

An Overview of the Complexities in Obesity: Limitations and Challenges

Sylvia R. Karasu

AMERICAN JOURNAL OF LIFESTYLE MEDICINE published online 4 October 2012

DOI: 10.1177/1559827612460832

The online version of this article can be found at:

<http://ajl.sagepub.com/content/early/2012/09/18/1559827612460832>

Published by:



<http://www.sagepublications.com>

Additional services and information for *American Journal of Lifestyle Medicine* can be found at:

Email Alerts: <http://ajl.sagepub.com/cgi/alerts>

Subscriptions: <http://ajl.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

>> [OnlineFirst Version of Record](#) - Oct 4, 2012

[What is This?](#)

Sylvia R. Karasu, MD

An Overview of the Complexities in Obesity: Limitations and Challenges

Abstract: *A recent survey, even one limited to human studies, found considerable “publication scatter” in that more than 250 different professional journals publish articles on obesity. Over the years, and particularly since the 1970s and 1980s when the so-called obesity epidemic began, there has been an explosion of clinical interest in a field that encompasses general medicine, pediatrics, surgery, psychiatry, and almost every subspecialty. And rightly so, since even by 2008, there were an estimated 1.46 billion adults worldwide who were overweight, and of these, 502 million were in the obese category, all of which translate into major public health consequences. Despite many highly publicized studies, why do we not have a greater understanding about obesity than we do? It is certainly not from a lack of trying. This article presents an overview of the limitations and challenges, that is, complexities, due to discrepant frameworks and diverse conceptualizations of obesity; potential flaws inherent in its clinical studies; and particularly, impediments due to difficulties in the measurement of body composition (and particularly adipose accumulation), food intake,*

and physical activity, as well as to notoriously inaccurate self-reporting by subjects. As a result, clinicians remain limited in issuing recommendations to their patients.

Keywords: obesity; weight control; exercise; diet; behavior

A recent survey, even one limited to human studies, found considerable “publication scatter” in that more than 250 different professional journals publish

there has been an explosion of interest in the field that encompasses general medicine, pediatrics, surgery, psychiatry, and almost every subspecialty. And rightly so since even by 2008, there were an estimated 1.46 billion adults worldwide who were overweight, and of these, 502 million in the obese category,³ all of which potentially translate into major health consequences. Despite highly publicized, well-conducted studies, such as those on diet and lifestyle,⁴ the importance of calories,⁵ a comparison

Over the years, and particularly since the 1970s and 1980s when the obesity “epidemic” began, there has been an explosion of interest in the field that encompasses general medicine, pediatrics, surgery, psychiatry, and almost every subspecialty.

articles on obesity --and of those fewer than 20% are found in the 3 leading obesity journals.¹ Over the years, and particularly since the 1970s and 1980s when the obesity “epidemic” began,²

of different diets,⁶ or the relationship of body mass index (BMI) to mortality,⁷ why do we not know more than we know? It is certainly not from a lack of trying.

DOI: 10.1177/1559827612460832. Manuscript received March 19, 2012; revised May 24, 2012; accepted June 15, 2012. From the Weill Cornell Medical College, New York, New York. Address correspondence to Sylvia R. Karasu, MD, 2 East 88th Street, Apt. 1B, New York, NY 10128; e-mail: sylkar@aol.com.

For reprints and permissions queries, please visit SAGE's Web site at <http://www.sagepub.com/journalsPermissions.nav>.

Copyright © 2012 The Author(s)

There are several reasons why the study of obesity lends itself to such complexity. Although there is no particular impediment that is specific to the study of obesity, clinicians may find it is the aggregate of uncontrolled and uncontrollable variables in all areas of clinical research that predisposes investigators to potential difficulties. Researchers called attention to some of these issues well over a decade ago,⁸ and though considerable work has been published since then, we are still facing many of these same problems that not only compromise the validity of a study but also make it difficult for clinicians to issue recommendations to their patients as they may relate to disease prevention.^{9,10} Borrowing from the language of social planning,¹¹ researchers have referred to the complexity of obesity as a “wicked problem” that includes no definitive formulation, complex (not binary) solutions, no immediate test of a solution, and so on.² We can even question whether research on diet and obesity in population-based studies is feasible at all.¹² This article will limit discussion to complexities due to different conceptual frameworks and diverse conceptualizations of obesity, potential flaws inherent in its clinical studies, and, particularly, impediments due to difficulties in the measurement of body composition (and particularly adipose tissue), food intake, and physical activity, as well as to notoriously inaccurate self-reporting by patients.

An Understanding of Clinical Bias

The major reason that a study has compromised validity is that it suffers from bias. In other words, validity is “the degree to which a study is free from bias.”¹³ More specifically, bias is any *systematic error* (as opposed to random error by chance) that can affect the design or implementation of a research study.¹⁴ Multiple biases may be operating simultaneously and are not, by any means, mutually exclusive.⁸ Only when researchers are content with a descriptive approach alone, without making recommendations or inferences, can they avoid bias analysis.¹⁵ Clinical bias can occur, of course, in all fields of scientific

research.^{14,16-20} For example, the testing of “non-prespecified hypotheses”—what is called the bias of “data dredging”¹⁶—led to detecting implausible and “spurious” associations linking astrological signs and health outcomes in a population of more than 10 million Canadians.²¹

Sackett,¹⁶ in his classic article, identified biases by the stage of research: conducting a literature review, specifying and selection of the sample population, execution of the experimental design, measurement of exposures and outcomes, analysis of the data, interpretation of the analysis, and publication of the results. He specifies, accordingly, about 55 different categories subsumed under these stages and defines bias as anything that “systematically deviates from the truth.”^{16,22} More recently, almost 75 different mechanisms have been delineated by which research bias can manifest itself.¹⁹ Biases do not necessarily have to compromise research studies if they are accounted for by statistical means,^{20,23-27} but can certainly compromise research if they are not appreciated and accounted for. A focus on bias has been called “a constant preoccupation among nutritional epidemiologists” and as such, researchers have to settle for “relative validity.”⁸

It was Cochrane in the conclusions to his classic treatise on health care written 40 years ago²⁸ who called attention to T. S. Eliot’s play in verse, *The Family Reunion*.²⁹ Cochrane urged clinicians to “abandon the pursuit” of the “margin of the impossible” and settle instead for what he called “reasonable probability.” Nowhere is this a more relevant and appropriate suggestion than in the study of obesity.

Complexities Due to Discrepant Frameworks of Obesity

One of the first complexities encountered in the study of obesity is in its conceptualization. Essentially, obesity is an excess accumulation of adipose (fat) tissue.³⁰ Simplistically, it results from an energy imbalance such that the amount of food taken in is greater than the amount of energy exerted (or calories expended). As such, it reflects the first

law of thermodynamics, the conservation of energy.^{31,32} Other than excess adipose tissue, though, there are no other “inevitable” or characteristic signs or symptoms present in everyone with obesity.^{33,34} In fact, this excess adipose tissue does not even accumulate in the same place in everyone with obesity: for some, it accumulates predominantly in the abdominal area and most dangerously around internal organs (ie, the so-called android distribution, because it is a pattern more common in men), whereas for others it accumulates predominantly more superficially and below the waist (ie, the so-called gynoid distribution, because it is a pattern more common in women).³⁵ Researchers cannot even agree on what really causes this accumulation of excess fat and even whether obesity is a disorder³⁶ or a disease at all.^{33,37} Years ago, it was even referred to as a “psychosomatic disorder,” though “multicausal” in origin.³⁸ More recently, it has been considered an impulse disorder³⁹ or even a “psychological disorder involving impulse control” and “reinforcement pathology.”⁴⁰

Obesity is, however, recognized as a disease, under the “Endocrine Nutritional, and Metabolic Diseases” category by the World Health Organization (WHO) in its *International Classification of Diseases (ICD 10)*.⁴¹ The concept of “obesity” as a disease is “controversial,” although obesity “meets all the criteria of a medical disease, including a known etiology, recognized signs and symptoms, and a range of structural and functional changes that culminate in pathological consequences.”⁴² Obesity, in fact, has variously been called a brain disease,^{43,44} a metabolic disease,⁴⁵ a genetic disease,⁴⁶⁻⁴⁸ a disease of inflammation,⁴⁹ a neurochemical disease,³¹ and even an infectious disease caused by a virus.⁵⁰ It has also been called a matter of “energy balance dynamics.”⁵¹ On the other hand, from an evolutionary perspective, some believe obesity is an example of “inappropriate adaptation”⁵² or even the “result of people responding normally to the obesogenic environments they find themselves in.”² Though there are “layered determinants” of obesity, the actual “physiology of energy balance is proxi-

Table 1.

