Surrogate disease markers as substitutes for chronic disease outcomes in studies of diet and chronic disease relations

Elizabeth A Yetley,[1] David L DeMets,[3] and William R Harlan Jr.[2]

[1]Office of Dietary Supplements and [2]Office of the Director, NIH, Bethesda, MD; and [3]Department of Biostatistics and Medical Informatics, University of Wisconsin, Madison, WI

ABSTRACT

Surrogate biomarkers for clinical outcomes afford scientific and economic efficiencies when investigating nutritional interventions in chronic diseases. However, valid scientific results are dependent on the qualification of these disease markers that are intended to be substitutes for a clinical outcome and to accurately predict benefit or harm. In this article, we examine the challenges of evaluating surrogate markers and describe the framework proposed in a 2010 Institute of Medicine report. The components of this framework are presented in the context of nutritional interventions for chronic diseases. We present case studies of 2 well-accepted surrogate markers [blood pressure within sodium intake and cardiovascular disease (CVD) context and low density lipoprotein–cholesterol concentrations within a saturated fat and CVD context]. We also describe additional cases in which the evidence is insufficient to validate their surrogate status. Guidance is offered for future research that evaluates or uses surrogate markers.   *Am J Clin Nutr* 2017;106: 1175–89.

Keywords: biomarker qualification, diet and health relations, surrogate disease markers, diet and chronic disease, surrogate disease marker qualification

INTRODUCTION

Chronic diseases are the most common cause of mortality and morbidity in developed populations. With large numbers of individuals at risk, relatively small benefits can have large population-based impacts. The identification of the benefits of diet and dietary components in reducing the risk of chronic diseases has a potential applicability for both public health and clinical contexts. However, progress has been limited by several scientific challenges including the difficulty of studying chronic disease outcomes that are characterized by long developmental times, the multifactorial nature of chronic disease development, the multiplicity of effects of diets and dietary components, and the difficulty of differentiating between causal and associative relations. One approach to reducing the duration and size of controlled trials that are designed to confirm a causal relation between dietary intakes and chronic disease risk is to qualify (i.e., validate) surrogate disease markers as substitutes for assessments of chronic disease events.

In this article, we consider information from 2 expert working groups to address the challenge of qualifying surrogate disease markers to bring efficiencies to studies of diet and chronic disease relations. The primary focus was based on a 2010 Institute of Medicine (IOM), US National Academy of Sciences, report on a generic framework for qualifying biomarkers and surrogate endpoints in chronic disease evaluations (1). The rationale for a current re-examination of this 2010 report was a 2017 US-Canadian working group report on Options for basing Dietary Reference Intakes (DRIs) on chronic disease endpoints (2). One of several challenges that were identified by this report addresses what types of endpoints are acceptable outcomes for evaluating the effect of a food substance on the risk of a chronic disease. One question is whether to use qualified surrogate-disease markers as substitutes for the actual measurement of a chronic disease event as defined by accepted diagnostic criteria. As a follow-up to this latter report, a recently available prepublication copy of a consensus study report of the National Academies of Sciences, Engineering and Medicine (National Academies) recommended that, "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 composition endpoints, when applicable" (3). The report also

recommended that "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." The key concepts in the 2010 biomarker report (1) are directly relevant to the recommendations for surrogate marker qualifications in the recent National Academies report (3) and can help to inform the challenges and processes that are involved in qualifying a surrogate disease marker for multiple policy, programmatic, and clinical purposes including for the development of DRIs that are based on chronic disease endpoints.

Our focus in this paper is on the criteria and process of qualifying potential surrogate markers for use as substitutes for actual assessments of chronic disease events as determined by accepted diagnostic criteria including composite endpoints. This limited focus is not intended to undermine or minimize other valuable and legitimate uses of biomarkers. The other uses that are outside the scope of this paper include, but are not limited to, assessing the risk or susceptibility to a chronic disease, serving as diagnostic tools, monitoring of patients and populations, having a prognostic and predictive value, elucidating mechanisms of action, and identifying safety concerns. Any biomarker can be excellent for one purpose but relatively useless for other purposes (4). Instead, our purpose is to describe a framework for evaluating potential surrogate markers within a nutrition and chronic disease context with particular emphasis on uses related to public health and prevention applications. To illustrate the use of the framework for surrogate biomarkers, we present examples of commonly accepted surrogate markers and their evolution to a qualified status. The changing recommendations of expert groups regarding these markers is documented to provide an appreciation of the increasing scientific rigor and more demanding scientific assessments of the evidence.

DEFINITIONS

Biomarkers are defined as characteristics that are objectively measured and evaluated as indicators of normal biological processes, pathogenic processes, or pharmacologic responses to an intervention (**Table 1**) (1). Examples of biomarkers include blood concentrations of nutrients, LDL-cholesterol and HDL-cholesterol concentrations, blood pressure (BP), enzyme concentrations, tumor size, genetic variations, and combinations of these measurements. Biomarkers are used to describe risk, exposures, treatment effects, and biological mechanisms. They are essential for research, clinical, monitoring, and regulatory applications.

A surrogate marker is one type of biomarker. (In this article, surrogate marker, surrogate biomarker, surrogate endpoint, and modifiable risk factor or marker are used interchangeably). It has been "deemed useful as a substitute for a defined, disease-relevant clinical endpoint" (Table 1) (1, 5, 6). The surrogate marker is on the major causal pathway between an intervention (e.g., diet or dietary component) and the outcome of interest (e.g., chronic disease) that captures the full effect of the intervention on the outcome. The measurement of a surrogate marker (e.g., LDL cholesterol) instead of the true clinical

TABLE 1
Definitions from the 2010 Institute of Medicine report[1]

Definition
Biomarkers
Biomarker: a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a[n] … intervention.
Risk biomarker: a biomarker that indicates a component of an individual's level of risk of developing a disease or level of risk of developing complications of a disease.
Surrogate endpoint: a biomarker that is intended to substitute for a clinical endpoint. A surrogate endpoint is expected to predict clinical benefit (or harm or lack of benefit or harm) based on epidemiologic, therapeutic, pathophysiologic, or other scientific evidence.
Biomarker evaluation framework
Analytic validation: assessing an assay and its measurement performance characteristics, determining the range of conditions under which the assay will give reproducible and accurate data. [It includes the biomarker's limits of detection, limits of quantification, reference (normal) value cutoff concentrations, and total imprecision at the cutoff concentration.]
Qualification: evidentiary process of linking a biomarker with biological processes and clinical endpoints. An assessment of available evidence on associations between the biomarker and the disease states, including data showing effects of interventions on both the biomarker and clinical outcomes.
Utilization: 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.
Other definitions
Chronic disease: 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.
Clinical endpoint: a characteristic or variable that reflects how a patient [or consumer] feels, functions, or survives.[2]
Fit for purpose: being guided by the principle that an evaluation process is tailored to the degree of certainty required for the use proposed.
Prognostic value: the ability to predict disease outcome or course using a specified patient measurement.
True endpoint: the endpoint for which a surrogate endpoint is sought.

[1] Definitions are taken directly from reference 1.
[2] The Institute of Medicine authors note that it is assumed that the characteristic or variable that is indicative of the clinical endpoint is accurately measured and has general acceptance by qualified scientists as a valid indicator of the underlying condition of interest.

outcome [e.g., cardiovascular disease (CVD) morbidity and mortality] can reduce the duration and sample-size requirements of interventional studies. Therefore, the substitution of surrogate markers for clinical outcomes can result in lower trial costs, more opportunities to test diverse approaches, and faster access to results. These benefits are particularly important in studies of chronic diseases that require large at-risk populations, have low clinical-outcome rates, and develop over long durations. Indeed, the use of surrogate markers to study relations between diet and chronic disease has been the basis for the DRIs that reflect chronic disease endpoints (2, 7) and for US Dietary Guidelines (8).

It is important to differentiate surrogate markers from risk biomarkers. The latter are biomarkers that indicate a component of an individual's level of risk of developing a disease or complications of a disease (1) (Table 1). A risk biomarker is

predictive of an outcome such as a chronic disease risk because of its correlation (e.g., association) with that disease. The stronger the association is, the greater is the value of the risk factor as a prognostic or diagnostic tool. However, the correlation of a risk biomarker with a chronic disease risk is not a sufficient basis for concluding that the risk biomarker is also a surrogate marker (6). To be qualified as a surrogate marker, the risk biomarker must be shown, among other things, to be on the causal pathway between the intervention of interest and the outcome of interest and to fully capture the effect of the intervention on the outcome (5). Therefore, although surrogate markers are also risk biomarkers, the converse is not valid. That is, risk biomarkers are not automatically also surrogate markers for a given diet–chronic disease relation. We discuss the use of these concepts and illustrate the concepts with examples.

## WHAT DETERMINES THE LIKELIHOOD THAT A SURROGATE MARKER WILL SUCCEED OR FAIL?

More than 20 y ago, Prentice (5) identified 2 criteria that are necessary for validating a surrogate marker: correlation and capture. First, the surrogate marker must have prognostic or predictive value relative to the chronic disease outcome (e.g., be correlated with). This criterion is commonly met. However, because it is correlative rather than causal, this criterion alone is not a sufficient basis to conclude that the biomarker is a surrogate marker. Instead, it is a risk factor that may or may not be on the causal pathway between dietary intake and chronic disease risk.

