The Inadmissibility of What We Eat in America and NHANES Dietary Data in Nutrition and Obesity Research and the Scientific Formulation of National Dietary Guidelines

Edward Archer, PhD; Gregory Pavela, PhD; and Carl J. Lavie, MD

Abstract

The Scientific Report of the 2015 Dietary Guidelines Advisory Committee was primarily informed by memory-based dietary assessment methods (M-BMs) (eg, interviews and surveys). The reliance on M-BMs to inform dietary policy continues despite decades of unequivocal evidence that M-BM data bear little relation to actual energy and nutrient consumption. Data from M-BMs are defended as valid and valuable despite no empirical support and no examination of the foundational assumptions regarding the validity of human memory and retrospective recall in dietary assessment. We assert that uncritical faith in the validity and value of M-BMs has wasted substantial resources and constitutes the greatest impediment to scientific progress in obesity and nutrition research. Herein, we present evidence that M-BMs are fundamentally and fatally flawed owing to well-established scientific facts and analytic truths. First, the assumption that human memory can provide accurate or precise reproductions of past ingestive behavior is indisputably false. Second, M-BMs require participants to submit to protocols that mimic procedures known to induce false recall. Third, the subjective (ie, not publicly accessible) mental phenomena (ie, memories) from which M-BM data are derived cannot be independently observed, quantified, or falsified; as such, these data are pseudoscientific and inadmissible in scientific research. Fourth, the failure to objectively measure physical activity in analyses renders inferences regarding diet-health relationships equivocal. Given the overwhelming evidence in support of our position, we conclude that M-BM data cannot be used to inform national dietary guidelines and that the continued funding of M-BMs constitutes an unscientific and major misuse of research resources.

When the facts change, I change my mind. What do you do, sir?
John Maynard Keynes

SUCCESS, FAILURE, AND CONFUSION IN NUTRITION RESEARCH

During the past century, our nation’s food supply and the nutritional status of Americans have improved to a level unparalleled in human history. Although this reality may be contrary to the popular belief that our modern diet is inherently inadequate, the data are clear. In the early 20th century, nutritional diseases such as pellagra, beriberi, rickets, and goiter were substantial public health challenges. In the United States alone, pellagra (a disease of niacin deficiency) claimed more than 100,000 lives and severely affected more than 3 million people. Yet in 2013, the Centers for Disease Control and Prevention’s Second National Report on Biochemical Indicators of Diet and Nutrition reported that nearly “80% of Americans (aged ≥6 y) were not at risk of deficiencies in any of the 7 vitamins examined—examinations via biomarkers (ie, vitamins A, B6, B12, C, D, E, and folate; emphasis added). In addition, approximately 90% of women of childbearing age (12-49 years) were not at risk for iron deficiency, and folate levels have increased by approximately 50% since the previous national report. As such, most of the US population is not at risk for nutritional deficiencies, and
Neither do they have nutritional deficiencies and associated diseases. Given these important improvements in diet-related health and recent work demonstrating that nongenetic evolution is the predominant driver of the diseases of excess (eg, obesity and type 2 diabetes mellitus), it can be posited that diet is no longer a major risk factor for disease for most Americans. If accurate, this hypothesis suggests that the billions of research dollars targeted for diet and nutrition—related health research are misdirected. Nevertheless, despite the important dietary milestones of the past century and the substantial increases in federal funding during the past 2 decades, research into human nutrition has been increasingly criticized. The genesis of these criticisms is the appalling track record of highly publicized nutrition claims derived from epidemiologic studies (eg, see the studies by Stampfer et al and Rimm et al) that consistently failed to be supported when tested using objective study designs. Young and Karr examined more than 50 nutritional claims from observational studies for a variety of dietary patterns and nutrient supplementation and found that “100% of the observational claims failed to replicate” and that 5 claims were statistically significant “in the opposite direction.” These outcomes and others suggest that as often as not, when epidemiologic nutrition claims are tested against objective research methods, the results are either inconclusive or indicative of a contrary outcome.