Complexities Due to Discrepant Frameworks of Obesity

How Obesity Has Been Conceptualized
Energy model: energy imbalance determined by behaviors and environments (first law of thermodynamics: calories in, calories expended)
Disease or disorder model:
Multicausal psychosomatic disorder
Psychological disorder involving impulse control and reinforcement pathology Brain disease Metabolic disease Genetic disease Disease of inflammation Neurochemical disease Infectious disease (eg, viral transmission)
Evolutionary model: inappropriate adaptation to toxic environment
Body diversity model (National Association to Advance Fat Acceptance): analogous to racial, ethnic, or sexual diversity (ie, obesity <i>not</i> a disease or disorder to be treated)
Multifactorial model: disorder resulting from complex interaction of genetic, perinatal, gender, developmental, environmental, neuroendocrinological, psychosocial, and behavioral factors

mally determined by behaviors and distally by environments.²² Those in the National Association to Advance Fat Acceptance believe obesity is a form of “body diversity that should be tolerated and respected,” analogous to diversity of ethnicity, race, or sexual preference.⁵³ Whatever model we use to define obesity, it is likely that the regulation of fat accumulation is extraordinarily complex, multifactorial, and determined by genetic, gender, perinatal, developmental, dietary, environmental, neural, and psychosocial factors⁴³ whereby “genetics loads the gun while the environment pulls the trigger” (Table 1).³¹ Furthermore, it is also likely we are dealing with the “obesities” rather than “obesity.”

Complexities Due to Clinical Study Design

The Sample Population

One of the major issues in the study of obesity, as in all research, is the choice of a sample population. A wrong sample

size can affect results: samples can be too large and prove anything, whereas they can be too small and prove nothing.¹⁶ The gold standard of research, of course, is the randomized controlled study.^{54,55} Neither the cohort study, in which 2 groups are identified and followed “forward in time,” nor the case–control study, in which cases are gathered, compared with a control group, and studied retrospectively (“direction of inquiry backward in time”), is as valid and free from bias as is the randomized controlled study.⁵⁶ A case series, with no control group, is the least free from bias and “prone to overinterpretation.”⁵⁶ Researchers in the field of obesity are confronted with several options: they can conduct large and sweeping community-based epidemiological studies with thousands of subjects (but fairly limited control over their subjects’ behavior) or they can draw their sample from smaller, more specific clinical populations. The most controlled of all human obesity studies are those conducted on an inpatient metabolic unit, but while researchers gain control, they for-

feit exposure to a real-life situation and must often have studies of short duration. Of course, when the sample population is limited (eg, specific race, gender, ethnic group, or age), not only is the pool of subjects limited but the generalizability of the study may also be limited. Obesity studies commonly focus on specific populations (eg, Caucasians, Europeans, and postmenopausal women).⁵⁷ For example, the original BMI guidelines for obesity were validated among a population of those of European descent and hence revisions might be warranted for those of non-European descent, such as those from China, South Asia, and so on.⁵⁸

Researchers, though, often opt for restricting the sample population because of the possibility of confounding. Confounding is essentially a “confusion of effects” such that the apparent effect of study becomes “mixed with” or “distorted” because of some extraneous factor or factors associated with the outcome.⁵⁹ A confounder is a “risk factor” for the outcome, but it is not affected by either the exposure or the outcome.^{59,60} Failure to account

for confounding can lead to either overestimation or underestimation of any effect, and the degree of confounding is more important than whether it is there at all.⁵⁹ Confounding describes an association that is true but potentially misleading, whereas bias creates an association that is not true.²⁰ In obesity studies, one of the most important confounders is *smoking*, but age, sex, and race can also be confounders.^{59,61}

Not controlling for smoking, for example, can have serious consequences for studies in obesity. Taking a smoking history, though, is much more complicated than it first appears. Misclassifications can result when researchers use a binary classification, such as “smoker or non-smoker.” For example, not only is it important to inquire about general smoking history but also about the duration and intensity of the smoking exposure (eg, when smoking began; age at cessation; brand of tobacco; whether cigarettes, cigar, pipe, even filtered or unfiltered; how much inhaling; etc). Even the category of “no smoking” may require further clarification.^{62,63} However, when smoking is not accounted for accurately, for example, studies particularly involving obesity and its effects on mortality may be severely compromised, leading to the so-called J-curve of mortality; the wrong conclusions can be drawn, namely, that increased mortality is not only in the obese but also among those who are considered normal weight or thin.^{7,64-67}

“Reverse causation,” also called “effect cause,” can occur in obesity studies.^{64,68} Here, an underlying (and maybe even unrecognized) disease is responsible for a low body weight such as in chronic “wasting diseases” (eg, end-stage kidney disease, congestive heart failure, AIDS, and many end-stage cancers). In these cases, increased mortality may give the false impression that obesity may even have a survival advantage.⁶⁴ For example, it has been suggested that obesity and its metabolic abnormalities, from an evolutionary perspective, may have been advantageous against the wasting, devastating disease, tuberculosis.⁶⁹ Obesity researchers cannot even agree on the definition of reverse causality and cannot rule out the possibility that

“bias due to preexisting illness may affect weight-mortality studies.”⁶⁸

Another bias typical of obesity studies is the “nonresponder bias”: in general, those who agree to participate in a study may, in fact, be different from those who do not.^{8,12,16,25,59} Some believe that “reduced participation” or “suboptimal control samples” are the “most common problem” when conducting population-based studies.²⁵ For example, biased results occurred when only 42% of a randomly selected group of more than 3600 Swedish subjects participated in research on cardiovascular risk,⁷⁰ and “the main limitation” of one study on BMI and its relationship to psychiatric disorders was a response rate of only 57%.⁷¹ Here, inferences were made that BMI and common mental problems were the same in responders and nonresponders, “but it is not possible to test the validity of this assumption.”⁷¹

Furthermore, those who volunteer for a study may also be different from the general population, that is, “volunteer bias.”¹⁶ For example, it has been reported that those who volunteer in obesity studies are less likely to have the metabolic syndrome, a serious cluster of symptoms often seen in obesity.¹² One of the most important and long-term studies in obesity research, for example, is the National Weight Control Registry, a study begun in the early 1990s to investigate successful dieters. Now following thousands of individuals, it began with an original group of 629 women and 155 men, all of whom were self-selected volunteers recruited through local and national media advertisements, mailings to weight loss programs, and so on, and not subject to any randomization and not at all even typical of the US population.^{72,73}

A very high attrition (dropout) rate is characteristic of many obesity studies, particularly those involving weight management or specific dietary changes (eg, comparison of low carbohydrate with low fat), not uncommonly as high or higher than 50%, even after only 1 year of follow-up.^{26,74,75} A very high dropout rate, for example, was noted in a study that compared diets by Atkins, Ornish, Weight Watchers, and Zone. This kind of bias cannot be easily corrected, even by statistical calculations.⁶⁴ Subjects may

withdraw from studies for a multitude of reasons, including a lack of motivation,²⁷ but sometimes for reasons completely unknown.¹⁶ Even in a study where the dropout rate was fairly low (13%), researchers found differences between those who dropped out and those who continued to participate over 2 years.⁷⁶ Obtaining high response rates with high-quality data retrieval has been called “the single largest obstacle to high-quality epidemiological research,” and the loss of follow-up of recruited participants is much more significant than the loss of a specific population initially because the rate of loss may be reflective of both disease and exposure.⁷⁷ When only 60% of subjects can be traced, studies are looked at skeptically, and even when 70% or 80% are traced, those numbers can still be too low to assure against bias if the loss to follow-up might be associated with both exposure and disease.⁶³

In obesity studies, “membership bias,” in which those who belong to a certain profession or engage in certain activity or even who are employed may be healthier or even more health conscious than the general population, can distort results. One of the most important studies in obesity, for example, is the longitudinal Nurses’ Health Study.⁷⁸ To what extent being in the health professions affects results is open to question. Along those lines, those subjects willing to engage in research may exhibit “clustering,” whereby certain habits, particularly about health—whether positive or negative—cluster together so that it is difficult to assess correlations. For example, smokers were more apt to eat red meat, engage in less physical exercise, and drink more sugared soft drinks.⁷⁹

Data Collection

Missing information, such as on questionnaires or in records, is also a common problem.⁸⁰ Data may be missing because it is normal, never measured, negative, or even measured but never recorded.¹⁶ Furthermore, participants who realize they are controls can choose to change their behavior (eg, eating more healthily or exercising) so that there is the possibility of the “bias of

contamination.⁸¹ A more general concern, however, is how often and when to observe. Cross-sectional observation versus longitudinal observation can lead to significantly different results. Weight fluctuations, even in the course of a day, are extraordinarily common. This can be particularly significant, for example, in studies of weight fluctuations (eg, yo-yo dieting) in which there are repeated patterns of weight gain and weight loss over time that may not be accurately assessed with the limited data that cross-sectional, one-time observation provides. Furthermore, subjects are often asked to remember patterns of weight loss that may have occurred many years earlier and subject to memory distortions.⁸²