The second criterion states that an intervention's entire effect on the clinical outcome should be explained by the intervention's effect on the surrogate marker (**Figure 1**A). That is, the surrogate marker should fully account for all of an intervention's effects on the outcome of interest and should be a proxy for the effect of an intervention on risk of this outcome. This second criterion is more difficult to achieve, particularly with chronic diseases that have multifactorial causes and nutrient intakes that can have multiple effects on the clinical outcome of interest.

A key issue in the use of surrogate markers is to better understand why many promising surrogate markers have failed to survive the rigors of confirmation trials (1, 6, 9–11). One potential reason is that the proposed surrogate marker or biomarker is not on the causal pathway between intake and the chronic disease although it is correlated with the chronic disease of interest (i.e., is a risk factor for the chronic disease) and is also correlated with intake of interest (Figure 1B). That is, the hypothesized surrogate marker is a confounding rather than a mediating factor in the relation between intake and the chronic disease outcome. Fairchild and McDaniel (12) noted that distinguishing between confounding compared with mediating models is an important distinction in understanding causal pathways. The situation illustrated in Figure 1B would result in a spurious correlation between intake and chronic disease risk because of the common relation of intake and risk to the measured biomarker. A reliance on this hypothesized surrogate marker to suggest a relation between intake and the chronic disease would be misleading.

For example, observational studies have shown associations between elevated homocysteine concentrations and increased risk of CVD [coronary heart disease (CHD), stroke, and related diseases] (**Figure 2**A). It has also been well documented that inadequate intakes of B vitamins result in elevated homocysteine concentrations. Therefore, researchers hypothesized
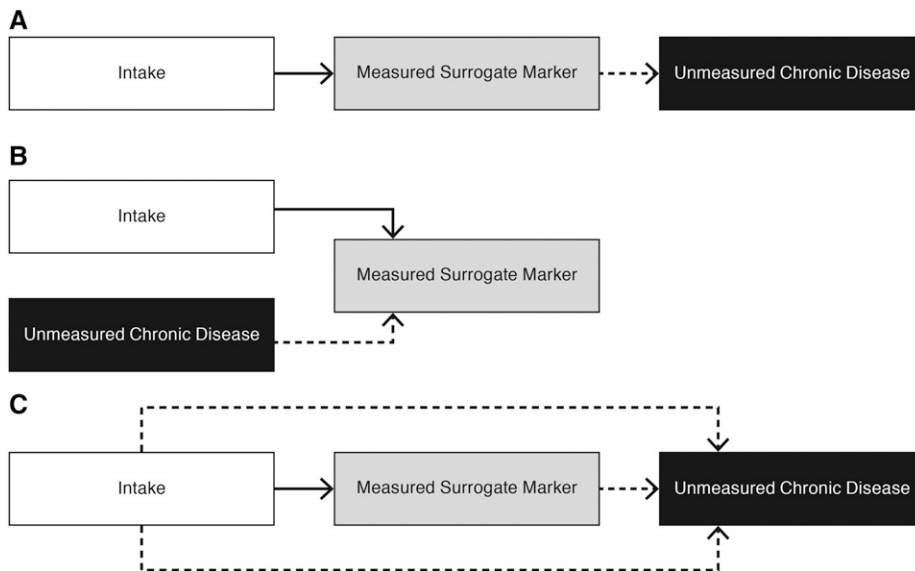

FIGURE 1 Possible scenarios for understanding whether a measured potential surrogate marker is a useful substitute for assessing the effect of intake on a chronic disease risk. Dashed arrows reflect unmeasured pathways. (A) The surrogate marker is a useful substitute for assessing the effect of intake on the development of a chronic disease because the surrogate marker is on the only pathway linking intake to the chronic disease risk and fully captures the causal effect of intake on the chronic disease outcome. (B) The surrogate marker is not a useful substitute for assessing the effect of intake on the development of a chronic disease because the surrogate marker does not mediate an effect of intake on the chronic disease risk; therefore, it is not on a causal pathway between intake and the chronic disease. The potential surrogate marker is a risk factor for the chronic disease but not a qualified surrogate marker. (C) The surrogate marker is not a useful substitute for assessing the effect of intake on the development of a chronic disease because the surrogate marker is on only one of several causal pathways between intake and the chronic disease. Therefore, the measured surrogate marker does not fully capture all of the effect of intake on the chronic disease. The unmeasured pathways may have greater or even opposite effects on the relation of intake to the chronic disease. Thus the measured surrogate marker would give misleading results.

that homocysteine would be a useful surrogate marker for relating insufficient B-vitamin intakes to an increased risk of CVD (Figure 2B). However, clinical trials have shown that lowering homocysteine concentrations with supplemental intakes of B vitamins had no effect on CVD risk (Figure 2C) (13). These results suggested that a higher concentration of homocysteine, although a risk factor for CVD and a result of inadequate intakes of B vitamins, was not on a causal pathway between this intake and this chronic disease risk (Figure 2D). In addition, these results suggested that the usefulness of elevated homocysteine concentrations as a possible diagnostic tool in identifying CVD risk would decrease in populations who consumed foods that were fortified with B vitamins or who consumed B vitamin–containing dietary supplements.

Surrogate markers can also fail to predict the clinical outcome of interest when multiple pathways exist between the intervention and the clinical outcome (i.e., the measured biomarker will not fully capture the effect of the intervention on the outcome) (Figure 1C) (1, 5, 6). If a study evaluates only one of these pathways by measuring the biomarker on this pathway, this measurement will not reflect the additional effect on the outcome of the unmeasured (alternate) pathways, thereby not qualifying the marker as a surrogate. Data on only one of these pathways can give misleading results, which is particularly the case if the unmeasured pathways have a more predominant or opposite effect than that of the measured pathway. An example of the likelihood of unmeasured pathways affecting an expected mediating effect of a potential surrogate marker on an intervention-outcome relation is LDL cholesterol within a context of a hormone replacement therapy (HRT) intervention and CVD [i.e., CHD (defined as myocardial infarction or coronary death) plus stroke] risk reduction. Because of the established benefit of LDL cholesterol on lowering CVD risk with statins (**Figure 3**A), it was commonly assumed that the observed HRT effect of lower concentrations of LDL cholesterol (as well as increased concentrations of HDL cholesterol) (14) would also result in a lower risk of CVD in HRT users. This assumption was the basis of the widespread use of HRT in common practice beyond symptom relief. Unexpectedly, in the Women's Health Initiative, the observed LDL-cholesterol reductions and increases in HDL cholesterol did not reduce the overall risk of CHD and significantly increased the risk of stroke (and therefore increased the risk of CVD) (Figure 3B) (15). The unexpected effects of HRT on CVD risk, despite the beneficial changes in LDL cholesterol, are likely related to unmeasured pathways in which the hormone can profoundly affect multiple organ systems and metabolic processes (Figures 1C, 3B). Thus, the failure of a surrogate marker that is qualified for one context (e.g., statins, LDL cholesterol, and CVD) to fully capture the effect of another intervention on this same outcome (i.e., HRT, LDL cholesterol, and CVD) can have unexpected results including unexpected adverse effects. Because chronic diseases have multifactorial causes, the measurement of only a single biomarker often fails because it does not capture all or most of the effects of the intervention on the disease. Understanding the disease mechanism of action can help to minimize the uncertainty that is inherent in the assumption that a proposed surrogate marker can predict all of an intervention's effects (1, 5).
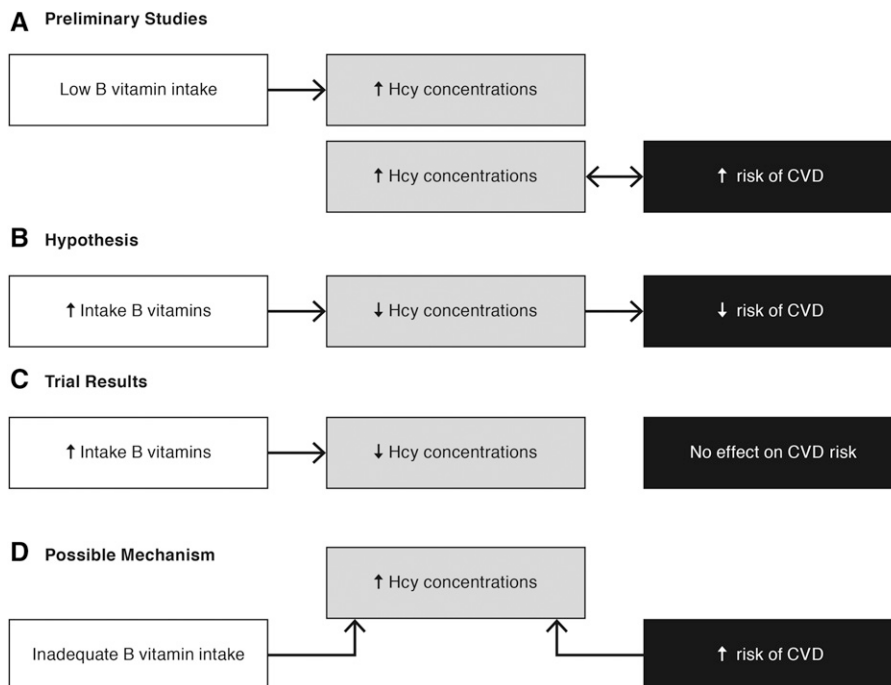
FIGURE 2 Example of a potential surrogate marker that did not perform as expected. (A) Low B-vitamin intakes produce elevated plasma Hcy concentrations. Observational studies showed an association between elevated Hcy concentrations and CVD risk. (B) Therefore, it was hypothesized that low vitamin B intakes would reduce the risk of CVD via a mediating influence of Hcy concentrations. (C) Results from phase III clinical trials showed that B-vitamin supplements reduced Hcy concentrations but had no effect on CVD risk. (D) One possible mechanism is that B-vitamin intake and CVD risk are correlated because of a common relation to Hcy concentrations. However, Hcy is not a mediating factor on a causal pathway between B-vitamin intake and CVD. Thus, elevated Hcy concentrations are a risk factor, but not a surrogate marker, for CVD. CVD, cardiovascular disease; Hcy, homocysteine; ↓, decreased; ↑, increased.
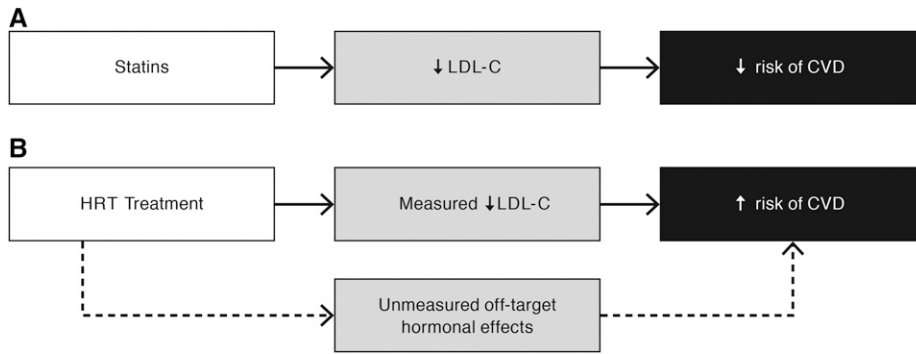
FIGURE 3 Generalizability of the surrogate marker from one intervention to a different intervention. (A) Multiple trials with different statins showed the usefulness of LDL-C as a surrogate marker for CVD risk. (B) However, when LDL-C was evaluated as a potential useful surrogate marker for CVD risk with HRT as the intervention, LDL-C did not accurately predict CVD risk and therefore was not a useful surrogate marker. Instead of LDL-C lowering being predictive of decreased CVD risk, it was associated with an increased risk of CVD likely because of the opposite and overwhelming effects of unmeasured pathways between the HRT treatment and CVD risk. Thus, LDL-C did not meet the criteria for surrogate markers because it did not fully capture the effect of the intervention on the outcome. CVD, cardiovascular disease; HRT, hormone replacement therapy; LDL-C, LDL cholesterol; ↓, decreased; ↑, increased.

