPCS approach, illustrated in Figure 7-2-1, Panel C, is that the goal of deriving quantities like the UL is a "target human dose" HD_M^I, defined as to estimated human dose (or intake) at which effects with magnitude *M* occur in the population with an incidence *I*, along with an associated CI.

By providing intake-response functions rather than "bright lines," changes in risk of adverse effects from changes in intake can be quantified. This may be particularly important in the scenario described in Chapter 7 in which increasing intake decreases chronic disease risk up to and perhaps beyond the UL. This type of "risk-benefit" comparison would be infeasible under the traditional "NOAEL divided by 100" approach, because there is no characterization of the gradient of the intake response over a wide enough range of doses. However, the approach to derive an HD_M^I would enable such comparison to be made much more easily. For additional details, see WHO/IPCS (2014) and Chiu and Slob (2015).

FIGURE 7-2-1 Panel A: Conceptual model for the RDA and the UL. Although conceptualized as a point at which there is exactly zero risk of adverse effects, in practice an unspecified level of residual risk cannot be ruled out due to how the UL has traditionally been derived. **Panel B:** Traditional approach to UL derivation involves dividing a NOAEL by "Uncertainty Factors" accounting for inter-species and inter-individual differences in toxicity. The result is an intake level that is identified as the UL, but without any characterization as to the residual risk or the expected intake-response function for adverse effects. **Panel C:** Conceptual illustration of the "target human dose" HD_M^I as a potential replacement for the approach in Panel B. The principles of extrapolating from experimental animal studies to human populations is the same as for the traditional UL, but the output of the approach is explicit as to the magnitude of effect M and incidence I at which this effect occurs, as well as the confidence interval of the estimate. For details, see WHO/IPCS (2014).

ANNEX 7-2 REFERENCES

Chiu, W. A., and W. Slob. 2015. A unified probabilistic framework for dose-response assessment of human health effects. *Environ Health Perspect* 123(12):1241-1254.

Crump, K. S. 1984. A new method for determining allowable daily intakes. *Fundam Appl Toxicol* 4(5):854-871.

EPA (U.S. Environmental Protection Agency). 2012. *Benchmark dose technical guidance*. Washington, DC: U.S. Environmental Protection Agency.

IOM (Institute of Medicine). 1997. *Dietary Reference Intakes for calcium, phosphorus, magnesium, vitamin D, and fluoride*. Washington, DC: National Academy Press.

Lehman, A. J., and O. G. Fitzhugh. 1954. 100-fold margin of safety. *Assoc Food Drug Off U.S.Q. Bull* 18:33-35.

WHO/IPCS (World Health Organization/International Program on Chemical Safety). 2014. *Guidance document on evaluating and expressing uncertainty in hazard characterization*. Geneva, Switzerland: World Health Organization International Program on Chemical Safety.

8

The Process for Establishing Chronic Disease Dietary Reference Intakes

Chapters 1, 2, and 3 briefly explain the history of the Dietary Reference Intakes (DRIs) in the United States and Canada, the process to establish them, and the challenges that nutrition researchers face when assessing the associations between nutrients and chronic diseases. Chapters 4, 5, 6, and 7 offer recommendations and guiding principles to address the methodological and conceptual challenges when recommending chronic disease DRIs.

Chronic disease, however, is only one of several types of indicators that are reviewed through the DRI process. Other key indicators are those for adequacy and toxicity, which are not the focus of this report but will also continue to be reviewed and assessed so that Estimated Average Requirements (EARs), Recommended Dietary Allowances (RDAs), and Tolerable Upper Intake Levels (ULs) can be established. The establishment of DRIs occurs with the collaboration of a number of stakeholders, and within that collaboration, each stakeholder has a specific role to contribute. For example, the Canadian and U.S. Steering Committees identify the nutrients and questions to be addressed. The systematic review team conducts the systematic review in consultation with a technical expert panel. The DRI committee—which ideally will include some members of the technical expert panel—receives and assesses the systematic review, makes conclusions based on the certainty in the evidence, and makes DRI recommendations based on one or more indicators (i.e., adequacy, toxicity, and chronic disease risk). The goal of this chapter is to address two questions related to specific aspects of the DRI process and for each, recommend one of the options offered in *Options for Basing Dietary Reference Intakes (DRIs) on Chronic Disease Endpoints: Report from a Joint US-/Canadian-Sponsored*

Working Group (i.e., the Options Report) (Yetley et al., 2017). The first question deals with how to incorporate the deliberations and recommendations regarding chronic disease DRIs into the already existing DRI process, which relies on the work of a National Academies of Science, Engineering, and Medicine (the National Academies) committee (i.e., DRI committee). The second question addresses the nature of the initial question that will drive the systematic review and ensuing tasks. This committee provides a justification for selecting one of the options to answer these two questions but notes that tasks such as deciding when and how to select a nutrient/chronic disease for review are out of the scope for this report, as mentioned in Chapter 1.

CONVENING A DRI COMMITTEE AND THE ROLE OF DRI COMMITTEES

As Chapter 2 explains, since 1938 and 1941, nutrient reference values have been issued in Canada and the United States, respectively. In the United States, the nutrient reference values (both the original RDAs and the more recent DRIs) were published by the Food and Nutrition Board (FNB) of the National Academy of Sciences. In the current DRI process—initiated in the early 1990s and then modified for the latest DRIs on calcium and vitamin D in 2011—the National Academies convenes an expert committee to develop DRIs for a group of related nutrients (see Table 2-1, Chapter 2). In addition to following the National Academies' policies and procedures, committees must comply with provisions of section 15 of the Federal Advisory Committee Act[1] so that U.S. government agencies can use the recommendations provided.

When they are appointed, DRI committees are given a statement of task, which is a description of the objective of the project and the specific charge to the committee. The statement of task is typically developed by the governments of the United States and Canada, which also have been the sole funders of the work of DRI committees, and in consultation with the FNB. DRI committees are composed of scientists in the relevant disciplines that are needed to complete the specific task (e.g., human nutrition, epidemiology, toxicology) and are convened based on a selection process that considers suggestions by stakeholders and potential conflicts of interests and biases following the policies of the National Academies. DRI committees are supported by FNB staff, who coordinate their work and ensure completion of the task and the publication of a report containing the DRIs as well as the scientific rationales, in the style of the National Academies reports. It generally takes 1 to 2 years for a DRI committee to complete

[1] Federal Advisory Committee Act, 5 U.S.C. § 15.

its work. As mentioned in Chapter 1, the committee developed its recommendations and guiding principles in the context of the process shown in Figure 1-2 in which formal systematic reviews are conducted by an outside contractor before the formation of the DRI committee. Current practice is that, for each set of related nutrients, one committee addresses all indicators (i.e., adequacy, toxicity, and, possibly, chronic disease outcomes) to establish DRIs. As Chapter 2 indicates, for a single nutrient one or more DRIs may be established, and many nutrients have an EAR and an RDA (or an AI) and a UL. If adequate data are available, the DRIs may incorporate chronic disease considerations (IOM, 2003). Chapter 2 lists all the DRIs established to date and the reports by the National Academies in which they have been published.

As Chapters 6 and 7 explain, future DRI committees that will consider chronic disease DRIs not only will have to assess whether a DRI can be recommended for a particular chronic disease, but also will have to balance harms and risks related to all relevant indicators. Although risks and harms also were considered in the past, the scenarios will likely become more complex as committees formally consider chronic disease indicators and more evidence becomes available. For example, future DRI committees may have to consider situations where one nutrient or food substance (NOFS) may be associated with more than one chronic disease but within different ranges of intake. Other scenarios can be anticipated in which a range of intakes for one nutrient may provide benefits for a chronic disease outcome but may overlap with the UL.

CONVENING OF DRI COMMITTEES: OPTIONS AND JUSTIFICATION

Box 8-1 shows the options presented in the Options Report with regard to the DRI process. Option 1 was to continue to use a single DRI development process and option 2 was to create two separate processes for developing DRIs. The committee recommends a variation of option 1, such that for each set of NOFSs under review, a single DRI "parent" committee would be formed but it could have two subcommittees focused on different questions. If two subcommittees are formed, the parent committee would then integrate the recommendations of each subcommittee into a single report.

The committee chose not to use option 2 because the Federal Advisory Committee Act rules, which apply to committees of the National Academies and are meant to minimize external influences, require that committee members and staff keep deliberations confidential. Two separate DRI processes would not allow for sufficient exchange of ideas between the two committees. Because of the need for coordination of the recommendations,

BOX 8-1
Process Components and Options

Options Report

Option 1: Continue to use a single DRI development process

This option continues to consider chronic disease endpoints in future DRI reviews but expands the types of reference values to clearly distinguish those based on classical nutrient adequacy from those based on chronic disease endpoints. This option makes the addition of CD_{XX} (where XX denotes the specific chronic disease) and UL_{CD} values or ranges a natural extension of the current process.

Option 2: Create two separate processes for developing DRIs

This option would create two separate but complementary, and possibly iterative or integrated, committees to develop reference values on the basis of chronic disease endpoints or deficiency diseases. The FNB or a government agency could appoint a new committee to establish reference values on the basis of chronic disease endpoints, or an existing group that is independent of the National Academies of Sciences, Engineering, and Medicine (e.g., expert panels from chronic disease societies or standing government advisory committees) could establish these reference values. This new reference-setting group would coordinate its activities closely with the current DRI process based on adequacy.

Committee's Recommendation 10

Because of the need for close coordination and exchange of ideas when setting Dietary Reference Intakes (DRIs) based on indicators of adequacy, toxicity, and chronic disease, one single National Academies of Sciences, Engineering, and Medicine parent committee should develop DRIs for the prevention of nutrient deficiencies and toxicities and for reducing the risk of chronic disease. Due to the need for different expertise and different methodological considerations, two subcommittees could be established at the discretion of the parent committee, for reviewing evidence on (1) adequacy and toxicity and (2) chronic disease, respectively.

and in particular the need to consider harms and benefits with regard to all health indicators, one DRI committee would be a better choice. This committee structure will allow for full exchange of ideas and deliberations among all members.

Because of the importance of drawing on historical knowledge and experience with the process, the National Academies would be the logical organization to coordinate the process. When enough scientific evidence for an update of the DRIs exists for a given set of NOFSs, and as requested

by sponsors,[2] the FNB would convene an ad hoc consensus committee (parent committee) that would be responsible for establishing the DRIs. A wide range of expertise would need to be represented on the committee to accommodate the interdisciplinary approach needed to set standards for reducing the risk of chronic disease as well as the narrower perspective of preventing nutrient deficiencies and toxicities. As Guiding Principle 9 states (see Chapter 6), "Particular elements of needed expertise will be guided by the general scientific question(s) and specific questions and will generally include nutrition science, scientific study design and analysis, public health, biostatistics, nutrition and chronic disease epidemiology, disease pathogenesis, and evidence review conduct."

It should be noted that not all nutrients have been associated with a chronic disease. Thus, some DRI committees would not have a chronic disease DRI subcommittee. When two subcommittees are formed, the FNB staff working with the parent committee would coordinate the work of the two subcommittees so that a full exchange and integration of ideas between the subcommittees could occur. Various strategies would be implemented to ensure this cross-fertilization of ideas, such as designating some members to participate in both subcommittees and the parent committee, organizing meetings that all members would attend, and sharing evidence tables (see Chapter 6) and other documents.

When two DRI subcommittees are formed, they would approach the development of DRIs for a nutrient or a group of nutrients as separate processes, though they would use the same formally conducted systematic evidence review to complete their work.[3] One subcommittee would recommend DRIs for all 22 life-stage groups that focus on the requirements for specific nutrients to ensure nutrient adequacy and prevent nutrient deficiency symptoms as well as prevent toxicities. The second subcommittee would examine the certainty of the scientific evidence on the relationship between the nutrients under review and chronic disease outcomes, keeping in mind the distinction that chronic disease DRIs are desirable but not essential. The evidence review and, ultimately, the certainty in the data regarding the intake-response relationship (see Chapter 7) would dictate whether it is necessary to recommend DRIs for one or more NOFSs, including NOFSs individually or in combination that may reduce the risk of a chronic disease of interest. If insufficient certainty existed in the evidence

[2] As described in Chapter 6, once the general nutrient and health outcomes of interest have been identified, an important next preliminary step is to conduct a scoping review, that is, to sample the scientific literature and determine whether sufficient studies exist to conduct a full systematic review.

[3] Although it is assumed that both subcommittees will have access to the same systematic review(s), this committee is not commenting on the process of setting nutrient reference values for adequacy or toxicity, per the statement of task.

to develop DRIs for reducing risk of chronic disease for the nutrients under review, the second subcommittee's process would stop. If sufficient evidence did exist to recommend chronic disease DRIs for any of the life-stage groups, all members would deliberate their conclusions about DRIs for all indicators (i.e., adequacy, toxicity, and chronic disease risk), consider benefits and harms, and recommend the chronic disease DRIs, incorporating all needed documentation and explanations in the report (see Guiding Principles). If recommended for any of the life-stage groups, these chronic disease DRIs could be defined as proposed in Chapter 7 (see Table 7-1), that is, as acceptable range of intake, range of beneficial increased intakes, or range of beneficial decreased intakes.

An alternative to the committee structure above would be to convene only one parent committee, with all the necessary expertise, but no subcommittees, to establish all DRIs. However, this committee concludes that the process to derive DRIs based on adequacy and toxicity would not be adequate to derive chronic disease DRIs. Although the general approach to establishing DRIs would be consistent with the same general risk framework (see Annex to Chapter 1), developing chronic disease DRIs presents unique challenges (see Chapter 3) that can only be adequately addressed by applying specific methodological and statistical models (see Chapter 7). In fact, although the same systematic review(s) might be used to compile the relevant information for all indicators, the type of evidence used as a basis for chronic disease DRIs will differ from the type of evidence used to establish DRIs for adequacy and toxicity because of the unique features of the relationships between nutrients and chronic diseases. A few examples are:

- When dietary factors influence the risk of chronic disease, DRIs will likely be expressed in ranges rather than a single number. In some cases, a DRI recommendation regarding chronic disease endpoints would most appropriately involve a ratio of nutrients.
- Multiple factors (including foods and nutrients) are associated with the risk of chronic disease. The current EAR, RDA, and UL models that focus on biological outcomes or functions for a single nutrient will not work well for chronic disease endpoints that have multiple risk factors.

STARTING-POINT ISSUES: OPTIONS AND JUSTIFICATION

As the Options Report indicates, the current starting point for establishing DRIs is individual or small groups of NOFSs (Option 1 in Box 8-2). However, given the number of NOFSs that might affect a chronic disease, the Options Report presented a second potential approach, whereby DRI

BOX 8-2
Starting-Point Issues and Options

Options Report
Option 1: Establish DRIs for individual or small groups of interrelated food substances

Option 2: Establish DRIs for multiple food substances on the basis of a chronic disease endpoint

This option requires a different paradigm from the one that DRI committees currently use. For each selected chronic disease, DRI committees would develop a reference value for all food substances that have a causal relation with the risk of that disease.

Committee's Recommendation 11

When sufficient evidence exists to develop chronic disease Dietary Reference Intakes for one or more nutrient or other food substances (NOFSs) that are interrelated in their causal relationships with one or more chronic diseases, a committee should be convened to review the evidence of their association with all selected diseases. Using a chronic disease as the starting point for the review is not recommended because balancing health risks and benefits for multiple NOFSs that are related to a single chronic disease endpoint will be a challenge in cases where the same NOFSs might be associated with more than one chronic disease.

committees would assess the evidence on all NOFSs related to a particular chronic disease.

The committee deliberated about the potential advantages of this second option as an alternative to the current process, including the fact that this option could be, in principle, more helpful for those who have higher risk of developing a particular chronic disease. However, the committee found that continuing with the current approach of recommending DRIs for individual or small groups of related NOFSs has advantages, at least for now. First, relatively little experience has been accrued in developing DRIs based on chronic disease, so it may be premature to change the current process before additional experience is gained. Second, the current scientific literature and study designs have tended to explore relationships substance by substance; that is, individual studies are more likely to document how a particular NOFS is related to various diseases rather than studying all the NOFSs that are related to a particular chronic disease. Although the evidence base on how dietary patterns relate to chronic diseases is growing, it is difficult to use this evidence to develop DRIs for each individual NOFS.

In addition, the 2015 Dietary Guidelines Advisory Committee[4] already addressed the relationship between dietary patterns and chronic disease risk. As the research findings increase, it may be possible to identify key nutrients within dietary patterns that are influencing chronic disease risk. When that occurs, this approach to review the evidence (or starting point) for recommending chronic disease DRIs could be reconsidered.

Finally, and perhaps most importantly, the fact that some NOFSs might contribute to more than one chronic disease implies that, if this second option were selected, balancing of harms and benefits would be more challenging, if not impossible. The committee would lack essential information if, for example, an NOFS range reduces the risk of the chronic disease in question but affects the risk of another disease in the opposite direction.

For the reasons noted above, the committee recommends option 1, that is, when sufficient evidence exists to develop chronic disease DRIs, DRIs should be recommended for one or more NOFSs that are interrelated in their causal relationships with chronic disease(s). As knowledge about relationships between NOFSs and chronic diseases advances, the Canadian and U.S. DRI Steering Committees might expand the specific questions to consider not only the effects of the NOFSs of interest themselves, but also NOFSs that act as effect modifiers. For example, questions about the potential effects of calcium, magnesium, and potassium in the association between sodium and blood pressure could be included in the systematic review protocol for DRIs (see also Chapter 7, "Were Interactions with Other Nutrients Considered?).

Since 1994, DRIs have served as the foundation for nutrition policies and guidance for individuals and groups in the United States and Canada. The *Dietary Guidelines for Americans* and the *Canadian Food Guide*, for example, draw on information from the DRI reports. DRIs have many other important uses, such as providing benchmarks for monitoring dietary intake of populations, evaluating the quality of government food assistance programs, planning foods and diets for military personnel, and planning and monitoring other nationwide health programs. However, in the past DRIs were based on reaching adequacy and minimizing the potential for toxicity. As populations have changed their diet habits and more information about the prevalence of chronic disease and its risk factors has become available, there is a need to explore how nutrients contribute to chronic

[4] The Dietary Guidelines Advisory Committees review the body of scientific and medical evidence in nutrition and prepare an Advisory Report for the Secretaries of the U.S. Department of Health and Human Services (HHS) and the U.S. Department of Agriculture (USDA) every 5 years. The Advisory Report provides an evidence base for HHS and USDA as the departments update the *Dietary Guidelines for Americans* (https://health.gov/dietaryguidelines/purpose.asp; accessed May 10, 2017).

disease and whether their specific levels can be determined to ameliorate the risk of chronic disease. Decisions about chronic disease DRIs, however, are fundamentally different from decisions concerning adequacy and toxicity DRIs. Differences lie in the nature of the health outcomes (e.g., the long-term nature of chronic disease), the scientific data available, and the methodologies required to analyze the data. Differences also exist in the expertise needed. For these reasons, it is necessary to develop some guiding principles and recommendations that take these differences into account. Although the United States and Canada have many years of experience in setting DRIs, integrating chronic disease as a focus is a fairly recent task; therefore the recommendations in this report should be revisited in the future as more practice and knowledge is gained. As chronic disease DRIs are set, there will be a need to develop guidance, possibly separate from the DRI reports themselves, on how these new DRIs could be used in dietary assessment and planning, especially in complex situations, such as a single NOFS with DRI ranges to lower risk of different chronic diseases.

REFERENCES

IOM (Institute of Medicine). 2003. *Dietary Reference Intakes: Applications in dietary planning*. Washington, DC: The National Academies Press.

Yetley, E. A., A. J. MacFarlane, L. S. Greene-Finestone, C. Garza, J. D. Ard, S. A. Atkinson, D. M. Bier, A. L. Carriquiry, W. R. Harlan, D. Hattis, J. C. King, D. Krewski, D. L. O'Connor, R. L. Prentice, J. V. Rodricks, and G. A. Wells. 2017. Options for basing Dietary Reference Intakes (DRIs) on chronic disease endpoints: Report from a joint US-/Canadian-sponsored working group. *Am J Clin Nutr* 105(1):249S-285S.

Appendix A

Open Session Agendas

The committee held two open sessions in Washington, DC, on October 13, 2016, and January 9, 2017. The open session agendas for the public meeting and a workshop are presented below:

The Development of Guiding Principles for the Inclusion of Chronic Disease Endpoints in Future Dietary Reference Intakes
Open Session

Committee on the Development of Guiding Principles for the Inclusion of Chronic Disease Endpoints in Future Dietary Reference Intakes

October 13, 2016

Keck Center of the National Academies
500 Fifth Street, NW, Washington, DC
Room 208

1:00 PM **Welcome Remarks**, Shiriki Kumanyika, Ph.D., M.P.H., M.S.W., *Committee Chair*

1:05 PM **Sponsor Perspectives on the Study**
Christopher Lynch, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health (NIH), U.S. Department of Health and Human Services (HHS)
Christine Taylor, Office of Dietary Supplements, NIH, HHS
Amanda McFarlane, Nutrition Research Division, Health Canada

1:45 PM **Questions from the Committee**

2:30 PM **Informal Q&A on the Dietary Reference Intake (DRI) Experience**
Christine Taylor, Office of Dietary Supplements, NIH, HHS
Linda Meyers, Office of Dietary Supplements, NIH, HHS

3:00 PM **Public Comments**

3:05 PM BREAK

The Development of Guiding Principles for the Inclusion of Chronic Disease Endpoints in Future Dietary Reference Intakes
A Public Workshop

Committee on the Development of Guiding Principles for the Inclusion of Chronic Disease Endpoints in Future Dietary Reference Intakes

January 9, 2017

National Academy of Sciences Building
2101 Constitution Avenue, NW Washington, DC
Lecture Room

8:30 AM **Welcome Remarks,** Shiriki Kumanyika, Ph.D., M.P.H., M.S.W., *Committee Chair*

SESSION I: CURRENT TOOLS FOR ASSOCIATING NUTRIENTS AND CHRONIC DISEASE

8:35 AM **Overview of the State of Biomarkers**
Marian Neuhouser, Ph.D., R.D., Fred Hutchinson Cancer Research Center

8:55 AM **Using All the Data: Bayesian Models and the Determination of Optimal Blood Folate Concentration for Neural Tube Defect Prevention**
Krista Crider, Ph.D., U.S. Centers for Disease Control and Prevention

9:15 AM **A Brief Overview of Absolute Risk Models and How They Can Be Used to Design Intervention Studies**
Ruth Pfeiffer, Ph.D., National Cancer Institute

9:45 AM **Case Study: Diet and Alzheimer's Disease**
Martha Clare Morris, Sc.D., Rush University

10:05 AM **Panel Discussion**
Facilitator: Patrick J. Stover, Ph.D., Cornell University

10:30 AM BREAK

SESSION II: EMERGING METHODOLOGIES FOR ESTABLISHING NUTRIENT-CHRONIC DISEASE RELATIONSHIPS

10:45 AM **Overview and Comparisons of Risk of Bias and Strength of Evidence Assessment Tools: Opportunities and Challenges in Applying DRIs**
Mei Chung, Ph.D., M.P.H., Tufts University

11:05 AM **A Probabilistic Hazard Characterization Framework for Addressing Uncertainty and Variability**
Weihsueh Chiu, Ph.D., Texas A&M University

11:25 AM **Pattern of Lipid Biomarkers to Predict Disease Risk**
Robert Clarke, Nuffield Department of Public Health

11:45 AM **Uses of Agent-Based Modeling to Inform Policy and Science in Chronic Disease-Brief Overview**
Ross Hammond, Ph.D., Brookings Institution

12:05 PM **Panel Discussion**
Facilitator: Susan Barr, Ph.D., R.D., University of British Columbia

12:30 PM BREAK

SESSION III: USES OF DRIs

1:30 PM **Uses of DRIs and How to Incorporate DRIs with Chronic Disease Endpoints into Those Uses**

- Perspectives on Chronic Disease Outcomes (and Issues) from the DRI Committee on Electrolytes and Water
  Larry Appel, M.D., Johns Hopkins University
- Use of DRIs in Federal Nutrition Programs
  David Klurfeld, Ph.D., U.S. Department of Agriculture
- Use of the DRIs at the U.S. Food and Drug Administration
  Paula Trumbo, Ph.D., U.S. Food and Drug Administration

2:45 PM **Panel Discussion**
Facilitator: Linda D. Meyers, M.S., Ph.D., Office of Dietary Supplements, NIH, HHS

3:15 PM **Public Comments**

3:30 PM **Adjourn**

Appendix B

Options Report

Options for basing Dietary Reference Intakes (DRIs) on chronic disease endpoints: report from a joint US-/Canadian-sponsored working group[1–3]

Elizabeth A Yetley,[4] Amanda J MacFarlane,[5] Linda S Greene-Finestone,[5] Cutberto Garza,[6–8] Jamy D Ard,[9] Stephanie A Atkinson,[10] Dennis M Bier,[11] Alicia L Carriquiry,[12] William R Harlan,[13] Dale Hattis,[14] Janet C King,[15–17] Daniel Krewski,[18] Deborah L O'Connor,[19,20] Ross L Prentice,[21,22] Joseph V Rodricks,[23] and George A Wells[24]*

[4]Office of Dietary Supplements, NIH, Bethesda, MD; [5]Bureau of Nutritional Sciences, Health Canada, Ottawa, Ontario, Canada; [6]Boston College, Chestnut Hill, MA; [7]Department of Global Health, George Washington University Milken Institute School of Public Health, Washington, DC; [8]Department of International Health, Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD; [9]Wake Forest School of Medicine, Wake Forest University, Winston-Salem, NC; [10]Department of Pediatrics, McMaster University, Hamilton, Ontario, Canada; [11]Children's Nutrition Research Center, Baylor College of Medicine, Houston, TX; [12]Department of Statistics, Iowa State University, Ames, IA; [13]Retired, Office of the Director, NIH, Bethesda, MD; [14]The George Perkins Marsh Institute, Clark University, Worcester, MA; [15]Children's Hospital Oakland Research Institute, Oakland, CA; [16]Department of Nutritional Sciences, University of California, Berkeley, Berkeley, CA; [17]Department of Nutrition, University of California, Davis, Davis, CA; [18]McLaughlin Centre for Population Health Risk Assessment, University of Ottawa, Ottawa, Ontario, Canada; [19]Department of Nutritional Sciences, University of Toronto; [20]The Hospital for Sick Children, Toronto, Ontario, Canada; [21]Fred Hutchinson Cancer Research Center; [22]School of Public Health, University of Washington, Seattle, WA; [23]Ramboll-Environ International Corporation, Arlington, VA; and [24]Department of Epidemiology and Community Medicine, University of Ottawa Heart Institute, Ottawa, Ontario, Canada

ABSTRACT

Dietary Reference Intakes (DRIs) are used in Canada and the United States in planning and assessing diets of apparently healthy individuals and population groups. The approaches used to establish DRIs on the basis of classical nutrient deficiencies and/or toxicities have worked well. However, it has proved to be more challenging to base DRI values on chronic disease endpoints; deviations from the traditional framework were often required, and in some cases, DRI values were not established for intakes that affected chronic disease outcomes despite evidence that supported a relation. The increasing proportions of elderly citizens, the growing prevalence of chronic diseases, and the persistently high prevalence of overweight and obesity, which predispose to chronic disease, highlight the importance of understanding the impact of nutrition on chronic disease prevention and control. A multidisciplinary working group sponsored by the Canadian and US government DRI steering committees met from November 2014 to April 2016 to identify options for addressing key scientific challenges encountered in the use of chronic disease endpoints to establish reference values. The working group focused on 3 key questions: *1*) What are the important evidentiary challenges for selecting and using chronic disease endpoints in future DRI reviews, *2*) what intake-response models can future DRI committees consider when using chronic disease endpoints, and *3*) what are the arguments for and against continuing to include chronic disease endpoints in future DRI reviews? This report outlines the range of options identified by the working group for answering these key questions, as well as the strengths and weaknesses of each option. *Am J Clin Nutr* 2017;105(Suppl):249S–85S.

Keywords: Dietary Reference Intakes, chronic disease, intake response, evidentiary challenges, evidence assessments

I. EXECUTIVE SUMMARY

Background

Dietary Reference Intakes (DRIs)[21] represent a common set of reference intake values used in Canada and the United States in planning and assessing diets of apparently healthy individuals and population groups. Past expert committees that developed these reference values took into consideration the deficiencies, inadequacies, and toxicities of nutrients and related food substances as well as relevant chronic disease outcomes. The increasing

[1] This is a report based on working group meetings held between November 2014 and April 2016 and a public workshop titled "Options for Consideration of Chronic Disease Endpoints for Dietary Reference Intakes (DRIs)" held at the NIH in Bethesda, MD, 10–11 March 2015.

[2] Supported by the Bureau of Nutritional Sciences, Health Canada; Office of Nutrition Policy and Promotion, Health Canada; the Social Determinants and Science Integration Directorate, Public Health Agency of Canada; the Office of Dietary Supplements, NIH; the Agricultural Research Service, USDA; the National Heart, Lung, and Blood Institute, NIH; the Center for Food Safety and Applied Nutrition, US Food and Drug Administration; and the National Center for Chronic Disease Prevention and Health Promotion, US CDC. This is a free access article, distributed under terms (http://www.nutrition.org/publications/guidelines-and-policies/license/) that permit unrestricted noncommercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

[3] The findings and conclusions in this article are those of the authors and do not necessarily represent the official views or positions of Health Canada, the US NIH, the USDA, the US Food and Drug Administration, or the US CDC.

*To whom correspondence should be addressed. E-mail: amanda.macfarlane@hc-sc.gc.ca.

[21] Abbreviations used: AHRQ, Agency for Healthcare Research and Quality; AI, Adequate Intake; AMDR, Acceptable Macronutrient Distribution Range; AMSTAR, A Measurement Tool to Assess Systematic Reviews; CD, chronic disease; CD_{cancer}, chronic disease risk reduction intake value for cancer; CD_{CVD}, chronic disease risk reduction intake value for cardiovascular disease; CVD, cardiovascular disease; DRI, Dietary Reference Intake; EAR, Estimated Average Requirement; FNB, Food and Nutrition Board; GRADE, Grading of Recommendations Assessment, Development, and Evaluation; RCT, randomized controlled trial; RDA, Recommended Dietary Allowance; ROBINS, Risk of Bias in Nonrandomized Studies; SIGN 50, Scottish Intercollegiate Guidelines Network 50; UL, Tolerable Upper Intake Level; UL_{CD}, chronic disease Tolerable Upper Intake Level.

First published online December 7, 2016; doi: 10.3945/ajcn.116.139097.

proportions of elderly citizens, the growing prevalence of chronic diseases, and the persistently high prevalence of overweight and obesity, which predispose to chronic disease, in Canada and the United States highlight the importance of understanding the impact of nutrition on chronic disease prevention and control, and on health promotion.