Longitudinal observation, on the other hand, also has its complexities, especially when subjects do not maintain experimental protocol. Many obesity studies involve long-term follow-up over months or even years such that noncompliance with the experimental design can become a problem, particularly over long-term follow-up. Over the 8 years of follow-up in the Women's Health Initiative Study, for example, the group randomized to a low-fat level of 20% could not maintain that level and hence biased findings (toward the null) regarding cholesterol and triglyceride levels.⁶⁴ In fact, it is never really possible in community studies to measure compliance with a prescribed diet, and this has been called "the fundamental flaw in obesity research."⁸³ Though random selection tends to control confounding in a study, when there is considerable nonadherence or noncompliance with the treatment protocol, even in large randomized studies, considerable nonrandom confounding can result.^{59,60}

One means of controlling for noncompliance is conducting a study in a laboratory setting rather than in a free-living environment. The laboratory setting has been criticized for not providing for "real meals, real people, real eating situations"⁸⁴ and has been called "artificial," particularly when the cost of food, short-term compensation in food intake over several days, and timing of eating, including diurnal rhythms, seasonal effects, and even differences

between daily and weekend eating patterns, are often overlooked in a lab setting.⁸⁵ Furthermore, the effects of alcohol on food consumption as well as the presence of other people are often not considered, and many of the factors, such as environmental, psychological, and social, that influence food intake are lost in the clinical research-controlled lab environment.⁸⁵

Observation in a laboratory setting, though, may mitigate against other forms of bias, such as the "obsequious bias," when subjects tell researchers what they think the researchers want to hear the "unacceptability bias" in which subjects may be embarrassed to admit to certain behaviors.¹⁶ In a free-living environment, these biases are seen with certain frequency, particularly when obesity studies depend on subjects' self-reporting⁸² (see below). Observation in a lab setting, though, where subjects know they are being observed (the so-called "Hawthorne effect"¹³), can also affect behavior. Even just having subjects see how much they are eating (eg, leaving dirty plates on the table) can affect how much they eat.⁸⁶

Meta-Analysis

Meta-analysis is a pooling or systematic review of multiple studies. The purpose of a meta-analysis is to identify patterns among study results and sources of disagreement among those results.⁸⁷ Because meta-analysis deals with considerable heterogeneity in design and statistical methods, even definitions of the problem can differ so much among studies so that an actual meta-analysis becomes impossible and only a qualitative approach is possible. For example, inconsistencies on the potential dangers of weight cycling were seen among studies because there was not even a consistent definition of what constituted a weight cycle,⁸⁸ and meta-analysis of randomized controlled trials to assess the effects of calcium supplementation on weight found so many discrepancies among the studies that they could not even conduct a proper meta-analysis and had to settle for a "narrative review"⁸⁹ (see Table 2).

Complexities Due to Measurement

Nobel Laureate Sir Henry Dale⁹⁰ said, "All true measurement is essentially comparative," and whenever there is measurement, there is always the possibility that there will be error—either by random chance or systematically.⁵⁹ One of the most essential impediments in obesity research is, in fact, measurement bias—whether of body composition, food and caloric intake, and/or physical activity and specifically exercise. This is sometimes categorized as "information bias."^{19,20,59}

Measurements of Adipose Tissue

Currently, we use BMI to define categories of obesity.^{91,92} Obesity is defined arbitrarily as a "threshold" and as such "a relatively small increase in average weight has had a disproportionate effect" on the actual incidence of obesity.⁹³ It is not clear how BMI became the general standard to measure obesity. Adolphe Quetelet, a Belgian mathematician and astronomer and the father of modern statistics, in the middle of the 19th century, established this ratio of weight in kilograms to height in meters squared.⁹⁴ BMI, though, did not become popular as a measure of obesity until recent years. Back in the early 1970s, though, Ancel Keys and his colleagues noted the "need for an index of relative body weight" and credited Quetelet for this ratio that they called for the first time "body mass index."⁹⁵ Earlier in the 20th century, when scales became available for home use, insurance companies gathered data on weight and its relationship to mortality.⁹⁶ These early measurements were highly inaccurate such that people were weighed with shoes and clothing and without any standardization. Even the categories of "small, medium, and large" build were determined by the subjective judgment of an examiner without any corroborating data.⁹⁶ Before the use of BMI classifications, the measurements of obesity were much less precise. For example, categories might include "overweight" and "percent overweight."⁹⁶ As a result,

Table 2.

Complexities Due to Clinical Study Design

<p>Sample populations: randomized controlled trials vs case studies</p> <ul style="list-style-type: none"> Large, community-based epidemiological studies (less control) Small, specific populations (less generalizable) Confounding and “reverse causation” (eg, smoking, chronic disease)
<p>Data collection</p> <ul style="list-style-type: none"> Missing information (eg, never measured, “normal,” never recorded) Cross-sectional vs longitudinal observation (eg, weight fluctuations even day-to-day and over time; memory distortions) Free-living environments vs laboratory setting (eg, natural vs artificial; Hawthorne effect: being observed changes behavior)
<p>Common biases in clinical obesity research</p> <ul style="list-style-type: none"> Nonresponder bias (those who don't participate may be different) Volunteer bias (those who volunteer may be different, eg, healthier) High attrition (dropout rate can be greater than 50%) Membership bias (those who belong to certain group or engage in certain activity may be different (eg, more health conscious) Clustering of behaviors (healthy or unhealthy), that is, difficult to assess individual correlations when behaviors seen together Noncompliance (nonadherence to protocol, especially over long-term) Contamination bias (eg, controls change behavior and adopt protocol being studied, especially if thought beneficial) Obsequiousness bias (subjects tell researchers what they think researchers want to hear) Unacceptability bias (subjects embarrassed by behavior and misrepresent what they actually do) Diagnostic vague bias (definitions of obesity change over time, complicating comparisons with previous or future studies)
<p>Meta-analysis</p> <ul style="list-style-type: none"> Inconsistent definitions across studies prevent pooling of data Heterogeneity among studies in design, population, and so on (eg, meta-analysis becomes impossible, and researchers must settle for qualitative “narrative review”)

as the definition of obesity has become more standardized (though still arbitrary and subject to potential change in the future), comparing older studies with more recent ones or even future ones can lead to what is called “diagnostic vague bias” in which the same condition can receive different diagnostic labels over time.¹⁶ Though BMI use began earlier,⁹⁷ it was only in the late 1990s that there were the guidelines established by the US Department of Health and Human Services and the WHO to measure overweight and obesity by the BMI categories that are in use today.^{98,99} The WHO had “convened a Consultation on Obesity” because of its concern about the comorbidities as well as the “social bias, prejudice, and discrimination” to which the obese are often subjected. Their conclusion was that BMI was a “coherent system” that should be adopted internationally.⁹⁹

Over the years since, use of BMI as a standard, however, has caused considerable controversy itself. Early on, researchers began questioning the new guidelines and felt that lowering the overweight threshold “stigmatized” too many people and was not justified on the basis of data on mortality.¹⁰⁰ Even today, researchers describe the classification of obesity as a BMI of 30 kg/m² or more as having “a certain degree of arbitrariness,” without genetic markers.⁴⁷ Furthermore, because BMI measures not only degrees of fatness but also muscle and skeletal mass, it may be inaccurate in those who are particularly muscular or in those who have lost muscle (eg, sarcopenia) typically in old age (ie, may underestimate BMI),^{101,102} and there is a need to make “adjustments” when calculating BMI not only in athletes¹⁰³ but also in particularly tall or short people, as height is part of the equation, and in children

younger than 16 years.¹⁰⁴ The practice of using BMI as a measurement of obesity has been called “obsolete,” resulting in a considerable “underestimation of the grave consequences of the obesity epidemic,”¹⁰⁵ as well as a “deeply flawed measure of fatness,”¹⁰⁵ a “surrogate measure” providing “misleading information,”³⁰ and only a “proxy” measure for body fat.¹⁰⁶ For example, when BMI was compared with more accurate measurements of total body fat, such as the use of deuterium water, BMI “was a poor surrogate for body fatness for both males and females.”¹⁰⁷ And because BMI is only an indirect measure of obesity, it did not discriminate between body fat and lean muscle in patients with coronary artery disease¹⁰⁸ and should not be used alone but rather only with other measurements such as a direct assessment of body composition and measurement of waist circumference.¹⁰⁹ Not all studies, though,

have been critical of the use of BMI.^{65,110} For example, the Prospective Studies Collaboration called it a “reasonably good measure of general adiposity.”⁶⁵