A FAILED RESEARCH PARADIGM

Epidemiologic studies suggest that almost any nutrient can be associated with a myriad of outcomes, as observed in Schoenfeld and Ioannidis’ article, “Is Everything We Eat Associated With Cancer?” With persistent cycles of specious nutrition claims in the media, it is not surprising that the public is confused and incredulous. Insofar as the provision of clear and consistent dietary guidelines for the consuming public is a goal of nutrition epidemiology, it has failed in decisively answering the simple question, “What should we eat?” Nowhere is this fact more evident than the shifting sands of opinion on the relative risks of fat, salt, cholesterol, and sugar. Five decades of controversy surrounding basic dietary guidelines and nutrition recommendations is a public acknowledgement of a failed research paradigm. The striking incongruence between the improvements in the nutritional status of the US population and the current state of confusion, controversy, and clinical failure of epidemiologic nutrition research could not be clearer and necessitates an examination of the validity and value of epidemiologic nutrition research.

PURPOSE OF THIS REVIEW

Memory-based dietary assessment methods (M-BMs) (eg, interviews, questionnaires, and surveys) are the dominant data collection protocols in national nutrition surveillance and government-funded epidemiologic nutrition research. Importantly, M-BM data are used to inform national nutritional policy and dietary guidelines. The recent Scientific Report of the 2015 Dietary Guidelines Advisory Committee (DGAC) stated explicitly that most of the DGAC data analyses used the M-BMs of the National Health and Nutrition Examination Survey (NHANES) dietary component, What We Eat in America (WWEIA). Although decades of unequivocal evidence demonstrate that the indirect, proxy estimates derived from M-BMs bear little relation to actual energy or nutrient consumption, the underlying assumptions regarding the validity of human memory and recall in dietary assessment have not been questioned. To the contrary, M-BM data are vigorously defended as valid and inherently valuable despite no empirical support for those assertions. Although the relationship between 2 different constructs may be expected to be weak, the trivial relationships between the proxy estimates (ie, self-reported energy intake [EI] and nutrient intake) and their referents (ie, actual EI and nutrient intake) are unacceptable. We assert that the explanatory and predictive failure of epidemiologic nutrition research is explained by its reliance on M-BMs, and, as such, the uncritical faith in the validity and value of M-BMs has wasted significant resources and constitutes the single greatest impediment to actual scientific progress in the fields of obesity and nutrition research.

The purpose of this review is to survey the explanatory and predictive failure of nutrition epidemiology in general with a focus on the WWEIA-NHANES data, and argue that
these failures are due to the reliance on M-BMs. First, we present evidence that the anecdotally derived proxy data produced by M-BMs bear little relation to actual EI or nutrient consumption.13,33,35-45 Second, we provide interdisciplinary evidence that human memory is an amalgam of constructive and reconstructive processes47-52 (eg, imagination53) that render the archival model of human memory54 and the naive assumption that recall provides literal, accurate, or precise reproductions of past events indisputably false.50,52,55-58 Third, M-BMs require respondents to undergo protocols59 and perform behaviors60 that mimic procedures known to induce false recall.50,52,53,60,61 Fourth, the subjective (ie, private, not publicly accessible) mental phenomena (ie, memories) from which M-BM data are derived are not subject to independent observation, quantification, falsification, or verification; as such, M-BM data are pseudoscientific and inadmissible in scientific research.62-66 Fifth, the failure to accurately and objectively measure and control for physical activity (PA), cardiorespiratory fitness (CRF), and other obvious confounders annuls inferences regarding diet-health relationships.