The probabilistic nature of predictions made with the use of biomarker data, including surrogate marker data, means that some uncertainty may be associated with the use of surrogate markers as a substitute for the clinical outcome of interest. One source of uncertainty is that potential safety concerns may not be detected with the shorter duration and smaller sample sizes in studies that rely on surrogate markers (1). For example, the Antihypertensive and Lipid Lowering Treatment to Prevent Heart Attack Trial tested the efficacy of 4 previously approved anti-hypertensive drugs, all of which had been shown in earlier studies to be effective BP-lowering agents. However, the arm for testing one of these drugs (i.e., α-adrenergic blocker) was terminated early because of a higher incidence of combined CVD events even with concurrent BP-lowering effects (16). Thus, the initial reliance on BP as a surrogate marker for CVD with the use of an α-adrenergic blocker drug as the intervention did not detect this additional serious adverse consequence.

In his 1989 paper, Prentice (5) expressed pessimism concerning the potential of generalizing from one intervention in which the use of a surrogate marker was shown to be useful to other interventions. For a surrogate marker that has been validated in one situation to be generalizable to another situation, the surrogate marker must be valid for each intervention-outcome relation. An example of this validation was previously shown whereby LDL cholesterol did not successfully predict the overall lack of effect of HRT on CHD (defined as myocardial infarction or coronary death) or an increased risk of stroke (i.e., CVD) (15). The generalizability of a given potential surrogate marker can also be affected by a number of context-of-use (fit-for-purpose) issues (1). That is, the generalizability from one context to an-other is more complex than simply evaluating the intake-outcome relation and also requires a consideration of context-of-use factors such as the population (e.g., healthy compared with diseased, age, and sex), purpose (e.g., prevention compared with treatment), timing (e.g., before or after menopause), or health condition of the participant (e.g., baseline risk factors). For example, even within the Women's Health Trial in which LDL cholesterol was consistently reduced with HRT, the generaliz-ability of results from one type of HRT to another type of HRT resulted in different trends in CHD (i.e., nonfatal myocardial infarction or coronary death) risk (i.e., increasing for estrogen

plus progestin and decreasing for estrogen alone in women aged 50–59 y) (**Table 2**) (17). Different age groups having the same type of HRT (i.e., estrogen alone) had different trends in a global index outcome with younger women trending toward beneficial results and older women trending toward increased risk. The same type of HRT (e.g., estrogen plus progestin) produced variable effects on a range of outcomes with increased trends for CHD, stroke, and pulmonary embolism risks and decreased trends for colorectal cancer and hip-fracture risks. Increased risk that is associated with the global index indicates that, overall, the risk:benefit ratio was not advantageous. In these examples in which CHD or CVD was the outcome, the use of LDL choles-terol as a surrogate for disease risk would have often resulted in misleading results. It is likely that, in some of these examples, LDL cholesterol was not on a causal pathway between the HRT and the outcome (e.g., hip fracture) (Figure 1B). In other cases, it is likely that unmeasured alternate pathways overpowered an effect of LDL cholesterol on the outcome of interest (e.g., stroke) (Figures 1C, 3B).

These limited examples underscore the need to consider context-of-use (fit-for-purpose) issues when generalizing from one context of use for a qualified surrogate marker to another context of use with an understanding that the context of use varies between interventions (including even subtle differences between interventions) and outcomes can also be affected by such factors as demographics and population characteristics. A new context of use (different purpose, different intervention, different outcome, and other demographic and health conditions) requires a critical evaluation as to whether it is appropriate to generalize the use of a surrogate marker from one context in which qualification has been shown to another context in which such qualification has not been confirmed.

However, despite the inappropriateness of LDL cholesterol to serve as a surrogate marker for CHD and CVD with HRTs, it is beneficial to recognize that a particular biomarker may be useful for one purpose (e.g., risk prediction) but not useful for another purpose (e.g., surrogate marker for a chronic disease). In the Women's Health Initiative, the LDL-cholesterol concentration at baseline was shown to be a useful marker for predicting sub-sequent CHD risk in participants (18). Women with lower LDL cholesterol compared with women with higher LDL cholesterol

TABLE 2
Examples of challenges in generalizing the usefulness of a surrogate marker that is qualified in one context to other contexts[1]

Intervention	Intermediate biomarker	Outcome
Different interventions, same outcome		
Statins	Reductions LDL cholesterol[2]	Lower CHD risk
HRTs (estrogen plus progestin and estrogen alone)	Reductions LDL cholesterol[3]	No overall effect on CHD risk
Different forms of HRT intervention, same outcome, 50- to 59-y-old women		
Estrogen plus progestin	Reductions LDL cholesterol	CHD: 1.34 (0.82, 2.19)[4]
Estrogen alone	Reductions LDL cholesterol	CHD: 0.60 (0.35, 1.04)
Different age groups, same intervention, same outcome (global index)		
Estrogen alone, age, y		
50–59	Reductions LDL cholesterol	0.84 (0.66, 1.07)
60–69	Reductions LDL cholesterol	0.99 (0.85, 1.15)
70–70	Reductions LDL cholesterol	1.17 (0.99, 1.39)
Different outcomes, same intervention		
Estrogen plus progestin	Reductions LDL cholesterol	CHD: 1.18 (0.95, 1.45)
		Stroke: 1.37 (1.07, 1.76)
		Pulmonary embolism: 1.98 (1.36, 2.87)
		Colorectal cancer: 0.62 (0.43, 0.89)
		Hip fracture: 0.67 (0.47, 0.95)
		Global index: 1.12 (1.02, 1.24)

[1] From reference 17. CHD, coronary heart disease (defined as nonfatal myocardial infarction or coronary death); HRT, hormone replacement therapy.

[2] LDL cholesterol is a qualified surrogate marker for statin interventions and their effects on CHD and cardiovascular disease outcomes.

[3] On the basis of evidence in the Women's Health Initiative, LDL cholesterol does not meet the criteria as a qualified surrogate marker for the outcomes evaluated via HRT interventions.

[4] HR; 95% CI in parentheses (all such values).

at baseline had lower risk of incident CHD as the study progressed. Thus, LDL cholesterol was a useful risk factor but not a useful surrogate marker.

IOM'S BIOMARKER-EVALUATION PROCESS

The IOM committee recommended a framework for evaluating biomarkers that has the following 3 interrelated, but conceptually distinct, critical components: *1*) analytic validity, *2*) evidentiary qualification, and *3*) utilization analysis (**Table 3**) (1). The framework is intended to bring consistency and transparency to a currently nonuniform process and to facilitate research development and confidence in findings by clarifying the process. The framework applies to biomarkers that are used singly or in combination. It is applicable to all types of biomarkers ranging from exploratory uses, for which less evidence is needed, to surrogate marker uses, for which strong evidence is required. As noted previously, our focus is limited to the qualification of a proposed surrogate marker or biomarker as a substitute for measuring chronic disease outcomes in diet- and nutrition-based studies.