One of the most problematic issues with use of BMI is that many studies employ use of subjects’ self-reports to calculate BMIs. Though there is some question regarding the accuracy of self-reports, most researchers suggest that people tend to underreport weight and overreport height.¹¹¹⁻¹¹³ Although self-reports are easier to collect, they “should not be used exclusively as an obesity surveillance tool.”¹¹⁴ Likewise, self-reports are more likely to be “underestimations” when people round off their measurements or do not even know their height, particularly as height may change with age,¹¹⁵ or even when they have certain diseases.¹¹⁶ Criticisms of self-reports of BMI as a measurement of obesity have been worldwide.^{82,113,117-126} Studies from Japan,¹¹⁷ the Netherlands,¹²⁰ Sweden,¹²¹ Australia,¹²⁴ France,¹²⁵ Canada,¹¹⁴ Greece,¹¹⁶ and Spain¹¹³ have reported on the inaccuracy of self-reports, the need for caution in the interpretation of findings, and even a need to make adjustments for discrepancies. In the United States, both the sensitivity and specificity of BMI “have been shown to be poor” and demonstrate “various deficiencies as a measure of obesity” when BMI is obtained through self-reports¹⁰² such that BMI by self-report is not interchangeable with BMI by actual measurement.¹²⁶ In fact, data from 2 waves of the NHANES (I and II; National Health and Nutrition Examination Survey) in a subgroup of healthy subjects who have never smoked concluded that bias and inconsistency produced by self-reported BMI data may actually account for discrepancies in published data regarding mortality and its relationship to BMI and “even small changes in BMI distribution in future studies could have dramatic effects on misclassification rates.”¹²⁶ Another difficulty is that measurement conditions, such as clothing worn, equipment used, instructions given, and even the time of measurement, are rarely, if ever, specified.⁸²

Of course, BMI has not been the only means of measuring obesity. Clinicians

have used calipers for skin fold thickness in various areas of the body, such as arm, scapula, back, hip, and so on. Though sometimes seen as a “comparatively simple and reasonably accurate assessment of body fatness,”¹²⁷ most believe it is the most inaccurate of all ways to measure body fat and may not only vary from examiner to examiner but on different examinations with the same examiner.¹²⁸ Researchers have also used measurement of both waist circumference and waist-to-hip ratios, but these, too, may depend on the skill of the examiner. For example, it is often difficult to locate the so-called natural waist—or smallest circumference—on an obese person so measurement can be taken at the level of the umbilicus. Although intrarater reliability was “acceptable,” measurement error is more likely to occur in the overweight and obese.¹²⁹ Some have suggested that waist circumference may be a useful adjunct to BMI errors from self-reports.¹³⁰ Others have found the use of waist circumference led to misclassification.¹³¹ A meta-analysis of more than 82 000 people in the United Kingdom found that use of self-reported BMI led to inconsistent results in relating mortality to the accumulation of body fat, but waist-to-hip measurements “showed the strongest association with mortality from cardiovascular disease,” as compared with either waist circumference or BMI.¹³² As a result of missing data, though, 25% of the original sample had to be eliminated, and results might not apply to other ethnic, more diverse samples.¹³² There is also controversy, though, over the use of the waist-to-hip measurement, for example, the waist-to-hip ratio has been described as “a superior measure of central obesity with low measurement error,”¹³³ but its use has been questioned, particularly since hip circumference measures both muscle, fat, and bone.¹³⁴

The most accurate (and reproducible) way of measuring body composition is by dual energy X-ray absorption, which is based on the fact that X-ray beams pass through bone, fat, and muscle differently.^{64,134} Though its use is limited because it is not portable (and hence impractical for large epidemiological

studies) and cannot be used on pregnant women, it uses the same machine employed for assessing bone density. As a result, those people being evaluated for osteoporosis can easily request a simultaneous evaluation of their body composition.⁶⁴ The so-called gold standard of measuring body composition, though, is underwater weighing, called densitometry, which uses the principle that fat is less dense than water.⁶⁴ Clearly, this is an unwieldy technique that cannot be used in large-scale studies or easily with children or the elderly. Finally, both computed tomography and magnetic resonance imaging can both measure body composition but are expensive and obviously require special equipment, and computed tomography exposes subjects to radiation.⁶⁴ It was British cardiologist Sir Thomas Lewis¹³⁵ who said that “there is a manifest tendency . . . for the medical profession to exaggerate the accuracy of its subjective methods of examination.”¹³⁵ Clearly, there is “no single measurement method that is error-free.”¹³⁶

Measurement of Food Intake and Caloric Consumption

As noted, our inability to measure accurately what people are really eating is the “fundamental flaw” in research in the field of obesity.⁸³ We are left, as a result, with “partly inaccurate information” and failing “in a fundamental task of science, accurately measuring the independent variable.”⁸³ This becomes so much more problematic in obesity research, as noted, because of social desirability: people can be embarrassed by their behaviors, especially about food (and alcohol), and misrepresent their intake. (ie, “unacceptability bias”¹⁶). Furthermore, much of this information relies on a subject’s recall, which can be notoriously inaccurate, even with the best of intentions. Though sometimes related to social embarrassment, underreporting of food intake may also be reflective of a poor memory or even a genuine lack of awareness regarding specific food items and actual amounts consumed.¹² Wansink⁸⁶ has described the phenomenon of “portion distortion,” seen not only in obese subjects but also in those of normal weight. Furthermore, studies of diet

are often limited by the use of “disappearance data” that are only indirectly limited to intake,¹³⁷ and the complexity of the human diet represents a “daunting challenge” to those studying a connection between diet and disease.¹³⁷ Dietary exposures can rarely be characterized as present or absent: individuals rarely make clear changes in their diet at identifiable periods of time. More typically, patterns evolve over years, and even though diets of individuals are often consistent over time, they are usually characterized by marked variation from day to day.¹³⁷

Food intake, though, can be measured by several means, including 24-hour recall, the most widely used dietary assessment method (and the basis for national nutrition surveys), food diaries for varying periods of time (often 3-7 days), and food-frequency questionnaires.^{64,137} With food diaries, subjects must be highly motivated to keep these records, but this effort may increase their awareness of (and hence alterations in) food intake. Information retrieval can also be by telephone or in-person interview. As researchers in all fields appreciate, use of the telephone has made randomization more complex, as cell phones (ie, area codes), voicemail, and other technological advancements do not necessarily identify a subject’s location.¹⁰² Both the food-frequency questionnaires and the 24-hour recall depend on memory, leading food writer Michael Pollan to wonder whether Marcel Proust could remember with precision all that he had eaten.¹³⁸ Furthermore, complications in dietary research can stem from the inherent biological complexity of nutrient–nutrient interactions, and since diet is often associated with health consciousness in general, the diets of those who participate may differ substantially from those who do not participate and hence bias samples.¹³⁷ Another problem with dietary studies is that the time between any change in diet and any expected change in incidence of disease is typically uncertain: even if an effect is not found, it may not be possible to rule out that follow-up was not long enough.¹³⁷ And, as noted, compliance often wanes over a long trial, particularly if treatment involves a real change in food intake, and sometimes the control

group chooses to adopt the prescribed diet of the treatment group, particularly if it is thought to be of benefit¹³⁷ (“bias of contamination”⁸¹).

The obesity literature is replete with references to inaccuracy in reporting of diet not only in obese subjects but particularly in the obese and often correlated with the degree of obesity as measured by BMI.^{8,32,139-147} It is a “major challenge” to link diet with health when subjects are “implausible reporters,” and even using statistical means to account for underreporting cannot determine “true validity.”²³ A review of both prospective and retrospective studies yielded underreporting discrepancies in food intake when measured against doubly labeled water.³² Underreporting has been linked not only with greater BMI but also with greater body dissatisfaction and lower income.¹⁴⁸ In general, the failure of obese people to lose weight while on a specified diet (what the subjects called “diet resistance”) may reflect both underreporting of caloric intake and overreporting of physical exercise rather than on any metabolic differences between the obese and the nonobese.¹³⁹ Furthermore, subjects, particularly the obese, can both underreport and undereat during the period of observation,³² often by 20%.^{32,85} The “eye–mouth gap” is the discrepancy between the food intake people believe they are eating and what they are actually eating.¹⁴⁹ Doubly labeled water or 24-hour urinary collection for nitrogen excretion can assess protein specifically in an attempt to validate dietary intake, but these methods are cumbersome and expensive and not suitable for large epidemiological studies.⁸ When there is overreporting of protein, it is suggestive that there is an underreporting of fat and carbohydrate, but there are no means of assessing specifically what nonprotein sources (eg, fat, carbohydrates, alcohol) are underreported by subjects.⁸ Underreporting leads to a “dual bias”—general underreporting of total caloric intake and underreporting for specific foods.¹² Furthermore, intensified public health campaigns regarding lowering fat and sugar intake may have led over time to even more inaccurate underreporting, even in those

who were not obese.¹⁰ Underreporting was also found in up to 45% of pregnant women, and those who tend to underreport tend to be less compliant in general with dietary recommendations for pregnant women.¹⁵⁰ Underreporting can also occur particularly in obese patients who are depressed.¹⁵¹

Measurement of Physical Activity

Perhaps even more difficult than measuring caloric intake or actual percentage of body fat is the measurement of physical activity. Our bodies burn calories through the digestion, absorption, and storage of food (ie, its thermogenic effect); through our resting metabolic rate; and any/all physical activity, the most variable of the 3 components.⁷³ Caloric expenditure by physical activity can vary by 3-fold from the extremely active to those who are sedentary.¹⁵²

There are 2 kinds of physical activity: *nonexercise physical activity thermogenesis*, that is, spontaneous movement of the body (including posture, fidgeting, sitting, standing, or even chewing gum), and *exercise*, that is, physical activity that is “purposeful” and planned specifically for maintaining health or fitness or for burning calories. Exercise can be measured by its intensity, its frequency, and its duration.⁷³ Many studies, though, particularly large epidemiological surveys, do not measure precisely and employ inaccurate self-reports that are sometimes merely estimates. The “Compendium of Physical Activities” lists thousands of activities in categories such as sports, occupation, home repair, self-care, and so on, all of which are given a value compared with sitting comfortably.^{153,154} No 2 people, though, perform an activity in exactly the same way so that these values are only approximations. In some studies, attempts to measure physical activity more accurately can be done using an Actigraph, an instrument that measures the intensity of movement.¹⁵⁵ Only in a lab setting, though, can we obtain accurate measures of actual physical activity. Most studies just report that their subjects did “moderate” exercise without a precise definition. There are even further difficulties in calculating

Table 3.