THE M-BMs OF NUTRITION EPIDEMIOLOGY

Self-reported Dietary Intake

The primary methods of data collection for nutrition epidemiologic research (eg, the WWEIA-NHANES) are M-BMs (eg, 24-hour dietary recalls [24HRs] and food frequency questionnaires [FFQs]11-33). For clarity, these methods do not directly or objectively measure EI or nutrient intake, and neither do they directly or objectively measure food and beverage consumption. The actual data derived from M-BMs are the a priori numeric values from nutrient databases that are assigned by researchers to the participants’ reports of their memories of past eating and drinking behaviors. In other words, nutrition researchers designate numeric values to whatever the respondents are willing or able to recall about what they think (or would like the researcher to think67) he or she consumed during the study period. Given the indirect, pseudoquantitative (ie, number-generating68) nature of M-BMs and the fact that the respondents’ reports of their memories are subject to intentional and unintentional distorting factors (eg, perceptual, encoding, and retrieval errors59, social desirability42; false memories25; and omissions48,49,76), it is not surprising that most conclusions drawn from these number-generating protocols have not been supported when subjected to rigorous objective examination.11,17

The Implausibility of M-BMs in Dietary Assessment

It is the natural tendency of the ignorant to believe what is not true. In order to overcome that tendency it is not sufficient to exhibit the true; it is also necessary to expose and denounce the false.

H. L. Mencken71, p124

Research into M-BMs reports a wide range of EIs that are not physiologically plausible (ie, incompatible with survival) and that do not accurately quantify the foods and nutrients consumed.11,33,35,36-40,42 Recently, we used multiple methods to ascertain the validity and plausibility of the NHANES and WWEIA-NHANES EI data from 1971 to 201033 and found that they had such severe systematic biases as to render them fatally flawed. Given that “[a] cross the 39-year history of the NHANES, [self-reported energy intake] data on the majority of respondents (67.3% of women and 58.7% of men) were not physiologically plausible”33 (Figure), we concluded that these data are not valid for any inferences regarding EI and the etiology of the obesity epidemic. A recent editorial in the British Medical Journal concurred and stated that the NHANES dietary data are “incompatible with life.”11, p7

In a previous report,33 we used 2 objective, physiologically based methods to determine misreporting: (1) Goldberg cutoff values44,45,72 (ie, reported EI [rEI] divided by basal metabolic rate [BMR]) and (2) the disparity between the Institute of Medicine total energy expenditure (TEE) equations73 and rEI via NHANES M-BMs. The 2 methods were in close agreement, demonstrating significant misreporting. The cutoff values we used (ie, rEI/BMR <1.35 and >2.40) were more generous than the rEI/BMR cutoff value of 1.50 suggested by Goldberg et al15 when using a single 24HR, and the BMR is “predicted from the Schofield equations” with a sample size of 300 or greater.15, p377 Given the reduced sensitivity of the cutoff values, we captured far fewer underreporters. As reported, when using the proposed cutoff value of 1.50,
underreporting increased to more than 70% for the entire NHANES sample and to approximately 77% and 85% for obese men and women, respectively. We also reported the large and significant disparity between rEI and the Institute of Medicine TEE: 467 and 554 kcal/d (17% and 30%) for obese men and women, respectively. In addition to underreporting, there was significant overreporting in all of the subpopulations (eg, normal, overweight, and obese men and women). One important caveat with the use of cutoff values is that the term plausible reporter is not synonymous with accurate reporter. Participants with high levels of PA may substantially underreport yet still be considered plausible reporters.

Given these results, we ask 4 questions, (1) What is the value of WWEIA-NHANES M-BM data if 70% to 80% of obese women’s self-reported EI is physiologically implausible and, therefore, incompatible with life (Figure)? (2) Given the extant objective data on the nutrition-related health status of Americans, why does the DGAC rely on the subjective M-BM data? (3) What is the “unrealized potential” and “utility” of these data when implausible overreporting and implausible underreporting are demonstrated in all of the subgroups? (4) Can statistical alchemy transform these implausible data into valid estimates of dietary consumption, or will it continue to spawn searches for machinations that generate numbers with improved correlations (ie, post hoc data manipulation) while ignoring the lack of validity?