The approaches that expert committees have used to establish the DRIs usually worked well when these groups considered classical nutrient deficiencies and/or toxicities. However, when committees concluded that there was sufficient evidence to base a reference value on a chronic disease endpoint, deviations from the frameworks that were initially developed for DRI use were often required. In some cases, committees were unable to establish reference values for intakes that affected chronic disease outcomes despite evidence that supported relations between intakes and chronic disease outcomes.

Current project

A multidisciplinary working group sponsored by Canadian and US government DRI steering committees met from November 2014 to April 2016 to identify key scientific challenges that past DRI committees encountered in the use of chronic disease endpoints to establish reference values. The working group focused its discussions on 3 key questions:

1) What are the important evidentiary challenges for selecting and using chronic disease endpoints in future DRI reviews?

2) What intake-response models can future DRI committees consider when using chronic disease endpoints?

3) What are the arguments for and against continuing to include chronic disease endpoints in future DRI reviews?

Currently, DRIs apply to apparently healthy populations, but changing demographics (e.g., an aging population) and health status (e.g., increasing rates of obesity) suggest a possible need for broader population coverage. Past DRIs generally focused on intakes achievable by dietary strategies, but the growing ability to modify intakes through fortification and supplementation is increasingly relevant to future DRI development. In addition to these evolving concerns, future DRI committees need to continue to take into account the broad and diverse uses of DRIs when considering options for DRIs, including those based on chronic disease endpoints.

The sponsors asked the working group to identify a (not necessarily exhaustive) range of options for answering each of the key questions and the strengths and weaknesses of each option, while keeping in mind current and future DRI contexts and uses. The sponsors did not ask the group to reach a consensus on which options have the highest priority. Final decisions about the feasibility and options for specific approaches for deriving DRIs on the basis of chronic disease outcomes will be made by a future DRI committee.

Judging the evidence

The DRI process includes 2 key scientific decisions: *1*) whether the available evidence supports a causal relation between the food substance of interest and a selected outcome and, *2*) if so, which DRIs are appropriate based on the available data. DRI committees make these decisions for both beneficial and adverse effects. In the current project, the outcome of interest is a chronic disease.

Challenges in evaluating the evidence

When a DRI committee assesses whether the intake of a given food substance is causally related to a chronic disease or attempts to determine the nature of an intake-response relation between a food substance and a chronic disease, it considers the characteristics of individual study designs and overarching issues that apply across different types of study designs. One of these overarching issues is the risk of bias, which depends on the design, conduct, and analysis of a study and is useful for assessing whether evidence is likely to support a conclusion about a causal relation. Randomized controlled trials (RCTs) when they are well conducted and have adequate statistical power can minimize or eliminate many sources of bias, whereas observational studies are more vulnerable to confounding and sample-selection bias. Causality can be directly assessed with RCTs but must be inferred or its likelihood assessed from observational studies.

In RCTs, the food-substance intervention is known. Randomization increases the likelihood that measurement error or bias associated with dietary intake assessment will be evenly distributed among the groups. In contrast, assessing relations between food substances and chronic diseases in observational studies is particularly challenging because the assessment of intake is most often based on self-reported dietary intakes, which are subject to systematic bias, particularly intakes of energy. Unlike RCTs, in which valid comparisons among randomly assigned groups are possible without the use of dietary-assessment data, the validity and usefulness of observational studies depend on the accuracy and precision of the dietary assessments these studies use. Systematic reviews and meta-analyses, when they are well designed, can provide useful and well-documented summaries of the evidence on a relation between food substances and chronic diseases. However, the use of data from such analyses also requires caution because these analyses have the same biases and confounding problems as the original studies.

Which outcome measures a DRI committee selects for assessing the causality of a relation between food substances and chronic diseases is also important. It is possible to measure the occurrence of a chronic disease of interest directly or indirectly. Confidence that an observed relation between a food substance and a chronic disease outcome is causal is greatest when a study directly measures the chronic disease event or incidence. An indirect measurement involves a substitute measure (e.g., a qualified surrogate disease marker such as LDL cholesterol or a nonqualified disease marker such as carotid intima-media thickness for coronary heart disease). Some uncertainty is associated with the use of qualified surrogate disease markers, and considerable uncertainty is associated with the use of nonqualified disease markers as outcome measures.

Tools for assessing the evidence

Tools are available to assess *1*) individual study quality and *2*) the overall strength of the totality of the evidence. Tools to assess individual study quality include the Bradford Hill criteria,

quality-assessment instruments, and risk-of-bias tools. Quality-assessment instruments, such as the Scottish Intercollegiate Guidelines Network 50 (SIGN 50) methodology, assess the quality of a study from conception to interpretation. Risk-of bias tools assess the accuracy of estimates of benefit and risk in RCTs and nonrandomized studies. Other tools evaluate the quality of systematic reviews and meta-analyses [e.g., A Measurement Tool to Assess Systematic Reviews (AMSTAR)] or provide criteria for grading the evidence [e.g., Grading of Recommendations Assessment, Development, and Evaluation (GRADE)]. For DRI applications, reviewers might need to add nutrition-specific measures to generic assessment tools when they evaluate relations between food substances and chronic diseases (e.g., information on baseline or background nutritional status, assay methods used to measure biomarkers).

Options for addressing evidence-related challenges

An early challenge in the DRI decision-making process is the identification of potentially useful measures (indicators) that reflect a health outcome associated with the food substance of interest. One option is to select an endpoint that is assessed as the chronic disease event (i.e., chronic disease defined by accepted diagnostic criteria) or by a qualified surrogate disease marker (e.g., LDL cholesterol for coronary heart disease). An alternative option would expand the types of outcome measures of chronic disease to include nonqualified disease markers. This would increase the number of relations between food substances and chronic disease outcomes for which committees could establish DRIs but is associated with considerable uncertainty as to whether the relation of the food substance and the chronic disease is causal.

Another challenge is to specify the acceptable level of confidence in the data that a DRI committee uses to establish causality. The level of confidence is based on the type of endpoint measured and the overall strength of the evidence. One option is to specify an acceptable level of confidence in (e.g., high or moderate) about the validity of the results that must be met before a reference value can be established. Another option is to use the actual level of certainty (e.g., high, moderate, or low) to describe the evidence associated with a given reference value. A final option is to let committees make this decision on a case-by-case basis.

Intake-response relations

Intake-response relations for classical nutrient requirements and adverse events associated with excessive intakes differ from those associated with chronic diseases. Traditional deficiency relations are based on absolute risk, in which an inadequate intake of the nutrient is both necessary and sufficient to cause a deficiency and an adequate intake is both necessary and sufficient to treat a deficiency. The intake-response relation between a nutrient and a deficiency disease is linear or monotonic within the range of inadequacy. In contrast, food substance–chronic disease relations are often expressed as relative risks, in which the baseline risk of a chronic disease is never zero and changes in intake may alter risk by relatively small amounts. In addition, reductions in relative risk are achievable through >1 intervention, which means that the food substance of interest may not be necessary or sufficient to increase or decrease the relative risk of the disease. The relation between a food substance and a chronic disease indicator can be diverse (e.g., linear, monotonic, or nonmonotonic). A single food substance can have a causal relation with >1 chronic disease, and intake-response curves for these different relations can differ.

Options for determining an acceptable level of confidence

Several options are available for determining the acceptable level of confidence in the data that a DRI committee uses to determine intake-response relations once it has data that establish a causal relation. One option is to require a high level of confidence by, for example, using RCTs with a chronic disease or qualified surrogate disease marker as the outcome measure. Another option is to accept a moderate level of confidence in the data, which would allow for inclusion of data on chronic disease outcomes or qualified surrogate markers of disease from observational studies. A third option is to "piece together" different relations in which the outcome marker of interest is a common factor when direct evidence of the outcome marker's presence on the causal pathway between the food substance and a chronic disease is lacking. Therefore, if data show a quantitative relation between a food-substance intake and the outcome marker of interest and other data show a quantitative relation between the outcome marker of interest and the chronic disease, this evidence could be combined to establish a quantitative reference intake value for the chronic disease risk, if the confidence in the data is at an acceptable level.

Options for types of reference values

If data for an acceptable level of confidence are available, a reference value based on chronic disease risk reduction can be determined. The challenges presented by the use of chronic disease endpoints to set reference values by using the traditional framework suggest the need for different types of reference values than are used for classical nutrient deficiencies and toxicities. For cases in which increasing intakes will reduce the risk of a chronic disease, one option is to estimate a chronic disease risk-reduction intake value [e.g., a chronic disease risk-reduction intake value, such as a chronic disease (CD) value for reduced cardiovascular disease (CVD) reduction, could be denoted as CD_{CVD}] that is specific to a chronic disease outcome and is based on data reported as relative rather than absolute risk. Within this type of approach, 3 possible adaptations are identified: *1*) set a single chronic disease value at a level above which higher intakes are unlikely to achieve additional risk reduction for a specified disease (i.e., point estimate), *2*) set multiple reference values in relation to the expected degree of disease risk reduction across a spectrum of intakes to give a "family of targeted reductions," or *3*) set multiple chronic disease–related values (e.g., CD_{CVD}, CD_{cancer}) if the food substance is related to multiple diseases at different intakes. Another option is to express reference intakes as ranges of beneficial intakes.

Options for the derivation of Tolerable Upper Intake Levels (ULs) include the use of either one or both traditional adverse events (i.e., toxicities) and chronic disease endpoints, depending on the nature and strength of available evidence. One option is to derive ULs on the basis of a threshold approach by using traditional adverse events, if the UL based on chronic disease risk would be higher than a UL associated with a traditional adverse

effect. A second option is to use chronic disease endpoints to set a UL in cases in which intakes associated with increased chronic disease risk are at a level below those associated with traditional adverse events. These values could be denoted as a chronic disease UL (UL_{CD}) to distinguish them from a traditional UL. For this second option, approaches analogous to the derivation of CD values (e.g., the development of 1 or multiple values for specified levels of relative risk) or a threshold approach (e.g., identifying the inflection point at which absolute or relative risk increases) could be used. When increased chronic disease risks are observed over a range of intakes and the intake-response curve shows an inflection point that supports a threshold effect, the inflection point could be set as a UL_{CD}. If there is no clear inflection point, then a single UL_{CD} value or a set of UL_{CD} values could be based on intakes that reduce risk at specified levels with the acknowledgment that it may not be possible to eliminate the targeted risk. Basing UL_{CD} values on risk reduction or minimization rather than risk elimination would further differentiate UL_{CD} values from traditional UL values. Such an option would entail the provision of adequate guidance to users with regard to their uses and application. A third option is to develop multiple values on the basis of both traditional adverse events and chronic disease endpoints with guidance provided to users with regard to the strengths and weaknesses of derived values, and examples of their appropriate uses. For all options, the feasibility of avoiding or minimizing the food substance in the diet must be considered when there is no threshold for risk.

Options for resolving overlaps between benefit and harm

Intake distributions for some food substances associated with disease risk reduction might overlap with intake distributions associated with adverse events, including higher chronic disease risk. Several descriptive options are proposed for dealing with this issue. One option is to ensure that no point estimate or range of beneficial intakes for chronic disease risk reduction extends beyond the intake at which the risk of adverse events, including chronic diseases, increases. A second option is to predetermine criteria related to the severity and prevalence of targeted chronic diseases and the degree of change in the risk of specified intakes required to set a reference value. A third option is to simply describe the nature of the evidence and the public health implications of benefits and risks across the full range of intakes in which inferences are reasonably possible together with remaining uncertainties. Users would choose an appropriate balance between benefit and harm for the population of concern.

Options for selecting an indicator or indicators and specifying intake-response relations

Several possible options are identified to address examples of challenges likely to be encountered when intake-response curves are based on chronic disease endpoints. One possible approach is to identify alternatives for addressing different types of outcome markers [e.g., chronic diseases defined by accepted diagnostic criteria (clinical diseases per se) compared with qualified surrogate disease markers and nonqualified disease markers] to derive intake-response relations. In this approach, several possible options are identified. One option is to select a single outcome indicator on the causal pathway, provided that it is sufficiently sensitive to quantify the relation between the food substance and the chronic disease. Another option is to integrate information from multiple indicators for a given chronic disease if they add substantially to the accuracy of the intake-response relation and reference value variation. A third option may be required when a single food substance is related to multiple chronic disease outcomes, each with a distinct intake-response relation. In this case, criteria for selecting appropriate endpoints or surrogate endpoints to establish intake-response relations, methods to integrate multiple endpoints, and methods to account for interindividual variability in the relations of interest need to be developed. Another option is to use a biological mode-of-action framework instead of a statistical approach in establishing quantitative reference intakes.

In applying these possible approaches, several factors that influence or confound quantitative intake-response relations need to be considered. The accuracy of intake-response relations is dependent on the accuracy of the measurements of intakes and outcomes. Systematic bias due to substantial underreporting (e.g., intakes, particularly energy intakes) is of particular concern. When available, the use of qualified and accurately measured biomarkers of nutrient and food-substance intakes may overcome biases in self-reported intakes. Another factor relates to the common problem of data being available on some, but not all, life-stage groups for which DRIs are established. Two options for dealing with this issue are identified, including limiting the establishment of DRI values based on chronic disease endpoints to populations that are identical or similar to the studied groups. Alternatively, extrapolation could be considered when sufficient evidence is available that specific intakes of a food substance can increase or decrease the risk of a chronic disease.

DRI process

Arguments for or against including chronic disease endpoints in future DRIs

Evidence-based reference intake values and/or recommendations with regard to food substances causally related to the chronic diseases are desirable from public health and clinical perspectives. Yet, despite the growing chronic disease burden and continued use of DRIs, substantial challenges persist related to both the paucity of sufficiently relevant and robust evidence for evaluating putative causal relations between intakes and a chronic disease and the often-poor fit of the current Estimated Average Requirement (EAR)/Recommended Dietary Allowance (RDA) and UL frameworks for deriving DRIs on the basis of chronic disease endpoints. There is a clear desire to include chronic disease endpoints in the DRIs; however, the challenges reviewed in this report underscore the fact that the broader incorporation of chronic disease endpoints requires more sophisticated approaches than those previously used. These must also include approaches to issues concerning processes and starting points.

Options for process components

The current DRI values were set by a process that reviews a group of essential nutrients and related food substances and clearly focuses on intakes required for health maintenance and chronic disease risk reduction. Two possible options for organizing future reviews and derivations of DRIs based on chronic

disease endpoints are identified. The first option is to continue incorporating chronic disease endpoint considerations in future DRI reviews but to expand the types of reference values that could be set, while clearly differentiating between values based on classical nutrient adequacy and chronic disease endpoints. A second option is to create 2 separate but complementary, and possibly iterative and/or integrated, processes for the development of reference values on the basis of chronic disease endpoints and/or deficiency diseases. For example, a review is initiated specifically to set DRIs on the basis of chronic disease endpoints or when an existing independent process could be used.

Options for starting point

The starting point of current DRI processes is individual food substances, and all pertinent outcomes related to varying intakes of given food substances are considered. If chronic disease endpoints are to be considered, one option is to focus on individual food substances or small groups of interrelated nutrients, an approach that is similar to the current DRI process. Conversely, another option is to focus on a specific chronic disease and its relation with multiple food substances.

Forthcoming tools

Examples are discussed of forthcoming tools and novel study designs with potential utility in overcoming anticipated hurdles, such as complexities related to multiple, interactive etiologies and longitudinal characteristics of chronic diseases. These include the identification and use of new dietary intake biomarkers, the potential for the use of Mendelian randomization studies to inform causality, the use of U-shaped dose-risk relation modeling based on severity scoring and categorical regression analysis, consideration of enhanced function endpoints, the use of systems science, and the application of principles subsumed under the umbrella of precision medicine.

Conclusions

The development of the DRIs has proven to be critical for the successful elimination of diseases of deficiency in Canada and the United States. If the DRI framework could be improved to more effectively incorporate chronic disease outcomes, the potential impact on public health would be even greater. The next steps are to assess the feasibility of including chronic disease endpoints in future DRI reviews, to evaluate the relevance and appropriateness of expanding DRIs to populations beyond those currently targeted, and to determine which of the options and/or their adaptations identified in this report may warrant inclusion in a future chronic disease DRI framework.

II. BACKGROUND

DRIs are a common set of reference intake values that the Canadian and US governments, individuals, and organizations use for planning and assessing the diets of apparently healthy individuals and populations (1–3). The Food and Nutrition Board (FNB) periodically convenes ad hoc expert committees to develop DRIs for specified food substances. DRIs are guides for achieving safe and adequate intakes of nutrients and other food substances from foods and dietary supplements. The DRI committees establish DRIs within a public health context for the prevention of nutrient deficiencies, for reduction in risk of other diseases, and for the avoidance of potential adverse effects of excessive intakes. DRIs are available for 22 groups based on age, sex, pregnancy, and lactation in apparently healthy populations. Future DRI committees might need to review whether the population coverage should be expanded to include morbidities of high prevalence.

Text Box 1
Food substances consist of nutrients that are essential or conditionally essential, energy nutrients, or other naturally occurring bioactive food components.

The definition of "food substances" for this report is provided in **Text Box 1**. Future DRI committees might find it useful to review and revise this definition.

Previous DRI committees have used the term "apparently healthy populations" as defined in **Text Box 2**.

There is no single uniform definition of "chronic disease" (4) and defining this concept for DRI evaluations, although highly relevant, is outside this project's scope. Future DRI committees will probably need to define this term. Existing definitions of this term differ with respect to whether a chronic disease requires medical attention, affects function, has multiple risk factors, or can be cured. There are many definitions of chronic disease, several examples of which are shown in **Text Box 3**.

History of nutrient intake reference values

The establishment of quantitative nutrient intake reference values in the United States and Canada began around 1940 with a single type of reference value in each country: *1*) the Recommended Nutrient Intakes, or RNIs, for Canadians and *2*) the RDAs for the United States (1). These values were the intakes of essential nutrients that the experts who developed them expected would meet the known nutrient needs of practically all healthy persons.

In 1994, an FNB committee recommended that future intake reference values reflect more explicit statistical constructs of distributions of requirements across individuals (7). As a result, DRI committees began deriving reference values from population-specific estimates of average requirements (EARs) and associated population variability (RDAs) (1, 3). This approach allowed DRI users to calculate the prevalence of inadequacy in populations and the probability of inadequacy in individuals (1, 8–10).

Text Box 2
DRIs are reference intakes for *apparently healthy populations*. DRI intake levels are not necessarily sufficient for individuals who are malnourished, have diseases that result in malabsorption or dialysis treatments, or have increased or decreased energy needs because of disability or decreased mobility (1).

Text Box 3
Examples of definitions of chronic diseases

WHO: Noncommunicable diseases, also known as chronic diseases, are not passed from person to person. They are of long duration and generally slow progression. The 4 main types of noncommunicable diseases are CVDs, cancers, chronic respiratory diseases, and diabetes (5).

US Department of Health and Human Services: Chronic illnesses are conditions that last ≥1 y and require ongoing medical attention and/or limit activities of daily living (4).

Institute of Medicine Biomarkers Committee: A chronic disease is a culmination of a series of pathogenic processes in response to internal or external stimuli over time that results in a clinical diagnosis or ailment and health outcomes (e.g., diabetes) (6).

The FNB committee also recommended adding a reference value that reflects an upper safe level of intake (UL) (7, 11). All DRI reports published after 1996 implemented these recommendations (**Table 1**). However, with the progressive implementation of the revised DRI process, the committees that produced these reports recognized that the EAR and RDA model and the UL model were inappropriate for some outcomes of interest. Therefore, DRI committees added new reference values, as follows: *1*) Adequate Intake (AI), *2*) Acceptable Macronutrient Distribution Range (AMDR), and *3*) Estimated Energy Requirement, or EER (Table 1).

In response to evolving science that suggests beneficial effects of diets and dietary components in reducing the risk of chronic disease (12), the 1994 FNB committee also recommended that DRI committees include reduction in the risk of chronic disease in the formulation of future reference values when sufficient data on efficacy and safety are available (7). All 7 subsequently published DRI reports placed a high priority on an evaluation of potential chronic disease endpoints for all of the nutrients they reviewed (13, 14). However, these panels based only a limited number of DRIs on chronic disease endpoints: dietary fiber and coronary heart disease, fluoride and dental caries, potassium and both hypertension and kidney stones, and sodium and CVD (15).

Uses of DRIs

The uses of reference intake values have expanded considerably beyond the original intent of helping governments plan and evaluate nutrition programs and policies. Uses now include general nutrition education and guidance for the public, dietary management of clinical patients, identification of research gaps and priorities, research design and interpretation, food product development, regulatory applications, and guidance for international and other organizational reference values.

The evolving range of diverse uses and users of reference intake values underscores the need for the transparent documentation of scientific decisions made by DRI committees and for reference intake values that lend themselves to a wide range of applications. DRI reports focus on the scientific and public health aspects of the intakes of nutrients and food substances, but they do not make policy recommendations, with one notable exception. The 1997 amendments to the US Food, Drug, and Cosmetic Act mandated that food manufacturers could use "authoritative statements" from certain scientific bodies, including the National Academies of Sciences, Engineering, and Medicine, as health claims on food labels in the US marketplace without undergoing usual US Food and Drug Administration review and authorization procedures (16). This latter policy is not operative in Canada.

TABLE 1
DRIs and their definitions[1]

DRIs	Definition
Based on 1994 Food and Nutrition Committee recommendations	
EAR	The average daily nutrient intake level that is estimated to meet the requirements of half of the healthy individuals in a particular life stage and sex group.
RDA	The average daily dietary nutrient intake level that is sufficient to meet the nutrient requirements of nearly all (97–98%) healthy individuals in a particular life stage and sex group.
UL	The highest average daily nutrient intake level that is likely to pose no risk of adverse health effects for almost all individuals in the general population. As intake increases above the UL, the potential risk of adverse effects may increase.
Added by DRI committees in 1994–2011	
AI	The recommended average daily intake level based on observed or experimentally determined approximations or estimates of nutrient intake by a group (or groups) of apparently healthy people that are assumed to be adequate; used when an RDA cannot be determined.
AMDR	The range of intakes of an energy source that is associated with a reduced risk of chronic disease, yet can provide adequate amounts of essential nutrients; expressed as a percentage of total energy intake.
EER	The average dietary energy intake that is predicted to maintain energy balance in a healthy adult of a defined age, sex, weight, height, and level of physical activity consistent with good health. In children and pregnant and lactating women, the EER includes the needs associated with the deposition of tissues or the secretion of milk at rates consistent with good health.

[1] From reference 1. AI, Adequate Intake; AMDR, Acceptable Macronutrient Distribution Range; DRI, Dietary Reference Intake; EAR, Estimated Average Requirement; EER, Estimated Energy Requirement; RDA, Recommended Dietary Allowance; UL, Tolerable Upper Intake Level.

Report overview

This report, in section III, provides an overview of the current project, whose purpose is to critically evaluate key scientific challenges in the use of chronic disease endpoints to establish reference intake values. Section IV describes the framework that the working group used as background information for this project. Sections V-A, V-B, and V-C describe options that the working group identified to assess evidentiary challenges related to determining whether relations between food substances and targeted chronic diseases are causal. Options for establishing intake-response relations between food substances and chronic disease endpoints are the focus of section VI. Section VII addresses considerations for future DRI committee processes, and section VIII discusses some forthcoming tools that could be applied to the establishment or application of DRI values based on chronic disease endpoints. Section IX offers a few conclusions and next steps.

III. CURRENT PROJECT

This section describes the rationale for this project as well as its objectives and key questions. Motivations for the project were well-established links between diet and health throughout the life course and the expectation that evidence-based changes in the intakes of food substances would enhance well-being and reduce disease risk. The broad application of reference intake values, increasing rates of chronic diseases among US and Canadian populations, growing financial and quality-of-life burdens represented by that dynamic, and shortcomings of the EAR/RDA and UL models provided additional reasons to undertake this effort.

Several US and Canadian government agencies are continuing DRI-related harmonization efforts initiated in the mid-1990s by jointly sponsoring the current project. These agencies convened a working group with a broad and diverse range of scientific and DRI experience (**Table 2**). The group had numerous discussions via conference calls and at a public workshop (17). The sponsors also solicited public comment on the working group deliberations.

The focus of the current project was on the relation between food-substance intakes and chronic disease endpoints. The working group applied elements of the traditional DRI-related context to its work: a prevention (public health) orientation, intakes that are achievable within a dietary context (and, in a few highly selected cases, through dietary supplements, such as folate supplements during pregnancy), and primary applicability to the apparently healthy population.

Objectives

One objective of this project was to critically evaluate key scientific issues involved in the use of chronic disease endpoints to establish reference intake values. A second objective was to provide options for future decisions about whether and/or how to incorporate chronic disease endpoints into the process for establishing DRI values. The sponsors asked the working group not to try to reach consensus on which options were best, but rather, to identify a range of options and their strengths and weaknesses. None of the options in this report excludes other possibilities, and the order of presentation or amount of space devoted to each option is not intended to convey relative priorities. Subsequent expert groups will make final decisions about future DRI approaches to chronic disease endpoints. The key scientific decisions that are the backbone of DRI development (**Table 3**) provided context for the working group's discussions.

The working group identified a (not necessarily exhaustive) range of options for answering each of 3 key questions and identifying strengths and weaknesses of each option, while keeping in mind current and future DRI uses. The key questions are listed in the following sections.

TABLE 2
Working group members and their institutions

Working group member	Institution
Jamy D Ard, MD	Associate Professor, Wake Forest School of Medicine, Wake Forest University
Stephanie Atkinson, PhD, FCAHS	Professor, Department of Pediatrics, McMaster University
Dennis M Bier, MD	Professor of Pediatrics, and Director, Children's Nutrition Research Center, Baylor College of Medicine
Alicia L Carriquiry, PhD	Distinguished Professor, Department of Statistics, Iowa State University
Cutberto Garza, MD, PhD (Chair)	Professor, Boston College, and Visiting Professor, George Washington University Milken Institute School of Public Health and Johns Hopkins University
William R Harlan, MD, FACP, FACPM, FAAFP, FAHA	Research Consultant (retired), NIH
Dale B Hattis, PhD	Research Professor, The George Perkins Marsh Institute, Clark University
Janet C King, PhD	Executive Director, Children's Hospital Oakland Research Institute, and Professor Emeritus, University of California, Berkeley and Davis
Daniel Krewski, PhD	Professor and Director, McLaughlin Centre for Population Health Risk Assessment, University of Ottawa
Deborah L O'Connor, PhD, RD	Professor, Department of Nutritional Sciences, University of Toronto, and Senior Associate Scientist, The Hospital for Sick Children
Ross L Prentice, PhD	Member, Public Health Sciences Division, Fred Hutchinson Cancer Research Center, and Professor of Biostatistics, University of Washington
Joseph V Rodricks, PhD, DABT	Principal, Ramboll-Environ International Corporation
George A Wells, PhD, MSc	Professor, Department of Epidemiology and Community Medicine, University of Ottawa Heart Institute

TABLE 3
DRI decisions and considerations[1]

1. **Causality: Is the relation between the food substance and the chronic disease or diseases causal?**
   a. Objective assessment of the relevance and robustness of available studies
   b. Clear identification of the putative benefit or increased risk ascribed to targeted food substance or substances (e.g., amelioration or exacerbation of absolute or relative risks, level of severity)
   c. Selection of candidate chronic disease outcomes (e.g., chronic disease event, surrogate disease marker, nonqualified outcome) that reflects targeted causal relations
   d. Delineation of uncertainties related to determination of causality
   e. Evaluation of challenges likely to be encountered because of the extrapolation of causality from studied to unstudied groups
2. **Intake-response relation: What is an appropriate DRI value (provided that causality has already been determined)?**
   a. Objective assessment of the relevance and robustness of available evidence
   b. Determination of the type of reference value that is most appropriate given the available data (e.g., mean ± variances, ranges) and user needs (e.g., planning or assessment for individuals or groups)
   c. Selection of candidate indicators for establishing an intake-response relation (i.e., endpoints for quantification)
      i. What are the complexities of the intake-response relation (e.g., linear, curvilinear, overlapping of benefit, or increased risk curves)?
      ii. What are the characteristics of possible indicators (e.g., chronic disease event or biomarker relative to the causal pathway between intake and the chronic disease)?
   d. Identification of statistical models or other approaches (e.g., statistical, population-derived) to quantify the relation
   e. Delineation of uncertainties in the available data
   f. Identification of adjustments that may be necessary (e.g., about bioavailability, bias in exposure, outcome measures)
   g. Evaluation of challenges likely to be encountered in the extrapolation of a reference intake value from studied to unstudied groups

[1] Evaluations of the effect of increasing intakes on both benefit (i.e., decreased risk of chronic disease) and safety (i.e., increased risk of chronic disease) as intakes increase are a core part of the DRI review process. Although DRI committees often review benefit and safety separately, the generic nature of the issues they must address in their review are likely to be the same for both types of review. This report focuses on the key questions related to causality and the intake-response relation. DRI, Dietary Reference Intake.

Key question 1: What are the important evidentiary challenges for selecting and using chronic disease endpoints in future DRI reviews?

The types of scientific evidence in the DRI-development process that are necessary to establish the essentiality of nutrients differ from the type of evidence needed to evaluate relations between food substances and chronic diseases (7). A key challenge is the limited availability of RCTs that are designed to establish that a food substance of interest is causally related to a given chronic disease outcome. A much larger body of evidence based on prospective cohort and other observational studies is available that shows associations between food substances and chronic diseases, but common study design limitations in such instances make it challenging to determine causality (18). The availability of studies that measured functional and other intermediate biomarkers (including qualified surrogate disease markers and nonqualified disease markers) of chronic disease risk has strengthened the ability to determine the utility of different study designs and endpoints for accurately predicting the impact of reference intakes on chronic disease outcomes (6).

The availability of recently developed evaluation tools and techniques (e.g., SIGN 50 methodology) (19) and grading tools (e.g., GRADE) (20) have enhanced the ability to assess the quality of individual studies and the overall strength of the totality of the available evidence. Although developers did not design and validate these types of tools for DRI applications (21), DRI committees can adapt them for DRI applications to help address the evidentiary challenges that are discussed more fully in sections V-A, V-B, and V-C.

A re-evaluation of the appropriateness of chronic disease endpoints and development of criteria for their use is timely because of the substantive knowledge base that has emerged in recent decades on relations between food substances and chronic diseases. The working group focused on options for addressing evidentiary challenges that future DRI committees must consider when they evaluate and select chronic disease endpoints.

Key question 2: What intake-response models can future DRI committees consider when using chronic disease endpoints?

The DRI intake-response relation models best equipped to deal with deficiency endpoints often are not appropriate for chronic disease endpoints (13, 22). For the purpose of this report, "intake" refers to intake exposure to a food substance. "Intake-response relation" refers to the impact on physiologic processes of a range of dietary intakes. Related challenges include difficulties in the use of nutrient-status indicators (e.g., serum nutrient concentrations) to estimate optimal intakes on the basis of chronic disease endpoints. In addition, it is often difficult to use the relative risk data commonly available on relations between food substances and chronic diseases to calculate a population average and variance, as is necessary for deriving EARs and RDAs. DRI committees have generally found AIs to be useful for deriving chronic disease endpoints, but DRI users have found AIs difficult to apply when assessing and planning diets for groups (13).