Complexities Due to Measurement

Measurement of body composition (ie, adipose tissue)

Body mass index: arbitrary, inaccurate, measures more than fat
 Skin calipers, waist circumference, waist-to-hip ratio: varies from examiner to examiner and even examination to examination
 Underwater weighing, DXA, CT, and MRI: unwieldy and restricted use; exposure to radiation with CT and DXA
 Self-reports: notoriously inaccurate; overestimation of height and underestimation of weight in most people
 Measurement conditions rarely specified (eg, clothing worn, time of measurement, equipment)

Measurement of food intake and caloric consumption

Food diaries, food-frequency questionnaires, 24-hour recall: poor information retrieval: “implausible reporting”; increasing awareness of diet leads to alterations
 Self-reports are inaccurate because of poor memory, embarrassment, and genuine lack of knowledge
 “Eye–mouth gap”: underreporting of total caloric intake and/or of specific foods in most, but particularly in obese
 Diet changes evolve over time: correlating incidence of disease with changes in diet difficult; how long is long enough to observe?
 Biological complexities of “nutrient–nutrient” interactions: foods eaten together in complex combinations
 Considerable variation in diet from day-to-day (eg, weekends vs week days, etc)

Measurement of physical activity

Inaccurate measurements due to crude instruments (eg, pedometer)
 Inability to capture “incidental” movements other than in respiratory chamber
 Considerable variation among people when performing activities so only approximations
 Inaccurate measurements due to self-report: overreport duration and intensity of exercise
 Inconsistencies in judging intensity (eg, moderate vs intense)

Abbreviations: DXA, dual-energy X-ray; CT, computed tomography; MRI, magnetic resonance imaging.

caloric expenditure during exercise: one has to consider not only the number of calories expended during an exercise but also consider (and subtract) the number of calories that might have been expended just by standing or sitting.¹⁵⁶

Exercise research suffers from other methodological problems, such as not being randomized controlled studies but merely observational with poor follow-up.¹⁵⁷ Studies of exercise and its role in psychiatric disorders found that the exact nature of the exercise recommended was not even specified, nor its intensity or sometimes even its dropout rate.¹⁵⁸ Intensity of exercise, for example, was also not measured (and considered a limitation)¹⁵⁸ in a study where subjects were asked to take a “brisk” walk for 30 minutes a day,¹⁵⁹ and the study population was too homogeneous to make generalizations to other populations and did not measure adherence over time.¹⁵⁹

One measure of physical activity is the pedometer that, once set to a person’s stride length, calculates how many steps

a person has taken within a day.^{154,159,160} A pedometer can be a “useful tool” for tabulating the amount of walking because it can provide immediate feedback,¹⁵⁹ but there is conflicting evidence regarding the accuracy of pedometers in actually capturing physical activity. When, for example, the pedometer was compared with measurements conducted in a respiratory chamber, it was “at best only a crude predictor” of physical activity, and because it does not record the duration or the intensity of the steps taken, it does not provide accurate enough information for calculating energy expenditure.¹⁶¹ Furthermore, pedometers cannot even accurately measure “stride length” as stride changes depending on the speed of walking¹⁶¹ and may be limited when comparing samples that are based on varying recommendations for physical activity.¹⁶²

As in studies involving food intake and body measurements, those involving physical activity are also subject to inaccuracy with self-report, particularly with the use of questionnaires.¹⁶³ As in

questionnaires tabulating food intake, even the order of the questions can significantly affect responses. “Subjective interpretations” involving the intensity of exercise may contribute to errors in classifying the intensity of an exercise and self-reports involving physical activity tend to overestimate physical activity levels when compared with “objective monitoring” as, for example, by accelerometry.¹⁶⁴ Likewise, a systematic review comparing direct measurements of physical activity with self-report data found considerable inaccuracies in self-reports, with both higher and lower levels reported, although self-report data can give information on an individual’s “perception” of an activity’s difficulty but not in “capturing all levels of activity.”¹⁶⁵ Direct measurement, though, may fail to capture “incidental daily movements” or even activities like swimming so that there is a need for “valid, accurate, and reliable measures” to assess physical activity, particularly as it relates to possible clinical interventions¹⁶⁵ (see Table 3).

Conclusion

Practitioners in all disciplines are familiar with obstacles that confront them and limit their expertise. The field of obesity unfortunately lends itself particularly well to the compounding of these difficulties. Although there is no impediment that is specific to obesity, clinicians may find themselves inadvertently thwarted by the aggregate of uncontrolled and uncontrollable variables that predispose them to potential and sometimes insurmountable challenges. These challenges include complexities due to discrepant frameworks and diverse conceptualizations of obesity, potential flaws inherent in its clinical studies, and particularly to problems in the measurement of body composition (and specifically adipose accumulation), food intake, and physical activity, as well as to notoriously inaccurate and misleading self-reporting by subjects. As a result, those who attempt to study and treat obesity are constantly on T. S. Eliot's "margin of the impossible."²⁹ Unfortunately, there are no straightforward solutions to these challenges, and clinicians often remain limited and even tentative in the recommendations they can offer their patients. In fact, given all these difficulties, we can marvel that we know as much as we do. Both researchers and clinicians alike, though, while striving for success, must remain cognizant of, and sensitive to, not only their patients' not so infrequent failures but also to their own as well. **AJLM**