Analytic validation is a necessary first step in the biomarker-evaluation process (Tables 1 and 3) (1). Biomarker measurements need to be accurate and reproducible across time, laboratories, and assay methodologies. They also need to possess adequate sensitivity and specificity for their intended use. In the nutrition field, several expert committees described the critical importance of accuracy and traceability to higher-order reference methods in the measurement of biomarkers of vitamin D and folate status (19, 20). The concerns and misleading results that have been raised by the inaccurate measurement of these biomarkers of nutritional status also have relevance for surrogate markers of

chronic diseases. Unless a biomarker can be accurately and reliably measured, researchers cannot assess the comparability of results across studies or across time when these studies are used to evaluate the strength of the evidence or to evaluate whether a biomarker is useful for its proposed use.

The second step of the IOM's biomarker-evaluation process (i.e., the qualification step) involves an objective and comprehensive evaluation of the relevant scientific literature to address questions about the prognostic value of the biomarker's relation with disease (Tables 1 and 3) (1). Increasingly, systematic reviews and meta-analyses are being used to inform this step (8, 21). The qualification step has 3 components beginning with an evaluation of the biomarker's prognostic value. This evaluation includes a description of the nature and strength of the evidence for a biomarker's association with a given clinical outcome including whether the biomarker predicts the clinical outcome of interest. Observational and small interventional studies can provide useful evidence. The second component is a review of the evidence showing that the intervention targeting the biomarker actually affects the clinical endpoint of interest. This component, which addresses whether the biomarker is on the causal pathway between the intervention and the clinical outcome, usually requires robust, adequately controlled, clinical trial data. Some of these studies might also offer mechanistic information about the pathway. The last component is a review of the evidence showing that the relation between the biomarker and clinical outcome persists across multiple interventions. If the relation persists, the biomarker is more likely to be generalizable across interventions and contexts of use.

The last step in the IOM biomarker evaluation framework is the utilization (i.e., fit-for-purpose or context-of-use) step (Tables

TABLE 3
Institute of Medicine biomarker-evaluation process[1]

Step	Description
1) Analytic validation	Analysis of available evidence on the analytic performance of an assay
	Measurements are accurate across time, laboratories, and assays
	Measurements possess adequate sensitivity and specificity for their intended use
2) 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
	An evaluation of the biomarker's prognostic value
	Evidence showing that the intervention targeting the biomarker actually affects the clinical endpoint of interest
	Evidence that the relation between the biomarker and clinical outcome persists across multiple interventions
3) Utilization	Contextual analysis that is based on the specific use proposed and the applicability of available evidence to this use, which includes a determination of whether the analytic validation and qualification conducted provide sufficient support for the use proposed
	Define the purpose and context of use for the intended biomarker
	Assess the potential benefits and harms of the biomarker for its proposed use
	Identify the tolerability of risk for the biomarker within the proposed purpose and context of use
	Assess the evidentiary status of the biomarker

[1] From reference 1.

1 and 3) (1). This step generally has 4 components as follows: *1*) define the purpose and context of use for the intended biomarker, *2*) assess the potential benefits and harms of the biomarker for its proposed use, *3*) identify the tolerability of risk of the biomarker within the proposed purpose and context of use, and *4*) assess the biomarker's evidentiary status. The IOM recommends that an expert scientific committee, using information from the analytic and evidentiary steps, be convened to analyze the utilization step because it requires scientific judgment as well as comprehensive evidence reviews. This step involves determining a biomarker's appropriateness within a particular context of use (e.g., whether the intended use of the biomarker is for prevention, treatment, or mitigation of a chronic disease; population characteristics, such as sociodemographic status and baseline nutritional and disease status; nature and dosages of the intervention; prevalence of morbidity and mortality associated with relevant conditions; and the risk benefit of the intervention in the specified context). A key component of this step is determining the biomarker's usefulness in the studied context and whether this experience can be generalized to other contexts of use (e.g., to other populations and to interventions with similar putative effects). For example, if the biomarker was studied in a treatment context, is it also applicable to a prevention context? If the biomarker was tested with a class of pharmaceutical agents, is it applicable to studies that use dietary interventions?

In general, the qualification and utilization steps of the IOM biomarker evaluation framework are informed by the conclusions from the analytic validation step (1). The qualification and utilization steps are interactive in nature in that the information that is needed for the utilization step is collected and organized in the qualification step, and the information needs of the utilization step determine what types of information need to be collated in the qualification step. By separating these 2 steps in the conceptual framework, the differences in investigative and analytic processes that are required to evaluate the evidence and define the context of use are clarified.

The IOM focused considerable attention on the validation and use of surrogate markers (1). Although the framework is generic in nature and can be applied to different types of interventions (e.g., different dietary components and drugs) and outcomes (e.g., indicators of nutritional status and surrogate markers for chronic diseases) and for different purposes (e.g., risk assessment, diagnostic, and surrogacy), the assessment and modification of a qualification decision for a particular application is context specific. Within our focus on the qualification of surrogate markers, a surrogate qualified for one context of use (such as one target population, purpose, intervention, or clinical outcome) is not necessarily applicable to other contexts of use. The qualification of the surrogate is needed for each context including for different dietary components. For example, the generalization of the appropriateness of a surrogate (e.g., LDL cholesterol) that is qualified for specific dietary intake (e.g., saturated fat) and a specific outcome (e.g., CVD) to other contexts (e.g., other nutrient intakes or other outcomes) would require independent qualification for these other applications. In addition, as noted previously, a biomarker cannot be qualified as a surrogate solely on the basis of its correlation with or prediction of a clinical outcome or on the basis of biologically plausible mechanisms and pathways. A causal relation must also be established, and the surrogate marker must fully capture the effect of the intervention on the outcome of interest. Therefore, a surrogate marker needs strong and compelling evidence for its intended use.

Evidence that is based on surrogate markers allows for inferences but is always associated with some degree of uncertainty. Moreover, when surrogate markers are used in place of true clinical endpoints, researchers should consider the likelihood of unintended effects (both beneficial and adverse). The trials that are often used to study diet and health relations, which frequently have limited outcome measures, short durations, and small sample sizes, are unlikely to detect these effects, particularly unanticipated adverse effects.

CASE STUDIES

Scientists who are involved in evaluating the evolving scientific perspectives on the use of surrogate markers have noted that the initial optimism about the promise of surrogate markers has been replaced by increasing caution (1, 5, 6, 9, 10, 22). Although some successes have occurred, research findings have often been disappointing. Contrary to the researchers' initial hypotheses, changes in proposed surrogate markers have often failed to predict expected beneficial clinical outcomes. In a few cases, potential surrogate markers that have shown benefit have unexpectedly failed to predict adverse health effects.

In this section, we discuss the surrogate marker status of several biomarkers with relevance to nutrition-based studies. We include the following 2 examples for which we have considerable historical experience on the basis of several decades of expert committee evaluations: *1*) sodium intake and CVD risk with BP as a surrogate marker, and *2*) SFA intake and CVD risk with LDL cholesterol as a surrogate marker. We also briefly discuss other examples that have been used in nutrition–chronic disease evaluations where there is either insufficient evidence or evidence against qualifying a particular surrogate marker within a given nutrition context of use.

Sodium intake, BP, and CVD risk

The 2010 IOM biomarker committee identified BP as an "exemplar surrogate endpoint for cardiovascular mortality and morbidity due to the levels and types of evidence that support its use" (1). The committee noted that a strong body of evidence for >75 hypertensive agents from 9 drug classes have shown consistent effects between BP lowering and reduced risk of CVD regardless of different mechanisms of action that are involved in the BP-lowering effect. The benefits were observed across different assessment variables (e.g., systolic alone, diastolic alone, and systolic and diastolic) and in diverse populations (e.g., different sexes, adult age groups, and races and ethnicities). Thus, BP lowering, per se, had a beneficial effect on CVD risk.

For several decades, expert committees that were convened by the IOM, the USDA, and the US Department of Health and Human Services and by the American Heart Association (AHA), the American College of Cardiology (ACC), and National Heart, Lung, and Blood Institute (NHLBI) of the NIH recommended that Americans reduce their sodium intakes as a public health preventive measure for reducing risk of CVD (8, 21, 23–25). Their evaluative processes in recent years were generally consistent with the 2010 IOM's biomarker evaluation framework approach. For example, recent evaluations of the scientific literature were robust in nature and relied on systematic reviews and meta-analyses as part of their evidentiary reviews (8, 21). Throughout their history, the evaluations included the use of expert scientific committees to make decisions as to the usefulness of BP as a surrogate marker within the context of sodium intake and CVD risk in the general population. For these committees, a major challenge in evaluating this relation was the difficulty of accurately estimating sodium intakes in observational studies or in achieving and maintaining targeted sodium intakes in clinical trials of long duration (26).

The first step in the biomarker-evaluation process is the accurate measurement of the surrogate marker of interest. Groups such as the AHA and the CDC have a long history of providing recommendations and guidelines for equipment and protocol issues that are required for accurate measures of BP (27, 28).

A comprehensive review of the available scientific literature is the second step in the biomarker-evaluation procedure. Several expert committees have conducted comprehensive reviews to evaluate the strength of the evidence relating sodium intake to CVD risk (Table 3) (8, 21, 24, 25, 29). The referenced committees that were convened after the 1989 IOM report cited strong and direct clinical trial evidence of a dose-dependent relation of sodium intake on BP levels. Drug trials that showed a direct effect of the gradient of BP lowering on CVD risk were considered relevant to the sodium-intake context on the basis of evidence linking sodium intakes and CVD risk in large, prospective cohort studies. Considering the totality of the evidence, these committees consistently showed that the evidence revealed a relation between sodium intake and CVD risk that was consistent with the known effects of sodium intake on BP.