DRI committees have also encountered challenges in basing ULs on chronic disease endpoints. These committees did find convincing evidence that higher intakes of several food substances were associated with increased risks of certain chronic diseases. However, the absence of an apparent threshold effect for the associated intake-response relations resulted in either failure to establish a UL or the establishment of an arbitrary UL on the basis of considerations other than the traditional model for establishing DRIs (23, 24). It is therefore important to identify other approaches and models for deriving quantitative reference values that are related to both benefits and risks of food-substance intakes for chronic disease outcomes.

Key question 3: What are the arguments for and against continuing to include chronic disease endpoints in future DRI reviews?

The 1994 FNB committee was concerned about the need to consider differences among relations between nutrients and diseases of deficiency compared with those between food substances and chronic diseases in decisions about whether to combine these 2 types of relations or to address them separately (7). Subsequent evaluations of the DRI process have continued to

question whether a single process or separate processes are most appropriate for this purpose (13, 22).

IV. CURRENT PROJECT FRAMEWORK

This section describes the framework that the working group used in its reviews and deliberations. Chronic diseases are the leading cause of death and disability in the United States and Canada, and they account for a major proportion of health care costs (25, 26). Globally, 38 million people die annually from chronic diseases, and almost three-quarters of these deaths occur in low- and middle-income countries (5). With changing demographics (e.g., aging populations) and increasing rates of overweight and obesity, public health concerns and costs related to chronic diseases are expected to increase further in the coming decades.

Published evidence shows that "healthy" dietary choices and lifestyles can help prevent or control several chronic diseases (27). The technological capabilities of assessing individual and population risks of chronic diseases and options for modifying foods and behaviors that affect diets are likely to expand. At the same time, the understanding of the development of chronic diseases through the life course is increasing.

The evaluation of relations between food substances and chronic diseases is complex, and a single conceptual model is unlikely to fit all cases. Chronic diseases are generally considered to be pathologic processes that are noncommunicable, of long duration, of slow progression, and of multifactorial etiologies, which, in turn, may be influenced by genetic backgrounds, age and sex, comorbidities, environments, lifestyles, and an increasing prevalence of obesity (5, 25). They represent a wide range of conditions, including heart disease, cancer, arthritis, diabetes, and macular degeneration. Chronic diseases have varying public health importance, severity, prevalence, and availability of effective treatments and prevention strategies. These diseases begin years before signs and symptoms become evident with the use of current diagnostic technologies. Complex factors interact to influence chronic disease progression, including interactions between food substances. In some cases, one factor (e.g., a particular food substance) may only exert an effect if other factors are also present or absent. Food-substance effects are often small in individuals but can have significant beneficial or detrimental effects on populations. Defining populations at risk of a chronic disease is also challenging because many diseases are associated with, or modified by, other morbidities (e.g., obesity is associated with several comorbidities in the elderly) and demographic characteristics (e.g., proportions of individuals aged ≥65 y and changing pharmaceutical uses).

Because the human diet is a complex mixture of interacting components that cumulatively affect health (28), isolating the effects on chronic disease risk of a single food substance or a small number of them can be challenging. In addition, the risks of chronic disease can be associated with either decreasing or increasing intakes of food substances (e.g., of fiber or saturated fat, respectively). The observed intake-response characteristics generally do not fit the threshold-based EAR/RDA and UL approaches that are based on absolute risk and that DRI committees use to set reference values for nutrient deficiencies and related toxicities (22).

Intake-response curves have varied shapes. Both high and low intakes of some substances may increase the risk of a chronic disease, and high and low intakes of the same food substance sometimes have overlapping effects [e.g., the intake-response curve for the decreasing effect of increasing fluoride intakes on dental caries overlaps with the intake-response curve for the effect of increasing fluoride intakes on fluorosis (29)]. Observational data suggest that a given food substance can be related to multiple chronic disease outcomes, and each relation can have its own distinctive intake-response curve (22, 30). These complexities indicate the need for a multidisciplinary approach to developing nutrient-specific and context-specific frameworks that involves scientists with a wide range of expertise.

It is useful to compare the reference value concepts traditionally used for nutrient requirements and toxicities with the concepts that pertain to chronic disease risk reduction (**Table 4**).

TABLE 4
Traditional and chronic disease endpoints for DRIs[1]

Issue	Eligibility for consideration	Focus	Characteristics	Expression of risk
Traditional endpoints	Food substances that are essential or conditionally essential or that are components of energy nutrients (e.g., fats, proteins, and carbohydrates).	Nutrient requirements	Adequate intakes are essential for preventing and treating deficiency diseases.	Average inflection point between adequate and inadequate intakes (EAR) of a group and its associated population variance (RDA).
		Nutrient toxicities	Intakes at some level above adequate intakes may pose the risk of adverse health effects.	Highest intake of a group that is unlikely to pose a risk of adverse effects and above which the risk of adverse effects increases (UL).
Chronic-disease endpoints	Naturally occurring food substances, including nutrients, for which changes in intake have been demonstrated to have a causal relationship to the risk of one or more chronic diseases.	↑Intakes of "beneficial" substances	With ↑ intakes, the relative risk ↓ compared with baseline intakes.	Relative risk (ratio of the probability of an event occurring in a group with higher intakes to the probability of an event in a comparison group with lower intakes).
		↓Intakes of "harmful" substances	With ↓ intakes, the relative risk ↓ compared with baseline intakes.	Relative risk (ratio of the probability of an event occurring in a group with lower intakes to the probability of an event in a comparison group with higher intakes).

[1] DRI, Dietary Reference Intake; EAR, Estimated Average Requirement; RDA, Recommended Dietary Allowance; UL, Tolerable Upper Intake Level; ↑, increased or increases; ↓, decreased or decreases.

Historically, the food substances for which expert panels established reference values tended to be essential or conditionally essential nutrients or those that supplied energy (31). With its inception, the DRI-development process broadened this concept to include food substances with documented effects on chronic disease risk (e.g., fiber, saturated fats, and *trans* fats). Today, there is considerable interest in expanding future DRIs to include other bioactive components with documented health effects (32–34). Although essential nutrients have a direct and specific effect on nutrient deficiencies, other food substances alone might be neither necessary nor sufficient to reduce disease risk. Even if research has established a causal relation between a food substance and a chronic disease outcome, the mechanisms of action are often unknown or poorly understood. Research results on chronic disease risks are often expressed as relative risks as opposed to the reporting of absolute risks that experts typically use to define nutrient requirements for essential nutrients. Although the evidence may be reported as relative risks, DRI decisions may also need to consider the relation of a food substance and chronic disease within an absolute risk context (35).

V-A. JUDGING THE EVIDENCE: EVIDENTIARY CHALLENGES

This section and the next 2 sections discuss ways to assess the strength of the evidence on causal relations between food substances of interest and targeted chronic diseases. This section focuses on study designs and related issues that affect the use of evidence to assess the causality of these relations in DRI evaluations.

The DRI process involves 2 key decisions: *1*) whether available evidence supports a causal relation between the food substance of interest and the chronic disease and, *2*) if so, what DRIs may be appropriately derived from the available data. DRI committees make these 2 key decisions for both beneficial and adverse effects as guided by 2 key questions and their component characteristics (Table 3). When DRI committees find causal relations between food substances and chronic diseases, they can then derive DRI values that are appropriate given the evidentiary base that supports the intake-response relations. Tolerance of uncertainty is likely to vary for decisions about beneficial compared with adverse effects and for decisions involving causal compared with intake-response relations.

Judging evidence to develop DRIs on the basis of chronic disease endpoints has been an evolutionary process that continues to present major challenges. The 1994 FNB committee noted that consideration of chronic disease endpoints often requires a different type of evidence than the evidence that committees have used for determinations of nutrient requirements on the basis of classical deficiency diseases (7). In the 6 DRI reports published between 1997 and 2005, the totality of the evidence from both observational and intervention studies, appropriately weighted, formed the basis for conclusions with regard to causal relations between food-substance intakes and chronic disease outcomes (23, 24, 29, 36–38). The 2011 DRI Committee on Calcium and Vitamin D stated that RCTs provided stronger evidential support over observational and ecologic studies and were therefore necessary for the committee to further consider a health-outcome indicator (14). This committee also considered whether evidence from published RCTs and high-quality observational studies was concordant and whether strong biological plausibility existed. The paucity of studies specifically designed to support the development of DRIs continues to be a challenge.

Overarching challenges

When a DRI committee considers the strength of the evidence for its decisions, it considers overarching challenges that apply across different types of study designs and specific study design characteristics. This section discusses 3 overarching challenges: sources of bias, selection of chronic disease outcome measures, and statistical issues.

Sources of bias

A bias consists of systematic (not random) errors in estimates of benefits or risks due to a study's design or in the collection, analysis, interpretation, reporting, publication, and/or review of data (39). Bias results in erroneous (as opposed to less precise) estimates of the effects of exposures (e.g., food substances) on outcomes (e.g., risk of chronic disease).

Evaluations of whether evidence likely supports a conclusion about causation often use risk-of-bias concepts. Risk of bias varies by study design (**Figure 1**) (40–43). At each ascending level in the pyramid in Figure 1, the quality of evidence is likely to improve (i.e., the risk of bias decreases) and the quantity of available studies usually declines. Within each level, however, quality varies by study design and implementation, which can blur the quality differences among hierarchies in the pyramid. Confidence in whether relations of interest are causally related generally increases toward the top of the pyramid.

Table 5 lists sources and types of bias that can affect nutrition studies. **Table 6** describes examples of criteria for assessing the risk of bias associated with different study types. It is possible to avoid or minimize some types of biases in the study design, conduct, and analysis stages by using, for example, double-blinding, management of confounding by matching and/or multivariable analyses, or assessment of objective exposure. A major source of bias in studies of relations between food substances and chronic diseases is the use of self-reported intake assessments (e.g., food-frequency questionnaires, 24-h recalls, or food diaries) (44). Zheng et al. (45) provided an example of the dominant influence that uncorrected nonrandom measurement error in energy intake estimates from self-reported diets may have on associations with risks of CVD, cancer, and diabetes in a cohort-study context.

Selection of chronic disease outcome measures

A second overarching challenge in evaluating the strengths and weaknesses of evidence relates to the selection of an outcome measure for assessing whether a relation between food substances and chronic diseases is causal and identifying an indicator for intake-response analysis. It is possible to measure a chronic disease outcome directly (e.g., as an incident event) or indirectly by using a substitute measure (e.g., a qualified surrogate disease marker or a nonqualified disease marker). The type of outcome measured affects the level of confidence in whether the relation between a food substance and chronic disease is causal. The selection of an indicator for deriving intake-response relations also depends on whether the indicator is on the causal pathway between the intake and the disease outcome.

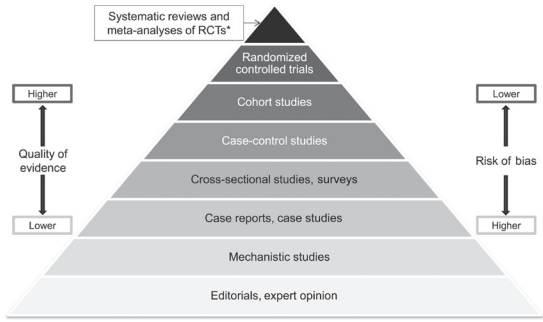


FIGURE 1 Hierarchy of evidence pyramid. The pyramidal shape qualitatively integrates the amount of evidence generally available from each type of study design and the strength of evidence expected from indicated designs. In each ascending level, the amount of available evidence generally declines. Study designs in ascending levels of the pyramid generally exhibit increased quality of evidence and reduced risk of bias. Confidence in causal relations increases at the upper levels. *Meta-analyses and systematic reviews of observational studies and mechanistic studies are also possible. RCT, randomized controlled trial.

For this report, the outcome of interest is a chronic disease. Ideally, the measured outcome in available studies consists of the incidence (event) of the chronic disease as determined by appropriate diagnostic criteria. Data on this type of outcome from an RCT provide the most direct assessment of a relation between a food substance and a chronic disease outcome and a high degree of confidence that the relation is causal (**Figure 2**).

The limiting factor is that studies that use a chronic disease outcome may not be available or even feasible, and DRI committees might then consider the use of a qualified surrogate disease marker or a nonqualified disease marker as the outcome measure. Most of these outcomes are biomarkers or are based on biomarkers, as defined in **Text Box 4**.

The types of outcomes that can substitute for direct measures of a chronic disease outcome can range from biomarkers close to the disease (e.g., blood pressure for CVD or LDL cholesterol for coronary heart disease) to those that are more distant from the disease (e.g., indicators of inflammation or immune function for CVD and cancer). One type of substitute disease outcome is the qualified surrogate disease marker, defined in **Text Box 5**, a short-term outcome measure that has the same association with the intake of a food substance as a long-term primary endpoint.

The use of a surrogate marker enables a more rapid determination of the effectiveness of changes in intake on the risk of the chronic disease. Achieving "surrogate" status requires strong evidence and a compelling context (6, 46). That is, the outcome measure must be qualified for its intended purpose (e.g., to show that changing the intake of a food substance can prevent or alter the risk of the chronic disease). A qualified surrogate marker has prognostic value (i.e., correlates with the chronic disease outcome), is on the causal pathway between the intake and the chronic disease, and substantially captures the effect of the food substance on the chronic disease. DRI committees have used LDL-cholesterol concentrations as a surrogate disease marker for coronary heart disease and blood pressure as a surrogate marker for CVD (15, 23, 24). The use of a surrogate marker instead of the incidence of a chronic disease can provide a reasonable basis, but not absolute certainty, for evaluating whether a relation between a food substance and a chronic disease is causal (Figure 2). The second type of substitute disease outcome is an outcome that has not been qualified as a surrogate disease marker, referred to in this report as a nonqualified disease marker as defined in **Text Box 6**.

An example of a nonqualified outcome for CVD is carotid intima-media thickness (47). A nonqualified outcome marker is associated with considerable uncertainty about whether the relation between a food substance and a chronic disease is causal (Figure 2).

Statistical issues

For any study design, careful interpretation of findings by experts is necessary to reach appropriate conclusions about the strength of the evidence. The use of inappropriate statistical methods (e.g., multiple statistical comparisons involving several outcomes and/or subpopulations without adjustment) can undermine the validity of conclusions. The primary outcome of an RCT and other study types is the endpoint for which the study is designed and powered and that investigators use to define inclusion and exclusion criteria. Secondary endpoints and post hoc endpoints might not have adequate statistical power, participants may not be appropriately randomized (in the case of RCTs), and participant inclusion and exclusion criteria might not be adequate for the analysis of secondary and post hoc outcomes. Importantly, reports on secondary and post hoc outcomes of RCTs and analyses of subsets of the trial cohort need to account for multiple tests of different trial hypotheses. Caution is therefore necessary in the use of secondary outcomes and post hoc analyses of RCTs or other study types when those outcomes were not part of the original study protocols.

Study designs

Past DRI committees have described how the known strengths and weaknesses of different study designs influenced their DRI

TABLE 5
Types of bias that can affect nutrition studies[1]

Bias due to confounding
- Confounding: error in the estimated effect of an exposure on an outcome due to the presence of a common cause of the outcome or to baseline differences between exposure groups in the risk factors for the outcome or because factors predicting the outcome (prognostic factors) are related to the exposure that the person experiences

Related terms
- Allocation bias: error in the estimate of an effect caused by the lack of valid random allocation of participants to the intervention and control groups in a clinical trial
- Others: selection bias, case-mix bias

Bias in selection of participants for the study
- Selection bias: systematic error resulting from participant-selection procedures and factors that influence participation, systematic differences between baseline characteristics of the groups compared, or exclusion of some participants from the analysis (i.e., some participants are excluded initially or during follow-up), thereby changing the association between the exposure and the outcome

Related terms:
- Sampling bias: systematic error due to the methods or procedures for selecting the sample (e.g., participants, scientific papers), includes errors due to sampling of a nonrandom population
- Others: inception bias, lead-time bias, immortal time bias

Bias in measurement of exposures: misclassification of exposure status or introduction of systematic bias by use of self-reported intake methodologies

Related terms:
- Dietary exposure assessment bias: error associated with the use of self-reporting tools for assessing dietary intakes
- Misclassification bias: systematic error due to inaccurate measurements or classifications of participants' exposure status; may be differential (related to the risk of the outcome) or nondifferential (unrelated to the risk of the outcome with an estimated effect that is usually biased toward the null)
- Recall bias: systematic error due to differences in accuracy of recall, particularly relevant to case-control studies because cases are more likely to recall potentially important events
- Others: observer bias, detection bias

Bias in measurement of outcomes: erroneous measurement or classification of outcomes

Related terms:
- Misclassification bias: systematic error due to inaccurate measurements or classifications of participants' outcome status
- Nondifferential measurement error: can be systematic (e.g., measurements that are all too high), which does not cause bias or affect precision, or can be random, which affects precision but does not cause bias
- Detection bias (also known as differential measurement error): systematic differences between groups in how outcomes are determined. This bias can occur when outcome assessors are aware of participants' exposure status and the outcome is subjective; the researchers use different methods to assess outcomes in different groups (e.g., questionnaires for the study group and medical records for the control group); or measurement errors are related to exposure status or a confounder of the exposure-outcome relation. Blinding of outcome assessors can help address this bias but is often not possible. Studies with self-reported outcomes have a higher risk of bias than those with clinically observed outcomes.
- Recall bias: see above

Bias in selection of reported findings
- Reporting bias: systematic differences between reported and unreported results

Related terms:
- Outcome-reporting bias: reporting on some, but not all, of the available outcome measures (e.g., reporting the most favorable results of multiple measurements or the results of the most favorable subscale of the many that are available)
- Analysis-reporting bias: investigators select results from exposure effects that they measured in multiple ways (e.g., multiple analyses with and without adjustment for different sets of potential confounders or use of a continuously scaled measure analyzed at different cutoffs)

Bias due to departures from intended exposures
- Performance bias: systematic differences between groups in care provided or in exposure to factors beyond the intended exposures
- Time-varying bias: change in the exposure over the follow-up period and postexposure prognostic factors that affect the exposure after baseline

Bias due to data missing not at random: can be due to attrition (loss to follow-up), missed appointments, incomplete data collection, or exclusion of participants from the analysis

Related terms:
- Attrition bias: systematic differences between groups in withdrawals from a study
- Selection bias: see above

Publication bias: result of the tendency for journals to publish articles with positive results, particularly if the articles report new findings, or of the tendency of authors to cite studies that conform to their or their sponsor's preconceived ideas or preferred outcomes

Conflict of interest from sponsor bias: may be incurred when there is financial conflict; sponsor participation in data collection, analysis, and interpretation of findings can compromise the validity of the findings. This may result from the choice of design and hypothesis, selective outcome reporting, inadequacy of reporting, bias in presentation of results, or publication biases.

[1] Data are from references 39 and 41–43. "Exposure" refers to the variable with the causal effect to be estimated (e.g., a food substance). In the case of a randomized controlled trial, the exposure is an intervention; "outcome" is a true state or endpoint of interest (e.g., a health condition). Lists of related terms are not intended to be exhaustive but to offer pertinent examples.

evaluations and decisions (14, 23, 24, 29, 36–38). Concurrently, evolving science provided new insights into how study designs can affect evaluations of relations between food substances and chronic diseases. Below, we integrate the perspectives of past DRI committees and newer science as to the potential usefulness of types of study designs for DRI contexts.

TABLE 6
Examples of criteria to assess the risk of bias by study type[1]

		Study type			
Type of bias	Criterion	RCT	Cohort study	Case-control study	Cross-sectional study
Bias due to confounding	Were relevant confounding factors prespecified and considered?	NA	✓	✓	✓
	Were study groups balanced with respect to the distribution of confounding factors?	NA	✓	✓	✓
	Were confounding factors taken into account in the design and/or analyses?	NA	✓	✓	✓
	Was the assignment of participants to study groups randomized?	✓	NA	NA	NA
	Was an adequate method of concealment of allocation to study groups used?	✓	NA	NA	NA
Bias in selection of participants for the study	Were the same inclusion and exclusion criteria used for all study groups?	✓	✓	✓	✓
	Was the likelihood that some participants might have the outcome before the exposure or intervention assessed and taken into account in the design and/or analysis?	✓	✓	✓	✓
	Was the percentage of eligible nonparticipants in each study group below an acceptable value?	✓	✓	✓	✓
Bias in measurement of exposures and interventions	Was the exposure or intervention status measured in an accurate and sufficiently precise way?	✓	✓	✓	✓
Bias due to departures from intended exposures and interventions	Were there systematic differences between study groups in the care provided and/or in exposures to factors beyond those intended by study design?	✓	✓	✓	✓
	Was the exposure or intervention status assessed more than once or in >1 way to help ensure fidelity to the study design?	✓	✓	✓	✓
Bias due to missing data	Was the percentage of participants dropping out in each study group below an acceptable value?	✓	✓	✓	✓
	Were missing data appropriately handled (e.g., intention-to-treat analysis, imputation)?	✓	✓	✓	✓
Bias in measurement of outcomes	Were all relevant outcomes measured in an appropriately accurate and sufficiently precise way (e.g., valid and reliable) and done consistently across all study participants?	✓	✓	✓	✓
	Was the length of follow-up among study groups in prospective studies the same, or in case-control studies were the times between exposures or interventions and targeted outcomes the same in cases and controls?	✓	✓	✓	✓
	Was the assessment of outcome made "blind" to exposure or intervention status or, when blinding was not possible, was there recognition that knowledge of exposure or intervention status could have influenced the assessment of the outcome or outcomes?	✓	✓	✓	✓
Bias in selection of the reported result	Were the prespecified outcomes partially reported or not reported because of the statistical significance or magnitude of the effect of the exposure or intervention?	✓	✓	✓	✓
	Is there evidence that the results from all participants, not only a subset, were analyzed or that all multiple-adjusted analyses, not only selected ones, were fully reported?	✓	✓	✓	✓

[1] NA, not applicable; RCT, randomized controlled trial; ✓, applicable to the study type.

RCTs

RCTs with a chronic disease event or qualified surrogate disease marker as the primary outcome. RCTs can minimize or eliminate the likelihood of some key types of bias when they use randomization, concealment, and double-blinding protocols and have adequate statistical power (14, 23, 24, 29, 36–38). It is possible to compare disease incidence among randomly assigned groups receiving different interventions (e.g., supplement compared with placebo) by using the so-called intention-to-treat analyses, without using any dietary-assessment data, thus avoiding the systematic biases associated with reliance on self-reported intakes to determine exposures in observational studies. Dietary assessments need only provide assurance that a trial has adequate precision (i.e., statistical power), and they can also provide useful background information for evaluating that adherence to interventions has been followed or to account for background intake when supplements are added. RCTs often allow testing of small effects that observational studies cannot reliably detect. RCTs usually are the only type of study that allows direct assessment of causation, although other approaches, such as Mendelian randomization, may offer an alternative in special situations (48–52).

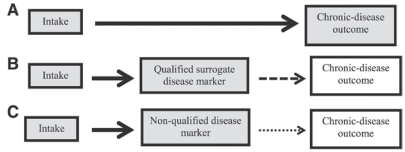


FIGURE 2 Conceptual framework for assessing causality on the basis of level of confidence that the intake–chronic disease relation is causal. Panel A: Direct assessment involving the measurement of both intake and chronic disease outcome (event or incidence); highest confidence that relation is causal. Panel B: Indirect assessment involving the measurement of a qualified surrogate disease marker as a substitute for a direct measurement of the chronic disease per se; provides a reasonable basis, but not absolute certainty, that the relation between the intake and the chronic disease is causal. Panel C: Indirect assessment involving the measurement of a nonqualified disease marker as a substitute for a direct measurement of the chronic disease; because this type of outcome measure lacks sufficient evidence to qualify as a substitute for the chronic disease of interest, there is considerable uncertainty as to whether the relation between the intake and the chronic disease is causal. Shaded boxes indicate variables and outcomes that are measured directly. Nonshaded boxes indicate variables or outcomes that are not measured but whose presence on the causal pathway is inferred. Arrows indicate a unidirectional, causal relation. This type of relation can be directly assessed by randomized controlled trials. If observational studies (e.g., prospective cohort studies) are being assessed, the observed relations are associations, not causal links. Solid bold arrows indicate a relation with high confidence. Dashed arrows indicate relations with some uncertainty. Lighter arrows indicate less certainty than bolder arrows. If any part of the causal pathway between intake and chronic disease outcome has uncertainty, then the entire causal pathway has uncertainty. "Qualified" biomarkers of outcome require strong evidence that their use as substitutes for unmeasured outcomes can accurately and reliably predict the outcome of interest. "Qualification" has a contextual basis in that the evidence about its use as a substitute for an unmeasured outcome needs to be relevant to the proposed use of the biomarker (e.g., relation between food-substance intake and a chronic disease). Intakes can be assessed directly or by measurement of qualified biomarkers of intake.

RCTs have the following limitations:

- The costs are typically high for outcomes based on chronic disease events.
- Persons agreeing to undergo randomization might be a select subset of the population of interest, which limits the generalizability of trial results.
- For practical reasons, RCTs usually measure only a single or limited intake range of one food substance or a few food substances.
- The study follow-up period is typically short relative to the period of food-substance exposure preceding the initiation of the study.
- Maintaining and reporting on intervention adherence can be challenging, particularly for diet-modification studies.
- Informed-consent procedures that indicate the study purpose (e.g., to evaluate the effect of vitamin D on bone health) may lead participants to choose to consume different foods and/or supplements independently of the study intervention.
- Blinding of study participants is difficult when interventions are based on dietary changes but is more achievable when the intervention consists of dietary supplements (e.g., to deliver micronutrients).

Over the past several decades, investigators designed several large RCTs in which the primary aim was to evaluate relations between food-substance intakes and chronic disease outcomes.

Text Box 4
A *biomarker* is "a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to [a]n … intervention" (6). ("Objectively" means reliably and accurately measured.)

Examples of completed studies include trials on the relations between the following:

- β-carotene and lung cancer (53–55);
- B vitamins and CVD (56);
- vitamin E and both CVD and prostate cancer (57, 58);
- salt and blood pressure (59);
- energy and fat (combined with physical activity) and diabetes (60); and
- a low-fat diet and breast and colorectal cancer (61, 62).

The DASH (Dietary Approaches to Stop Hypertension)-Sodium trial (59) confirmed the hypothesis that sodium-intake reductions result in lower blood pressure, and the Diabetes Prevention Trial showed that diet and physical activity changes could reduce diabetes incidence (60). However, other trials either found that the food substances of interest [B vitamins and risk of CVD (56) and vitamin E and risk of CVD (58)] had no significant effect or an unexpected adverse effect on the chronic disease outcomes studied [risk of lung cancer for β-carotene (53, 54), risk of prostate cancer for vitamin E (63)].

RCTs with nonqualified disease markers as primary outcomes. Similar to RCTs that use chronic disease events or qualified surrogate markers as primary outcomes, well-designed and conducted trials that rely on nonqualified outcomes can also reduce the possibility of outcome bias. Moreover, because nonqualified disease markers often change within relatively short times after an intervention is introduced and can be readily measured, such studies require less time to produce effects and often have adequate statistical power with smaller samples than studies that target clinical disease events (e.g., cardiovascular events). As a result, well-designed RCTs that use nonqualified disease markers can be less costly than those that measure clinical disease events. The use of nonqualified disease markers to measure relations between food substances and chronic diseases is relatively common, and many more studies use such outcomes than RCTs with a clinical event or a qualified surrogate disease marker as the primary outcome. However, substantial uncertainty about whether a relation between food substances and chronic diseases is causal frequently limits the usefulness of nonqualified disease markers because of the lack of evidence that shows that these outcome measures are accurate and reliable indicators for the risk of the chronic disease of interest (Figure 2) (6, 18, 46, 64, 65). Several publications noted the need for caution in the use of these types of trials to establish causal relations between food substances and chronic disease events (14, 18).

Observational studies

Cohort studies. An extensive body of evidence from observational studies suggests that changes in intakes of some food substances can beneficially or adversely alter the risk of certain

Text Box 5
A *surrogate disease marker* (also known as a surrogate marker, surrogate endpoint, or surrogate disease outcome marker) predicts clinical benefit (or harm, or lack of benefit or harm) based on epidemiologic, therapeutic, pathophysiologic, or other scientific evidence (6). A surrogate disease marker is qualified for its intended purposes.

Text Box 6
A *nonqualified disease marker* (also known as an intermediate disease outcome marker or intermediate endpoint) is a possible predictor of a chronic disease outcome but lacks sufficient evidence to qualify as an accurate and reliable substitute for that outcome.

chronic diseases. The increasing availability of large cohort studies with long follow-up periods has increased the use of cohort studies in recent evaluations of relations between food substances and chronic disease events.

Ideally, investigators collect data from cohort studies prospectively to more optimally control the type and quality of data collected. The prospective acquisition of dietary data is particularly important because recall of past dietary intakes is subject to considerable uncertainty.

Cohort studies have several advantages for supporting the development of DRIs on the basis of chronic disease outcomes (14, 23, 24, 29, 36–38):

- Study results are frequently directly relevant to noninstitutionalized humans.
- Study populations can be large and diverse.
- Follow-up can occur over many years or even many decades.
- A range of intakes can be associated with a range of relative risks.
- Temporal relations between intakes and outcomes are less uncertain than with cross-sectional or case-control studies.

The challenges in the use of cohort studies for DRI purposes include the following:

- Prospective cohort studies are more vulnerable to confounding and selective reporting bias than are RCTs (13, 22, 40).
- Statistical adjustments may decrease but cannot totally eliminate the likelihood of confounding.
- Evidence on the causal nature of relations between exposures and outcomes cannot be directly assessed and therefore must be inferred, thus increasing uncertainty as to the validity of the results.
- Variations in food-substance intakes may be limited in homogenous cohorts, making it difficult to identify intake differences between subgroups.
- Relative risk trends often have small effects, although small effects on diseases of sufficient prevalence or severity can be substantial at the population level.
- The reliability and interpretation of observed associations depend directly on the quality of the dietary exposure assessment; systematic bias in self-reported intakes is particularly problematic.
- Factors other than variations in intakes of the food substance of interest can affect comparisons of results across time (e.g., long-term follow-up in a given cohort or comparison of studies conducted at different time periods). For example, the increasing use of statins and aspirin can affect assessments of coronary heart disease over time. Increasing intakes of fortified foods and supplements can overwhelm the effect of the food substance of interest. Investigators must account for the confounding effects of these temporal changes when evaluating relations between food substances and chronic diseases that span long time periods.
- The use of a single diet assessment in some prospective cohort studies assumes that no variation in dietary intake occurred over time.

Other types of observational studies. Other types of observational studies include case reports, ecologic and cross-sectional studies (including surveys), and case-control studies. These types of studies played an important role in generating early hypotheses about relations between nutrients and chronic diseases (12).