References

- Baier LA, Wilcznski NL, Haynes RB. Tackling the growth of the obesity literature: obesity evidence spreads across many journals. *Int J Obes (Lond)*. 2010;34:1526-1530.
- Swinburn BA, Sacks G, Hall KD, et al. The global obesity pandemic: shaped by global drivers and local environments. *Lancet*. 2011;378:804-814.
- Finucane MM, Stevens GA, Cowan MJ, et al; Global Burden of Metabolic Risk Factors of Chronic Diseases Collaborating Group (Body Mass Index). National, regional, and global trends in body-mass index since 1980: systematic analysis of health examination surveys and epidemiological studies of 960 country-years and 9.1 million participants. *Lancet*. 2011;377:557-567.
- Mozaffarian D, Hao T, Rimm EB, Willett WC, Hu FB. Changes in diet and lifestyle and long-term weight gain in women and men. *N Engl J Med*. 2011;364:2392-2404.
- Sacks FM, Bray GA, Carey VJ, et al. Comparison of weight-loss diets with different compositions of fat, protein, and carbohydrates. *N Engl J Med*. 2009;360:859-873.
- Gardner CD, Kiazand A, Alhassan S, et al. Comparison of the Atkins, Zone, Ornish, and LEARN diets for change in weight and related risk factors among overweight premenopausal women: the A to Z Weight Loss Study; a randomized trial. *JAMA*. 2007;297:969-977.
- Berrington de Gonzalez A, Hartge P, Cerhan JR, et al. Body-mass index and mortality among 1.46 million white adults. *N Engl J Med*. 2010;363:2211-2219.
- Lissner L, Heitmann BL, Lindroos AK. Measuring intake in free-living human subjects: a question of bias. *Proc Nutr Soc*. 1998;57:333-339.
- Heitmann BL, Frederiksen P. Imprecise methods may both obscure and aggravate a relation between fat and breast cancer. *Eur J Clin Nutr*. 2007;61:925-927.
- Heitmann BL, Lissner L, Osler M. Do we eat less fat, or just report so? *Int J Obes Relat Metab Disord*. 2000;24:435-442.
- Rittel HWJ, Webber MM. Dilemmas in a general theory of planning. *Policy Sci*. 1973;4:155-169.
- Lissner L, Heitmann BL, Bengtsson C. Population studies of diet and obesity. *Br J Nutr*. 2000;83(suppl 1):S21-S24.
- Porta M, ed. *A Dictionary of Epidemiology*. 5th ed. Oxford, England: Oxford University Press, 2008; 252.
- Tripepi G, Jager KJ, Dekker FW, Wanner C, Zoccali C. Bias in clinical research. *Kidney Int*. 2008;73:148-153.
- Greenland S, Lash TL. Bias analysis. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2008:345-380.
- Sackett DL. Bias in analytic research. *J Chron Dis*. 1979;32(1-2):51-63.
- Cofield SS, Corona RV, Allison DB. Use of causal language in observational studies of obesity and nutrition. *Obes Facts*. 2010;3:353-356.
- Hsu JL, Banerjee D, Kushner WG. Understanding and identifying bias and confounding in the medical literature. *South Med J*. 2008;101:1240-1245.
- Delgado-Rodriguez M, Llorca J. Bias. *J Epidemiol Community Health*. 2004;58:635-641.
- Adebiyi AO. Bias: a review of current understanding. *Afr J Med Med Sci*. 2010;39:241-248.
- Austin PC, Mamdani MM, Juurlink DN, Hux JE. Testing multiple statistical hypotheses resulted in spurious associations: a study of astrological signs and health. *J Clin Epidemiol*. 2006;59:964-969.
- Murphy EA. *The Logic of Medicine*. 2nd ed. Baltimore, MD: Johns Hopkins University Press; 1997:345.
- Mendez MA, Popkin BM, Buckland G, et al. Alternative methods of accounting for underreporting and overreporting when measuring dietary intake-obesity relations. *Am J Epidemiol*. 2011;173:448-458.
- Martin GS, Tapsell LC, Batterham MJ, Russell KG. Relative bias in diet history measurements: a quality control technique for dietary intervention trials. *Public Health Nutr*. 2002;5:537-545.
- Pandeya N, Williams GM, Green AC, Webb PM, Whiteman DC. Do low control response rates always affect the findings? Assessments of smoking and obesity in two Australian case-control studies of cancer. *Aust N Z J Public Health*. 2009;33:312-319.
- Couper MP, Peytchev A, Strecher VJ, Rothert K, Anderson J. Following up non-respondents to an online weight management intervention: randomized trial comparing mail versus telephone. *J Med Internet Res*. 2007;9(2):e16.
- Eysenbach G. The law of attrition. *J Med Internet Res*. 2005;7(1):e11.
- Cochrane AL. *Effectiveness & Efficiency: Random Reflections on Health Services*. London, England: Royal Society of Medicine Press; 2004:85.
- Eliot TS. *The Family Reunion: The Centenary Edition—1888-1988*. Orlando, FL: Harcourt Brace; 1988:33-34.
- Prentice AM, Jebb SA. Beyond body mass index. *Obes Rev*. 2001;2:141-147.
- Bray GA. Obesity is a chronic, relapsing neurochemical disease. *Int J Obes Relat Metab Disord*. 2004;28:34-38.
- Trabulsi J, Schoeller DA. Evaluation of dietary assessment instruments against doubly labeled water, a biomarker of habitual energy intake. *Am J Physiol Endocrinol Metab*. 2001;281:E891-E899.
- Heska S, Allison DB. Is obesity a disease? *Int J Obes Relat Metab Disord*. 2001;25:1401-1404.
- Allison DB, Downey M, Atkinson RL, et al. Obesity as a disease: a white paper on evidence and arguments commissioned by the Council of the

- Obesity Society. *Obesity (Silver Spring)*. 2008;16:1161-1177.
35. Vague J. The degree of masculine differentiation of obesities: a factor determining predisposition to diabetes, atherosclerosis, gout, and uric calculous disease. *Am J Clin Nutr*. 1956;4:20-34.
 36. Stunkard AJ. Current views on obesity. *Am J Med*. 1996;100:230-236.
 37. Sturm R. The effects of obesity, smoking, and drinking on medical problems and costs. *Health Aff (Millwood)*. 2002;21:245-253.
 38. Kaplan HI, Kaplan HS. The psychosomatic concept of obesity. *J Nerv Men Dis*. 1957;125:181-201.
 39. Sutin AR, Ferrucci L, Zonderman AB, Terracciano A. Personality and obesity across the adult life span. *J Pers Soc Psychol*. 2011;101:579-592.
 40. Carr KA, Daniel TO, Lin H, Epstein LH. Reinforcement pathology and obesity. *Curr Drug Abuse Rev*. 2011;4:190-196.
 41. World Health Organization. *International Classification of Diseases (ICD-10-CM)*. <http://www.ICD10data.com/ICD10CM/Codes/E00-E89/E65-E68E66-E66.9>. Accessed November 30, 2011
 42. Aronne LJ, Nelinson DS, Lillo JL. Obesity as a disease state: a new paradigm for diagnosis and treatment. *Clin Cornerstone*. 2009;9(4):9-25.
 43. Levin BE. The drive to regain is mainly in the brain. *Am J Physiol Regul Integr Comp Physiol*. 2004;287:R1297-R1300.
 44. Volkow ND, O'Brien CP. Issues for DSM V: should obesity be included as a brain disorder? *Am J Psychiatry*. 2007;164:708-710.
 45. MacLean PS, Higgins JA, Jackman MR, et al. Peripheral metabolic responses to prolonged weight reduction that promote rapid, efficient regain in obesity-prone rats. *Am J Physiol Regul Integr Comp Physiol*. 2006;290:R1577-R1588.
 46. Maes HH, Neale MC, Eaves LJ. Genetic and environmental factors in relative body weight and human adiposity. *Behav Genet*. 1997;27:325-351.
 47. O'Rahilly S, Farooqi IS. Genetics of obesity. *Philos Trans R Soc Lond B Biol Sci*. 2006;361:1095-1105.
 48. Mark AL. Dietary therapy for obesity: an emperor with no clothes. *Hypertension*. 2008;51:1426-1434.
 49. Wisse BE. The inflammatory syndrome: the role of adipose tissue cytokines in metabolic disorders linked to obesity. *J Am Soc Nephrol*. 2004;15:2792-2800.
 50. Atkinson RL. Could viruses contribute to the worldwide epidemic of obesity? *Int J Pediatr Obes*. 2008;3(suppl 1):37-43.
 51. Hall KD, Sacks G, Chandramohan D, et al. Quantification of the effect of energy balance on bodyweight. *Lancet*. 2011;378:826-837.
 52. Power ML, Schulkin J. *The Evolution of Obesity*. Baltimore, MD: Johns Hopkins University Press; 2009:11.
 53. Saguy AC, Riley KW. Weighing both sides: morality, mortality, and framing contests over obesity. *J Health Polit Policy Law*. 2005;30:869-921.
 54. Rothman KJ, Greenland S, Lash TL. Types of epidemiologic studies. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2008:87-99.
 55. Hill AB, Hill ID. *Bradford Hill's Principles of Medical Statistics*. 12th ed. London: Edward Arnold; 1991.
 56. How to read clinical journals: IV. To determine etiology or causation. *Can Med Assoc J*. 1981;124:985-990.
 57. Gallagher KI, Jakicic JM, Kiel DP, Page ML, Ferguson ES, Marcus BH. Impact of weight-cycling history on bone density in obese women. *Obes Res*. 2002;10:896-902.
 58. Razak F, Anand SS, Shannon H, et al. Defining obesity cut points in a multiethnic population. *Circulation*. 2007;115:2111-2118.
 59. Rothman KJ, Greenland S, Lash TL. Validity in epidemiologic studies. In Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2008:128-147.
 60. Glymour MM, Greenland S. Causal diagrams. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2008:183-209.
 61. Flegal KM, Graubard BI, Williamson DF, Gail MH. Sources of differences in estimates of obesity-related deaths from first National Health and Nutrition Examination Study (NHANES I) hazard ratios. *Am J Clin Nutr*. 2010;91:519-527.
 62. Rothman KJ, Greenland S, Poole C, Lash TL. Causation and causal inference. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2008:5-31.
 63. Rothman KJ, Greenland S. Cohort studies. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2008:100-110.
 64. Hu, FB. *Obesity Epidemiology*. New York, NY: Oxford University Press; 2008.
 65. Prospective Studies Collaboration. Body-mass index and cause-specific mortality in 900,000 adults: collaborative analyses of 57 prospective studies. *Lancet*. 2009;373:1083-1096.
 66. Zajacova A, Dowd JB, Burgard SA. Overweight adults may have the lowest mortality: do they have the best health? *Am J Epidemiol*. 2011;173:430-437.
 67. Greenberg JA. Correcting biases in estimates of mortality attributable to obesity. *Obesity*. 2006;14:2071-2079.
 68. Flegal KM, Graubard BI, Williamson DF, Cooper RS. Reverse causation and illness-related weight loss in observational studies of body weight and mortality. *Am J Epidemiol*. 2011;173:1-9.
 69. Roth J. Evolutionary speculation about tuberculosis and the metabolic and inflammatory processes of obesity. *JAMA*. 2009;301:2586-2588.
 70. Strandhagen E, Berg C, Lissner L, et al. Selection bias in a population survey with registry linkage: potential effect on socioeconomic gradient in cardiovascular risk. *Eur J Epidemiol*. 2010;25:163-172.
 71. McCrea RL, Berger YG, King MB. Body mass index and common mental disorders: exploring the shape of association and its moderation by age, gender, and education. *Int J Obes (London)*. 2011;36:414-421.
 72. Klem ML, Wing RR, McGuire MT, Seagle HM, Hill JO. A descriptive study of individuals successful at long-term maintenance of substantial weight loss. *Am J Clin Nutr*. 1997;66:239-246.
 73. Karasu SR, Karasu TB. *The Gravity of Weight: A Clinical Guide to Weight Loss and Maintenance*. Washington, DC: American Psychiatric Publishing; 2010.
 74. Dansinger ML, Gleason JA, Griffith JL, Selker HP, Schaefer EJ. Comparison of the Atkins, Ornish, Weight Watchers and Zone diets for weight loss and heart disease risk reduction: a randomized trial. *JAMA*. 2005;293:43-53.
 75. Nordmann AJ, Nordmann A, Briel M, et al. Effects of low-carbohydrate vs low-fat diets on weight loss and cardiovascular risk factors: a meta-analysis of randomized controlled trials. *Arch Intern Med*. 2006;166:285-293.
 76. Alonso A, Segui-Gómez M, de Irala J, Sánchez-Villegas A, Beunza JJ, Martínez-González MA. Predictors of follow-up and assessment of selection bias from drop-outs using inverse probability weighting in a cohort of university graduates. *Eur J Epidemiol*. 2006;21:351-358.
 77. Hartge P, Cahill J. Field methods in epidemiology. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2008:492-410.