The last step in the biomarker-evaluation process is the utilization or fit-for-purpose (i.e., context-of-use) step. The committees defined their context of use as the development of nutrition recommendations and guidelines with applicability to a prevention context (reduced risk of CVD) by using a dietary intervention (i.e., reduced sodium intakes). The committees noted the following public health significance of the issues under review: *1*) 30% of US adults have high BP, and the estimated lifetime risk of developing hypertension in the United States is 90%; and *2*) >90% of adults have sodium intakes greater than the 2300-mg Tolerable Upper Intake Level with a mean of 3400 mg/d (8, 30). The committees deemed the BP-lowering effects of lower sodium intakes to be of sufficient magnitude to benefit persons with normotensive, prehypertensive, and hypertensive statuses as well as younger and older adult-age groups, males and females, and race-ethnic groups (21). Therefore, these committees concluded that, for the purpose of recommendations for the general public and for clinicians involved in managing CVD risk in patients, the available scientific evidence was sufficiently compelling to recommend reductions in the currently high sodium intakes in the United States. In making these recommendations, including considerations of the context of use (fit for purpose), 2 committees specifically stated that they recognized the usefulness of BP as a surrogate endpoint for evaluating the relation between sodium intake and risks of CVD and stroke (8, 25). A third working group identified BP as a "modifiable risk factor" for CVD (21). [Modifiable risk factors are considered surrogate endpoints because they are both modified by the intervention of interest and have an attributable effect on reducing risk of the outcome of interest (31, 32)].

Although the cited committees deemed the strength of the evidence in support of a relation between sodium intake and CVD risk within the population-based public health context to be sufficient to make recommendations for sodium-intake reductions, the evidence for defining optimal ranges of sodium intake was less clear (**Table 4**) (8, 24, 25). Multiple expert scientific committees over several decades and up to the current time have confirmed that the recommendation to reduce currently high sodium intakes (e.g., mean: 3440 mg/d) to lower intakes (e.g., <2300 mg/d or, more recently, by 1000 mg/d) is both safe and beneficial. However, questions have arisen as to whether very low sodium intake may adversely affect blood lipids, insulin resistance, renin, and aldosterone concentrations and may potentially increase risk of CVD and stroke

risk (29). Two evaluations of this topic concluded that the evidence is insufficient to determine whether there is a safety concern with very low sodium intakes (8, 25, 30).

As noted previously, the 2010 IOM biomarker committee defined the use of BP as a surrogate marker for CVD risk as an "exemplar surrogate endpoint for cardiovascular mortality and morbidity due to the levels and types of evidence that support its use." Over several decades, the evolving science has strengthened the use of BP as a surrogate marker within the context of sodium reduction as a public health recommendation for reducing risk of CVD, and multiple expert committees have consistently reconfirmed this recommendation (Table 4). Strong trial evidence relating sodium intakes to BP concentrations in conjunction with supporting evidence from observational studies linking sodium intakes and BP to CVD risk have informed its appropriateness for this context of use. Sodium reductions from currently high intakes to recommended intakes (i.e., <2300 or 2400 mg/d) have repeatedly been deemed both safe and effective. Future evaluations will likely expand their focus to more precisely define optimal intakes and to clarify potential modifying effects of other dietary components (e.g., potassium and water) (24, 33). Within the biomarker evaluation framework, the question of whether the usefulness of BP as a surrogate marker for CVD risk is generalizable to other nutrient intakes would require a similar type of surrogate marker qualification process.

SFA intakes, LDL-cholesterol concentrations, and CVD risk

The 2010 biomarker committee evaluated LDL cholesterol as a surrogate marker for CVD as one of their case studies (1). The committee evaluated this topic by using a qualification process that followed their biomarker evaluation framework including the use of the fit-for-purpose step. The committee concluded that "there is a high probability that lowering LDL for several interventions decreases risk of CVD and LDL, although not perfect, is one of the best biomarkers for CVD." LDL cholesterol is directly involved in the atherosclerotic disease process. The strength of the association between LDL cholesterol and CVD risk is based on findings that have shown that *1*) increases in LDL cholesterol result in increases in CVD risk even when other risk factors are present; and *2*) greater decreases in LDL cholesterol result in greater beneficial effects on CVD risk than do smaller decreases. However, as noted previously, although trials with different types of statin drugs have consistently shown beneficial effects of LDL-cholesterol lowering on CVD risk, similar LDL-cholesterol–lowering effects with other interventions (e.g., HRT) have not always been correlated with improved patient outcomes (1, 15, 17). The IOM biomarker committee concluded that "the strength of LDL-cholesterol as a surrogate endpoint is not absolute due to the heterogeneity of cardiovascular disease processes, the heterogeneity of LDL-cholesterol-lowering drug effects, and the heterogeneity of LDL-cholesterol particles themselves" (1).

Several nutrition-based expert committees have provided evaluations on the relation of SFA, LDL-cholesterol concentrations, and CVD risk (8, 21, 34). These evaluations had many similarities to the 2010 biomarker evaluation framework.

As noted previously, the first step in the biomarker evaluation framework is analytic validation. Issues that are related to the accurate and precise measurement of LDL cholesterol have been routinely addressed by 2 agencies. The National Cholesterol Education Program published recommendations for the accurate and precise measurement of LDL cholesterol (35), and the CDC conducts a program for the certification of manufacturers' clinical diagnostic products (36).

Expert committee reports between 1989 and 2015 conducted extensive scientific reviews to evaluate the relation between SFA intakes, LDL-cholesterol concentrations, and risk of CHD or CVD (**Table 5**). The reports' findings consistently supported recommendations that Americans reduce their intakes of SFA (8, 21, 34). Over time, the strength of the evidence relating SFA intakes to LDL-cholesterol concentrations as well as the evidence relating SFA intakes to CVD outcomes increased in strength and consistency. Recently, the ACC/AHA/NHLBI (21) report noted that "favorable effects on lipid profiles are greater when saturated fat is replaced by polyunsaturated fatty acids, followed by monounsaturated fatty acids, and then carbohydrates." Subsequently, the 2015 Dietary Guidelines Advisory Committee (DGAC), which had access to additional data, concluded that there was strong and consistent evidence from randomized controlled trials and the statistical modeling of prospective cohort studies that the replacement of SFAs with unsaturated fats, especially PUFAs, significantly reduces LDL cholesterol and reduces the risk of CVD events and coronary mortality (8). However, the committee also noted that replacing SFA with total carbohydrates, while reducing LDL cholesterol, appeared not to be associated with CVD benefits. Carbohydrate replacements of SFA adversely affected the lipid profiles of subjects (i.e., increased triglyceride concentrations and reduced HDL-cholesterol concentrations, which are factors that are related to increased risk of CVD). There was insufficient evidence to determine whether the type of carbohydrates (e.g., sugars and dietary fibers) has similar or different effects on CVD risk. The DGAC 2015 findings are illustrative of the need for evaluating and qualifying surrogate markers within their context of use.

These findings are consistent with the conclusion of the 2010 IOM biomarker committee that "the strength of LDL-cholesterol as a surrogate endpoint is not absolute due to the heterogeneity of cardiovascular disease processes, the heterogeneity of LDL-cholesterol-lowering drug effects, and the heterogeneity of LDL-cholesterol particles themselves" (1). With the SFA, LDL cholesterol, and CVD relation, there are heterogeneities at every step. There is variability in the different types of SFAs and their effects on LDL cholesterol (34, 37). There are multiple macronutrient-replacement possibilities when SFA intakes are reduced, and the different replacements have variable effects on CVD risk (8, 21). LDL-cholesterol particles are heterogeneous with variable risk factors (1). CVD is a multifactorial disease, and not all effective interventions exert their effects via reductions in LDL cholesterol. With evolving science, these issues will likely be clarified by future expert committees. As always, caution is warranted in attempting to generalize qualification decisions for the SFA-based surrogate status of LDL cholesterol to other nutrient intakes and CVD risk.