Case reports and case studies are descriptive studies of outcomes in individuals or small groups who are exposed to a food substance but are not compared with a control group or groups. Cross-sectional studies and ecologic studies examine a food substance and a health condition in a population at 1 point in time. In a cross-sectional study, investigators examine the food substance and health condition in each individual in the sample. In an ecologic study, investigators examine the variables of interest at an aggregated or group level, sometimes resulting in errors in association (known as "ecological fallacy"). Case-control studies are retrospective in that they enroll patients with and without a given condition and attempt to assess whether the 2 groups of participants had different exposures to a food substance or substances.

A major limitation of these types of studies is their inability to establish the temporal relation between the intake of a food substance and the appearance of a chronic disease. These types of studies remain useful for hypothesis generation, but their utility for setting DRI values is limited. Like prospective cohort studies, they are vulnerable to confounding.

Special challenges for observational studies

Measurement error in intake assessments. Unlike RCTs, observational studies, including cohort studies, require accurate dietary assessments for their validity and usefulness. A major challenge for observational studies in evaluating relations between food substances and chronic diseases is the difficulty in obtaining accurate estimates of food-substance intakes when using self-reported data (13, 66). Self-reported intake estimates result in substantial underestimation bias for energy and protein intakes, especially among overweight and obese individuals (66, 67). These systematic biases can severely distort intake-response curves. Random errors in assessing intake may also attenuate the relation between intakes of a food substance and chronic disease risk, making it difficult to detect such a relation if it exists.

Cohort studies can minimize both systematic and random aspects of intake measurement error bias by estimating food-substance intakes with the use of a biomarker of intake or dietary recovery (see **Text Box 7**) in addition to, or in place of,

self-reported intakes. However, other important sources of bias (e.g., confounding) may remain. Currently, only a small number of established biomarkers of food-substance intake (e.g., doubly labeled water for energy expenditure assessments and 24-h urinary nitrogen for assessing protein intake) satisfy the classical measurement error criteria for recovery biomarkers (67). However, these only assess intake over short periods. Biomarker-calibrated intake assessments hold promise for minimizing systematic and random errors in intake measurements, but the field needs qualified biomarkers for additional dietary components before their use can substantially affect nutritional epidemiology research (45).

Attribution to a food substance. A second challenge is the difficulty of attributing an observed effect to the food substance of interest (13, 22). In observational studies, investigators usually calculate the amounts of food substances that participants consume from self-reports of food and supplement intakes. Interactions between food substances make it difficult to determine whether an observed association between the calculated intake of a specific food substance is a causal factor or simply a marker of another food component or components within the dietary pattern.

Statistical approaches

The Rubin potential-outcomes framework is an example of a statistical approach that potentially may enhance the usefulness of observational studies by producing approximate inferences about causal links in the absence of random allocation of subjects to treatments when candidate data sets include a large number of covariates (including key covariates) and the key covariates have adequate overlap of their distributions between experimental and control groups (68). The rationale is that the covariates incorporated in the analyses might include potential confounders. However, there is no way to guarantee that all confounders were measured in an observational study, and it is possible that ≥1 important confounders are missing. Researchers need to validate these approaches for future applicability to diet and health studies.

Systematic reviews and meta-analyses

Systematic reviews. A systematic review is the application of scientific strategies to produce comprehensive and reproducible summaries of the relevant scientific literature through the systematic assembly, critical appraisal, and synthesis of all relevant studies on a specific topic. Ideally, scientists with expertise in systematic reviews (e.g., epidemiologists) collaborate with subject matter experts (e.g., nutritionists) in the planning of the review. The subject matter experts can refine the key scientific questions and the study inclusion and exclusion criteria that will guide the review, ideally with the involvement of an experienced research librarian. The systematic review experts then abstract the data and summarize their findings, generally with duplication of key screening or data-abstraction steps. Once the review is in draft form, the review team solicits peer reviews from qualified experts in the subject matter and in systematic review methodology. This approach maintains scientific rigor and independence of the systematic review while maximizing the likelihood that the review will be relevant to subject matter experts and users. This was the process that the 2011 DRI committee used for its systematic review on calcium and vitamin D (69).

Text Box 7
An *intake biomarker* (or dietary recovery biomarker) is usually a measure of metabolite recovery in urine or blood used to objectively assess the intake of a food substance over a prescribed period.

The advantages of systematic reviews include the following:

- The process is characterized by an organized and transparent methodology that locates, assembles, and evaluates a body of literature on a particular topic by using a set of specific criteria.
- The inclusion of all relevant research and the use of a priori criteria for judging study quality minimize study-related biases and enhance transparency.
- Non–content experts who search, assemble, and analyze the appropriate literature minimize the potential for study selection bias with assistance from content experts in refining the key scientific questions and in developing the inclusion and exclusion criteria.
- It is possible to apply the methodology, which was developed for RCTs, to other study types as long as those conducting the review appropriately account for biases in the analysis and interpretation of the data.

Systematic reviews also have several disadvantages, including the following:

- Researchers have not agreed on or validated selection, evaluation, and analytic criteria that are uniquely applicable to studies of relations between food substances and chronic diseases.
- The quality of published reviews can vary by *1*) the degree of adherence to consensus methods and reporting standards and *2*) the rigor applied to measures of variables related to food substances (e.g., baseline intakes and status, the effect of biases in intake estimates and biomarker assays) in the reviewed studies. Deficits can lead to the possible omission of critical information, inappropriate conclusions, and/or unbalanced dependence on expert opinion; and each of these can increase the likelihood of bias or misinterpretation (21).
- Systematic reviews will carry forward the biases (e.g., energy-based intake underestimates) of the included studies.
- Reporting and publication biases can be problematic, particularly if those conducting the reviews do not adequately account for these issues. The use of a range of effect estimates, such as ORs or relative risks, or tallies of positive and negative studies to summarize data can also lead to misleading results due to publication bias (70). Public solicitation is one approach to identify unpublished research (i.e., gray literature) for comparison to published data to help assess the potential impact of publication bias (21).

Meta-analysis. Meta-analysis uses statistical methods to combine the results of several studies to increase statistical power, improve effect estimates, and resolve disagreements and uncertainties among studies. These analyses compare and contrast

study results to identify consistent patterns and sources of disagreement.

Meta-analysis has several advantages, including the following:

- It can appropriately weight quantitative relations between food substances and chronic diseases by the precision of individual studies, yielding an overall estimate of the benefits or risks with greater precision than can be achieved with individual studies.
- It can identify differences in relations between food substances and chronic diseases across studies.

Meta-analysis has several disadvantages, including the following:

- Meta-analysis techniques might not be appropriate when considerable heterogeneity exists across the set of studies to be combined (70). Heterogeneity across studies is commonly related to factors such as differences in intake assessment, intervention protocols, population characteristics, outcome measures, and analytic procedures (70).
- Meta-analyses carry forward biases that are present in the included studies (e.g., systematic bias in energy intake estimates).
- Pooled-effect estimates can be misleading without consideration of study quality, a strong methodologic grasp of the meta-analysis techniques, extensive content knowledge of the research question, and commitment to impartial application of the approach (70).
- Reporting bias may be a problem, because less beneficial treatment effects are more often observed in unpublished than in published trials. Studies not published in English or not indexed in publication databases (e.g., Medline or Cochrane Central) might have different treatment effects than more readily available studies (70, 71).

Meta-analyses and systematic reviews can provide succinct, useful summaries of the available literature that are relevant to the research question of interest. However, the results will still require careful interpretation by experts to reach appropriate conclusions, including conclusions about causation and possible biases.

Systematic reviews and meta-analysis for nutrition-related topics

The use of systematic reviews and meta-analyses for nutrition-related topics is relatively recent (21, 72). The 2011 DRI review on vitamin D and calcium was the first to use these types of studies within a DRI context (14, 69). WHO and European Micronutrient Recommendations Aligned nutrition-related applications also use these studies (73, 74).

A relatively recent approach is for reviews to include both observational and trial data on the same relation between food substances and chronic diseases (75–77). This approach facilitates direct comparisons of results between these study designs. It is then possible to evaluate the strengths and weaknesses of each study type for a given relation between food substances and chronic diseases.

Animal and mechanistic studies

In the past, animal and mechanistic studies played an important role in establishing the essentiality of nutrients, although similar results in humans were generally necessary to confirm the findings (7, 12, 31). These studies have also been important in traditional toxicologic evaluations of environmental contaminants and food ingredients when ethical considerations precluded human testing (11). DRI committees have found that animal and mechanistic studies provided important supporting information on biological mechanisms and pathogenesis, but these committees generally did not consider such studies adequate to infer causality or to derive intake-response curves for DRIs (14, 23, 24, 29, 36–38). Moreover, until recently, few animal models were available that could adequately simulate relations between food substances and human chronic diseases.

Other evidence-related challenges

Evaluations of relations between food substances and chronic diseases pose a number of challenges in addition to those mentioned above, including those discussed below (13).

Extrapolations from studied to unstudied groups

DRI committees set reference values for 22 life-stage groups on the basis of age, sex, pregnancy, and lactation. These values are intended for apparently healthy populations. Yet, most available research does not readily fit this framework. Committees therefore need to consider whether to generalize results from studied to unstudied groups. For example, this challenge can arise when attempting to extrapolate results from the following groups:

- persons with diagnosed chronic diseases to persons without such diagnoses;
- persons with metabolic disorders that affect a substantial proportion of the general population (e.g., obesity) to healthier populations;
- one life-stage or sex group to a different life-stage or sex group (e.g., from older adults to children or from young women to pregnant females) (13); and
- a population with a single ethnic origin to a population with ethnic diversity.

Interactions between study variables

The following interactions of the food substance of interest with other study variables may make it difficult to isolate the effect of the food substance on the targeted chronic disease:

- food substance and food-substance interactions (e.g., between sodium and potassium and vitamin D and calcium);
- food substances and physiologic characteristics (e.g., responsiveness to a food substance in smokers and nonsmokers or in lean and obese individuals); and
- food substances and environmental characteristics (e.g., socioeconomic status).

Effects on responsiveness to dietary intervention and effect sizes

Various inherited and acquired subject characteristics and contextual factors may influence responsiveness to exposures of interest.

- Differences in baseline characteristics, including baseline nutritional status
- Variations in gene polymorphisms
- Duration of the observation and/or intervention

- The amount, timing, context, and nature of the food-substance exposure

These challenges can affect studies in different ways. For example, they can highlight biologically important interactions that DRI committees need to take into account when setting reference values. However, they can also lead to residual confounding not accounted for by covariate adjustment. These issues can also lead to erroneous, misleading findings that form part of the knowledge base and can misinform interpretations or comparisons of study results. Past reviews of the use of chronic disease endpoints in DRI contexts have not identified effective strategies for addressing these challenges (13, 22).

V-B. JUDGING THE EVIDENCE: TOOLS FOR ASSESSING THE EVIDENCE

This section describes tools to assess the quality of individual studies and the overall nature and strength of the evidence.

Tools for assessing the quality of individual studies

Bradford Hill criteria

The Bradford Hill criteria are a guide to making causal inferences (78) (**Table 7**). The National Research Council's 1989 report on diet and health (12) and the first 6 DRI reports (23, 24, 29, 36–38) used these criteria. As with most assessment tools, these criteria do not address dietary intake measurement issues [e.g., poor correlation of subjective measures of intake with objective measures (67)], which are fundamental to considerations of causality and intake-response relations.

Study quality-assessment tools

The main types of tools for evaluating evidence from RCTs and observational studies are as follows: *1*) quality-assessment instruments that assess the quality of a study from conception to interpretation as a whole and *2*) risk-of-bias schema that assess the accuracy of estimates of benefit and risk (**Table 8**) (40). There is also a move toward conducting quality assessments at the outcome level. Within a particular study, for example, quality may be higher for subsets of outcomes, or blinding may be more important to one outcome than another.

After evaluating published quality-assessment instruments, Bai et al. (79) recommended the use of SIGN 50 methodology; versions are available for cohort studies, case-control studies, and RCTs (19). SIGN 50 uses the following 5 domains to assess the quality of data from cohort and case-control studies: comparability of subjects, assessment of exposure or intervention, assessment of outcome measures, statistical analysis, and funding. For RCTs, important domains are random allocation, adequate concealment of participant assignment to groups and blinding to treatment allocation, comparability of groups, no differences between groups except for the treatment, and assessment of outcome measurement. On the basis of these criteria, a study's overall assessment may be judged to be of high quality overall (has little or no risk of bias and conclusions are unlikely to change after further studies are done), acceptable (some study flaws with an associated risk of bias and potential for conclusions to change with further studies), or low quality (substantial flaws in key design aspects and likely changes in conclusions with further studies). The advantages of SIGN 50 are that it is simple and includes key criteria for quality, good guidance is available for its application and interpretation, and there is extensive experience with its use. Disadvantages are that it is not outcome specific, not sufficiently inclusive of study characteristics that are relevant to food substance and dietary studies, and its assessment of bias domains is considered superficial according to some experts (19).

Risk-of-bias tools that are specific to study type are available to assess degree of bias (40). They provide a systematic way to organize and present available evidence relating to the risk of bias in studies and focus on internal validity. The Cochrane Collaboration's risk-of-bias tool can be used to assess risk of bias for randomized studies (41). Domains of this tool include random-sequence generation; allocation concealment; blinding of participants, personnel, and outcome assessors; incomplete outcome data; selective outcome reporting; and other sources of bias (41). A risk-of-bias tool for nonrandomized studies, Risk of Bias in Nonrandomized Studies (ROBINS), is similar to the Cochrane risk-of-bias tool for randomized studies (42). Advantages of ROBINS are that it can be outcome specific, it provides a detailed assessment of bias domains, and good guidance is available for its application and interpretation (42). Disadvantages are that it is complex, not sufficiently inclusive of study characteristics that are relevant to food substances and dietary patterns, and there is little experience with its use.

Development of a quality-assessment instrument that is specific to food substances

It is possible to develop a quality-assessment instrument that is specific to food substances by adding food-substance–specific aspects of quality to currently available algorithms for quality assessment for use in conjunction with a general study-quality tool (e.g., SIGN 50 or AMSTAR) (21, 40). Food-substance quality-assessment instruments could take into account covariates, confounders, and sources of error that are especially relevant to food substances. For intervention studies, these additional items could be the nature of food-substance interventions, doses of the food-substance interventions, and baseline food-substance exposures in the study population (21). For observational studies, food-substance–specific quality factors might be methods or instruments used to assess intakes of food-substance exposures, ranges or distributions of the food-substance exposures, errors in assessing food-substance exposures, and potential impacts of errors from assessing food-substance exposures on the food-substance–outcome association (21). Other food-substance–specific items are assessment of dietary intakes, including longitudinal patterns, and mapping of dietary intakes to food-substance intakes (40). The need for quality assessments related to food-substance exposure in observational studies speaks to the dominant effect of random and nonrandom intake errors on assessments of magnitude, and even direction, of intake-response associations.

Food substances and dietary applications

The Agency for Healthcare Research and Quality (AHRQ) has produced systematic evidence reviews of associations between food substances and health outcomes (e.g., for vitamin D and calcium) (69, 76, 77, 80) that experts have used to develop DRIs

TABLE 7
Bradford-Hill criteria and application by the Institute of Medicine[1]

Bradford-Hill criteria (78)	Diet and health (12) and DRI reports (14, 23, 24, 29, 36–38)
Strength: effect sizes (not statistical significance)	Yes
Consistency: consistency across study types, locations, populations, study times, and other factors	Yes
Specificity: Is there likely one cause for the effect? Is the association specific to a particular population, context, or outcome and not observed in other populations, contexts, or outcomes?	Yes
Temporality: cause before effect with appropriate delay	Yes
Biological gradient: dose-response relation (could be curvilinear with a dose-response relation in part of the curve)	Yes
Biological plausibility: Is the nutrient of interest a biologically plausible cause of the beneficial effect?	Yes
Coherence: Does cause-and-effect interpretation of data seriously conflict with generally known facts and laboratory evidence?	No
Analogy: Is it possible to judge by analogy?	No
Experiment: Is there experimental evidence from human and/or animal and in vitro studies that is consistent with the associational findings?	No, with the exception of the 2011 DRI report (14)

[1] DRI, Dietary Reference Intake.

and for other applications. In a recent food-substance review to assess the quality or risk of bias of individual studies, the AHRQ (77) used the Cochrane risk-of-bias tool for RCTs that identifies biases related to selection, performance, detection, attrition, reporting, and other factors (41). For observational studies, the AHRQ used questions from the Newcastle Ottawa Scale (81). In addition, the review included food-substance–specific questions to address the uncertainty of dietary-assessment measures (21, 72).

Tools for assessing the overall quality of the evidence

Tools for systematic reviews and meta-analyses

The AMSTAR 2007 tool is useful for the development of high-quality systematic reviews and meta-analyses of RCTs (79, 82, 83). Its methodologic checklist addresses 7 domains: the study question, search strategy, inclusion and exclusion criteria, data extraction method, study quality and validity, data synthesis, and funding. The overall assessment can be high quality, acceptable quality, or low quality. AMSTAR2, for nonrandomized studies, is under development (http://amstar.ca/Developments.php). It will also include confounding and reporting-bias domains. Methodologic checklists for nonrandomized studies are available (84). The newly released ROBIS (Risk of Bias in Systematic Reviews) tool, which is similar to ROBINS, can assess risk of bias in both nonrandomized and randomized studies (85).

The process for evaluating the quality of studies for inclusion in systematic reviews and meta-analyses involves assessing the quality or risk of bias of each candidate study, assembling all of the assessments into a summary table or figure, and assessing the overall study quality or risk of bias (41). No formal tool to determine overall quality is currently available.

GRADE criteria for evidence grading

Methods of judging evidence of causation can vary from binary yes-or-no decisions to ranked approaches. Many systematic reviews use the GRADE (20) criteria, which are in the latter group. GRADE uses evidence summaries to systematically grade the evidence on the basis of risk of bias or study limitations, directness, consistency of results, precision, publication bias, effect magnitude, intake-response gradient, and influence of

TABLE 8
Study types and tools for quality assessment and risk of bias[1]

	Quality-assessment tools	Risk-of-bias tools
Systematic review of RCTs	AMSTAR (http://amstar.ca/)	ROBIS (http://www.robis-tool.info)
Systematic review of nonrandomized studies	AMSTAR2 (in development; http://amstar.ca/Developments.php)	ROBIS (http://www.robis-tool.info)
RCT	SIGN 50 RCT (http://www.sign.ac.uk/methodology/checklists.html)	Cochrane Collaboration risk-of-bias tool (http://handbook.cochrane.org/chapter_8/8_5_the_cochrane_collaborations_tool_for_assessing_risk_of_bias.htm)
Cohort study	SIGN 50 cohort (http://www.sign.ac.uk/methodology/checklists.html)	ROBINS (http://ofmpub.epa.gov/eims/eimscomm.getfile?p_download_id=526737)
Case-control study	SIGN 50 case-control (http://www.sign.ac.uk/methodology/checklists.html)	ROBINS (http://ofmpub.epa.gov/eims/eimscomm.getfile?p_download_id=526737)
Cross-sectional study	SIGN 50 cohort or case-control (http://www.sign.ac.uk/methodology/checklists.html)	ROBINS for cross-sectional studies is in development

[1] AMSTAR, A Measurement Tool to Assess Systematic Reviews; RCT, randomized controlled trial; ROBINS, Risk of Bias in Nonrandomized Studies; ROBIS, Risk of Bias in Systematic Reviews; SIGN 50, Scottish Intercollegiate Guidelines 50.

residual plausible confounding and "antagonistic bias." The latter refers to bias that can result in underestimates of an observed effect. As noted previously, evidence based on observational studies will generally be appreciably weaker than evidence from RCTs and other intervention trials due to the likelihood of confounding and various biases, in particular dietary measurement bias. GRADE also considers study quality in its algorithms. There may be cases for which evidence from observational studies is rated as moderate or even high quality, because extremely large and consistent estimates of an effect's magnitude increase confidence in the results. GRADE users assess and grade the overall quality of evidence for each important outcome as high, moderate, low, or very low. Users describe recommendations as weak or conditional (indicating a lack of confidence in the option) or strong (indicating confidence in the option) (20).

Food-substance applications

The AHRQ uses the AHRQ Methods Guide to grade the strength of the evidence for each outcome in a systematic review (86). The AHRQ explores differences in findings between observational and intervention studies as well as their risks of bias to offer possible explanations for interstudy disparities. The AHRQ summarizes ratings of the strength of the evidence in evidence profile tables that describe the reasoning for the overall rating. This approach builds on the GRADE method by requiring information on reporting biases (publication bias, outcome-reporting bias, and analysis-reporting bias). It incorporates the domains included in GRADE—the study limitations (risk of bias), consistency, directness, precision, intake-response association, strength of association, and plausible uncontrolled confounding—that would diminish an observed effect. AHRQ evidence reviews use additional guidance for scoring consistency and precision, grading bodies of evidence by study type, addressing high-risk-of-bias studies, and other topics. AHRQ evidence reviews grade the strength of the evidence as high, moderate, low, or insufficient, indicating the level of confidence in the findings.

Weighing the evidence

Establishing causality requires a careful evaluation of the weight of the evidence on causal associations between exposures and outcomes. This step can be complex, particularly in the presence of multiple sources of information, not all of which are consistent or of equal relevance or reliability. Systematic reviews can summarize the available evidence in a comprehensive and reproducible manner (87). However, they do not evaluate the weight of the evidence, which various DRI decisions require. Although the Bradford Hill criteria for evaluating causal associations provide useful general guidance on weighing the evidence on causality, more detailed guidance can also be helpful in some circumstances. The International Agency for Research on Cancer, for example, has a detailed scheme for identifying agents that can cause cancer in humans based on a careful evaluation of the available human, animal, and mechanistic data (88). An option for purposes of the committee's charge is to develop an analogous scheme for assessing relations between food substances and chronic disease endpoints.

A review of 50 "weight-of-evidence" frameworks identified 4 key phases for assessments: *1*) defining the causal question and developing criteria for study selection, *2*) developing and applying criteria for the review of individual studies, *3*) evaluating and integrating evidence, and *4*) drawing conclusions on the basis of inferences (89). This review identified important attributes of a broadly applicable weight-of-evidence framework, although the authors did not develop such a framework.

Applicability to food-substance studies

The US National Research Council (90) identified systematic review, quality assessment, and weight of evidence as key components of a qualitative and quantitative risk-assessment paradigm (**Figure 3**). Each of these activities is also directly relevant to the establishment of DRIs, especially for those that are based on chronic disease endpoints. As with any synthesis of information on a population health risk issue, there is a need to carefully evaluate the available information and weigh the available evidence for causality in reaching conclusions about the association between food substances and chronic disease endpoints.

V-C. JUDGING THE EVIDENCE: OPTIONS FOR ADDRESSING EVIDENCE-RELATED CHALLENGES

This section identifies the challenges related to 2 DRI-based evidentiary decisions involved in assessing whether a food substance is causally related to a chronic disease. The first challenge deals with the type of endpoint (outcome or indicator) that is best suited to these DRI decisions. The second challenge addresses the desired level of confidence in the available evidence that the food substance and chronic disease relation is valid. The decisions about which options to implement to address these evidence-related challenges need to be made in an integrative manner because decisions about how to address one challenge have implications for the nature of and responses to the other challenge.

Options for selecting chronic disease endpoints

An early step in the decision-making process associated with the development of a DRI value is the identification of potentially useful measures (indicators) that reflect a health outcome—in this case a chronic disease outcome—associated with the intake of the food substance of interest (15). If a DRI reference value is to be based on a chronic disease outcome, what types of indicators are appropriate to use in making these decisions? Studies vary in the type of outcome measured, ranging from direct measures of the chronic disease based on generally accepted diagnostic criteria to indirect assessment by using either a qualified surrogate marker of the chronic disease outcome or a nonqualified disease marker (Figure 2) (15). Guidance on selection of an indicator based on a chronic disease outcome would inform decisions as to whether newer types of DRI values specifically focused on chronic disease outcomes are more appropriate than are more traditional reference values (Table 4) (see section VI on intake-response relations). In addition, it would clarify applications for some major users of DRIs (e.g., regulatory, policy) for whom clear differentiation between chronic disease and functional endpoints is important for legal and programmatic purposes.

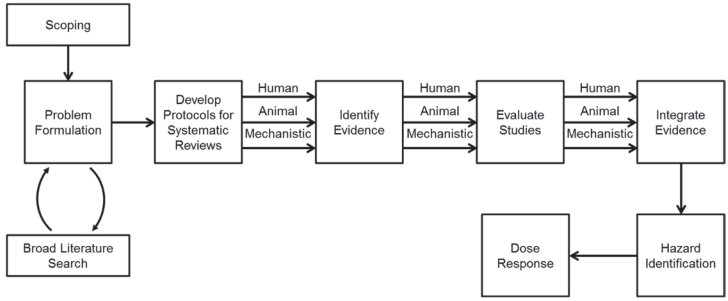


FIGURE 3 Framework for evidence integration. Adapted from reference 90 with permission.

Option 1: Endpoint (outcome) is the incidence of a chronic disease or a qualified surrogate disease marker

The first option is to only accept study endpoints that are assessed by a chronic disease event as defined by accepted diagnostic criteria, including composite endpoints, when applicable, or by a qualified surrogate disease marker. These types of endpoints are associated with higher levels of confidence that the food-substance and chronic disease relation is causal than are nonqualified disease markers (Figure 2). However, few RCTs designed to evaluate the relation of food substances to chronic diseases have used a chronic disease event as the outcome measure. In addition, only a few qualified surrogate markers of chronic disease are available for evaluations of the relation between food substances and chronic disease outcomes. The process of qualifying a surrogate disease marker for evaluating food-substance and chronic disease relations requires sound science and expert judgment (6). Much of the evidence in which outcomes are assessed as a chronic disease event comes from observational studies, and uncertainty is greater about whether relations are causal with data from observational studies than from RCTs (Figure 1). In addition, some of the evidence would likely come from RCTs with chronic disease outcomes assessed by qualified surrogate disease markers. These outcome measures would provide a reasonable basis, but not absolute certainty, that the relation between the food substance and chronic disease is causal (Figure 2). Depending on the level of confidence deemed acceptable for chronic disease–based DRI decisions about causation and intake-response relations (see options on level of confidence below), the use of this option could result in either a small body of evidence if high levels of confidence in the validity of the relation are deemed necessary (e.g., causality is based on the availability of RCTs with chronic disease or qualified surrogate disease outcomes) or a larger body of evidence if lower levels of confidence are acceptable (e.g., causality is inferred from observational studies with outcomes based on chronic disease events or qualified surrogate disease markers).

Option 2: Endpoint (outcome) may include nonqualified disease markers

To implement this option, a DRI committee would also accept studies with outcomes that are possible predictors of the chronic disease of interest but have not been qualified as surrogate disease markers because they lack sufficient evidence for this purpose. Examples of potential biomarkers of chronic disease risk include brain atrophy as the combination of low $A\beta_{42}$ and high T-tau and P-tau levels for Alzheimer disease risk, endothelial dysfunction for atherosclerosis risk, and certain polymorphisms for neural tube defects. A large evidence base is available on relations between food substances and nonqualified chronic disease markers. However, DRI committees have rarely chosen these types of outcomes to establish a DRI value on the basis of a chronic disease endpoint (15).

Compared with option 1, this option increases the number of relations between food substances and chronic disease outcomes for which committees could establish DRIs. However, considerable uncertainty exists about whether decisions about causal relations on the basis of nonqualified disease markers are valid (6). The use of such outcome measures could therefore lead to a loss of confidence in the DRI process.

Options for acceptable levels of confidence that the relation is causal

The overall level of confidence deemed appropriate for DRI decisions on the relation between a food substance and a chronic disease is dependent on an integrated consideration of the type of endpoint that a DRI committee accepts (i.e., a chronic disease event, qualified surrogate disease marker, or a nonqualified disease marker) and the overall evidence rating of the totality of the evidence (**Table 9**). Establishing whether the evidence is sufficient to proceed with making a chronic disease–related DRI decision involves an evaluation of the level of confidence deemed appropriate to determine that the relation of the food substance of interest and the chronic disease is valid.

Option 1: Require a high level of confidence

The first option is to require a high level of confidence (e.g., level A; Table 9) that a proposed relation is causal. This level of confidence likely requires at least some evidence from high-quality RCTs in which the measured outcome is a chronic disease event or qualified surrogate disease marker.

A major advantage of this option is that it provides a robust basis for DRI decisions and therefore conclusions about the relation are unlikely to change substantially when new findings become available, although conclusions would probably need minor modifications to integrate the new evidence. This option would enhance both user and consumer confidence by reducing the likelihood of major changes in DRI decisions over time. Initially, DRI committees could only use this approach to establish DRIs on the basis of a few relations between food substances and chronic diseases because of the limited number of high-quality studies with primary chronic disease outcomes that are currently available or likely to become available in the near future.

Past experience shows the value of this option. For example, consistent results from several observational studies and evidence of biological plausibility suggested that β-carotene reduces the risk of lung cancer, vitamin E lowers the risk of both CVD and prostate cancer, and B vitamins reduce the risk of CVD. However, subsequent large clinical trials failed to support these initial conclusions (53–58). Therefore, conclusions of benefit based almost exclusively on strong and consistent evidence from observational studies would have been overturned by the subsequent availability of evidence from large RCTs.

Option 2: Use level B evidence

A second option is to also include level B evidence (defined in Table 9) as a basis for DRI decisions about causation. This level of evidence suggests a moderate degree of confidence that the relation of interest is causal, but new findings could change the DRI decision. This approach allows committees to establish DRI reference values for more topics than in option 1 that are related to chronic diseases. However, early conclusions based on strong observational evidence and trials that used nonqualified outcomes often need to change because of the conflicting results of subsequent RCTs, as the examples for option 1 show. This option therefore has a risk of a loss of confidence in DRI decisions.

Option 3: Use actual level of certainty

The third option is to identify the actual level of certainty [e.g., levels A, B, C, or D, as defined in Table 9, or GRADE levels of high, moderate, low, or very low (insufficient)] for each DRI reference value based on a chronic disease endpoint. The advantage of this approach is that it provides more information than do options 1 and 2 about the scientific evidence that supports a given relation between a food substance and a chronic disease endpoint. A disadvantage is that DRI values may become separated from grading scores as they are used and applied, thus inadvertently suggesting that all DRI values are based on evidence of similar strength. Decisions about this option would benefit from evidence that shows that users take the evidence grades into account when they use such DRI reference values.

Option 4: Make decisions on a case-by-case basis

The fourth option is to make decisions about the strength of evidence appropriate to support a conclusion about the relation between a given food substance and a chronic disease endpoint on a case-by-case basis. This option maximizes flexibility for DRI committees and can enable them to consider other factors (e.g., the public health importance of the relation). However, a major disadvantage is that this option could lead to inconsistency among DRI reviews, which could reduce the confidence of users in DRI reference values. This approach is also inconsistent with the grading-of-evidence approach that many health professional organizations and government agencies are now using.