78. Hu FB, Willett WC, Li T, Stampfer MJ, Colditz GA, Manson JE. Adiposity as compared with physical activity in predicting mortality among women. *N Engl J Med*. 2004;351:2694-2703.
79. Schulze MB, Manson JE, Ludwig DS, et al. Sugar-sweetened beverages, weight gain, and incidence of type 2 diabetes in young and middle-aged women. *JAMA*. 2004;292:927-934.
80. Wang M, Fitzmaurice GM. A simple imputation method for longitudinal studies with non-ignorable responses. *Biom J*. 2006;48:302-318.
81. Sackett DL. Commentary: measuring the success of blinding in RCTs: don't, must, can't or needn't? *Int J Epidemiol*. 2007;36:664-665.
82. Gorber SC, Tremblay M, Moher D, Gorber B. A comparison of direct vs self-report measures for assessing height, weight, and body mass index: a systematic review. *Obes Rev*. 2007;8:307-326.
83. Winkler JT. The fundamental flaw in obesity research. *Obes Rev*. 2005;6:199-202.
84. Meiselman HL. Methodology and theory in human eating research. *Appetite*. 1992;19:49-55.
85. de Castro JM. Eating behavior: lessons from the real world of humans. *Nutrition*. 2000;16:800-813.
86. Wansink B. *Mindless Eating: Why We Eat More Than We Think*. New York, NY: Bantam Books; 2006:37-40.
87. Greenland S, O'Rourke K. Meta-analysis. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2008:652-682.
88. Brownell KD, Rodin J. Medical, metabolic, and psychological effects of weight cycling. *Arch Intern Med*. 1994;154:1325-1330.
89. Trowman R, Dumville JC, Hahn S, Torgerson DJ. A systematic review of the effects of calcium supplementation on body weight. *Br J Nutr*. 2006;95:1033-1038.
90. Sir Dale H. Measurement in medicine: introduction. *Brit Med Bull*. 1951;7:261-263.
91. Devlin MJ, Yanovski SZ, Wilson GT. Obesity: what mental health professionals need to know. *Am J Psychiatry*. 2000;157:854-866.
92. Kalarchian MA, Marcus MD, Levine MD, et al. Psychiatric disorders among bariatric surgery candidates: relationship to obesity and functional health status. *Am J Psychiatry*. 2007;164:328-334.
93. Friedman JM. A war on obesity, not the obese. *Science*. 2003;299:856-858.
94. Rössner S. Adolphe Quetelet (1796-1874). *Obes Rev*. 2007;8:183.
95. Keys A, Fidanza F, Karvonen MJ, Kimura N, Taylor HL. Indices of relative weight and obesity. *J Chronic Dis*. 1972;25:329-343.
96. Pai MP, Paloucek FP. The origin of the "ideal" body weight equations. *Ann Pharmacother*. 2000;34:1066-1069.
97. Himes JH, Bouchard C, Pheley AM. Lack of correspondence among measures identifying the obese. *Am J Prev Med*. 1991;7:107-111.
98. Clinical guidelines on the identification, evaluation, and treatment of overweight and obesity in adults: the evidence report. National Institutes of Health. *Obes Res*. 1998;6(suppl 2):51S-209S.
99. Obesity: preventing and managing the global epidemic. Report of a WHO consultation. *World Health Organ Tech Rep Ser*. 2000;894:i-xii, 1-253.
100. Strawbridge WJ, Wallhagen MI, Shema SJ. New NHLBI clinical guidelines for obesity and overweight: will they promote health? *Am J Public Health*. 2000;90:340-343.
101. Bouchard C, Pérusse L, Rice T, Rao DC. Genetics of human obesity. In: Bray GA, Bouchard C, eds. *Handbook of Obesity: Etiology and Pathophysiology*. 2nd ed. New York, NY: Marcel Dekker; 2004:157-200.
102. Rothman KJ. BMI-related errors in the measurement of obesity. *Int J Obes (Lond)*. 2008;32(suppl 3):S56-S59.
103. Nevill AM, Winter EM, Ingham S, Watts A, Metsios GS, Stewart AD. Adjusting athletes' body mass index to better reflect adiposity in epidemiological research. *J Sports Sci*. 2010;28:1009-1016.
104. Deurenberg P, Weststrate JA, Seidell JC. Body mass index as a measure of body fatness: age- and sex-specific prediction formulas. *Br J Nutr*. 1991;65:105-114.
105. Kragelund C, Omland T. A farewell to body-mass index? *Lancet*. 2005;366:1589-1591.
106. Burkhauser RV, Cawley J. Beyond BMI: the value of more accurate measures of fatness and obesity in social science research. *J Health Econ*. 2008;27:519-529.
107. Piers LS, Soares MJ, Frandsen SL, O'Dea K. Indirect estimates of body composition are useful for groups but unreliable in individuals. *Int J Obes Relat Metab Disord*. 2000;24:1145-1152.
108. Romero-Corral A, Somers VK, Sierra-Johnson J, et al. Diagnostic performance of body mass index to detect obesity in patients with coronary artery disease. *Eur Heart J*. 2007;28:2087-2093.
109. Svendsen OL. Should measurement of body composition influence therapy for obesity? *Acta Diabetol*. 2003;40(suppl 1):S250-S253.
110. Mora S, Yanek LR, Moy TF, Fallin MD, Becker LC, Becker DM. Interaction of body mass index and Framingham risk score in predicting incident coronary disease in families. *Circulation*. 2005;111:1871-1876.
111. Danubio ME, Miranda G, Vinciguerra MG, Vecchi E, Rufo F. Comparison of self-reported and measured height and weight: implications for obesity research among young adults. *Econ Hum Biol*. 2008;6:181-190.
112. Shields M, Gorber SC, Tremblay MS. Estimates of obesity based on self-report versus direct measures. *Health Rep*. 2008;19:61-76.
113. Gil J, Mora T. The determinants of misreporting weight and height: the role of social norms. *Econ Hum Biol*. 2011;9:78-91.
114. Elgar FJ, Stewart JM. Validity of self-report screening for overweight and obesity: evidence from the Canadian Community Health Survey. *Can J Public Health*. 2008;99:423-427.
115. Taylor AW, Dal Grande E, Gill TK, et al. How valid are self-reported height and weight? A comparison between CATI self-report and clinic measurements using a large cohort study. *Aust N Z J Public Health*. 2006;30:238-246.
116. Yannakoulia M, Panagiotakos DB, Pitsavos C, Stefanadis C. Correlates of BMI misreporting among apparently healthy individuals: the ATTICA study. *Obesity (Silver Spring)*. 2006;14:894-901.
117. Wada K, Tamakoshi K, Tsunekawa T, et al. Validity of self-reported height and weight in a Japanese workplace population. *Int J Obes (Lond)*. 2005;29:1093-1099.
118. Himes JH, Hannan P, Wall M, Neumark-Sztainer D. Factors associated with errors in self-reports of stature, weight, and body mass index in Minnesota adolescents. *Ann Epidemiol*. 2005;15:272-278.
119. Ezzati M, Martin H, Skjold S, Vander Hoorn S, Murray CJL. Trends in national and state-level obesity in the USA after correction for self-report bias: analysis of health surveys. *J R Soc Med*. 2006;99:250-257.
120. Visscher TL, Viet AL, Kroesbergen IH, Seidell JC. Underreporting of BMI in adults and its effect on obesity prevalence estimations in the period 1998-2001. *Obesity (Silver Spring)*. 2006;14:2054-2063.
121. Nyholm M, Gullberg B, Merlo J, Lundqvist-Persson C, Råstam L, Lindblad U. The validity of obesity based on self-reported weight and height: implications for population studies. *Obesity (Silver Spring)*. 2007;15:197-208.
122. DelPrete LR, Caldwell M, English C, Banspach SW, Lefebvre C. Self-reported and measured weights and heights