HDL cholesterol and CVD risk

The HDL-cholesterol concentration has long been accepted as a risk factor for CVD. As such, it has been commonly relied on to help interpret study results. For example, both the AHA/ACC

TABLE 4
Examples of expert committee evaluations of the sodium, blood pressure, and CVD relation[1]

Expert committee, year (ref)	Sodium and BP	BP and CVD	Sodium and CVD	Recommendation/conclusion
Diet and Health, 1989 (23)	Strong epidemiologic evidence Supportive animal data	High BP is a major CVD risk factor	—	BP levels are strongly and positively correlated with habitual intake of salt. Limit total daily intake of salt (sodium chloride) to ≤6 g.
Dietary Reference Intakes, 2005 (24)	Rigorous dose-response trials	Persuasive results: observational studies and drug trials Well-accepted public health tenet	Animal models: strong association Observational studies: association	There is a progressive, direct effect of dietary sodium intake on BP in nonhypertensive and hypertensive individuals and a direct relation between BP and risk of CVD and end-stage renal disease.
Sodium Intake in Populations, 2013 (25)	Intervention studies: dose-response between sodium intake and BP in normotensive and hypertensive individuals	Strong support for high BP and higher risk of CVD BP as a surrogate endpoint for risk of CVD and stroke is widely recognized and accepted	Consistent evidence for association between excessive sodium intakes and increased risk of CVD Inconsistent evidence for beneficial or adverse effects of intakes <2300 mg/d	Evidence of direct health outcomes: a positive relation between higher sodium intakes and risk of CVD, which are consistent with the known effects of sodium intake on BP. There is inconsistent and insufficient evidence to conclude that decreasing sodium intakes to <2300 mg/d either increases or decreases risk of CVD outcomes or all-cause mortality.
AHA/ACC: Lifestyle Management to Reduce CVD Risk, 2013 (21)	Strong and consistent clinical trial data of relation (high)[2] Reduce sodium to ~2400 mg/d (moderate)[3] Reduce sodium intake by ~1000 mg/d (high)	Considered BP a modifiable risk factor for CVD prevention and treatment[4]	Observational data: higher sodium intake is associated with greater risk for fatal and nonfatal stroke and CVD (low)[5]	Advise adults who would benefit from BP lowering to decrease sodium intake.
Dietary Guidelines Advisory Committee, 2015 (8)	Strong evidence linking sodium and BP Moderate evidence on amounts of sodium intakes	—	Moderate evidence linking sodium to CVD (on the basis of an updated review of the 2013 IOM and AHA/ACC evidence reviews) Grade not assignable that intakes <2300 mg/d increase or decrease CVD risk; limited evidence that lowering sodium intake by 1000 mg/d might lower CVD risk by 30%	Consume <2300 mg Na/d; recommendation included considerations of evidence on BP as a surrogate indicator of CVD risk.

[1] AHA/ACC, American Heart Association/American College of Cardiology; BP, blood pressure; CVD, cardiovascular disease; IOM, Institute of Medicine; ref, reference.

[2] High strength-of-evidence grade is based on results from well-designed and well-executed randomized clinical trials. There is high certainty about the estimate of the effect. Further research is unlikely to change confidence in the estimate of the effect (8, 21).

[3] Moderate strength-of-evidence grade reflects evidence from randomized clinical trials with minor limitations, well-designed and well-executed non-randomized controlled studies, and well-designed and well-executed observational studies (8, 21). There is moderate certainty about the estimate of the effect. Further research may have an impact on confidence in the estimate of the effect and may change the estimate.

[4] Modifiable risk factors and markers are also sometimes called surrogate endpoints and may serve as surrogates for the incidence of chronic diseases (31). Modifiable risk factors are modified by the intervention of interest and can be used to estimate population-attributable risk that can be attributed to a particular risk factor for a particular chronic disease (32).

[5] Low strength-of-evidence grade reflects evidence from randomized clinical trials with major limitations, nonrandomized controlled studies and observational studies with major limitations, and uncontrolled clinical observations without an appropriate comparison group. There is low certainty about the estimate of the effect. Further research is likely to have an impact on confidence in the estimate of the effect and is likely to change the estimate (8, 21).

and the DGAC compared the relative beneficial effects of different SFA replacements on LDL cholesterol by also looking at concurrent changes in lipid profiles (e.g., HDL cholesterol and triglycerides) (8, 21). Recently, the Risk Assessment Work Group that was sponsored jointly by the ACC, the AHA, and the NHLBI/NIH used a systematic review methodology and sex- and

TABLE 5
Examples of expert committee evaluations of the saturated fat, LDL cholesterol, and CVD relation[1]

Expert committee, year (ref)	SFAs and LDL cholesterol	LDL cholesterol and CVD	SFAs and CVD	Recommendation or conclusion
Diet and Health, 1989 (23)	Clinical, animal, and epidemiologic studies: higher intakes of SFAs are related to higher LDL cholesterol	Higher LDL cholesterol leads to atherosclerosis and higher risk of CVD	Observational studies: CHD rates and population risk are strongly related to average LDL cholesterol, and LDL cholesterol is strongly influenced by SFA intakes	Reduce SFA to <10% of calories.
DRIs, 2002/2005 (34)	Many studies: higher intake of SFAs result in higher LDL cholesterol concentrations SFAs differ in metabolic effects on LDL cholesterol	Positive linear relation between serum LDL cholesterol and CHD risk or mortality from CHD	Observational studies: association between SFAs and risk of CHD	There is a positive linear trend between total SFA intake, LDL cholesterol concentrations, and higher risk of CHD. UL was not set because relation is linear across all intakes; a curvilinear relation with a threshold dose is a requirement for applying the UL model in setting DRIs.
AHA/ACC: Lifestyle Management to Reduce CVD Risk, 2013 (21)	Strong evidence: reductions in LDL cholesterol are achieved when SFAs are reduced from 14–15% to 5–6% of calories Lower SFAs reduce LDL cholesterol regardless of whether SFAs are replaced by carbohydrate, MUFAs, or PUFAs, but there are more favorable lipid profiles with replacement by PUFAs	Considered LDL cholesterol a modifiable risk factor for CVD prevention and treatment[2]	Not reviewed: outside scope of guidelines	Advise adults who would benefit from LDL-cholesterol lowering to aim for a dietary pattern that achieves 5–6% of calories from SFAs.
Dietary Guidelines Advisory Committee, 2015 (8)	Strong and consistent evidence from RCTs shows that replacing SFAs with unsaturated fats, especially PUFAs, significantly results in lower LDL cholesterol Replacing SFAs with carbohydrate results in lower LDL cholesterol but higher triglycerides and lower HDL cholesterol	—	Strong and consistent evidence from RCTs and statistical modeling in prospective cohort studies: replacing SFAs with PUFAs results in lower risks of CVD and coronary mortality Consistent evidence from prospective cohort studies: higher SFA intake compared with total carbohydrates is not associated with CVD risk	Consume <10% of calories per day from saturated fat; shift food choices that are high in SFAs to foods that are high in PUFAs and MUFAs.

[1] AHA/ACC, American Heart Association/American College of Cardiology; CHD, coronary heart disease; CVD, cardiovascular disease; DRI, Dietary Reference Intake; RCT, randomized controlled trial; ref, reference; UL, Tolerable Upper Intake Level.
[2] Modifiable risk factors and markers are also sometimes called surrogate endpoints and may serve as surrogates for the incidence of chronic diseases (31). Modifiable risk factors are modified by the intervention of interest and can be used to estimate population-attributable risk that can be attributed to a particular risk factor for a particular chronic disease (32).

race-specific proportional hazards models to develop and validate a quantitative risk-assessment tool (38). The Work Group included HDL cholesterol as one of the 6 covariates in their new risk-assessment algorithm. Thus HDL cholesterol is an accepted risk factor for CVD risk because it indicates a component of an individual's level of risk of developing a disease (Table 1).

However, as noted previously, a risk factor may not meet the criteria for qualification as a surrogate marker (Table 1) (1, 6). As a correlate of a disease, a risk factor has a predictive value. But this value alone is not sufficient to show that the factor meets the criteria for a surrogate marker that it must be on the causal pathway between an intervention and disease outcome and must also fully capture the effect of the intervention on the outcome. The 2010 biomarker committee evaluated HDL cholesterol as a surrogate marker as part of their case study on LDL cholesterol and CVD risk (1). The committee concluded that, "although a low level of HDL may signal a higher CHD risk than a moderately high LDL-cholesterol level, HDL has not yet qualified as a surrogate endpoint for CVD risk because there is no

evidence that HDL-raising interventions can improve outcomes." The biological complexities of HDL cholesterol and its relation to CVD risk are not well understood. For example, the results of several completed trials dashed expectations that increases in HDL cholesterol would be a useful surrogate marker for CVD risk in drug trials (39–42). In one trial, a drug (torcetrapib) raised HDL cholesterol but unexpectedly increased the risk of death (39). In another trial, high-dose niacin (1500–2000 mg/d) in patients who were treated with simvastatin but who had residual low HDL-cholesterol and high triglyceride concentrations showed the expected favorable improvements in HDL cholesterol and triglycerides, but there was no evidence of a clinical benefit or reduction in CVD events (40). There was also a nonsignificant trend toward an increased risk of ischemic stroke in the niacin-treated group. Despite the expected changes in biomarker concentration in these studies, the use of HDL cholesterol as a surrogate marker for CVD risk did not accurately predict a clinical benefit or harm.

Surrogate markers for cancer risk

The reduction of cancer risk through nutrient-intake changes has been a topic of considerable research and public health interest for many years. Because of relatively low risk of developing cancer, large clinical trials of long duration are needed to evaluate diet and cancer relations. Therefore, the use of intermediate markers of possible cancer-preventive efficacy has been of interest. Although researchers are actively investigating many biomarkers for use as surrogate endpoints in clinical trials for the prevention of cancer, attempts to validate surrogate markers for cancer risk have often been disappointing (9, 11, 43). Mayne et al. (11) described the results of several nutrition intervention trials in which results that were based on premalignancy endpoints for head and neck cancers suggested a benefit. However, subsequent phase III trials with head and neck cancer outcomes showed no benefit of the nutritional interventions although, unexpectedly, also showed increased adverse effects (e.g., mortality and risk of some other cancers). Schatzkin and Gail (9) discussed the challenges of the use of adenomatous polyps as surrogates for colorectal cancer. The timing of the intervention is important because recurrent adenomas occur early in the tumorigenic sequence. Thus, results of adenoma-recurrence trials can be misleading if the intervention being tested is introduced later in the neoplastic process. Also, the biological heterogeneity of adenomas can be a problem. Only a relatively small proportion of adenomas develop into cancer. Thus, it is possible to show significant reductions in the pool of innocent adenomas without affecting the bad adenomas or vice versa. Therefore, there trial results would likely produce misleading information. Consistent with this assessment, the intervention group who consumed a low-fat, high-fruit and -vegetable diet in the Women's Health Initiative clinical trial of a nutritional intervention showed a lower incidence of self-reported polyps or adenomas, but there was no evidence of a trend toward lower colorectal cancer risk over a mean 8.1-y study period (44). Although the use of adenomatous polyps in primary prevention trials has not been formally validated, this does not mean that adenomatous polyps cannot be useful for other purposes in cancer-prevention studies. For example, these biomarkers can provide useful information to identify individuals according to their cancer risk for trial-selection

and -stratification purposes to determine the optimal dose or exposure and to prioritize decisions on moving from phase II to phase III trials.