VI. INTAKE-RESPONSE RELATIONS

Once a DRI committee establishes a causal relation between the intake of a food substance and the risk of $\geq$1 chronic disease, it must determine the intake-response relation so that it can establish a DRI. Ultimately, the reference value and how users can apply it depend on the decisions that the committee made when it established the intake-response relation between a chronic disease indicator and the observed intakes of a food substance. A number of conceptual challenges have made it difficult to apply the traditional DRI framework to chronic disease endpoints, including how risk is expressed for chronic diseases, the multifactorial nature of chronic diseases, and the diversity of intake-response relations between food substances and chronic diseases. This section describes options for defining an acceptable level of confidence in the data that a DRI committee uses to determine intake-response relations after establishing causality, the types of reference values that could be set, and the types of indicators that could be used to set reference values and for avoiding overlap between beneficial intakes and intakes associated with harm.

Conceptual challenges

Previous committees have based DRIs on the intakes necessary to avoid classical nutritional deficiencies (i.e., EARs and RDAs) and unsafe intakes associated with toxicities or adverse events (i.e., ULs) (**Figure 4**, Table 1). Intake-response relations between traditional endpoints for nutrient requirements (i.e., deficiency diseases) and adverse events are often different from those between food substances and chronic disease endpoints (Table 4).

Absolute compared with relative risk

Previous DRI committees based their reference values on direct evidence from human studies that measured both intakes and outcomes, which allowed committees to develop quantitative intake-response relations on the basis of absolute risk, which is the risk of developing a given disease over time. At "low intakes," these essential nutrients have intake-response-relation characteristics in which the known health risks, which are diseases of deficiency for essential nutrients, occur at very low intakes and can affect up to 100% of a population at a specified life stage, and the risk declines with increasing intakes. Inadequate intakes of essential nutrients are necessary to develop diseases of deficiency. The risk of a disease of deficiency is 0% when intakes are adequate, and an adequate level of intake is necessary to treat a deficiency disease. For example, chronically inadequate intakes of vitamin C are necessary and sufficient to develop scurvy, and the entire population is at risk of scurvy when intakes are inadequate. An adequate level of intake of vitamin C is necessary and sufficient to reverse the deficiency. At "high intakes," it is assumed that these essential nutrients cause adverse health effects, including toxicity (Figure 4). As

TABLE 9
Level of confidence in DRI decisions[1]

	Overall evidence rating based on evidence review		
Chronic disease endpoint	High	Medium	Low
Chronic disease event	Level A	Level B	Levels C or D
Qualified surrogate disease marker	Levels A or B	Levels B or C	Levels C or D
Nonqualified outcome	Level C	Levels C or D	Level D

[1] Level A: highest degree of confidence that results are valid (e.g., "high"); level B: some uncertainty about validity of results (e.g., "moderate"); level C: considerable uncertainty about validity of results (e.g., "low"); level D: substantial uncertainty about validity of results (e.g., "insufficient"). DRI, Dietary Reference Intake.

with inadequate intakes, the absolute risk of an adverse effect from excessive intakes is represented as increasing from 0% to 100% with increasing intakes of the nutrient. All members of a population are assumed to be at risk of the adverse effect at sufficiently high intakes.

In contrast, DRI values based on chronic disease endpoints have been based on relative risk, which is risk in relation to another group. Past DRI committees used data from observational studies, which contain the biases described earlier in this report, primarily to calculate the relations between food substances (essential or otherwise) and chronic diseases because only a limited number of relevant RCTs are available in the published literature. The risk of the chronic disease based on observational and intervention studies is usually reported as relative to a baseline risk and is therefore not absolute. The baseline risk is never 0% or 100% within a population, and it can vary by subgroup [e.g., those with high blood pressure and/or high LDL cholesterol have a higher risk of CVD death than do those with lower blood pressure and LDL-cholesterol concentrations (91, 92)]. The intake of a given food substance might alter the risk of a disease by a small amount (e.g., $<10\%$) compared with the baseline risk, but these changes could be very important from a public health perspective depending on the prevalence of the chronic disease (e.g., a 5% reduction in a highly prevalent disease could have a meaningful public health impact), severity, impact on quality of life, cost, and other factors. Conversely, the intake of a given food substance might alter the relative risk by a large amount compared with baseline risk, but changes in absolute risk could be small and have a less meaningful impact on public health (35).

Interactions of multiple factors

The pathogenesis of chronic disease is complex and often involves the interaction of multiple factors, in contrast to traditional endpoints that commonly are associated with interactions of fewer factors. Intakes of a group of food substances might contribute to the risk of a chronic disease, for example. The magnitude of risk might vary by intake, and several factors (e.g., behaviors or physiologic characteristics) might influence the risk. Furthermore, ≥ 1 food substances might be associated with >1 chronic disease. Finally, although a given food substance might contribute independently to the development of a chronic disease, changes in intake might not be necessary or sufficient to increase or decrease the risk of the chronic disease due to the complex interacting factors in the disease's pathogenesis.

Shape of the intake-response curve

The shape of the intake-response relation curve can vary depending on whether the relation is between an essential nutrient and a deficiency disease or between a food substance and a chronic disease endpoint. The intake-response relation between a nutrient and a deficiency disease is often depicted as linear or monotonic within the range of inadequacy, whereas the relation between a food substance and a chronic disease indicator can be more diverse (e.g., linear, monotonic, or nonmonotonic). Nonmonotonic intake-response relation curves can be U-shaped, J-shaped, or asymptotic. Furthermore, a single food substance can have a causal relation with >1 chronic disease, and the intake-response curves for each relation can differ (30, 93). The effect of a nutrient intake on chronic disease risk might be saturable in some cases. **Figure 5** shows examples of diverse intake-response relations between a food substance and a chronic disease or diseases.

DRI committees must take into account the statistical fit of the intake-response curve to the available data and its adherence or relevance to underlying biological mechanisms when determining the shape of the intake-response curve for a food substance and a chronic disease outcome. Deriving intake-response curves when single food substances affect multiple chronic diseases can be particularly challenging. Future DRI committees will need to determine whether to apply available statistical methods or to develop new ones to address these challenges (13). Ideally, future expressions of reference values will include estimates of uncertainties and interindividual variability.

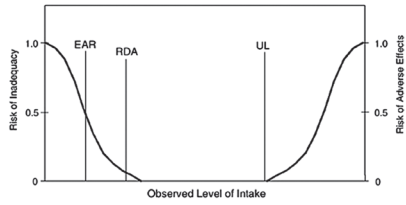


FIGURE 4 Relations of intakes and adverse effects of substances that are nutritionally necessary. EAR, Estimated Average Requirement; RDA, Recommended Dietary Allowance; UL, Tolerable Upper Intake Level. Reproduced from reference 7 with permission.

Examples of diverse intake-response relations between food substances and chronic disease endpoints that show the complexity of ensuring the statistical fit of the intake-response curve to the data include the following:

- The relative risk of coronary heart disease has a linear intake-response relation to fiber intakes and no apparent threshold for the beneficial effect. The DRI committee based the AI for fiber on the mean intake associated with the highest relative effect (24).
- The relation between the risk of dental caries and fluoride intake appears to have an inflection point and a critical value for statistically detectable risk reduction (dental caries prevention), but the range of intakes associated with benefit overlaps with the range of intakes associated with harm (fluorosis) (29).
- Omega-3 fatty acids and multiple chronic diseases, as suggested by results from observational studies, have several intake-response relations, depending on the chronic disease (30).

DRIs based on intake-response relations involving chronic diseases

DRI users include a wide range of organizations (e.g., health professional groups and societies and government agencies), many of which rely on DRI values to make decisions and to develop policies for their organization. These varied user groups have requested information to help them interpret findings in DRI reports (13). These groups have also asked DRI committees to present the information in a way that supports flexible applications while informing users of the nature of the available evidence and public health implications.

The approach to setting DRI values would be enhanced by transparency. Clear descriptions of the scientific and public health characteristics of the benefits and risks of the intake of a food substance are also valuable. For example, for each benefit and risk, descriptions could include the strength of the evidence, the sizes and characteristics of groups at risk, and the likelihood and severity of the risks. Users could then evaluate these descriptions to decide how to apply the DRIs in ways that address their organizational mission and decision-making framework.

Acceptable level of confidence in the intake-response data

Several options are available for determining the acceptable level of confidence in the data that a DRI committee uses to determine intake-response relations once it has data that establish a causal relation.

Option 1: Require a high confidence level

One option is to require a high level of confidence by, for example, using RCTs with a chronic disease event or a qualified surrogate disease marker as the outcome measure (Table 9). This approach typically requires usable intake-response data from RCTs, which is probably impractical because most RCTs have only 1 intervention dose or a limited number of doses. This option could result in failure to establish a DRI even though the data have established a causal relation. The use of this option is therefore unlikely to be optimal for public health because no reference value, or even a reasonable estimate of one, would be available for a documented relation between a food substance and a chronic disease.

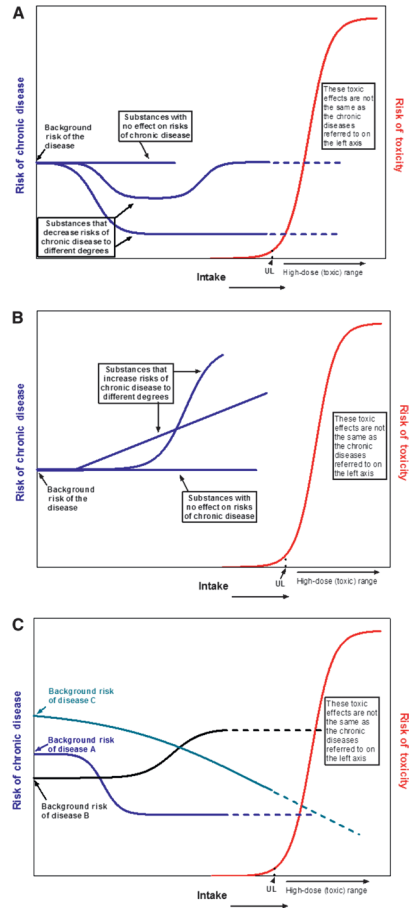


FIGURE 5 Intake-response relations between the intake of a food substance and chronic disease risks can vary. The intake of a food substance could decrease (A) or increase (B) chronic disease risk. The intake of a food substance could be independently related to multiple chronic diseases that show different and overlapping dose-response relations (C). The relation or relations between the intake of the food substance and chronic disease or diseases might not be monotonic. The background risk of a given chronic disease is not zero. "Substances" could be individual food substances or groups of interacting substances. UL, Tolerable Upper Intake Level.

Option 2: Accept a moderate confidence level

Another option is to accept a moderate level of confidence in the data for decisions about intake-response relations (Table 9). DRI committees could then expand the types of data being considered to include high-quality observational studies with outcomes based on chronic disease events or qualified surrogate disease markers.

These data would likely be associated with some uncertainty (Table 9). For example, systematic biases in intake estimates are likely to affect intake-response data from observational studies. Intake-response data from intervention trials would likely lack some details on baseline intakes, making total intake estimates difficult. DRI committees would need to determine how much and what type of uncertainty are acceptable.

Option 3: Piecemeal approach

A third option is to "piece together" different relations in which the biomarker of interest is a common factor when direct evidence of the biomarker's presence on the causal pathway between the food substance and a chronic disease is lacking. For example, if data show a quantitative relation between a food-substance intake and the biomarker of interest and other data show a quantitative relation between the biomarker of interest and the chronic disease, this evidence could be combined to establish a quantitative reference intake value for the chronic disease risk. This option has the advantage of relying on a wider breadth of the available evidence than the first 2 options and likely would enable DRI committees to consider more nutrient–chronic disease relations, but the approach would be fraught with uncertainties. Among its major disadvantages is its heavy reliance on expert judgments, which limit objectivity in its application.

Different types of reference values

Because of the conceptual issues discussed earlier in this section, reference values based on chronic disease endpoints likely need to be different from the traditional reference values for essential nutrients. Because many food substances share metabolic pathways, DRI committees could consider joint DRI values for groups of related food substances. Similarly, because a single dietary source might supply >1 food substance, DRI committees could base reference values on groups of food substances to prevent harm (e.g., to minimize the risk that limiting the intake of 1 food substance will produce undesirable changes in intakes of other food substances). If a DRI committee uses a variety of chronic disease endpoints or a family of targeted food-substance-intake reductions to establish reference intake values, this process is likely to be strengthened by enhanced transparency and the estimation of associated uncertainties. Providing information on how benefits and risks are weighted would also likely assist users in their applications of derived values.

The impact of DRI values would likely be strengthened if their potential uses are considered in their derivation. Previous DRI committees have identified differences in the applicability and use of different types of reference values for planning and assessment in groups and individuals (8, 10). For example, the AI has limited applicability to dietary assessments of groups (13). As DRI committees consider possible approaches to establish reference values for chronic disease endpoints, how the different types of reference values could meet user needs and how users could apply these values will remain critical considerations.

Types of reference values associated with benefit

Option 1: Establish chronic disease risk-reduction intake values (e.g., CD_{CVD}). DRI committees could modify the traditional EAR/RDA approach to estimate the mean intakes of individuals and the interindividual variability associated with specified disease risk reductions. This option is conceptually very similar to the traditional EAR/RDA approach, but the definitions and interpretations of reference values based on chronic disease endpoints are different from those based on classical deficiency endpoints. This option uses relative risks and requires knowledge of baseline disease prevalence, whereas the traditional approach is based on absolute risks and is independent of baseline prevalence. The mean intake values and associated variances for given magnitudes of risk reduction give valuable information on the "typical" person and population variability. These values could, therefore, be useful for assessing population and group prevalence. Several adaptations of this option are possible, depending on the nature of the available data.

Adaptation 1 is to set a single chronic disease risk-reduction (CD) value at the level above which higher intakes are unlikely to further reduce the risk of a specific disease. Such values would be similar to traditional EAR/RDA values in that they would be a point estimate with some known variation (Figure 4). An advantage of this kind of reference value is its similarity to the traditional EAR/RDA, which could help users understand the value as well as its use and application. Furthermore, this approach requires a high level of evidence and an understanding of the uncertainty around the value, which could maximize confidence in the value and its uses. The data required to establish a single CD value of this type are probably very limited. However, the possibility of developing this type of value may guide research.

Adaptation 2 is to establish multiple reference values on the basis of the expected degree of disease-risk reduction across a spectrum of intakes to yield a family of targeted reductions for a given chronic disease outcome and potentially for a variety of disease indicators with distinct intake-response relations to the disease. If a DRI committee uses this adaptation, it may find it useful to consider such factors as the severity and prevalence of outcomes. For a given distribution of intake within the population that has a given mean and some variability, a DRI committee could establish the expected risk reduction and identify an expression of uncertainty. Multiple values could be established on the basis of >1 level of risk reduction.

Future DRI committees could establish reference values for different degrees of disease risk reduction and for different groups with different risk levels within a population. An advantage of this adaptation is that it gives users flexibility to choose reference values that meet their needs and are suitable for the risk profiles of individuals or groups to whom they apply the reference values. However, users could be confused about when and how to apply the different values. For this reason, the use of this adaptation requires careful attention to implementation guidance.

Adaptation 3 is for food substances that have causal relations at different intake levels to multiple chronic diseases. This adaptation

involves establishing different reference values for different diseases (e.g., CD_{CVD}, CD_{cancer}). In addition, for each relation with a different chronic disease, a DRI committee could identify a family of targeted risk reductions to establish multiple CD values, each of which would be associated with a specific degree of risk reduction.

An advantage of this option is that DRI committees could establish CD point estimates for specified risk reductions for ≥1 chronic disease, which would provide flexibility to both committees and users. This adaptation may make it easier than the other 2 adaptations for users to understand when (e.g., for a life stage with a higher risk of the disease) and to which population or populations (e.g., those at higher risk of a given disease) to apply the values. Establishing reference values for multiple chronic diseases requires the same level of evidence, or an equivalent kind of evidence, for each disease to ensure that committees can develop all values and users can apply them with the same level of confidence. A disadvantage is that establishing several values could confuse users about their appropriate application. DRI committees could minimize this confusion by developing appropriate guidance on how to implement the values.

Option 2: Identify ranges of beneficial intakes. In some cases, available data might be adequate only for deriving an intake range that can reduce the relative risk of a chronic disease to a specified extent. One end of this intake range is close to the point at which risk begins to decline or increase, depending on the relation, and the other end extends as far as the available evidence permits. The DRI committee could establish the range so that it does not increase the risk of adverse health effects (Figure 5A, B; see also section entitled "Options for resolving overlaps between benefit and harm" below and Figure 5C).

These reference values have a purpose similar to the estimated risk-reduction intake value for a chronic disease (option 1), except that data for making point estimates or for estimating interindividual variation are not available, making a point estimate impossible to develop. Advantages of this option are that the level of evidence it requires is less stringent than that required for option 1 and it provides flexibility to users. A disadvantage is that the value is associated with lower confidence but users might apply it with confidence if they are unaware of its limitations. The use of a range to assess the prevalence of beneficial intakes in a population might also be challenging. Users would need clear guidance on how to apply these kinds of values. This approach could incorporate the AMDRs because the AMDRs represent a range of intakes associated with macronutrient adequacy. Future committees could be charged to review how users could apply such an approach to intakes of macronutrients or their constituents (e.g., a protein compared with a specific amino acid).

ULs and reduction in chronic disease risk

The UL (Figure 4) is the highest average daily intake level likely to pose no risk of adverse health effects for nearly all people in a particular group (7). The UL is not a recommended level of intake but rather the highest intake that people can tolerate without the possibility of ill effects (7).

DRI committees have based most ULs on (often limited) evidence of toxicity or adverse events at a high nutrient intake level. Past DRI committees used a threshold model to calculate ULs, in which the intake-response relation has an inflection (threshold) point (11). Because of the paucity of evidence, most ULs were not based on chronic disease endpoints, although DRI committees tried to do so for a few nutrients (e.g., saturated and *trans* fats as well as sodium) with limited success (13). A key reason why basing ULs on chronic disease endpoints is so challenging is that the traditional UL definition is based on an intake level associated with no increase in absolute risk, whereas most data related to chronic disease risk are expressed as relative risk. When the interval between intakes associated with benefit and harm is wide and intakes associated with benefit do not overlap with those associated with harm (see below), options for setting the UL include the use of 1 or both traditional adverse events and chronic disease endpoints, depending on the nature and strength of the available evidence.

Option 1: Base ULs on the traditional threshold model. One option is to continue to base ULs on the traditional threshold model when UL values based on chronic disease endpoints are higher than those based on traditional adverse effects. The advantages of this approach are that it allows the DRI committee to evaluate and consider the evidence available for setting a UL on the basis of chronic disease risk, while also allowing the committee to set a traditional UL, which has an established process and its limitations and applications are well understood. However, many traditional ULs are based on (very) limited data. Therefore, a disadvantage is that this option could prevent a DRI committee from establishing a UL on the basis of chronic disease risk (UL_{CD}) that is higher than the intake levels associated with a traditional adverse effect regardless of the evidence available to support the UL or public health implications of the chronic disease. To date, DRI committees have not set any EAR or RDA at intakes higher than the traditional UL to ensure safe intakes across the population. This has been the case even if the intake of a food substance has a beneficial effect on chronic disease risk that is continuous above the UL. A more detailed discussion of the issue of overlapping beneficial and risk curves is given below under the section entitled "Overlaps between benefits and harms."

Option 2: Base UL_{CD} on intakes associated with chronic disease risk. When the risk of a chronic disease increases at an intake below the traditional or current UL, a DRI committee could base a UL on chronic disease endpoints by using approaches analogous to the derivation of CD values (e.g., the development of 1 or multiple values for specified levels of relative risk reduction) or a threshold approach (e.g., identifying the inflection point at which absolute or relative risk increases). These values could be denoted as a chronic disease UL (UL_{CD}) to distinguish them from a traditional UL. The UL_{CD} would be set at a level below which lower intakes are unlikely to achieve additional risk reduction for a specified disease. The traditional UL definition would have to be expanded to include intakes associated with changes in relative risk (in contrast to absolute risk) of an adverse effect. Because the UL_{CD} is based on changes in the relative risk of the chronic disease, intakes below the UL_{CD} might reduce but not necessarily eliminate disease risk, reflecting the multifactorial nature of chronic diseases.

An advantage of basing ULs on chronic disease endpoints is that it maximizes public health benefit. In addition, this approach is straightforward, and users could apply such a UL in a similar manner to a traditional UL. Elimination or limits on intake (e.g.,

of sodium) might be challenging to achieve. Nonetheless, the availability of the type of intake-response information in Figure 5B might be useful for analyzing dietary patterns that minimize risk. This information could also be useful for identifying the magnitude of chronic disease risk reduction achievable with various intake levels. Intakes associated with acute toxicity could still be documented in DRI reports, and intakes below a traditional UL would eliminate the risk of toxicity.

Adaptation 1 is to base a family of UL values on multiple levels of risk reduction, multiple endpoints (adverse effects and/or chronic diseases), or both. A strength of this adaptation is that it acknowledges that the relation between food substances and chronic diseases might be continuous and not have a threshold. Another strength of this approach is its flexibility in the application of the various UL values. A disadvantage is that the determination of a desirable target risk reduction could be challenging. A DRI committee could define the intakes necessary to achieve specific risk reductions. However, users might need to identify the level of risk, or a range of risk levels, that is appropriate for their application. Clear guidance on the application of the UL_{CD} values would be critical.

Overlaps between benefits and harms

Traditionally, the intakes of food substances associated with benefits and risks are separated by an interval that is large enough to prevent overlaps between the ranges associated with benefits and those associated with harms. Therefore, DRI committees have not typically needed to balance the risks and benefits of various intakes. However, intakes of some food substances associated with disease-risk reduction may overlap with intakes associated with adverse events, including increased chronic disease risk (Figure 5C). In some cases, the benefits and risks are associated with overlapping intakes of the same food substance (e.g., the same range of fluoride intakes is associated both with reductions in dental caries and fluorosis). In other cases, the intake level of 1 food substance that is associated with a reduced risk of a chronic disease could result in changes in intake levels of other food substances and thus of their associated benefits or risks (e.g., reducing intakes of a naturally occurring food substance to recommended levels might require drastic changes in dietary patterns that could reduce intakes of co-occurring essential nutrients found in the foods that contain the food substance being reduced, thereby potentially decreasing intakes of these other nutrients to levels below requirements). This could lead to inadvertent imbalances in the dietary pattern. It is also useful to consider whether risks and benefits occur within the same population or in disparate groups, such as in men or women, children or adults, high-risk populations or not, and so on. Several options for addressing overlap issues are described below.

Option 1: Avoid overlap between beneficial intakes and intakes associated with adverse events

One option is to ensure that no point estimate or range of beneficial intakes for chronic disease risk reduction extends beyond the intake at which the risk of adverse events, including chronic diseases, increases. An advantage of this option is that the UL is easily interpretable and applicable because it does not require users to balance benefits and risks. A disadvantage is that it does not acknowledge possible benefits above the intake level associated with adverse events or reflect the limitations in the evidence and uncertainty factors that DRI committees have used to establish many ULs. These committees did not balance benefits and risks when they weighed the evidence on various potential endpoints (e.g., risk of a chronic disease or an acute adverse event).

Option 2: Establish criteria related to severity and risk of chronic disease

Another option is to establish criteria for ULs on the basis of the minimum level of severity and prevalence of targeted chronic diseases and the degree of risk reduction associated with specified intakes. The DRI committee would apply analogous information on the nature of candidate adverse outcomes when establishing ULs. The advantage of this approach is that it allows DRI committees to evaluate the weight of the evidence for all endpoints. This approach limits considerations of risks and benefits to those that are biological, avoiding the need to take into consideration nonbiological factors (e.g., health care costs or quality of life). A challenge is integrating and interpreting the evidence for all endpoints.

Option 3: Describe the nature of the evidence

Another option is to simply describe the nature of the evidence (e.g., type of evidence, quality, strength) and the public health implications of benefits and risks for the full range of intakes for which inferences are reasonably possible, along with remaining uncertainties. Ultimately, users would choose an appropriate balance between benefits and harms for their population of interest. An advantage is that this option allows the DRI committee to make science-based evaluations of the evidence and allows maximum flexibility for users in choosing the appropriate mix of risks and benefits for particular groups or scenarios when they apply the reference values. A challenge is the need to develop clear guidance on interpreting different types of reference values and on appropriate and inappropriate applications (8, 10). A disadvantage is that users might apply such reference values inappropriately.

Selecting indicators and specifying intake-response relations

This section focuses on options for addressing challenges associated with intake-response curves that are based on chronic disease endpoints, including the selection of indicators, confounding, and extrapolation to other age and sex groups.

Qualified surrogate disease markers and nonqualified disease markers

Chronic diseases often have a range of indicators of variable prognostic value (e.g., metabolic perturbations, proteomic changes, or changes in a physiologic function). These indicators are sometimes also on the causal pathway between a food substance and the risk of the disease (Figure 2). Although identifying the intake-response relations between food substances and targeted chronic disease events is highly desirable, this goal is not always achievable. Therefore, DRI committees would use indicators on the causal pathway between a food substance and a chronic disease to establish reference intake values. Qualified surrogate

disease markers provide the strongest evidence of a relation between the intake of a food substance and risk of a chronic disease. However, the list of qualified surrogate disease markers is short (6). The use of a nonqualified disease marker of intake-response is associated with higher uncertainty. Supporting mechanistic data for all indicators are desirable. Once a DRI committee has determined that the relation between the intake of a food substance and a chronic disease risk is causal, it has several options for selecting an indicator to quantify putative intake-response relations.

Option 1: Choose a single outcome indicator on the causal pathway. One option is to select a single outcome indicator that is on the causal pathway, provided that it is sufficiently sensitive to quantify the relation between a food substance and a chronic disease. An advantage is that because this option is similar to the current approach for setting DRI values, it is straightforward and clearly understood. A disadvantage is that it could lead DRI committees to discard other valid indicators that describe other relevant intake-response relations. In addition, the indicator that the committee chooses might not accurately portray the relation between a food substance and an endpoint that is relevant to groups with diverse genetic backgrounds or diverse health habits.

Option 2: Use multiple indicators of a chronic disease. A second option is to integrate information from multiple indicators of a given chronic disease that add substantially to the accuracy of the intake-response relation and the development of a reference value. The advantage of this approach is that it allows multiple, and possibly different, intake-response relations between the intake of a food substance and a chronic disease endpoint to be integrated. The use of this option helps the DRI committee understand variations in uncertainties about a given intake estimate associated with a chronic disease that is measured by multiple indicators. A challenge is that integrating information from multiple indicators could be complex, and the use of this option might require the development and validation of new statistical models.

Option 3: Use of multiple indicators for multiple diseases. DRI committees might need to use a third option when a single food substance has different intake-response relations with multiple chronic diseases (30). In this situation, the committee might need to develop criteria for selecting appropriate disease indicators to establish multiple intake-response relations, methods to integrate multiple endpoints, and approaches to account for the inevitable interindividual variability in the relations of interest. A committee might develop different reference values for each disease endpoint. The advantage of this approach is that it takes into consideration the full landscape of evidence on a given food substance. This option also gives users flexibility in applying the reference values that are most relevant to individuals or populations of interest. A disadvantage is that if a DRI committee establishes a single, integrated reference value by using this approach, this value might not be consistent with increasingly attractive approaches that fall under the category of precision or personalized medicine. A challenge is the need to develop a methodology to integrate this kind of evidence.

If a DRI committee uses a substitute outcome (qualified surrogate or nonqualified disease marker) to establish a DRI value, committees typically need to evaluate the evidence that supports the putative relations between surrogate outcomes and the intake-chronic disease relation. In most cases, data that show the relation between a qualified surrogate disease marker and a chronic disease would be available from existing sources.

Biomarkers of intake

To implement each of these options, DRI committees can use available intake biomarkers in place of or complementary to dietary intake data to determine intake-response relations. The advantage of doing so is that it minimizes nonrandom errors and biases linked to self-reported dietary intakes. A challenge is that qualified biomarkers of intake are not available for many food substances. In qualifying these as surrogate intake markers, any inherent errors or biases, as described in the section entitled "Factors that influence or confound intake-response relations" below, may need to be taken into account.

Mode-of-action frameworks

The use of a statistical approach is 1 way to establish the intake-response relation between a food substance and a chronic disease (13). However, disease pathogenic processes are often gradual and cumulative. DRI committees can use a number of indicators on causal pathways between food-substance intakes and risks of chronic diseases to determine intake-response relations. For this reason, biological approaches that use a mode-of-action framework can provide information on relations between food substances and chronic diseases. Such a framework takes into account the role of biological mechanisms in establishing quantitative reference intakes. The application of a biological framework requires knowledge of the key molecular events, biological systems, and biological pathways that a food-substance intake modifies (94). This approach has been proposed for the development of ULs, although it could also be useful for establishing relations between food substances and chronic disease endpoints (94).

Issues to consider when applying biological frameworks

Issues for DRI committees to consider include whether chronic disease risks respond to food-substance intakes and the potential severity of ultimate biological effects if prevention does not occur at an early, more modifiable phase of the disease process. The intake-response relation between a food substance and key events in the pathogenesis of a chronic disease might not be linear at all intakes or at all periods of exposure. Cumulative processes might proceed continuously, but incrementally over time and might surpass a threshold of reversibility. Examples of reversible events include enzyme inhibition and modest losses of readily replaceable cell types. Examples of events that are irreversible or difficult to reverse include the loss of cells that do not typically proliferate, such as many types of neurons whose loss causes chronic neurological diseases. Individuals in a population have a continuous distribution of these processes or events. The reversible or irreversible nature of key intermediate events in the causal chain affects whether targeted effects might respond to nutritional interventions and the dynamics of these potential benefits. Knowledge of these key events and their impact on intake-response relations could inform the establishment of DRI values based on chronic disease endpoints.

Factors that influence or confound intake-response relations

The accurate description of intake-response relations between food substances and chronic diseases depends heavily on the accuracy of the measurements of intakes and disease outcomes. Many methods produce inaccurate and inconsistent estimates of intakes of diverse food substances. For example, food-frequency questionnaires tend to be more biased than 24-h recalls or records when measuring energy intakes (44). Even the "gold standard" weighed food records result in underreporting of energy intakes (95). In addition, the use of outdated or flawed food-composition databases can introduce errors, and food-substance bioavailability can vary by food matrix or source (e.g., the bioavailability of naturally occurring folates differs from that of folic acid). To add to this complexity, some food substances (e.g., short-chain fatty acids, vitamin D, vitamin K, and folate) have nondietary sources, including the microbiome and metabolic processes that contribute to exposure but that dietary intake estimates cannot quantify. Random measurement errors often attenuate observed associations between intakes and chronic disease risks. Systematic assessment biases can also distort these associations, particularly if they are based on dietary self-reports. Nonrandom errors in estimates of dietary intake bias the regression relation in that the intercept is overestimated and the slope is underestimated. Such errors can result in highly variable and often biased intake measures that serve as the basis for substantial underestimates of intake and/or distortions of intake-response relations. For these reasons, biases related to measurement error require attention in the calculation of intake-response relations.