- of participants in community-based weight loss programs. *J Am Diet Assoc.* 1992;92:1483-1486.
123. Tell GS, Jeffery RW, Kramer FM, Snell MK. Can self-reported body weight be used to evaluate long-term follow-up of a weight-loss program? *J Am Diet Assoc.* 1987;87:1198-1201.
 124. Flood V, Webb K, Lazarus R, Pang G. Use of self-report to monitor overweight and obesity in populations: some issues for consideration. *Aust N Z J Public Health.* 2000;24:96-99.
 125. Niedhammer I, Bugel I, Bonenfant S, Goldberg M, Leclerc A. Validity of self-reported weight and height in the French GAZEL cohort. *Int J Obes Relat Metab Disord.* 2000;24:1111-1118.
 126. Keith SW, Fontaine KR, Pajewski NM, Mehta T, Allison DB. Use of self-reported height and weight biases the body mass index-mortality association. *Int J Obes (Lond).* 2011;35:401-408.
 127. Khaled MA, Kabir I, Goran MI, Mahalanabis D. Bioelectrical impedance measurements at various frequencies to estimate human body composition. *Indian J Exp Biol.* 1997;35:159-161.
 128. Rowe DA, Dubose KD, Donnelly JE, Mahar MT. Agreement between skinfold-predicted percent fat and percent fat from whole-body bioelectrical impedance analysis in children and adolescents. *Int J Pediatr Obes.* 2006;1:168-175.
 129. Wang CY, Liu MH, Chen YC. Intrarater reliability and the value of real change for waist and hip circumference measures by a novice rater. *Percept Mot Skills.* 2010;110(3, pt 2):1053-1058.
 130. Sullivan R, Johnson WD, Katzmarzyk PT. Waist circumference is an independent correlate of errors in self-reported BMI. *Obesity (Silver Spring).* 2010;18:2237-2239.
 131. Sebo P, Beer-Borst S, Haller DM, Bovier PA. Reliability of doctors' anthropometric measurements to detect obesity. *Prev Med.* 2008;47:389-393.
 132. Czernichow S, Kengne AP, Stamatakis E, Hamer M, Batty GD. Body mass index, waist circumference and waist-to-hip ratio: which is the better discriminator of cardiovascular disease mortality risk? Evidence from an individual participant meta-analysis of 82,864 participants from nine cohort studies. *Obes Rev.* 2011;12:680-687.
 133. Dhaliwal SS, Welborn TA. Measurement error and ethnic comparisons of measures of abdominal obesity. *Prev Med.* 2009;49:148-152.
 134. Willett W. *Nutritional Epidemiology* (Monographs in Epidemiology and Biostatistics, Vol. 30). 2nd ed. New York, NY: Oxford University Press; 1998.
 135. Lewis T. Diseases of the auriculo-ventricular valves. In: *Diseases of the Heart*. New York, NY: Macmillan; 1937:129-142.
 136. Gallagher D, Song MY. Evaluation of body composition: practical guidelines. *Prim Care.* 2003;30:249-265.
 137. Willett WC. Nutritional epidemiology. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2008:580-597.
 138. Pollan M. *In Defense of Food: An Eater's Manifesto*. New York, NY: Penguin; 2008:76.
 139. Lichtman SW, Pisarska K, Berman ER, et al. Discrepancy between self-reported and actual caloric intake and exercise in obese subjects. *N Engl J Med.* 1992;327:1893-1898.
 140. Samaras K, Kelly PJ, Campbell LV. Dietary underreporting is prevalent in middle-aged British women and is not related to adiposity (percentage body fat.). *Int J Obes Relat Metab Disord.* 1999;23:881-888.
 141. Schoeller DA, Bandini LG, Dietz WH. Inaccuracies in self-reported intake identified by comparison with doubly labelled water method. *Can J Physiol Pharmacol.* 1990;68:941-949.
 142. Braam LA, Ocké MC, Bueno-de-Mesquita HB, Seidell JC. Determinants of obesity-related underreporting of energy intake. *Am J Epidemiol.* 1998;147:1081-1086.
 143. Lafay L, Basdevant A, Charles MA, et al. Determinants and nature of dietary underreporting in a free-living population: the Fleurbaix Laventie Ville Santé (FLVS) Study. *Int J Obes Relat Metab Disord.* 1997;21:567-573.
 144. Voss S, Kroke A, Klipstein-Grobusch K, Boeing H. Obesity as a major determinant of underreporting in a self-administered food frequency questionnaire: results from the EPIC-Potsdam Study. *Z Ernahrungswiss.* 1997;36:229-236.
 145. Heitmann BL, Lissner L. Dietary underreporting by obese individuals: is it specific or non-specific? *BMJ.* 1995;311:986-989.
 146. Mendez MA, Wynter S, Wilks R, Forrester T. Under- and overreporting of energy is related to obesity, lifestyle factors and food group intakes in Jamaican adults. *Public Health Nutr.* 2004;7:9-19.
 147. Caan B, Ballard-Barbash R, Slattery ML, et al. Low energy reporting may increase in intervention participants enrolled in dietary intervention trials. *J Am Diet Assoc.* 2004;104:357-366.
 148. Scagliusi FB, Ferrioli E, Pfrimer K, et al. Characteristics of women who frequently under report their energy intake: a doubly labelled water study. *Eur J Clin Nutr.* 2009;63:1192-1199.
 149. Tataranni PA, Ravussin E. Energy metabolism and obesity. In: Wadden TA, Stunkard AJ, eds. *Handbook of Obesity Treatment*. New York, NY: Guilford Press; 2002:42-72.
 150. McGowan CA, McAuliffe FM. Maternal nutrient intakes and levels of energy underreporting during early pregnancy. *Eur J Clin Nutr.* 2012;66:906-913.
 151. Kretsch MJ, Fong AK, Green MW. Behavioral and body size correlates of energy intake underreporting by obese and normal-weight women. *J Am Diet Assoc.* 1999;99:300-306.
 152. Levine JA. Nonexercise activity thermogenesis: liberating the life-force. *J Intern Med.* 2007;262:273-287.
 153. Ainsworth BE, Haskell WL, Whitt MC, et al. Compendium of physical activities: an update of activity codes and MET intensities. *Med Sci Sports Exerc.* 2000;32(9 suppl):S498-S504.
 154. Howley ET, Franks BD. *Fitness Professional's Handbook*. 5th ed. Champaign, IL: Human Kinetics; 2007:483-496.
 155. Matthews CE, Chen KY, Freedson PS, et al. Amount of time spent in sedentary behaviors in the United States, 2003-2004. *Am J Epidemiol.* 2008;167:875-881.
 156. LaForge R. Key considerations for metabolic syndrome and diabetes prevention programs: exercise determinants of weight loss. Online clinical articles of the National Lipid Association, 2006. <http://www.lipid.org/clinical/tlc/1000002.php>. Accessed February 10, 2009.
 157. Daley A. Exercise and depression: a review of reviews. *J Clin Psychol Med Settings.* 2008;15:140-147.
 158. Stathopoulou G, Powers MB, Berry AC, Smits JAJ, Otto MW. Exercise interventions for mental health: a quantitative and qualitative review. *Clin Psychol: Sci Practice.* 2006;13:179-193.
 159. Hultquist CN, Albright C, Thompson DL. Comparison of walking recommendations in previously inactive women. *Med Sci Sports Exerc.* 2005;37:676-683.
 160. Tudor-Locke C, Hatano Y, Pangrazi RP, Kang M. Revisiting "how many steps are enough?" *Med Sci Sports Exerc.* 2008;40(7 suppl):S537-S543.
 161. Kumahara H, Tanaka H, Schutz Y. Are pedometers adequate instruments for assessing energy expenditure? *Eur J Clin Nutr.* 2009;63:1425-1432.
 162. De Cocker K, Cardon G, De Bourdeaudhuij I. Validity of the inexpensive Stepping Meter in counting steps in free living conditions: a pilot study. *Br J Sports Med.* 2006;40:714-716.

163. Mackay LM, Schofield GM, Schluter PJ. Validation of self-report measures of physical activity: a case study using the New Zealand Physical Activity Questionnaire. *Res Q Exerc Sport*. 2007;78:189-196.
164. Dishman RK, Rooks CR, Thom NJ, Motl RW, Nigg CR. Meeting U.S. Healthy People 2010 levels of physical activity: agreement of 2 measures across 2 years. *Ann Epidemiol*. 2010;20:511-523.
165. Prince SA, Adamo KB, Hamel ME, Hardt J, Gorber SC, Tremblay M. A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. *Int J Behav Nutr Phys Act*. 2008;5:56.