SUGGESTED WAY FORWARD

Over the past several decades, studies that were designed to confirm hypothesized relations between dietary intakes and chronic disease outcomes have resulted in both successes and failures. For example, the AHA/ACC and DGAC concluded that there is strong evidence that following plans such as the Dietary Approaches to Stop Hypertension Trial dietary pattern, the USDA Food Patterns, or the AHA diet recommendations is effective in lowering LDL cholesterol and BP and in modifying risk of CVD (8, 21). Conversely, phase III trials, such as those that have been used to evaluate hypothesized relations of B vitamins to CVD, vitamin E to CVD and prostate cancer, or β carotene to lung cancer, did not find the dietary interventions to be effective in reducing risk of the chronic disease of interest and, in several cases, unexpectedly observed increased risk of chronic disease outcomes with the dietary intervention (2). Other disciplines have experienced similar results. For example, the Food and Drug Administration identified 22 examples of drug, vaccine, and medical device products in which promising phase II clinical trial results were not confirmed in phase III clinical testing (45). Results included failures to confirm effectiveness (14 products), safety (1 product), and both safety and effectiveness (7 products). Bikdeli et al. (46) tracked the results of CVD-outcome trials on interventions that were tested with surrogate outcomes between 1990 and 2011. Nearly 50% of the positive surrogate trials that were subsequently tested in clinical outcome trials were not validated. Of those negative surrogate trial results that were subsequently followed by phase III clinical trials, most remained negative. Thus, surrogate markers were more effective for excluding likely benefit than for identifying it. More than 20 y of experience with the successes and failures of proposed or not adequately validated surrogate markers have underscored the need for confidence in their qualification (1, 4).

How do we improve the process of identifying and confirming potential surrogate markers for evaluating hypothesized relations between dietary intakes and chronic disease–risk-reduction outcomes in the future? Without question, the use of qualified surrogate disease markers can provide scientific and fiscal efficiency to studies of diet and health relations. A generic framework such as the sequence of scientific discovery outlined in the IOM biomarker evaluation framework could provide clarity of means, goals, and approaches as well as facilitate public and private research collaborations in the development, validation, and use of these biomarkers. For example, the framework could be useful in identifying the questions that would need to be addressed in contributing to an integrated database that is necessary for qualifying a surrogate marker and showing the validity of a dietary intake–CVD relation. The framework could help to clarify different uses and appropriate interpretations of biomarkers within a specific intake–chronic disease context. Preliminary research conducted by the private sector and academic research groups can contribute to evidence on assay validation for potential surrogate markers, the delineation of mechanisms of action, the identification of potential safety concerns, the evaluation of dosage conditions and sample-size estimations, and

risk predictions and the durability of effects. The framework can provide a systematic and transparent basis for identifying research needs and for collating and evaluating research results from both the private and public sectors.

The biomarker evaluation framework can also help to clarify questions as to the generalizability, from one nutrient to another nutrient, of a qualified surrogate marker decision for a given outcome. Are there classes of dietary interventions for which surrogate markers that are qualified for one nutritional component can be extended to other nutritional components within that class? If so, what criteria and approaches are relevant for these types of decisions? The framework can also be useful in clarifying how factors that modify the intake–surrogate marker–outcome context of use should be addressed within the surrogate marker qualification process.

Historically, clinical researchers often initiated large trials of diet and health relations without the sequential preclinical, phase I, and phase II testing that is commonplace for drugs. As a consequence, the phase III clinical trials that have tested nutritional or dietary interventions have had an unnecessarily high rate of failures because their designs often were not fully informed by preliminary findings. Currently, although both the private and public sectors often support smaller preliminary studies, the results are rarely coordinated and integrated into a single database for use in justifying and designing large confirmatory trials. These data are needed to design the large, long-term, prospective cohort studies and phase III clinical trials that frequently require public sector or foundation support and that are necessary to confirm the use of hypothesized surrogate markers.

Once the data are available to identify potential surrogate markers, it is necessary to confirm them in phase III trials. However, this is often not done. For example, Bikdeli et al. (46) noted that, of the identified surrogate trials that showed superiority of the tested intervention in phase II trials, fewer than one-third of them were subsequently evaluated in phase III trials. At the same time, evolving research using BP and LDL cholesterol has both strengthened and refined the usefulness of these biomarkers within particular dietary intake–CVD contexts of use. To enhance the likelihood that future large clinical trials will succeed, the use of the biomarker evaluation framework to organize and integrate the available results from well-designed, shorter studies could help to build enthusiasm in funding agencies for phase III trials. The use of the framework could also enhance the likelihood of success of future large trials by aiding in the design and conduct of the trials.

CONCLUSIONS

The IOM defines a generic biomarker evaluation framework with the following 3 steps: analytic validation, qualification, and utilization (1). This framework can be used to qualify and validate surrogate markers for studying the relations between diet and clinical outcomes. The IOM notes that the use of a surrogate marker depends on the quality of evidence that supports its use and the context in which the biomarker is intended to be used. We illustrate the application of the IOM's generic framework to the evaluation of the relations between diet and health outcomes.

The scientific process is both cumulative and self-correcting. These characteristics are evident when considering the use of surrogate markers in evaluating relations between interventions and outcomes. In >2 decades of accumulating evidence, the number of failures to identify and validate surrogate markers has exceeded the number of successes, and the early enthusiasm of many scientists for surrogate markers has been replaced with caution. We need to consider the cumulative experiences that have been gained in past evaluations and uses of surrogate markers in nutrition research and identify strategies for moving the science forward in as efficient and scientifically sound a manner as possible. Note that, in the 2 successful case studies described (i.e., sodium, BP and CVD; and SFAs, LDL cholesterol, and CVD), the science linking intakes to clinical outcomes through surrogate markers was evolving in important ways over time. The studies were increasingly more rigorously developed with respect to the sample size and duration to capture effects on well-defined clinical outcomes. Moreover, with the increasing use of systematic reviews and meta-analyses, the assemblage of study results was more complete and involved a critical and more quantitative assessment. This type of evolving science resulted in strengthening and refining prior recommendations and, although not previously mentioned, also resulted in a withdrawal of the long-standing recommendation to reduce intakes of dietary cholesterol because of the failure of newer evidence to support this recommendation (8). With evolving science and accumulating experiences with potential surrogate markers, in conjunction with the IOM biomarker-evaluation conceptual framework, our understanding and wise use of surrogate markers are likely to improve.

REFERENCES

1. Institute of Medicine. Evaluation of biomarkers and surrogate endpoints in chronic disease [Internet]. Washington (DC): The National Academies Press; 2010 [cited 2017 Jun 23]. Available from: https://www.nap.edu/read/12869/chapter/1.
2. Yetley EA, MacFarlane AJ, Greene-Finestone LS, Garza C, Ard JD, Atkinson SA, Bier DM, Carriquiry AL, Harlan WR, Hattis D, et al. Options for basing Dietary Reference Intakes (DRIs) on chronic disease endpoints: report from a joint US-/Canadian-sponsored working group. Am J Clin Nutr 2017;105:249S–85S.
3. National Academies of Sciences, Engineering, and Medicine. Guiding principles for developing Dietary Reference Intakes based on chronic disease [Internet]. Washington (DC): The National Academies Press; 2017 [cited 2017 Aug 19]. Available from: http://www.nationalacademies.org/hmd/Reports/2017/guiding-principles-for-developing-dietary-reference-intakes-based-on-chronic-disease.aspx.
4. Califf RM. Warning about shortcuts in drug development. J Am Heart Assoc 2017;6:e005737.
5. Prentice RL. Surrogate endpoints in clinical trials: definition and operational criteria. Stat Med 1989;8:431–40.

6. Fleming TR, DeMets DL. Surrogate end points in clinical trials: are we being misled? Ann Intern Med 1996;125:605–13.

7. 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 [Internet]. Washington (DC): The National Academies Press; 2007 [cited 2017 Jun 23]. Available from: https://www.nal.usda.gov/sites/default/files/fnic_uploads//Selection_Endpoints_Determining_EARs_AIs_ULs.pdf.

8. Dietary Guidelines Advisory Committee. Scientific report of the 2015 Dietary Guidelines Advisory Committee. Advisory report to the Secretary of Health and Human Services and the Secretary of Agriculture [Internet]. Washington (DC): US Department of Health and Human Services; US Department of Agriculture;2015 [cited 2017 Jun 23]. Available from: https://health.gov/dietaryguidelines/2015-scientific-report/PDFs/Scientific-Report-of-the-2015-Dietary-Guidelines-Advisory-Committee.pdf.