It is possible to overcome these shortcomings by using qualified biomarkers of food-substance intakes and exposures. Biomarkers of intake (not status) can mitigate or correct the biases associated with self-reported dietary intake data, but they are only available for a few food substances (e.g., urinary nitrogen is a biomarker for protein) (44, 67). Metabolomics is a promising approach for identifying new biomarkers that reflect nutritional intake or exposure in biological samples, such as blood or urine. Researchers can use these biomarkers, such as metabolite production by diet-dependent gut microbiota, to track dietary intakes and other exposures (96, 97). Finding biomarkers of long-term intakes is likely to be particularly challenging, however.

In determining intake-response relations, clinical events (e.g., CVD outcome such as stroke or heart attack) or qualified surrogate disease markers are nearly always preferable as the outcome measure because they can provide a moderate to high level of confidence in the reference values depending on the quality of the evidence (Table 9). However, intake-response data from RCTs based on clinical endpoints are seldom available and are often impractical to obtain. In additon, there are a limited number of qualified surrogate disease markers. Chronic disease outcome measures are more readily available from observational studies, but these study designs are subject to the systematic biases associated with self-reported intakes. Thus, calculations of intake-response relations might need to use less-than-ideal outcome data and/or be derived from observational studies. In these cases, DRI committees could consider and describe the associated uncertainties. The use of these types of outcome measures and study designs requires accurate and consistent measurement of chronic disease indicators and knowledge of assay biases. Random errors in the measurement of the dependent variable, which could be a disease or its biomarker, distort or obscure intake-response relations but do not necessarily bias them.

Extrapolation of intake-response data

DRI committees have often extrapolated intake reference values from a single life-stage or age and sex group to other life-stage and age and sex groups in the absence of group-specific data, with the primary aim of preventing deficiency diseases (13). Relations between food substances and chronic diseases may differ substantially by life stage, physiologic state, and time since exposure. Therefore, the extrapolation of DRI values based on chronic disease endpoints might be more challenging than of those based on deficiency disease criteria.

A framework to develop DRI values on the basis of chronic disease endpoints would benefit from the a priori development of criteria for appropriate use of imputation (and/or extrapolation). In developing such criteria, differences in background risk in subpopulations will be useful to consider. Figure 5A, B depicts a single risk background as the starting point from which food-substance intakes can modify chronic disease risk. However, background risk levels differ by population. Such differences probably alter intake-response relations between food substances and chronic diseases.

Option 1: Establish reference intake values only for similar populations

One option is to establish DRI values on the basis of chronic disease endpoints only for populations that are similar to studied groups. This differs from setting traditional DRI values for essential nutrients for which a value was set for all groups. The advantage is that the basis for the recommendation is very strong because of the limited chance of added error or uncertainty due to extrapolation. The disadvantage is that the reference values for health benefit or risk would apply only to selected subgroups even if they benefit others.

Option 2: Allow extrapolation when sufficient evidence is available

A second option is to allow extrapolation when sufficient evidence shows that specific intakes of a food substance can increase or decrease the risk of a chronic disease. An advantage is the option's potential to extend reference values to unstudied populations. A disadvantage is that the science supporting extrapolation is weak, and this option could lead to the perception that a given intake is associated with a health effect when direct evidence of such an association does not exist. DRI committees need reliable methods to extrapolate the effects of a food substance on a chronic disease. One possible approach is to incorporate baseline variances by assuming the central tendency of the population while taking changes in demographic characteristics (e.g., in age or body weight) into consideration. A potentially useful approach is to define the population distribution of susceptibilities to different chronic diseases in relation to food-substance intakes over time periods that are specific to individual causal processes. DRI committees can use such

distributions to conduct analyses that juxtapose changes in population benefits and risks that are likely to result from defined changes in dietary intakes.

VII. DRI PROCESS

Inclusion of chronic disease endpoints in future DRIs

Substantial challenges persist in basing reference intake values on chronic disease endpoints (22). One challenge is the paucity of sufficiently relevant and robust evidence for evaluating causality in suspected relations between food substances and chronic diseases. A second challenge is the frequent inappropriateness of the present EAR/RDA and UL frameworks for deriving DRIs on the basis of chronic disease endpoints. This situation is counterbalanced by improved tools to assess the quality of available evidence, increasingly transparent and rigorous approaches for synthesizing evidence, and new evidence likely to become available in the future. The promise of major technological advances and emerging scientific knowledge support the need for continuing attention to this area (see sections V-A and V-B) and the continued explicit consideration of chronic disease endpoints in DRI deliberations.

The approach (see section IV) a DRI committee chooses for establishing DRIs depends on the nature of the available evidence and/or the targeted endpoint or endpoints. Although a nutrient deficiency has a single direct cause (i.e., an inadequate intake), chronic diseases have multiple causes. Furthermore, a food substance can have multiple biological effects that are or are not on a disease's causal pathway, a food substance might modify the risk of >1 chronic disease, and intakes might have different effects at different life stages. In addition, absolute and often more immediate risks characterize classical relations between food substances and deficiencies, whereas relations between food substances and chronic diseases are most often reported as relative risks. The resultant pathology often becomes evident only after prolonged relevant exposures.

Such differences between nutrient deficiency diseases and chronic disease risk reduction require distinct definitions for reference values. These differences also have implications for the interpretation of reference values. Reference values for chronic disease relations usually reflect "optimal" intakes, whereas those for nutrient deficiencies are based on intake requirements to prevent deficiencies (98). DRI committees must often express reference values for chronic diseases as reductions in specific relative risks that vary by intake.

Ideally, future DRI values will be more applicable to specific population groups and more relevant to diverse settings, and they will better target chronic disease risks. The challenges that we have reviewed in the earlier sections underscore the fact that the broader incorporation of chronic disease endpoints into DRIs requires more sophisticated approaches than those that DRI committees have previously used. This section describes procedural issues pertaining to how to accommodate chronic disease endpoints into future DRI review processes.

Process components and options

The process to establish the current DRI values consisted of reviewing a group of related food substances that clearly focus on essential nutrients. When DRI committees selected indicators for setting reference values for adequacy or benefit and for potential increases in risk of harm, they considered both classical nutrient deficiency and chronic disease endpoints. The endpoint they selected depended primarily on the strength of the available evidence. DRI committees estimated reference values for adequacy (i.e., an EAR/RDA or AI) and increased risk (i.e., UL) across life-stage groups often by using extrapolations. Summary tables of reference values for adequacy did not identify whether selected endpoints were based on classical nutrient deficiency or chronic disease endpoints. ULs were based on measures of toxicity. Users had to consult the supporting text to determine the nature of the indicators or endpoints used. The continued need for reference intake values based on either classical nutrient endpoints or chronic disease risk and the attendant challenges suggest ≥2 options.

Option 1: Continue to use a single DRI development process

One option is to continue considering chronic disease endpoints in future DRI reviews but to expand the types of reference values to clearly distinguish those based on classical nutrient adequacy from those based on chronic disease endpoints (Table 4). This option makes the addition of CD_{XX} (where XX denotes the specific chronic disease) and UL_{CD} values or ranges a natural extension of the current process. A major advantage is that DRI committees would continue to use a single process to develop all reference values for individual food substances (or small groups of food substances). To enhance the usability of this option, future reference value summary tables could clearly describe the nature of the health indicator that the DRI committee used to establish each of the dietary reference values (e.g., EAR based on a disease of deficiency, CD_{XX} based on a chronic disease, UL based on traditional toxicities, UL_{CD} based on a chronic disease). The simultaneous review and establishment of all values related to ≥1 food substances would ensure consistency, when appropriate, among the multiple endpoints that a committee used. This approach also allows committees to suggest how to apply the various values (e.g., the populations to which these values apply under given conditions).

The challenges to continuing to include reference values based on chronic disease risk reduction within the DRI process result from experiences of DRI committees in applying the present framework to chronic disease endpoints and the expanding understanding of the pathophysiology of diseases of interest. Therefore, this option requires an expanded set of approaches for setting reference values (as described in section VI) because the current EAR/RDA and UL models often do not work well for chronic disease endpoints. This option also likely requires the development of criteria and approaches for addressing the types of evidence available for evaluating relations between food substances and chronic diseases, as described in section V-A.

An advantage is that this option would integrate multiple disciplines because future DRI committees would need a broader range of expertise than previous DRI committees, bringing an interdisciplinary approach to the setting of DRI values. However, reaching consensus among experts with different experiences, subject matter knowledge, and public health perspectives could be more challenging than narrower approaches.

Option 2: Create 2 separate processes for developing DRIs

A second option is to create 2 separate but complementary, and possibly iterative or integrated, committees to develop reference values on the basis of chronic disease endpoints or deficiency diseases. The FNB or a government agency could appoint a new committee to establish reference values on the basis of chronic disease endpoints, or an existing group that is independent of the National Academies of Sciences, Engineering, and Medicine (e.g., expert panels from chronic disease societies or standing government advisory committees) could establish these reference values. This new reference-setting group would coordinate its activities closely with the current DRI process based on adequacy.

As the volume of data grows over time, a challenge of incorporating chronic disease endpoints into current DRI processes will be appointing expert groups that can adequately address the challenges of analyzing all of the relevant evidence and calculating intake-response relations on the basis of either classical nutrient deficiencies or relations with chronic diseases for specific nutrients or groups of nutrients. Therefore, an advantage is that this option could allow committees to focus on the literature and challenges associated only with classical nutrient risks or with chronic diseases. Close coordination of these 2 committees would enhance the likelihood of consistent approaches to reference value development for the same food substances while engaging individuals with the most relevant expertise and, subsequently, more relevant audiences in the implementation of these reference values.

However, coordinating 2 separate committees is more complex than the current approach. The content experts who can best evaluate chronic disease risk might not be familiar with DRI processes and applications. A major disadvantage is that if coordination is not successful, the risk is high of developing contradictory reference values because committees could use different methods and frameworks. In addition, coordinating >1 panel will be more time-consuming and costly than the current structure. Another major disadvantage is that a single, internationally recognized authoritative body (i.e., the FNB) manages the current DRI process. As a result, the deliberations and decision-making processes of DRI committees are independent of vested interests, which enhance the integrity and status of their decisions. Assigning 2 separate, but coordinated, committees to develop reference values for the same food substances might not achieve the desired level of independence and integrity.

Starting-point issues and options

The starting point of current DRI processes is individual food substances, and DRI committees consider all pertinent outcomes related to varying intakes of given food substances. A possible alternative is to start with a chronic disease or diseases and then identify all food substances with established effects on that disease.

Option 1: Establish DRIs for individual or small groups of interrelated food substances

Advantages of this option are that it is consistent with current DRI approaches and that some key uses (e.g., for regulatory purposes) involve individual food substances. A disadvantage is the difficulty of separating the effects of individual food substances from those of diets and dietary patterns when addressing chronic disease relations. Most of the available evidence comes from observational studies with the strong potential to confound relations between multiple food substances and targeted chronic diseases.

Option 2: Establish DRIs for multiple food substances on the basis of a chronic disease endpoint

This approach requires a different paradigm than the one that DRI committees currently use. For each selected chronic disease, DRI committees would develop a reference value for all food substances that have a causal relation with the risk of that disease. This approach could probably accommodate interactions between food substances more easily than option 1. Because this approach is different from the current DRI approach, it would require a separate process or a major revamping and expansion of the current process. Developing DRIs in this way would be more complex and probably more expensive. Such a process would probably also require some a priori criteria to limit the number of chronic diseases that DRI committees consider to a manageable number. As a result, DRI committees might be unable to address some chronic diseases for which evidence of benefit of certain food-substance intakes exists (e.g., lutein and reduced risk of macular degeneration) but that do not receive a high-enough priority rating for the committee to consider them. This process could also compete with existing approaches to chronic disease prevention, such as the processes that chronic disease societies use to develop guidelines for disease prevention, which could lead to inconsistent recommendations.

VIII. FORTHCOMING TOOLS

The challenges identified by the working group led them to briefly consider examples of forthcoming tools and novel study designs of potential future utility in overcoming anticipated hurdles (e.g., addressing complexities related to multiple, interactive etiologies and longitudinal characteristics of chronic diseases). Neither the tools nor the study designs that we considered are under development specifically for establishing food-substance reference values. We viewed these examples as potentially adaptable to future DRI processes that focus on relations between food substances and chronic diseases and that represent research opportunities (**Table 10**).

Biomarker-based dietary assessment

As noted above, most of the literature on dietary factors in relation to chronic disease is based on observational studies that use self-report tools for individual dietary assessment. These intake-assessment tools are known to be associated with substantial underestimation biases, particularly for energy intakes (66, 67). However, for a few nutritional variables, there is an established biomarker of short-term intake; the most notable examples are a doubly labeled water biomarker for energy (99) and a urinary nitrogen biomarker for protein (100). The self-report data do not align well with these biomarkers, especially for energy, where correlations are mainly in the range of 0.0–0.2 [e.g. (67)]. Furthermore, when studies used these biomarkers to correct (calibrate) associations between energy consumption and chronic disease, they found strong positive associations for

TABLE 10
Opportunities for research related to basing DRIs on chronic disease endpoints

Report section or subsection	Topic
Biomarkers of intake	Evaluation and qualification of biomarkers of long-term food substance intake
Selection of chronic disease outcome measures	Evaluation and qualification of biomarkers of chronic diseases
Tools for assessing the evidence	Development of tools for integrating food substance–specific quality of evidence criteria with generic criteria for evaluating study quality
Options for acceptable levels of confidence that the relation is causal	Development of level of evidence criteria for setting different types of DRI values on the basis of chronic disease endpoints
Shape of the intake-response curve	Development of statistical approaches and theoretical paradigms for integrating diverse relations between food substances and chronic diseases, such as U-shaped intake-risk modeling
Forthcoming tools	Further evaluation and consideration of criteria for assessing the utility of Mendelian randomization for setting DRI values
	Further evaluation and consideration of systems science in setting DRI values
	Further evaluation and consideration of the usefulness of evolving concepts and understanding of precision medicine in setting DRI values developed specifically for individuals

[1] DRI, Dietary Reference Intake.

prominent vascular diseases, cancers, and diabetes (with some caveat about the need to use BMI for intake assessment in the calibration procedure) that are not evident without biomarker calibration (45). This experience suggests that a concerted research effort to develop qualified surrogate intake biomarkers for additional dietary substances (e.g., the use of metabolomics profiles in urine and blood) could create important opportunities to strengthen information on associations between diet and chronic disease outcomes for use in future DRIs and for other purposes. This approach could also allow observational study researchers to reduce their dependence on dietary self-report intake data and instead measure qualified biomarkers—for example, in stored biospecimens—to analyze prospective cohort data in a case-control mode.

Mendelian randomization and causality

The impact of factors, such as confounding and reverse causation, cannot be underestimated when data derived from observational studies are considered. However, in the absence of RCT data, Mendelian randomization may provide useful information for making causal inferences about observed associations between a food substance and a chronic disease in an observational study. This method uses genetic variants within a population that modify the relation between an "exposure" and a phenotype. DRI committees could use genetic variants that modify the status or metabolism of a food substance to assess its relation to chronic disease risk with consideration of the limitations of this approach. For example, studies have examined the relations between gene variants that modify circulating 25-hydroxyvitamin D concentrations and the risk of several chronic diseases, including multiple sclerosis (51) and CVD (50), all-cause mortality (48), and surrogate endpoints, such as hypertension (49). Other studies have examined the association between gene variants that modify circulating triglycerides and coronary artery disease (101) or HDL cholesterol and type 2 diabetes (102). These studies might offer an alternative or complementary approach for inferring causality in specific situations.

U-shaped dose-risk relations

Researchers have modeled the U-shaped exposure-risk relation for copper by using severity scoring and categorical regression analysis to develop a single intake value that balances the risk of deficiency with that of adverse events, including toxicity (103). This approach could simultaneously fit multiple endpoints (e.g., deficiency, chronic disease, and excess) to a U-shaped or J-shaped intake-response curve that maximizes benefit and minimizes the probability of an adverse outcome due to either excess or inadequate intake of a food substance. The bottom of the U-shaped curve for copper minimizes the total risk of an adverse outcome due to excess or deficiency (or both), and this curve provides a possible benchmark for establishing dietary reference intake values for food substances with U-shaped intake-response relations. Confidence limits around the value might also be useful in establishing an allowable range of intakes (103).

An advantage is that this approach integrates the risk of multiple endpoints, including those that are beneficial or adverse, related to the intake of a nutrient while enabling a single best estimate of the exposure that minimizes overall risk. A challenge is the likely lack of accurate data on intakes that result in deficiency, chronic disease, and/or toxicity in different population groups or of the ability to integrate all endpoints. In addition, the categorization of endpoints and the use of scoring criteria to categorize outcome severity are subjective, which may result in bias. Some information or data are "hidden" in the model, which reduces transparency. This approach could limit flexibility in the application of (multiple) reference values associated with a single endpoint (or variety of endpoints) and the use of these values in personalized medicine because it results in the development of a single optimized value or range. However, the approach might be valuable for setting optimized intake values for food substances that have a narrow range or no range between maximal benefit and minimal harm.

Enhanced function-based DRIs

The options in this document focus on the risk of chronic disease. However, it might be possible to apply these or similar

options to the relation between food substances and enhanced function, possibly within the normal range. Examples of endpoints include enhanced cognition and endothelial elasticity. However, because DRI-based conclusions serve as authoritative statements for health claims on food labels, it would need to be clear that reference values based on enhanced function do not necessarily reduce chronic disease risk and are outside the context of chronic disease risk reduction. The interpretation and use of these values could be challenging. Concerns similar to those about biomarkers of chronic disease endpoints apply to biomarkers of enhanced function. In addition, the concept of enhanced function might be more similar to the concept of nutrient adequacy than to that of chronic disease risk.

Systems science

Systems science is an interdisciplinary field that focuses on the nature of simple to complex systems that aims to develop interdisciplinary foundations that are applicable in a variety of areas, such as biology, medicine, and nutrition. The relations between nutrition and disease are complex and bidirectional (104). For example, many infectious diseases cause malnourishment even when the food and nutrient supply is consistent with current reference intakes. Another example is that malnourished and overnourished obese people are more susceptible to many diseases, including infectious diseases. Systems science could potentially integrate the multitude of factors that influence mechanistic relations between a food substance and a chronic disease, including such variables as compromised immune function, reduced epithelial integrity, an altered microbiome, oxidative stress, and other functions. Equally important is that systems science might enable the more effective inclusion of longitudinal aspects of relations between diet and chronic disease across life stages. Comprehensive system frameworks would be necessary that link dietary patterns and intakes of specific food substances to food-substance absorption, metabolism, bioactivity, excretion, tissue uptake, and function along with a variety of metabolic and functional health endpoints and food-substance to food-substance interactions. This tool also could accommodate the added complexity of environmental and behavioral factors that influence diet-disease risk relations. If successful, this approach would improve the ability to recommend what, when, and how to eat and what to prioritize to influence an individual's health status.

Application of chronic disease–based DRIs in precision medicine

Precision medicine focused on prevention involves interventions targeted to the needs of an individual on the basis of his or her genetic, biomarker, phenotypic, or psychosocial characteristics (105). Although clinicians do not apply precision medicine widely, genetic testing for polymorphisms associated with the risk of a disease (e.g., cancer) is increasingly available, and some specific therapies for treating these diseases exist. Examples of the application of precision medicine to the risk of a chronic disease associated with a food substance already exist. For example, genetic polymorphisms associated with varied responsiveness to statins for the treatment of CVD are available and now influence the choice of diet therapy to combine with drug therapy (106). Such approaches are not new to nutrition (e.g., dietary recommendations are available for highly penetrant and severe monogenic traits, phenylketonuria, and thalassemia, as well as more complex conditions, such as type 2 diabetes). New technologies that enable increasingly precise targeting of diet-based recommendations are likely to influence future DRI values and frameworks and help solve current challenges related to the use of chronic disease endpoints.

IX. CONCLUSIONS

The development of the DRIs has been critical for the successful (near) elimination of diseases of deficiency in Canada and the United States. If the DRI framework could be expanded to more effectively include chronic disease outcomes, the potential impact on public health would be even greater. This report identified the evidence-related and intake-response-relation challenges that have hampered the inclusion of chronic disease endpoints in the derivation of DRIs with the use of a traditional framework and approach. The report presents several potential options to address those challenges. The next step will be to make decisions about the feasibility of including chronic disease endpoints in future DRI reviews and to determine which options and/or their adaptations warrant inclusion in guiding principles for basing DRI values on chronic disease endpoints.

Traditional DRIs have always been based on adequacy for the apparently healthy population. However, when DRI values are based on chronic disease endpoints, the target population or populations might be narrower (e.g., individuals with high blood pressure or obesity). Although beyond the scope of this report, further consideration of how to define target populations when DRIs are based on reduction in chronic disease risk may be needed.

The report also highlights several research opportunities that are key to the derivation of future DRIs based on chronic disease endpoints (Table 10). Among the most salient examples of those opportunities are the need for qualified biomarkers of long-term intakes for a large array of nutritional variables (i.e., nutrients and other food substances), tools specifically designed to assess the quality of evidence required for setting DRIs, and novel statistical and other analytic methods for integrating diverse relations linking specific food components to multiple outcomes of interest.

We thank Joyce Merkel for help with the bibliography and editing and Debbie Berlyne for her expertise and patience in technical editing. We acknowledge the support and encouragement of the Joint Canada/US DRI Working Group.

The authors' responsibilities were as follows—EAY, AJM, and LSG-F: designed and provided oversight of the project and drafting of the report; CG: served as chair of the working group and helped with planning the project, drafting the manuscript, and editing the report; JDA, SAA, DMB, ALC, WRH, DH, JCK, DK, DLO, RLP, JVR, and GAW: served as members of the working group; and all authors: contributed to and were responsible for the final content of the manuscript. Working group duties included participation in conference calls and the public workshop, provision of expert input throughout the process, and conceptualizing, planning, and writing of the report. None of the authors declared a conflict of interest.

REFERENCES

1. Food and Nutrition Board, Institute of Medicine. Dietary Reference Intakes: the essential guide to nutrient requirements. Washington (DC): National Academies Press; 2006 [cited 2016 May 4]. Available from: https://fnic.nal.usda.gov/sites/fnic.nal.usda.gov/files/uploads/DRIEssentialGuideNutReq.pdf.
2. Taylor CL. Framework for DRI development components "known" and components "to be explored" background paper. Food and Nutrition Board, Institute of Medicine; 2008 [cited 2016 May 4]. Available from: https://fnic.nal.usda.gov/sites/fnic.nal.usda.gov/files/uploads/Framework_DRI_Development.pdf.
3. Murphy SP, Yates AA, Atkinson SA, Barr SI, Dwyer J. History of nutrition: the long road leading to the Dietary Reference Intakes for the United States and Canada. Adv Nutr 2016;7:157–68.
4. Goodman RA, Posner SF, Huang ES, Parekh AK, Koh HK. Defining and measuring chronic conditions: imperatives for research, policy, program, and practice. Prev Chronic Dis 2013;10:E66.
5. WHO. Noncommunicable diseases. Media centre fact sheet. 2015 [cited 2016 Nov 1]. Available from: http://www.who.int/mediacentre/factsheets/fs355/en/.
6. Institute of Medicine Committee on Qualification of Biomarkers and Surrogate Endpoints in Chronic Disease; Board on Health Care Services; Board on Health Sciences Policy; Food and Nutrition Board. In: Micheel CM, Ball JR, editors. Evaluation of biomarkers and surrogate endpoints in chronic disease. Washington (DC): National Academies Press; 2010 [cited 2016 May 4]. Available from: https://iom.nationalacademies.org/Reports/2010/Evaluation-of-Biomarkers-and-Surrogate-Endpoints-in-Chronic-Disease.aspx.
7. Institute of Medicine, Food Nutrition Board. How should the Recommended Dietary Allowances be revised? Washington (DC): National Academies Press; 1994 [cited 2016 May 4]. Available from: https://fnic.nal.usda.gov/sites/fnic.nal.usda.gov/files/uploads/How_Should_RDAs_Be_Revised.pdf.
8. Institute of Medicine Subcommittee on Interpretation and Uses of Dietary Reference Intakes; Standing Committee on the Scientific Evaluation of Dietary Reference Intakes. Dietary Reference Intakes: applications in dietary planning. Washington (DC): National Academies Press; 2003 [cited 2016 May 4]. Available from: https://fnic.nal.usda.gov/dietary-guidance/dri-nutrient-reports/applications-dietary-planning#overlay-context=dietary-guidance/dietary-reference-intakes/dri-reports.
9. Carriquiry AL. Assessing the prevalence of nutrient inadequacy. Public Health Nutr 1999;2:23–33.
10. Institute of Medicine. Dietary Reference Intakes: applications in dietary assessment. Washington (DC): National Academies Press; 2000 [cited 2016 May 4]. Available from: http://fnic.nal.usda.gov/dietary-guidance/dri-reports/applications-dietary-assessment.
11. Institute of Medicine, Food Nutrition Board. Dietary Reference Intakes: a risk assessment model for establishing upper intake levels for nutrients. Washington (DC): National Academies Press; 1998 [cited 2016 May 4]. Available from: https://fnic.nal.usda.gov/sites/fnic.nal.usda.gov/files/uploads/DRI_Risk_Assessment_Model.pdf.
12. National Research Council Committee on Diet Health. Diet and health: implications for reducing chronic disease risk. Washington (DC): National Academies Press; 1989.
13. Sheffer M, Taylor CL; Rapporteurs; Planning Committee for Dietary Reference Intakes Review Workshop, Food and Nutrition Board Institute of Medicine. The development of DRIs 1994-2004: lessons learned and new challenges: workshop summary. Washington (DC): National Academies Press; 2008 [cited 2016 May 4]. Available from: https://fnic.nal.usda.gov/sites/fnic.nal.usda.gov/files/uploads/Development_DRIs_1994-2004_Lessons_Learned.pdf.
14. Institute of Medicine Committee to Review Dietary Reference Intakes for Vitamin D and Calcium. In: Ross AC, Taylor CL, Yaktine AL, Del Valle HB, editors. Dietary Reference Intakes for calcium and vitamin D. Washington (DC): National Academies Press; 2011 [cited 2016 May 4]. Available from: https://fnic.nal.usda.gov/dietary-guidance/dri-nutrient-reports/vitamin-d-and-calcium.
15. Cheney M. Selection of endpoints for determining EARs/AIs and ULs. IOM/FNB Workshop on Dietary Reference Intakes: the development of DRIs 1994-2004; lessons learned and new challenges. Washington (DC): National Academies Press; 2007 [cited 2016 May 4]. Available from: https://fnic.nal.usda.gov/sites/fnic.nal.usda.gov/files/uploads/Development_DRIs_1994-2004_Lessons_Learned.pdf.
16. US Food and Drug Administration. Label claims for conventional foods and dietary supplements. 2013 [cited 2016 Mar 1]. Available from: http://www.fda.gov/Food/IngredientsPackagingLabeling/LabelingNutrition/ucm111447.htm.
17. US Dietary Reference Intakes Subcommittee; Canadian Federal Dietary Reference Intakes Steering Committee. Options for consideration of chronic disease endpoints for Dietary Reference Intakes (DRIs). 2016 [cited 2016 Apr 29]. Available from: https://ods.od.nih.gov/News/DRI_Workshop_March_10-11_2015.aspx.
18. Mayne ST, Ferrucci LM, Cartmel B. Lessons learned from randomized clinical trials of micronutrient supplementation for cancer prevention. Annu Rev Nutr 2012;32:369–90.
19. Scottish Intercollegiate Guidelines Network (SIGN) 50 methodology. Critical appraisal: notes and checklists. Checklist 2: randomised controlled trials, methodology. Checklist 3: cohort studies, methodology. Checklist 4: case-control studies. 2014 [cited 2015 Jul 21]. Available from: http://www.sign.ac.uk/methodology/checklists.html.
20. GRADE) Working Group. Grading of recommendations assessment, development and evaluation (GRADE) [cited 2015 Jul 22]. Available from: http://www.gradeworkinggroup.org/.
21. Chung M, Balk EM, Ip S, Raman G, Yu WW, Trikalinos TA, Lichtenstein AH, Yetley EA, Lau J. Reporting of systematic reviews of micronutrients and health: a critical appraisal. Am J Clin Nutr 2009;89:1099–113.
22. Institute of Medicine Food and Nutrition Board. IOM/FNB planning meeting. Framework Dietary Reference Intake development—chronic disease indicators. Washington (DC): Institute of Medicine Food and Nutrition Board; 2009 [cited 2016 May 4]. Available from: https://fnic.nal.usda.gov/sites/fnic.nal.usda.gov/files/uploads/Framework_DRI_Development_Chronic_Disease_Indicators.pdf.
23. Institute of Medicine Panel on Dietary Reference Intakes for Electrolytes and Water; Standing Committee on the Scientific Evaluation of Dietary Reference Intakes; Food and Nutrition Board. Dietary Reference Intakes for water, potassium, sodium, chloride and sulfate. Washington (DC): National Academies Press; 2004 [cited 2016 May 4]. Available from: https://fnic.nal.usda.gov/dietary-guidance/dri-nutrient-reports/water-potassium-sodium-chloride-and-sulfate#overlay-context=dietary-guidance/dri-reports/vitamin-c-vitamin-e-selenium-and-carotenoids.
24. Institute of Medicine Standing Committee on the Scientific Evaluation of Dietary Reference Intakes. Dietary Reference Intakes for energy, carbohydrate, fiber, fat, fatty acids, cholesterol, protein, and amino acids. Washington (DC): National Academies Press; 2005 [cited 2016 May 4]. Available from: https://fnic.nal.usda.gov/dietary-guidance/dri-nutrient-reports/energy-carbohydrate-fiber-fat-fatty-acids-cholesterol-protein#overlay-context=dietary-guidance/dietary-reference-intakes/dri-reports.
25. Centers for Disease Prevention and Control. Chronic disease overview. 2015 [cited 2016 Nov 1]. Available from: http://www.cdc.gov/chronicdisease/overview.
26. Statistics Canada. The 10 leading causes of death, 2011. 2015 [cited 2016 Mar 1]. Available from: http://www.statcan.gc.ca/pub/82-625-x/2014001/article/11896-eng.htm.
27. UN General Assembly. Sixty-sixth session agenda item 117, resolution adopted by the general assembly [without reference to a main committee (a/66/l.1)] 66/2. Political declaration of the high-level meeting of the general assembly on the prevention and control of noncommunicable diseases. 2012. [cited 2016 Nov 1]. Available from: http://www.who.int/nmh/events/un_ncd_summit2011/political_declaration_en.pdf.
28. Maki KC, Slavin JL, Rains TM, Kris-Etherton PM. Limitations of observational evidence: implications for evidence-based dietary recommendations. Adv Nutr 2014;5:7–15.
29. Institute of Medicine Standing Committee on the Scientific Evaluation of Dietary Reference Intakes. Dietary Reference Intakes for calcium, phosphorus, magnesium, vitamin D, and fluoride. Washington (DC): National Academies Press; 1997, [cited 2016 May 4]. Available from: https://fnic.nal.usda.gov/dietary-guidance/dri-nutrient-reports/calcium-phosphorus-magnesium-vitamin-d-and-fluoride#overlay-context=dietary-guidance/dietary-reference-intakes/dri-reports.
30. Mozaffarian D, Rimm EB. Fish intake, contaminants, and human health: Evaluating the risks and the benefits. JAMA 2006;296:1885–99.

31. National Research Council Subcommittee on the Tenth Edition of the Recommended Dietary Allowances. Recommended Dietary Allowances. 10th ed. Washington (DC): National Academies Press; 1989.
32. Flock MR, Harris WS, Kris-Etherton PM. Long-chain omega-3 fatty acids: time to establish a Dietary Reference Intake. Nutr Rev 2013;71: 692–707.
33. Lupton JR, Atkinson SA, Chang N, Fraga CG, Levy J, Messina M, Richardson DP, van Ommen B, Yang Y, Griffiths JC, et al. Exploring the benefits and challenges of establishing a DRI-like process for bioactives. Eur J Nutr 2014;53(Suppl 1):1–9.
34. Ellwood K, Balentine DA, Dwyer JT, Erdman JW Jr., Gaine PC, Kwik-Uribe CL. Considerations on an approach for establishing a framework for bioactive food components. Adv Nutr 2014;5:693–701.
35. Manson JE, Kaunitz AM. Menopause management—getting clinical care back on track. N Engl J Med 2016;374:803–6.
36. Institute of Medicine (US) Standing Committee on the Scientific Evaluation of Dietary Reference Intakes and its Panel on Folate, Other B Vitamins, and Choline. Dietary Reference Intakes for thiamin, riboflavin, niacin, vitamin B6, folate, vitamin B12, pantothenic acid, biotin, and choline [Internet]. Washington (DC): National Academies Press (US); 1998 [cited 2016 Nov 3]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK114310/.
37. Institute of Medicine Panel on Dietary Antioxidants Related Compounds. Dietary Reference Intakes for vitamin C, vitamin E, selenium, and carotenoids. Washington (DC): National Academies Press; 2000 [cited 2016 May 4]. Available from: https://fnic.nal.usda.gov/dietary-guidance/dri-nutrient-reports/vitamin-c-vitamin-e-selenium-and-carotenoids#overlay-context=dietary-guidance/dri-reports/vitamin-vitamin-k-arsenic-boron-chromium-copper-iodine-iron-manganese.
38. Institute of Medicine Panel on Micronutrients. Dietary Reference Intakes for vitamin A, vitamin K, arsenic, boron, chromium, copper, iodine, iron, manganese, molybdenum, nickel, silicon, vanadium, and zinc. Washington (DC): National Academies Press; 2001 [cited 2016 May 4]. Available from: https://fnic.nal.usda.gov/dietary-guidance/dri-nutrient-reports/vitamin-vitamin-k-arsenic-boron-chromium-copper-iodine-iron#overlay-context=dietary-guidance/dri-reports/thiamin-riboflavin-niacin-vitamin-b6-folate-vitamin-b12-pantothenic.
39. Porta M, Greenland S, Hernán M, dos Santos Silva I, Last M. A dictionary of epidemiology. 6th ed. New York: Oxford University Press; 2014.
40. Wells GA. Approaches for evaluating evidence–that is not derived from randomized controlled trials (videocast). Workshop on Options for Consideration of Chronic Disease Endpoints for Dietary Reference Intakes (DRIs), March 10–11, 2015 [cited 2016 Nov 1]. Available from: http://videocast.nih.gov/summary.asp?Live=15713&bhcp=1.
41. Higgins JP, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD, Savovic J, Schulz KF, Weeks L, Sterne JA. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011; 343:d5928.
42. Sterne JAC, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, Henry D, Altman DG, Ansari MT, Boutron I, et al. ROBINS-I: a tool for assessing risk of bias in non-randomized studies of interventions. BMJ 2016;355:i4919.
43. Viswanathan M, Ansari MT, Berkman ND, Chang S, Hartling L, McPheeters LM, Santaguida PL, Shamliyan T, Singh K, Tsertsvadze A, et al. Assessing the risk of bias of individual studies in systematic reviews of health care interventions [Internet]. Agency for Healthcare Research and Quality Methods Guide for Comparative Effectiveness Reviews. Mar 2012. AHRQ Publication No. 12-EHC047-EF. [cited 2016 Nov 3]. Available from: http://effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?productid=998&pageaction=displayproduct.
44. Freedman LS, Commins JM, Moler JE, Willett W, Tinker LF, Subar AF, Spiegelman D, Rhodes D, Potischman N, Neuhouser ML, et al. Pooled results from 5 validation studies of dietary self-report instruments using recovery biomarkers for potassium and sodium intake. Am J Epidemiol 2015;181:473–87.
45. Zheng C, Beresford SA, Van Horn L, Tinker LF, Thomson CA, Neuhouser ML, Di C, Manson JE, Mossavar-Rahmani Y, Seguin R, et al. Simultaneous association of total energy consumption and activity-related energy expenditure with risks of cardiovascular disease, cancer, and diabetes among postmenopausal women. Am J Epidemiol 2014;180:526–35.
46. Prentice RL. Surrogate endpoints in clinical trials: definition and operational criteria. Stat Med 1989;8:431–40.
47. Peters SA, den Ruijter HM, Bots ML, Moons KG. Improvements in risk stratification for the occurrence of cardiovascular disease by imaging subclinical atherosclerosis: a systematic review. Heart 2012;98:177–84.
48. Afzal S, Brondum-Jacobsen P, Bojesen SE, Nordestgaard BG. Genetically low vitamin D concentrations and increased mortality: Mendelian randomisation analysis in three large cohorts. BMJ 2014;349:g6330.
49. Vimaleswaran KS, Cavadino A, Berry DJ, Jorde R, Dieffenbach AK, Lu C, Alves AC, Heerspink HJ, Tikkanen E, Eriksson J, et al. Association of vitamin D status with arterial blood pressure and hypertension risk: a Mendelian randomisation study. Lancet Diabetes Endocrinol 2014;2:719–29.
50. Brøndum-Jacobsen P, Benn M, Afzal S, Nordestgaard BG. No evidence that genetically reduced 25-hydroxyvitamin D is associated with increased risk of ischaemic heart disease or myocardial infarction: a Mendelian randomization study. Int J Epidemiol 2015; 44:651–61.
51. Mokry LE, Ross S, Ahmad OS, Forgetta V, Smith GD, Leong A, Greenwood CM, Thanassoulis G, Richards JB. Vitamin D and risk of multiple sclerosis: a Mendelian randomization study. PLoS Med 2015;12: e1001866. Erratum in: PLoS Med 2016;13(3):e1001981.
52. Haycock PC, Burgess S, Wade KH, Bowden J, Relton C, Davey Smith G. Best (but oft-forgotten) practices: the design, analysis, and interpretation of Mendelian randomization studies. Am J Clin Nutr 2016;103:965–78.
53. Omenn GS, Goodman G, Thornquist M, Grizzle J, Rosenstock L, Barnhart S, Balmes J, Cherniack MG, Cullen MR, Glass A, et al. The beta-Carotene And Retinol Efficacy Trial (CARET) for chemoprevention of lung cancer in high risk populations: smokers and asbestos-exposed workers. Cancer Res 1994;54:2038s–43s.
54. Alpha Tocopherol Beta Carotene Cancer Prevention Study Group. The effect of vitamin E and beta carotene on the incidence of lung cancer and other cancers in male smokers. N Engl J Med 1994;330: 1029–35.
55. Hennekens CH, Buring JE, Manson JE, Stampfer M, Rosner B, Cook NR, Belanger C, LaMotte F, Gaziano JM, Ridker PM, et al. Lack of effect of long-term supplementation with beta carotene on the incidence of malignant neoplasms and cardiovascular disease. N Engl J Med 1996;334:1145–9.
56. Clarke R, Halsey J, Bennett D, Lewington S. Homocysteine and vascular disease: review of published results of the homocysteine-lowering trials. J Inherit Metab Dis 2011;34:83–91.
57. Lippman SM, Klein EA, Goodman PJ, Lucia MS, Thompson IM, Ford LG, Parnes HL, Minasian LM, Gaziano JM, Hartline JA, et al. Effect of selenium and vitamin E on risk of prostate cancer and other cancers: the Selenium and Vitamin E Cancer Prevention Trial (SELECT). JAMA 2009;301:39–51.
58. Bleys J, Miller ER III, Pastor-Barriuso R, Appel LJ, Guallar E. Vitamin-mineral supplementation and the progression of atherosclerosis: a meta-analysis of randomized controlled trials. Am J Clin Nutr 2006;84:880–7; quiz: 954–5.
59. Sacks FM, Svetkey LP, Vollmer WM, Appel LJ, Bray GA, Harsha D, Obarzanek E, Conlin PR, Miller ER III, Simons-Morton DG, et al; DASH-Sodium Collaborative Research Group. Effects on blood pressure of reduced dietary sodium and the Dietary Approaches to Stop Hypertension (DASH) diet. N Engl J Med 2001;344:3–10.
60. Orchard TJ, Temprosa M, Goldberg R, Haffner S, Ratner R, Marcovina S, Fowler S. The effect of metformin and intensive lifestyle intervention on the metabolic syndrome: the Diabetes Prevention Program randomized trial. Ann Intern Med 2005;142:611–9.
61. Beresford SA, Johnson KC, Ritenbaugh C, Lasser NL, Snetselaar LG, Black HR, Anderson GL, Assaf AR, Bassford T, Bowen D, et al. Low-fat dietary pattern and risk of colorectal cancer: the Women's Health Initiative randomized controlled dietary modification trial. JAMA 2006;295:643–54.
62. Prentice RL, Caan B, Chlebowski RT, Patterson R, Kuller LH, Ockene JK, Margolis KL, Limacher MC, Manson JE, Parker LM, et al. Low-fat dietary pattern and risk of invasive breast cancer: the Women's Health Initiative randomized controlled dietary modification trial. JAMA 2006; 295:629–42.

63. Klein EA, Thompson IM Jr., Tangen CM, Crowley JJ, Lucia MS, Goodman PJ, Minasian LM, Ford LG, Parnes HL, Gaziano JM, et al. Vitamin E and the risk of prostate cancer: the Selenium and Vitamin E Cancer Prevention Trial (SELECT). JAMA 2011;306:1549–56.
64. Schatzkin A, Gail M. The promise and peril of surrogate end points in cancer research. Nat Rev Cancer 2002;2:19–27.
65. Prentice RL. Surrogate and mediating endpoints: current status and future directions. J Natl Cancer Inst 2009;101:216–7.
66. Prentice RL, Mossavar-Rahmani Y, Huang Y, Van Horn L, Beresford SA, Caan B, Tinker L, Schoeller D, Bingham S, Eaton CB, et al. Evaluation and comparison of food records, recalls, and frequencies for energy and protein assessment by using recovery biomarkers. Am J Epidemiol 2011;174:591–603.
67. Freedman LS, Commins JM, Moler JE, Arab L, Baer DJ, Kipnis V, Midthune D, Moshfegh AJ, Neuhouser ML, Prentice RL, et al. Pooled results from 5 validation studies of dietary self-report instruments using recovery biomarkers for energy and protein intake. Am J Epidemiol 2014;180:172–88.
68. Rubin DB. For objective causal inference design trumps analysis. Ann Appl Stat 2008;2:808–40.
69. Chung M, Balk EM, Brendel M, Ip S, Lau J, Lee J, Lichtenstein A, Patel K, Raman G, Tatsioni A, et al. Vitamin D and calcium: a systematic review of health outcomes. Evid Rep Technol Assess (Full Rep) 2009;183:1–420.
70. Satija A, Hu FB. Big data and systematic reviews in nutritional epidemiology. Nutr Rev 2014;72:737–40.
71. Egger M, Ebrahim S, Smith GD. Where now for meta-analysis? Int J Epidemiol 2002;31:1–5.
72. Lichtenstein AH, Yetley EA, Lau J. Application of systematic review methodology to the field of nutrition. J Nutr 2008;138:2297–306.
73. Tovey D. The role of the Cochrane Collaboration in support of the WHO nutrition guidelines. Adv Nutr 2014;5:35–9.
74. Van't Veer P, Grammatikaki E, Matthys C, Raats MM, Contor L. EURRECA—framework for aligning micronutrient recommendations. Crit Rev Food Sci Nutr 2013;53:988–98.
75. Theodoratou E, Tzoulaki I, Zgaga L, Ioannidis JP. Vitamin D and multiple health outcomes: umbrella review of systematic reviews and meta-analyses of observational studies and randomised trials. BMJ 2014;348:g2035.
76. Newberry SJ, Chung M, Booth M, Maglione MA, Tang AM, O'Hanlon CE, Wang DD, Okunogbe A, Huang C, Motala A, et al. Omega-3 fatty acids and maternal and child health: an updated systematic review [Internet]. Evidence Report/Technology Assessment No. 224. (Prepared by the RAND Southern California Evidence-based Practice Center under Contract No. 290-2012-00006-I.) AHRQ Publication No. 16(17)-E003-EF. Rockville (MD): Agency for Healthcare Research and Quality; October 2016 [cited 2016 Nov 3]. Available from: https://www.effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productid=2320.
77. Balk EM, Adam GP, Langberg V, Halladay C, Chung M, Lin L, Robertson S, Yip A, Steele D, Smith BT, et al. Omega-3 fatty acids and cardiovascular disease: an updated systematic review [Internet]. Evidence Report/Technology Assessment No. 223. (Prepared by the Brown Evidence-based Practice Center under Contract No. 290-2015-00002-I.) AHRQ Publication No. 16-E002-EF. Rockville (MD): Agency for Healthcare Research and Quality; August 2016 [cited 2016 Nov 3]. Available from: https://www.effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productid=2261.
78. Hill AB. The environment and disease: association or causation? Proc R Soc Med 1965;58:295–300.
79. Bai A, Shukla VK, Bak G, Wells G. Quality assessment tools project report. Ottawa (Canada): Canadian Agency for Drugs and Technologies in Health; 2012.
80. Newberry SJ, Chung M, Shekelle PG, Booth MS, Liu JL, Maher AR, Motala A, Cui M, Perry T, Shanman R, et al. Vitamin D and calcium: a systematic review of health outcomes (update). Rockville (MD): Agency for Healthcare Research and Quality; 2014. Evidence Reports/Technology Assessments No.: 217 [cited 2016 Nov 1]. Available from: http://www.ncbi.nlm.nih.gov/books/NBK253540/.
81. Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M, Tugwell P. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. 2015 [cited 2015 Aug 10]. Available from: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.
83. Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, Porter AC, Tugwell P, Moher D, Bouter LM. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol 2007;7:10.
83. Shea BJ, Hamel C, Wells GA, Bouter LM, Kristjansson E, Grimshaw J, Henry DA, Boers M. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. J Clin Epidemiol 2009;62:1013–20.
84. Wells GA, Shea B, Higgins JP, Sterne J, Tugwell P, Reeves BC. Checklists of methodological issues for review authors to consider when including non-randomized studies in systematic reviews. Res Synth Methods 2013;4:63–77.
85. Whiting P, Savovic J, Higgins JP, Caldwell DM, Reeves BC, Shea B, Davies P, Kleijnen J, Churchill R. ROBIS: a new tool to assess risk of bias in systematic reviews was developed. J Clin Epidemiol 2016;69:225–34.
86. Berkman ND, Lohr KN, Ansari MT, Balk EM, Kane R, McDonagh M, Morton SC, Viswanathan M, Bass EB, Butler M, et al. Grading the strength of a body of evidence when assessing health care interventions: an EPC update. J Clin Epidemiol 2015;68:1312–24.
87. Wang MD, Gomes J, Cashman NR, Little J, Krewski D. A meta-analysis of observational studies of the association between chronic occupational exposure to lead and amyotrophic lateral sclerosis. J Occup Environ Med 2014;56:1235–42.
88. International Agency for Research on Cancer. Preamble to the IARC monographs on the evacuation of carcinogenic risks to humans. Lyon (France): World Health Organization; 2006 [cited 2016 Nov 1]. Available from: http://monographs.iarc.fr/ENG/Monographs/vol87/mono87-4.pdf.
89. Rhomberg LR, Goodman JE, Bailey LA, Prueitt RL, Beck NB, Bevan C, Honeycutt M, Kaminski NE, Paoli G, Pottenger LH, et al. A survey of frameworks for best practices in weight-of-evidence analyses. Crit Rev Toxicol 2013;43:753–84.
90. National Research Council. Review of EPA's integrated risk information system (IRIS) process. Washington (DC): National Academies Press; 2014.
91. Sundström J, Arima H, Jackson R, Turnbull F, Rahimi K, Chalmers J, Woodward M, Neal B. Effects of blood pressure reduction in mild hypertension: a systematic review and meta-analysis. Ann Intern Med 2015;162:184–91.
92. Satoh M, Ohkubo T, Asayama K, Murakami Y, Sakurai M, Nakagawa H, Iso H, Okayama A, Miura K, Imai Y, et al. Combined effect of blood pressure and total cholesterol levels on long-term risks of subtypes of cardiovascular death: evidence for cardiovascular prevention from observational cohorts in Japan. Hypertension 2015;65:517–24.
93. Schöttker B, Haug U, Schomburg L, Kohrle J, Perna L, Muller H, Holleczek B, Brenner H. Strong associations of 25-hydroxyvitamin D concentrations with all-cause, cardiovascular, cancer, and respiratory disease mortality in a large cohort study. Am J Clin Nutr 2013;97:782–93.
94. Ross AC, Russell RM, Miller SA, Munro IC, Rodricks JV, Yetley EA, Julien E. Application of a key events dose-response analysis to nutrients: a case study with vitamin A (retinol). Crit Rev Food Sci Nutr 2009;49:708–17.
95. Green TJ, Allen OB, O'Connor DL. A three-day weighed food record and a semiquantitative food-frequency questionnaire are valid measures for assessing the folate and vitamin B-12 intakes of women aged 16 to 19 years. J Nutr 1998;128:1665–71.
96. Jenab M, Slimani N, Bictash M, Ferrari P, Bingham SA. Biomarkers in nutritional epidemiology: applications, needs and new horizons. Hum Genet 2009;125:507–25.
97. Wu GD, Compher C, Chen EZ, Smith SA, Shah RD, Bittinger K, Chehoud C, Albenberg LG, Nessel L, Gilroy E, et al. Comparative metabolomics in vegans and omnivores reveal constraints on diet-dependent gut microbiota metabolite production. Gut 2016;65:63–72.
98. Beaton G. Letter to Scott Grundy, member of the Institute of Medicine's Standing Committee on the Scientific Evaluation of Dietary Reference Intakes. 1997 [cited 2016 May 3]. Available from: http://jn.nutrition.org/content/suppl/2015/05/13/jn.115.211185.DCSupplemental/nutrition211185SupplementaryData1.pdf.

99. Schoeller DA. Recent advances from application of doubly labeled water to measurement of human energy expenditure. J Nutr 1999;129: 1765–8.
100. Bingham SA. Urine nitrogen as a biomarker for the validation of dietary protein intake. J Nutr 2003;133(Suppl 3):921S–4S.
101. Sarwar N, Sandhu MS, Ricketts SL, Butterworth AS, Di Angelantonio E, Boekholdt SM, Ouwehand W, Watkins H, Samani NJ, Saleheen D, et al. Triglyceride-mediated pathways and coronary disease: collaborative analysis of 101 studies. Lancet 2010;375:1634–9. Erratum in: Lancet 2010;376(9735):90.
102. Haase CL, Tybjaerg-Hansen A, Nordestgaard BG, Frikke-Schmidt R. HDL cholesterol and risk of type 2 diabetes: a Mendelian randomization study. Diabetes 2015;64:3328–33.
103. Milton B, Farrell P, Birkett NB, Krewski D. Modeling U-shaped exposure-response relationships for agents that demonstrate toxicity due to both excess and deficiency. Risk Anal 2016 Apr 4 (Epub ahead of print; DOI: 10.1111/risa.12603).
104. Hammond RA, Dube L. A systems science perspective and transdisciplinary models for food and nutrition security. Proc Natl Acad Sci USA 2012;109:12356–63.
105. Jameson JL, Longo DL. Precision medicine–personalized, problematic, and promising. N Engl J Med 2015;372:2229–34.
106. Yu CY, Theusch E, Lo K, Mangravite LM, Naidoo D, Kutilova M, Medina MW. Hnrnpa1 regulates Hmgcr alternative splicing and modulates cellular cholesterol metabolism. Hum Mol Genet 2014;23: 319–32.

Appendix C

Acronyms and Abbreviations

AHRQ	Agency for Healthcare Research and Quality
AI	Adequate Intake
AICR	American Institute for Cancer Research
AMDR	Adequate Macronutrient Distribution Range
AMSTAR	A Measurement Tool to Assess Systematic Reviews
AR	Army Regulation
ARI	Acceptable Range of Intakes
BMC	bone mineral content
BMD	bone mineral density
BMI	body mass index
BOND	Biomarkers of Nutrition for Development
CCHS	Canadian Community Health Survey
CD	chronic disease
CDC	U.S. Centers for Disease Control and Prevention
CDER	Center for Drug Evaluation and Research
CHD	coronary heart disease
CI	confidence interval
CLA	conjugated linoleic acid
CVD	cardiovascular disease
DALY	disability-adjusted life year
DFE	dietary folate equivalent
DGA	*Dietary Guidelines for Americans*

DGAC	Dietary Guidelines Advisory Committee
DHA	docosahexaenoic acid
DNA	deoxyribonucleic acid
DRI	Dietary Reference Intake
DV	daily value
EAR	Estimated Average Requirement
EER	Estimated Energy Requirement
EPA	eicosapentaenoic acid
EPC	Evidence-based Practice Center
FDA	U.S. Food and Drug Administration
FFQ	food frequency questionnaire
FNB	Food and Nutrition Board
GRADE	Grading of Recommendations, Assessment, Development and Evaluation
HHS	U.S. Department of Health and Human Services
HR	hazard ratio
ICD	*International Classification of Diseases*
IHD	ischemic heart disease
IOM	Institute of Medicine
LDL	low-density lipoprotein
LOAEL	Lowest-Observed-Adverse-Effect Level
MDRI	Military Dietary Reference Intake
MREs	Meals Ready-to-Eat
MUFA	monounsaturated fatty acid
NCD	noncommunicable disease
NEL	Nutrition Evidence Library
NEL BAT	Nutrition Evidence Library Bias Assessment Tool
NHANES	National Health and Nutrition Examination Survey
NICE	National Institute for Health and Care Excellence
NIH	National Institutes of Health
NLEA	Nutrition Labeling and Education Act
NOAEL	No-Observed-Adverse-Effect Level
NOFS	nutrient or other food substance
NOS	Newcastle-Ottawa Scale

NTD	neural tube defect
NTP	National Toxicology Program
OHAT	Office of Health Assessment and Translation
OR	odds ratio
OIS	optimal information size
PAH	polycyclic aromatic hydrocarbon
PICO	population, intervention, comparator, and outcome
PICOTS	population, intervention, comparators, outcomes, timing, and setting
PREDIMED	Prevención con Dieta Mediterránea
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analysis
PRISMA-P	Preferred Reporting Items for Systematic Reviews and Meta-Analyses for Protocols
PROSPERO	prospective register of systematic reviews
PSA	prostate-specific antigen
PUFA	polyunsaturated fatty acid
R-AMSTAR	Revised Assessment of Multiple Systematic Reviews
RBC	red blood cell
RBDI	Range of Beneficial Decreased Intakes
RBII	Range of Beneficial Increased Intakes
RCT	randomized controlled trial
RDA	Recommended Dietary Allowance
RDIB	Range where Decreased Intake is Beneficial
RIIB	Range where Increased Intake is Beneficial
RNI	Recommended Nutrient Intake
ROBINS-I	Risk of Bias in Non-randomized Studies of Interventions
ROBIS	Risk of Bias in Systematic Reviews Tool
RR	relative risk
RRR	relative risk reduction
SCD	sudden cardiac death
SEER	surveillance, epidemiology, and end results
SFA	saturated fatty acid
SR	systematic review
TEP	technical expert panel
TFA	*trans* fatty acid

UF	uncertainty factor
UL	Tolerable Upper Intake Level
USDA	U.S. Department of Agriculture
USPSTF	U.S. Preventive Services Task Force
UV	ultraviolet
WCRF	World Cancer Research Fund
WHO	World Health Organization

Appendix D

Definitions

Absolute risk of a disease: The risk of developing the disease over a time period. It can be expressed as a ratio (e.g., a 1 in 10 risk of developing a certain disease during a lifetime) or percentage (e.g., 10 percent risk, or a 0.1 risk).

Acceptable Macronutrient Distribution Range (AMDR): A range of usual intakes for a macronutrient that is associated with reduced risk of chronic disease while providing adequate intakes of essential nutrients. An AMDR is expressed as a percentage of total energy intake.

Accuracy: Closeness of a measured or computed value to its "true" value, where the "true" value is obtained with perfect information. Owing to the natural heterogeneity and stochastic nature of many biologic and environmental systems, the "true" value may be an integrated average over a defined time period.

Adequate Intake (AI): The average daily nutrient intake observed in an apparently healthy sex and age group. It is based on experimentally derived intake levels or observations of mean nutrient intakes by a group of apparently healthy people who are maintaining a defined criterion of adequacy. When available evidence is not sufficient to determine the EAR for a nutrient, an AI is set. It is not certain where an AI level of intake fits relative to an actual nutrient requirement, as no Estimated Average Requirement (EAR) or Recommended Dietary Allowance (RDA) has been specified for these

nutrients. It is generally believed that the AI would be equal to or exceed the RDA (if one existed).

Analytical validation: Assessing assays and measurement performance characteristics and determining the range of conditions under which the assays will give reproducible and accurate data.

Apparently healthy population: The general population, excluding individuals who are malnourished, have diseases that result in malabsorption or dialysis treatments, or who have increased or decreased energy needs because of disability or decreased mobility. For the purposes of this report, it is recognized that the "apparently healthy population" potentially encompasses a diverse group of individuals with many different health conditions, such as individuals who have other chronic conditions such as obesity, hypertension, or diabetes.

Bayesian statistical methods: Statistical models with the unique feature of requiring the specification of prior distributions for any unknown parameters. These prior distributions are as integral to a Bayesian approach to statistical modeling as the expression of probability distributions.

Bias: A systematic error or deviation in results or inferences from the truth. The main types of bias arise from systematic differences in the groups that are compared (**selection bias**), exposure to other factors apart from the intervention of interest (**performance bias**), withdrawals or exclusions of people entered into a study (**attrition bias**), or inaccuracies in the dietary intake or outcome assessment methodologies (**ascertainment bias**). Systematic reviews of studies may also be particularly affected by **reporting bias**, where a biased subset of all the relevant data is available. **Risk of bias** (internal validity) is the evaluation of systematic error due to limitations in the study design or execution. More rigorously designed (better quality) randomized controlled trials are more likely to yield results that are closer to the truth than less rigorous designs.

Bioavailability: The efficiency with which a dietary component is used systematically through normal metabolic pathways. It is expressed as a percentage of intakes that is capable of being absorbed by the intestine and made available either for metabolic use or storage. It is influenced by dietary and host factors.

Bioequivalence: The comparison of two or more products with respect to their bioavailability.

Biomarker: A particular measurement sampled from a biological system or organism. It may take many forms, including an anatomic depiction (e.g., brain imaging), a physiological process (e.g., the glomerular filtration rate of the kidney or an electroencephalographic tracing of brain activity), an indicator of dietary intake (e.g., blood vitamin B12 levels), psychological or cognitive functions (e.g., remembering nouns from a recited list), or an indicator of the presence of a disease (e.g., high levels of blood enzymes indicating liver inflammation). All biomarkers have the same general potential problems: measurement error, variation over time and space, and difficulties in biological interpretation. In research and clinical medicine, biomarkers have important uses in understanding biological processes and in predicting the risk, presence, severity, response to, adverse effects of treatment, and outcomes of diseases. More general information on biomarkers is available in the report *Evaluation of Biomarkers and Surrogate Endpoints in Chronic Disease* (IOM, 2010).[1]

Calibration of a self-reported dietary intake method: The process of using a suitable intake biomarker in an attempt to correct a self-reported intake assessment for measurement error. Calibration equations are typically developed by regressing biomarker intake values on corresponding self-reported values and possibly other study participant characteristics.

Case-control study: An observational study that identifies "cases" based on a diagnosis of a disease or identification of risk factors. "Controls" are those who are without the disease or risk factor. A case-control study compares characteristics of the cases to those of the controls to determine what risk factors may account for who does or does not get the disease being studied. This design is particularly useful where the outcome is rare and past exposure can be validly measured. Measures of past exposure obtained after diagnosis (retrospective case-control studies) are more likely subject to biases that compromise validity than when measures obtained substantially before diagnosis, as in "nested" case-control studies.

Certainty (as it relates to judgments about evidence): The extent to which one can be confident that an estimate of effect is correct.

Chronic disease: The culmination of a series of pathogenic processes in response to internal or external stimuli over time that results in a clinical diagnosis or ailment and health outcomes. Also known as noncommu-

[1] IOM (Institute of Medicine). 2010. *Evaluation of biomarkers and surrogate endpoints in chronic disease*. Washington, DC: The National Academies Press.

nicable diseases; they are not passed from person to person. They are of long duration and generally slow progression. The main types of chronic diseases are cardiovascular diseases, cancers, chronic respiratory diseases, and diabetes.

Clinical endpoint: A characteristic or variable that reflects how an individual feels, functions, or survives. The value of an endpoint increases in relation to the degree to which it conveys information about the effect of an intervention on an individual's experience of life. Endpoints can be conceptualized in a spectrum. At one end are endpoints defined by biomarkers alone that have less relationship to an individual's experience; in the middle are clinical events that depend on biomarkers as part of the definition; further along the spectrum are endpoints that are more closely related to events that affect an individual's life. At the other end of the spectrum are the clearest clinical endpoints, such as death.

Cohort study: An observational study in which a defined group of people (the cohort) is without the disease of interest at the time of cohort enrollment and is followed over time, often for many years. The disease outcomes of people in the cohort are compared, to examine people who were exposed or not exposed (or exposed at different levels) to a particular factor (exposure) of interest. A **prospective** cohort study assembles participants and follows them into the future. A **retrospective** (or historical) cohort study identifies subjects from past records and follows them from the time of those records to the present.

Concentration biomarkers: Biomarkers that assess concentrations or relative percentages of nutrients or other food substances in the blood, urine, or other tissues (e.g., serum folate concentration) and can be used as an estimate of the intake of such a nutrient or other food substance.