9. Schatzkin A, Gail M. The promise and peril of surrogate end points in cancer research. Nat Rev Cancer 2002;2:19–27.

10. Schatzkin A. Promises and perils of validating biomarkers for cancer risk. J Nutr 2006;136:2671S–2S.

11. 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.

12. Fairchild AJ, McDaniel HL. Best (but oft-forgotten) practices: mediation analysis. Am J Clin Nutr 2017;105:1259–71.

13. 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.

14. Hsia J, Otvos JD, Rossouw JE, Wu L, Wassertheil-Smoller S, Hendrix SL, Robinson JG, Lund B, Kuller LH, Women's Health Initiative Research Group. Lipoprotein particle concentrations may explain the absence of coronary protection in the women's health initiative hormone trials. Arterioscler Thromb Vasc Biol 2008;28:1666–71.

15. Rossouw JE, Prentice RL, Manson JE, Wu L, Barad D, Barnabei VM, Ko M, LaCroix AZ, Margolis KL, Stefanick ML. Postmenopausal hormone therapy and risk of cardiovascular disease by age and years since menopause. JAMA 2007;297:1465–77.

16. ALLHAT Collaborative Research Group. Major cardiovascular events in hypertensive patients randomized to doxazosin vs chlorthalidone: the antihypertensive and lipid-lowering treatment to prevent heart attack trial (ALLHAT). JAMA 2000;283:1967–75.

17. Manson JE, Chlebowski RT, Stefanick ML, Aragaki AK, Rossouw JE, Prentice RL, Anderson G, Howard BV, Thomson CA, LaCroix AZ, et al. Menopausal hormone therapy and health outcomes during the intervention and extended poststopping phases of the Women's Health Initiative randomized trials. JAMA 2013;310:1353–68.

18. Bassuk SS, Manson JE. Menopausal hormone therapy and cardiovascular disease risk: utility of biomarkers and clinical factors for risk stratification. Clin Chem 2014;60:68–77.

19. Yetley EA, Pfeiffer CM, Schleicher RL, Phinney KW, Lacher DA, Christakos S, Eckfeldt JH, Fleet JC, Howard G, Hoofnagle AN, et al. NHANES monitoring of serum 25-hydroxyvitamin D: a roundtable summary. J Nutr 2010;140:2030S–45S.

20. Yetley EA, Pfeiffer CM, Phinney KW, Fazili Z, Lacher DA, Bailey RL, Blackmore S, Bock JL, Brody LC, Carmel R, et al. Biomarkers of folate status in NHANES: a roundtable summary. Am J Clin Nutr 2011;94:303S–12S.

21. Eckel RH, Jakicic JM, Ard JD, de Jesus JM, Houston Miller N, Hubbard VS, Lee IM, Lichtenstein AH, Loria CM, Millen BE, et al.. 2013 AHA/ACC guideline on lifestyle management to reduce cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol 2014;63:2960–84.

22. Prentice RL. Surrogate and mediating endpoints: current status and future directions. J Natl Cancer Inst 2009;101:216–7.

23. Institute of Medicine. Diet and health: implications for reducing chronic disease risk [Internet]. Washington (DC): The National Academies Press; 1989 [cited 2017 Jun 23]. Available from: https://www.nal.usda.gov/sites/default/files/fnic_uploads//Diet_Health_Implications_Reducing_Chronic_Disease_Risk.pdf.

24. Institute of Medicine. Dietary Reference Intakes for water, potassium, sodium, chloride, and sulfate [Internet]. Washington (DC): The National Academies Press; 2005 [cited 2017 Jun 23]. Available from: https://www.nap.edu/read/10925/chapter/1.

25. Institute of Medicine. Sodium intake in populations: assessment of evidence [Internet]. Washington (DC): The National Academies Press; 2013 [cited 2017 Jun 23]. Available from: https://www.nap.edu/catalog/18311/sodium-intake-in-populations-assessment-of-evidence.

26. Titze J. Estimating salt intake in humans: not so easy! Am J Clin Nutr 2017;105:1253–4.

27. National Health and Nutrition Examination Survey (NHANES). Physician examination procedures manual [Internet]. Atlanta (GA): Centers for Disease Control and Prevention; 2015 [cited 2017 Jun 23]. Available from: https://wwwn.cdc.gov/nchs/data/nhanes/2015-2016/manuals/2015_Physician_Examination_Procedures_Manual.pdf.

28. Pickering TG, Hall JE, Appel LJ, Falkner BE, Graves J, Hill MN, Jones DW, Kurtz T, Sheps SG, Roccella EJ. Recommendations for blood pressure measurement in humans and experimental animals. Part I: blood pressure measurement in humans: a statement for professionals from the Subcommittee of Professional and Public Education of the American Heart Association Council on High Blood Pressure Research. Hypertension 2005;45:142–61.

29. Strom BL, Anderson CA, Ix JH. Sodium reduction in populations: insights from the Institute of Medicine committee. JAMA 2013;310:31–2.

30. Gunn JP, Barron JL, Bowman BA, Merritt RK, Cogswell ME, Angell SY, Bauer UE, Frieden TR. Sodium reduction is a public health priority: reflections on the Institute of Medicine's report, sodium intake in populations: assessment of evidence. Am J Hypertens 2013;26:1178–80.

31. Fabian CJ, Kimler BF. Incorporating biomarkers in studies of chemoprevention. Adv Exp Med Biol 2016;882:69–94.

32. Goldstein LB, Adams R, Alberts MJ, Appel LJ, Brass LM, Bushnell CD, Culebras A, Degraba TJ, Gorelick PB, Guyton JR, et al. Primary prevention of ischemic stroke: a guideline from the American Heart Association/American Stroke Association Stroke Council: cosponsored by the Atherosclerotic Peripheral Vascular Disease Interdisciplinary Working Group; Cardiovascular Nursing Council; Clinical Cardiology Council; Nutrition, Physical Activity, and Metabolism Council; and the Quality of Care and Outcomes Research Interdisciplinary Working Group. Stroke 2006;37:1583–633.

33. King JC, Reimers KJ. Beyond blood pressure: new paradigms in sodium intake reduction and health outcomes. Adv Nutr 2014;5:550–2.

34. Institute of Medicine. Dietary Reference Intakes for energy, carbohydrate, fiber, fat, fatty acids, cholesterol, protein, and amino acids [Internet]. Washington (DC): The National Academies Press; 2002/2005 [cited 2017 Jun 23]. Available from: https://www.nap.edu/read/10490/chapter/1.

35. Bachorik PS, Ross JW. National Cholesterol Education Program recommendations for measurement of low-density lipoprotein cholesterol: executive summary. Clin Chem 1995;41:1414–20.

36. Cholesterol Reference Method Laboratory Network. LDL cholesterol certification protocol for manufacturers. National reference system for cholesterol [Internet]. Atlanta (GA): Cholesterol Reference Method Laboratory Network; 2006 [cited 2017 Jun 23]. Available from: https://www.cdc.gov/labstandards/pdf/crmln/MFRLDLJune2006final.pdf.

37. Bier DM. Saturated fats and cardiovascular disease: interpretations not as simple as they once were. Crit Rev Food Sci Nutr 2016;56:1943–6.

38. National Heart, Lung, and Blood Institute. Assessing cardiovascular risk. Systematic evidence review from the Risk Assessment Work Group, 2013[Internet]. Bethesda (MD): National Heart, Lung and Blood Institute, National Institutes of Health; 2013 [cited 2017 Jun 23]. Available from: https://www.nhlbi.nih.gov/health-pro/guidelines/in-develop/cardiovascular-risk-reduction/risk-assessment.

39. Barter PJ, Caulfield M, Eriksson M, Grundy SM, Kastelein JJP, Komajda M, Lopez-Sendon J, Mosca L, Tardif J-C, Waters DD, et al. Effects or torcetrapib in patients at high risk for coronary events. N Engl J Med 2007;357:2109–22.

40. AIM-HIGH Investigators, Boden WE, Probstfield JL, Anderson T, Chaitman BR, Desvignes-Nickens P, Koprowicz K, McBride R, Teo K, Weintraub W. Niacin in patients with low HDL cholesterol levels receiving intensive statin therapy. N Engl J Med 2011;365: 2255–67.

41. Giugliano RP. Niacin at 56 years of age – time for an early retirement? N Engl J Med 2011;365:2318–20.

42. Briel M, Ferreira-Gonzalez I, You JJ, Karanicolas PJ, Akl EA, Wu P, Blechacz B, Bassler D, Wei X, Sharman A, et al. Association between change in high density lipoprotein cholesterol and cardiovascular disease morbidity and mortality: systematic review and meta-regression analysis. BMJ 2009;338:b92.

43. Dunn BK, Akpa E. Biomarkers as surrogate endpoints in cancer trials. Semin Oncol Nurs 2012;28:99–108.

44. 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.

45. Food and Drug Administration. 22 case studies where phase 2 and phase 3 trials had divergent results [Internet]. Silver Spring (MD): Food and Drug Administration; 2017 [cited 2017 Jun 23]. Available from: https://www.fda.gov/downloads/AboutFDA/ReportsManuals-Forms/Reports/UCM535780.pdf.

46. Bikdeli B, Punnanithinont N, Akram Y, Lee I, Desai NR, Ross JS, Krumholz HM. Two decades of cardiovascular trials with primary surrogate endpoints: 1990–2011. J Am Heart Assoc 2017;6:pii: e005285.