Confidence interval: A measure of the uncertainty around the main finding of a statistical analysis. Estimates of unknown quantities, such as the relative risk comparing an experimental intervention with a control, are usually presented as a point estimate and a 95 percent confidence interval. This means that if someone were to keep repeating a study in other samples from the same population, 95 percent of the calculated confidence intervals from those studies would include the true underlying value. Wider intervals indicate less precision; narrow intervals, greater precision.

Confounding factor: A variable that is correlated (directly or inversely) to both the dependent variable and independent variable.

Cross-sectional study: An observational study that analyzes data collected from a population, or a representative subset, at a specific point in time—that is, cross-sectional data.

Deficiency disease: An illness associated with an insufficient supply of one or more essential dietary constituents.

Disease marker (or biomarker of effect): A biomarker that may predict clinical benefit (or harm or lack of benefit or harm) based on epidemiologic, therapeutic, pathophysiologic, or other scientific evidence. Includes both surrogate disease markers and non-qualified disease markers.

Dietary Reference Intakes (DRIs): A set of nutrient-based reference values established under the National Academies of Sciences, Engineering, and Medicine that are used for planning and assessing diets of apparently healthy individuals and groups.

Epigenetics: The study of stable heritable traits (or "phenotypes") that cannot be explained by changes in DNA sequence.

Essential nutrient: A substance that is required for normal physiological functioning that cannot be synthesized in the body or cannot be synthesized in sufficient amounts to meet needs and thus must be provided in the diet.

Estimated Average Requirement (EAR): The usual daily intake of a nutrient that is expected to meet the requirement of half of healthy individuals in a group defined by life-stage and sex. The requirement is based on a specific indicator of adequacy.

Estimated Energy Requirement (EER): A calculated level of energy intake that is estimated to maintain energy balance that incorporates weight, height, physiological state (i.e., pregnancy) and level of energy expenditure.

Evidence profile: Presentation of detailed information about the quality of evidence assessed and the summary of findings for each of the included outcomes. It presents information about the body of evidence (e.g., number of studies), the judgments about the underlying quality of evidence, key statistical results, and the quality of evidence rating for each outcome. Guideline panels (e.g., Dietary Reference Intake committees) are expected to review evidence profiles to ensure that members agree about the judgments underlying the quality assessments.

Evidentiary qualification: Assessment of available evidence on associations between a biomarker and disease states, including data showing effects of interventions on both the biomarker and clinical outcomes.

Food substances: Nutrients that are essential or conditionally essential, energy nutrients, or other naturally occurring bioactive food components.

Guideline panel: A panel of a knowledgeable, multidisciplinary group of experts and representatives from key affected groups that are charged with developing clinical practice guidelines. Standards for panel composition and managing members' conflicts of interests exist and should be followed as closely as possible. In the Dietary Reference Intake (DRI) process, a DRI committee is equivalent to the guideline panel in the Clinical Practice Guideline process.

GRADE (Grading of Recommendations, Assessment, Development and Evaluation): A method of assessing the certainty in evidence and the strength of recommendations in health care. It provides a structured and transparent evaluation of the importance of outcomes of alternative management strategies, acknowledgment of patients and the public values and preferences, and comprehensive criteria for rating down or up the certainty in evidence.

Hazard characterization: A description, preferably quantitative, of the relationship between a dose of a hazard and its effect.

Heterogeneity: The variation in study outcomes within the body of evidence for a particular outcome. It can be due to variability in participants, outcomes, or interventions, or intake response (clinical heterogeneity) or to variability in methods used, such as blinding, participant recruitment, or data collected (methodological heterogeneity).

Imprecision: A measurement of random error that often occurs when studies within the body of evidence for a particular outcome have a small sample size and the number of events is also small, resulting in a wide 95 percent confidence interval around the estimate of the effect.

Inconsistency: Unexplained heterogeneity or variability in the body of evidence for a particular outcome.

Indicator (of adequacy or toxicity): Clinical endpoints, surrogate endpoints, biomarkers, or risk factors for a chronic disease that may serve as the basis for estimating nutrient intake requirements or excessive levels of nutrient intake that might result in adverse health effects.

Indirectness: A situation that occurs when in the body of evidence for a particular outcome, studies do not directly compare the interventions of interest, apply the intervention to the population of interest, or measure the important outcomes.

Intake-response relationship: The relationship between levels of intake of a nutrient or food substance and a measure of chronic disease. If sufficient data exist, an intake-response relationship may be characterized quantitatively and may lead to a chronic disease Dietary Reference Intake.

Meta-analysis: A systematic review technique that uses statistical methods to quantitatively combine the results of similar studies in an attempt to allow inferences to be made from the sample of studies and be applied to a population of interest.

Metabolomics: The scientific study of chemical processes involving metabolites. Specifically, metabolomics is the "systematic study of the unique chemical fingerprints that specific cellular processes leave behind" (i.e., the study of their small-molecule metabolite profiles).

Monte Carlo simulation: A computerized mathematical technique that allows people to account for risk in quantitative analysis and decision making. It furnishes the decision maker with a range of possible outcomes and the probabilities they will occur for any choice of action. The technique is used by professionals in such widely disparate fields as finance, project management, transportation, the environment, and public health.

Neural tube defects: Birth defects of the brain, spine, or spinal cord that occur in the first month of pregnancy, often before a woman even knows that she is pregnant. The two most common neural tube defects are spina bifida and anencephaly. In spina bifida, the fetal spinal column does not close completely. There is usually nerve damage that causes at least some paralysis of the legs. In anencephaly, most of the brain and skull do not develop. Babies with anencephaly are usually either stillborn or die shortly after birth. Another type of defect, Chiari malformation, causes the brain tissue to extend into the spinal canal. The exact causes of neural tube defects are not known.

Non-qualified disease marker: A possible biomarker of effect that predicts a chronic disease outcome but lacks adequate evidence to be suitable as an accurate and reliable substitute for that outcome. Also known as an intermediate disease outcome marker or intermediate endpoint.

Observational study: A study in which the investigators do not intervene, but simply observe a study population. Changes or differences in characteristics or exposures are studied in relation to changes or differences in other characteristic(s) (e.g., whether or not they died), without action by the investigator. This study design has a greater risk of selection bias and ascertainment bias than do experimental studies. Cross-sectional studies, cohort studies, and case-control studies are types of observational studies.

Outcome: A term, used synonymously with "endpoints," that refers to the clinical results of a particular illness(es), often after particular therapeutic interventions. With regard to Dietary Reference Intakes, the outcome might be a change in disease incidence (primary prevention of coronary disease) but also can be improvement of the clinical outcome of patients who have already sustained a heart attack (secondary prevention).

PICO: A technique used in evidence-based practice to frame and answer a clinical or a health care–related question. The PICO framework is also used to develop literature search strategies. The PICO acronym stands for population (P), intervention (I), comparator (C), and outcome (O).

Precision: The quality of a measurement that is reproducible in amount or performance. Measurements can be precise in that they are reproducible, but can be inaccurate and differ from "true" values when biases exist. Measurement error can also affect precision. In risk assessment outcomes and other forms of quantitative information, precision refers specifically to variation among a set of quantitative estimates of outcomes.

Primary prevention: An effort to prevent the onset of specific diseases before they occur through risk reduction, by altering behaviors or exposures that can lead to disease (e.g., smoking cessation), or by enhancing resistance to the effects of exposure to a disease agent (e.g., immunization). Primary prevention reduces the incidence of disease by addressing disease risk factors or by enhancing resistance.

Publication bias: A systematic under-estimation or over-estimation of the underlying beneficial or harmful effect due to the selective publication of studies.

Quasi-experiment: Experimental research designs that test causal hypotheses of an intervention. In contrast to a randomized controlled trial, a quasi-experiment lacks random assignment, and assignment to conditions (e.g., treatment versus no treatment or comparison condition) is by means of self-selection or administrator selection. Quasi-experimental designs

identify a comparison group that is as similar as possible to the treatment group in terms of baseline (pre-intervention) characteristics.

Random error: The difference between assessments of a variable or variables collected from one administration of an instrument compared to a long-term average based on multiple administrations of an instrument.

Randomized controlled trial: An experimental study in which two or more interventions are compared by being randomly allocated to participants. In most trials, one intervention is assigned to each individual but sometimes assignment is to defined groups of individuals (e.g., in a household, worksite, or a community) or interventions are assigned within individuals (e.g., in different orders or to different parts of the body).

Recommended Dietary Allowance (RDA): The usual daily intake level that is sufficient to meet the nutrient requirements of 97 to 98 percent of healthy individuals in the specified life-stage and sex group. If the requirements in a specified group are normally distributed, the RDA is equivalent to the EAR plus two standard deviations.

Recovery biomarkers: Biomarkers that measure a nutrient of food substance intake and output that can be "recovered" and measured quantitatively (e.g., doubly labeled water or urinary nitrogen from 24-hour urine collections).

Relative risk: In statistics and epidemiology, relative risk or risk ratio (RR) is the ratio of the probability of an event occurring (e.g., developing a disease, being injured) in an exposed group to the probability of the event occurring in a comparison, non-exposed group.

Review of the totality of the evidence: In the context of setting chronic disease Dietary Reference Intakes (DRIs), it refers to evaluating the evidence about whether a chronic disease DRI should be developed, including the systematic review evidence profiles, quantitative characterization of the intake-response, consideration of relationships with various chronic diseases, potential overlapping benefits and harms, need and appropriateness of extrapolation to other populations, and other relevant evidence.

Risk assessment: The process that serves to estimate the risk to a given target organism, system, or population, including the identification of attendant uncertainties following exposure to a particular agent. Risk assessment encompasses four steps: hazard identification, hazard characterization, exposure assessment, and risk characterization.

Risk factors: Variables that predict outcomes and can be biomarkers and social and environmental factors. The value of a risk factor depends on the degree to which it can predict an event.

Risk identification: The determination that a substance with hazardous properties is present, but also more generally refers to the identification of the type and nature of adverse effects that an agent can cause in an organism, system, or given population.

Risk management: A set of actions that entail identifying foreseeable hazards and their associated risks, assessing the risks, controlling the risks, and monitoring and reviewing the risk management process.

Secondary prevention: Efforts to reduce the impact of a disease or injury that has already occurred. This is done by detecting and treating disease or injury as soon as possible to halt or slow its progress, encouraging personal strategies to prevent re-injury or recurrence (e.g., dietary behaviors).

Surrogate disease marker: A biomarker of effect that predicts clinical benefit (or harm, or lack of benefit or harm) based on epidemiologic, therapeutic, pathophysiologic, or other scientific evidence that is qualified for its intended purposes. Also known as a surrogate marker, surrogate endpoint, or surrogate disease outcome marker.

Synthesis of evidence: An evaluation of a body of evidence collected in a systematic manner and using quantitative and qualitative synthesis strategies. Standards for methods to synthesize the evidence include the use of consistent language to characterize the level of certainty in the estimates of the effect and the use of criteria to evaluate the body of evidence (i.e., risk of bias, consistency, precision, directness, and publication bias), including specific criteria for evaluating bodies of evidence of observational studies (i.e., dose-response association, plausible confounding, and size of the effect).

Systematic error (also known as bias): A type of error that results in measurements that consistently depart from the true value in the same direction. It affects the sample mean as well as percentiles and can result in incorrect estimates and conclusions. In contrast to random error, data affected by systematic error are biased, and this type of error cannot be reduced or eliminated by taking repeat measures.

Systematic review: A scientific investigation that focuses on a specific question and that uses explicit, planned scientific methods to identify, select,

assess, and summarize the findings of similar but separate studies. It may or may not include a quantitative synthesis (meta-analysis) of the results from separate studies.

Systematic review team: A group of experts contracted specifically to conduct a systematic review.

Technical Expert Panel: A group of subject-matter experts who serve as consultants to the systematic review team in scientific matters related to the questions of interest.

Tolerable Upper Intake Level (UL): The highest usual daily nutrient intake level that is likely to pose no risk of adverse effects to nearly all healthy individuals in the specified life-stage and sex group.

Uncertainty: Lack or incompleteness of information. Quantitative uncertainty analysis attempts to analyze and describe the degree to which a calculated value may differ from the true value; it is sometimes expressed as probability distributions. Uncertainty depends on the quality, quantity, and relevance of data and on the applicability and relevance of models and assumptions.

Uncertainty factor (UF): In toxicology, one of several factors used in calculating the reference dose from experimental data. The UF is intended to account for (1) the variation in sensitivity among humans, (2) the uncertainty in extrapolating from one population to another, (3) the uncertainty in extrapolating data obtained in a study that covers less than the full life of the exposed animal or human, and (4) the uncertainty in using Lowest-Observed-Adverse-Effect Level data rather than No-Observed-Adverse-Effect Level data.

Utilization analysis: Contextual analysis based on the specific use proposed and the applicability of available evidence to this use. This includes a determination of whether the validation and qualification conducted provide sufficient support for the use proposed.

Validation of a biomarker: The action of checking or proving the accuracy of some measure. Validity can sometimes be established by conducting controlled human feeding studies in a population of interest. Each participant is provided a diet over a defined time period, and potential biomarkers in pertinent biofluids (e.g., urine or serum/plasma) are examined for correlation with actual intake of the nutrient or food substance of interest. Biomarkers meeting criteria (e.g., correlation ≥ 0.6) may provide useful

objective measures of intake in the population from which feeding study participants were drawn.

Validation of a self-reported dietary intake method: A process to establish validity by comparing the self-reported measurement with an objective measure of intake (e.g., quantitative recovery biomarkers such as doubly labeled water assessment of short-term energy intake, or urinary nitrogen assessment of protein intake). It should be noted that objective intake measures, such as quantitative recovery biomarkers, are not available for all nutrients or food substances.

Variability: True differences in attributes due to heterogeneity or diversity. Variability is usually not reducible by further measurement or study, although it can be better characterized. Two important sources of variability are biological variability (inter-individual differences, i.e., attributable to genetic differences and influenced by environmental factors) and analytical variability (i.e., associated with analysis of dietary component).

Appendix E

Biographical Sketches of Committee Members

Shiriki K. Kumanyika, Ph.D., M.S., M.P.H. (*Chair*), is Professor of Epidemiology Emeritus at the University of Pennsylvania Perelman School of Medicine and Research Professor in the Department of Community Health & Prevention at the Dornsife School of Public Health at Drexel University. She founded and continues to chair the African American Collaborative Obesity Research Network, which now has its national office at the Dornsife School. Elected to the National Academy of Medicine in 2003, Dr. Kumanyika is a member of the National Academies of Sciences, Engineering, and Medicine's Roundtable on Obesity Solutions and the Steering Committee for the Vital Directions initiative and has chaired or served on several other National Academies committees and the Food and Nutrition Board. She served on two U.S. Dietary Guidelines Advisory Committees and on the World Cancer Research Fund Expert Panel on Diet, Nutrition, and Cancer Prevention and is past president of the American Public Health Association. Her current service includes membership on the U.S. Centers for Disease Control and Prevention Task Force on Community Preventive Services, the World Health Organization Nutrition Guidance Expert Advisory Group Subgroup on Diet and Health, and the Lancet Commission on Obesity. Dr. Kumanyika has a Ph.D. in human nutrition from Cornell University, an M.S. in social work from Columbia University, and an M.P.H. from Johns Hopkins University.

Cheryl A. M. Anderson, Ph.D., M.P.H., is Associate Professor in the Department of Family and Preventive Medicine at the University of California, San Diego. Before this appointment she was an assistant professor in the

Department of Epidemiology at the Johns Hopkins Bloomberg School of Public Health in Baltimore. Dr. Anderson's research centers on nutrition-related issues in chronic disease prevention in minority and underserved populations. Dr. Anderson is the principal investigator of the National Heart, Lung, and Blood Institute (NHLBI)-funded study of the effects of dietary sodium and potassium intake on subclinical and clinical cardiovascular disease. She is a co-investigator on the National Institute of Diabetes and Digestive and Kidney Diseases–funded national, multicenter Chronic Renal Insufficiency Cohort Study, which aims to identify risk factors and mechanisms of progressive renal disease and cardiovascular events in individuals with chronic kidney disease, and is a co-investigator on the NHLBI-funded Optimal Macronutrient Intake (OMNI)-Carb study, a randomized feeding study that compares the effects of type (glycemic index) and amount of carbohydrate on cardiovascular risk factors. Dr. Anderson is principal investigator of a study testing a unique biomarker (using carbon isotopic data) of intake of sweets (funded by an Innovation Grant Award from the Johns Hopkins Bloomberg School of Public Health). Before her appointment at Johns Hopkins, Dr. Anderson was an Instructor of Epidemiology at the University of Pennsylvania School of Medicine, Center for Clinical Epidemiology and Biostatistics. Dr. Anderson served on two Institute of Medicine committees—Committee on Strategies to Reduce Sodium Intake and Committee on Use of Dietary Supplements by Military Personnel. She currently serves on the National Academies of Sciences, Engineering, and Medicine's Committee on Consequences of Sodium Reduction in Populations. She has a B.S. from Brown University, an M.P.H. from the University of North Carolina at Chapel Hill, and an M.S. in epidemiology and Ph.D. in nutritional sciences from the University of Washington School of Public Health and Community Medicine.

Susan I. Barr, Ph.D., R.D., is Professor Emeritus of Food Nutrition and Health at the University of British Columbia. Her research interests relate to how women's cognitions about food, eating, and body weight may have physiological implications for their health. Dr. Barr also has an interest in dietary policy and was involved in the development of the Dietary Reference Intakes and has also been a member of Health Canada committees working on revision of Canada's Food Guide and on dietary sodium reduction. Additionally, she has an interest in dietary practices and dietary survey data. Recent work in this area has examined the contribution of breakfast to nutrient adequacy of Canadians, and the perceptions and practices of Canadians with regard to milk product intake. Dr. Barr has a Ph.D. in human nutrition from the University of Minnesota.

Kathryn G. Dewey, Ph.D., is Distinguished Professor in the Department of Nutrition and Director of the Program in International and Community Nutrition at the University of California, Davis. Her research focuses on maternal and child nutrition in both low-income and higher-income populations, particularly infant and young child feeding, growth during infancy and early childhood, micro- and macronutrient status of infants and young children, maternal nutrition during pregnancy and lactation, risk factors for early lactation difficulties, and the short- and long-term consequences of interventions to improve nutrition of mothers and their children. She has conducted clinical and community-based research in Bangladesh, Costa Rica, Ghana, Guatemala, Honduras, Malawi, Mexico, Peru, and the United States. Her professional service includes consultation for the World Health Organization, United Nations Children's Fund, Pan American Health Organization, National Institutes of Health, and the March of Dimes, scientific advisory committees for the Bill & Melinda Gates Foundation and the UK Medical Research Council, and serving as President of the Society for International Nutrition Research and of the International Society for Research on Human Milk and Lactation. She has a Ph.D. in biological sciences from the University of Michigan.

Gordon Guyatt, M.Sc., M.D., is a Distinguished Professor in the Department of Health Research Methods, Evidence, and Impact, McMaster University, and one of the founders of evidence-based medicine. He has played a key role in more than 30 major clinical studies (including both large-scale observational and randomized controlled trials) and has extensive expertise in study methodology. As co-founder and co-chair of the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) working group he has been intimately involved in the development and evolution of the GRADE approach. Dr. Guyatt's research interests include the dissemination of concepts of evidence-based medicine to health workers and health care consumers; the methodology of clinical practice guidelines and medical decision making; systematic review methodology; and ascertaining patients' values and preferences. Dr. Guyatt has published more than 1,000 peer-reviewed papers that have been cited more than 95,000 times. Dr. Guyatt has been a leading exponent of evidence-based approaches to clinical practice, having coined the term "evidence-based medicine" in 1990. He has an M.S. and an M.D. from McMaster University.

Janet C. King, Ph.D., is a Senior Scientist of the Children's Hospital Oakland Research Institute and Professor Emeritus of Nutrition at the University of California, Berkeley, and Davis. Throughout a long and distinguished career, Dr. King has made substantive contributions to the body of human

nutrition research, application, and policy development. In recognition of her national and international reputation, she was elected to the National Academy of Medicine in 1994, and in 2007, she was inducted into the U.S. Department of Agriculture (USDA) Research Hall of Fame. She directed the USDA Western Human Nutrition Research Center at the University of California, Davis (1995-2002) and chaired the Department of Nutritional Sciences, University of California, Berkeley (1988-1994). Dr. King's research focuses on metabolic adjustments to changes in nutrient intakes in humans; she is especially interested in metabolism and nutrient utilization of pregnant and lactating women and cellular and whole body zinc functions. Dr. King's impact on the field of human nutrition extends well beyond her research accomplishments. For example, she chaired the USDA/U.S. Department of Health and Human Services (HHS) 2005 Dietary Guidelines Advisory Committee. When Dr. King was the Chair of the National Academies of Sciences Food and Nutrition Board in 1994, a new paradigm for the Dietary Reference Intakes was established. She recently chaired a United Nations University, Food and Agriculture Organization, World Health Organization Joint Committee on Dietary Harmonization and currently serves as Director of the United Nations International Consultative Group on Zinc. Dr. King has published more than 250 scientific papers, review articles, and book chapters. She has trained more than 65 graduate students, postdoctoral fellows, and visiting scientists.

Marian L. Neuhouser, Ph.D., R.D., is Full Member in the Cancer Prevention Program, Division of Public Health Sciences at the Fred Hutchinson Cancer Research Center in Seattle, Washington. She is also Core Faculty in Nutritional Sciences and Affiliate Professor of Epidemiology, both in the School of Public Health, University of Washington. Dr. Neuhouser is a nutritional epidemiologist whose primary research focus is nutrition and energy balance and their relationship to cancer prevention and cancer survivorship. She has broad experience and leadership in large clinical trials, including the Women's Health Initiative and the Prostate Cancer Prevention Trial, small-scale controlled dietary interventions, and large observational cohorts. In addition, a portion of Dr. Neuhouser's research portfolio is focused on methods to improve diet and physical activity assessment and numerous aspects of health disparities, which links together nutrition, energy balance and cancer risk. Dr. Neuhouser was a member of the 2015 Dietary Guidelines Advisory Committee and served as Vice-President of the American Society for Nutrition (2015-2016), after which she became President (2016-2017).

Ross L. Prentice, Ph.D., is Member and former Director of the Public Health Sciences Division at the Fred Hutchinson Cancer Research Center,

and Professor of Biostatistics at the University of Washington. His research focuses on chronic disease population science and disease prevention, and on related methodology developments. His statistical research areas include failure time data analysis methods; cohort study design and analysis methods; the use of biomarkers to address measurement error issues, especially in diet and physical activity epidemiology; surrogate outcome methods and limitations; and genomic and proteomic methods. He has served as Principal Investigator (PI) of the Clinical Coordinating Center for the Women's Health Initiative from its inception in 1992 to 2011 (dual PI 2008-2011), which involves a multifaceted randomized controlled trial and cohort study among 161,000 postmenopausal U.S. women. The results of the trial have markedly changed clinical practice in the use of postmenopausal hormones. Dr. Prentice has received the COPSS Award and the Fisher Lecture Award from the "Joint Statistical Societies," the Research Excellence in Epidemiology and Prevention Award from the American Association for Cancer Research (AACR) and American Cancer Society, and the AACR Team Science Award. He is also a member (1990) of the National Academy of Medicine.

Joseph Rodricks, Ph.D., is a Founding Principal of Ramboll Environ. An expert in toxicology and risk analysis, Dr. Rodricks has consulted for hundreds of manufacturers and government agencies and for the World Health Organization in the evaluation of health risks associated with human exposure to chemical substances of all types. Before Environ, Dr. Rodricks served 15 years as a scientist at the U.S. Food and Drug Administration; in his last 4 years, he served as Associate Commissioner for Health Affairs. His experience extends from pharmaceuticals, medical devices, consumer products and foods, to occupational chemicals and environmental contaminants. He has served on the National Academies of Sciences, Engineering, and Medicine's Board on Environmental Studies and Toxicology and on 30 boards and committees of the National Academies, including the committees that produced the seminal works *Risk Assessment in the Federal Government: Managing the Process* (1983) and *Science and Decisions: Advancing Risk Assessment* (2009). Dr. Rodricks also served for 7 years on the Institute of Medicine Subcommittee on Upper Reference Levels of Nutrients. Dr. Rodricks has nearly 150 scientific publications and has received honorary awards from three professional societies for his contributions to toxicology and risk analysis. Dr. Rodricks earned his Ph.D. in biochemistry from the University of Maryland, College Park, and was a postdoctoral scholar at University of California, Berkeley.

Patrick J. Stover, Ph.D., is Professor and Director of the Division of Nutritional Sciences at Cornell University. He is also director of the World

Health Organization Collaborating Centre on Implementation Research in Nutrition and Global Policy at Cornell University, and Past-President of the American Society for Nutritional Sciences. Dr. Stover's research interests focus on the biochemical, genetic, and epigenetic mechanisms that underlie the relationships between folic acid and human pathologies, including neural tube defects and other developmental anomalies, cardiovascular disease, and cancer. Specific interests include the regulation of folate-mediated one-carbon metabolism and cellular methylation reactions, molecular basis of the fetal origins hypothesis, development of mouse models to elucidate mechanisms of folate-related pathologies, and nuclear one-carbon metabolism. In 2016, he was elected as a member of the National Academy of Sciences, and in 2014 was elected as a Fellow of the American Association for the Advancement of Science. In 2014, he received the State University of New York Chancellor's Award for Excellence in Scholarship and Creative Activities, the Osborne and Mendel Award for outstanding recent basic research accomplishments in nutrition from the American Society for Nutrition, and a MERIT award from the National Institute of Diabetes and Digestive and Kidney Diseases. In 1996, he received the Presidential Early Career Award for Scientists and Engineers, the highest honor bestowed by the U.S. government on outstanding scientists and engineers beginning their independent careers. He has been selected as an Outstanding Educator four times by Cornell Merrill Presidential Scholars. Dr. Stover served two terms on the National Academies of Sciences, Engineering, and Medicine's Food and Nutrition Board (FNB) and he served on the FNB Nutrigenomics Workshop Planning Group. Dr. Stover received his Ph.D. in biochemistry and molecular biophysics from the Medical College of Virginia.

Katherine L. Tucker, Ph.D., is Professor of Nutritional Epidemiology in the Department of Biomedical and Nutritional Sciences at the University of Massachusetts Lowell. She holds adjunct appointments at the University of Massachusetts Medical School and the Friedman School of Nutrition Science and Policy at Tufts University. Dr. Tucker has contributed to more than 300 articles in scientific journals. Her research focuses on dietary intake and risk of chronic disease, including osteoporosis, cognitive decline, obesity, metabolic syndrome, and heart disease. She is the Director of the National Heart, Lung, and Blood Institute (NHLBI)-funded Center on Population Health and Health Disparities, studying the roles of diet, health behavior, stress, and genetic predisposition in relation to chronic conditions in Puerto Rican adults. She currently serves as a scientific adviser for, and leads a Vanguard data analysis center with, the Jackson Heart Study. She is the Editor in Chief of *Advances in Nutrition*, the international review journal of the American Society of Nutrition (ASN), and was a co-editor of the 11th edition of the textbook *Modern Nutrition in Health and Disease*.

Dr. Tucker is currently a member of the National Academies of Sciences, Engineering, and Medicine's Food and Nutrition Board and previously served on Institute of Medicine committees to review the Child and Adult Care Food Program meal requirements and the implications of dioxin in the food supply. Dr. Tucker received her Ph.D. in nutritional sciences from Cornell University.

Robert B. Wallace, M.D., M.Sc., is the Irene Ensminger Stecher professor of epidemiology and internal medicine at the University of Iowa Colleges of Public Health and Medicine. He has a variety of public health experiences. He was an Epidemic Intelligence Service Officer with the U.S. Centers for Disease Control and Prevention. He has conducted many population health studies as well as clinical trials, focusing on the prevention and control of chronic illnesses and other disabling conditions of older persons. These have included neurological conditions, fracture, cancers, coronary disease, mental illnesses, and the health of older women. He has continuing experience with community interventions related to the prevention of falls and motor vehicle injuries in older persons. He was a member of the U.S. Preventive Services Task Force, and the National Advisory Council on Aging of the National Institute on Aging at the National Institutes of Health (NIH). He is an elected member of the National Academy of Medicine and has been a past chair of the National Academies of Sciences, Engineering, and Medicine's Board on Population Health and Board on the Health of Select Populations, and he has had substantial experience with National Academies studies and panels. He is currently involved in several actively funded research projects by NIH, including several related to nutritional issues. He received his M.S. in epidemiology from the State University of New York at Buffalo and his M.D. from the Northwestern University School of Medicine.

Weihsueh A. Chiu, Ph.D. (consultant to the committee), is a professor in the Department of Veterinary Integrative Biosciences in the College of Veterinary Medicine and Biomedical Sciences at Texas A&M University. Before joining the university, he worked at the U.S. Environmental Protection Agency (EPA) for more than 14 years, most recently as chief of the Toxicity Pathways Branch in the Integrated Risk Information System (IRIS) Division of the National Center for Environmental Assessment. His research has focused on human health risk assessment, particularly with respect to toxicokinetics, mechanisms of toxicity, physiologically based pharmacokinetic modeling, dose-response assessment, and characterizing uncertainty and variability. Dr. Chiu led the development of EPA's 2011 IRIS assessment of trichloroethylene, which pioneered the use of probabilistic methods for characterizing uncertainty and variability in toxicokinetics and dose

response. He is currently Chair-elect of the Dose-Response Specialty Group of the Society for Risk Analysis. He has served on several National Academies of Sciences, Engineering, and Medicine committees, including the Committee on Predictive-Toxicology Approaches for Military Assessments of Acute Exposures and the Committee on Endocrine-Related Low-Dose Toxicity. Dr. Chiu received a Ph.D. in physics from Princeton University.

Appendix F

Disclosure of Conflict of Interest

The conflict-of-interest policy of the National Academies of Sciences, Engineering, and Medicine (www.nationalacademies.org/coi) prohibits the appointment of an individual to a committee like the one that authored this Consensus Study Report if the individual has a conflict of interest that is relevant to the task to be performed. An exception to this prohibition is permitted only if the National Academies determine that the conflict is unavoidable and the conflict is promptly and publicly disclosed.

When the committee that authored this report was established a determination of whether there was a conflict of interest was made for each committee member given the individual's circumstances and the task being undertaken by the committee. A determination that an individual has a conflict of interest is not an assessment of that individual's actual behavior or character or ability to act objectively despite the conflicting interest.

Dr. Joseph Rodricks was determined to have a conflict of interest because he provides consulting services related to the safety of food ingredients to food companies.

The National Academies determined that the experience and expertise of the individual was needed for the committee to accomplish the task for which it was established. The National Academies could not find another available individual with the equivalent experience and expertise who did not have a conflict of interest. Therefore, the National Academies concluded that the conflict was unavoidable and publicly disclosed it through the National Academies Current Projects System (www8.nationalacademies.org/cp).