The Pervasiveness of Implausible Results
The conclusions drawn by our study and the recent British Medical Journal editorial are, in fact, supported by many decades of evidence demonstrating that M-BMs have severe, intractable systematic biases that render the data implausible and, therefore, invalid. Research with “… motivated … well-educated, non-smoking Caucasians (ie, respondents less likely to misreport) demonstrated that compared with doubly labeled water, a biomarker for TEE, self-reported dietary intake was significantly misestimated. Men underreported EI 12% to 14% using the average of...
two 24HRs and 31% to 36% using FFQs. Women underreported by 16% to 20% using the average of two 24HRs and by 34% to 38% using FFQs. Contrary to the oft-repeated statement that additional self-reports improve precision and accuracy, the second administration of the 24HR “showed greater underreporting.” The results are in agreement with our analyses of the NHANES in which the mean estimates for the second 24HR in every NHANES wave from 2001 to 2010 exhibited significantly greater levels of underreporting than the first. We agree with the authors of the Observing Protein and Energy Nutrition study when they wrote, “[w]e measure energy so poorly …” and “[t]he 24HR … may be particularly problematic in the obese.” These words echo statements on underreporting from 60 years ago.

Recently, some of the strongest proponents of M-BMs have provided additional data that clearly demonstrate the futility of the continued use of these methods. In the paper by Freedman et al., the pooled, squared average correlation between true EI and self-reported EI were similar to our results using NHANES data, ranging from 0.04 to 0.10. This suggests that the measurement noise (ie, error) is more than 9 times greater than the signal (ie, valid information) derived from M-BMs. Nevertheless, an important finding from the Observing Protein and Energy Nutrition study that Freedman et al. overlook in their analyses is that despite the fact that the second administration of the 24HR “showed greater underreporting,” the correlations between true and reported EI increased. This demonstrates an increase in precision with a concomitant reduction in the accuracy of the estimate. These results clearly support our position that M-BM data “offer an inadequate basis for scientific conclusions” and, more importantly, that statistical machinations, however sophisticated, cannot overcome the systematic recall bias that renders all inferences suspect.

The phenomenon of misreporting is not limited to US epidemiologic studies or specific populations. The European Prospective Investigation Into Cancer and Nutrition study is one of the largest epidemiologic studies in the world and found strong evidence of systemic underreporting across all study sites, with approximately 10% to 14% of survey respondents being “extreme underreporters.” This nonuniformity of misreporting leads to macronutrient- and micronutrient-specific errors, which alter nutrient to EI ratios in an unpredictable and nonquantifiable manner. This simple fact renders energy adjustments fallacious.

Examinations of Dietary Patterns via M-BMs
It is well-established that specific macronutrients, foods, beverages, and food groups (eg, protein, fat, carbohydrate, alcohol, sugar, and vegetables) are subject to differential misreporting that significantly affects subsequent estimates of EI. Because EI is the foundation of dietary consumption and all nutrients must be consumed within the quantity of food and beverages needed to meet minimum energy requirements, it is a logical and analytic truth that dietary patterns (ie, macronutrient and micronutrient consumption; eg, protein, carbohydrate, fat, vitamins, and minerals) are differentially and unpredictably misreported when total rEI is physiologically implausible. For example, both macronutrient and micronutrient composition are significantly altered in underreporters, with reported fat and carbohydrate consumption often lower and reported protein, fruit, and vegetable intakes higher. In other words, participants qualitatively and quantitatively misreport owing to both unintentional (eg, forgetting and false memories) and intentional (eg, health-related perceptions) factors. This nonuniformity of misreporting leads to macronutrient- and micronutrient-specific errors, which alter nutrient to EI ratios in an unpredictable and nonquantifiable manner. This simple fact renders energy adjustments fallacious.
and demonstrates that the assumption that M-BM data can be used to examine patterns of diet or dietary composition is not logically valid.

THE VALIDITY OF HUMAN MEMORY AND RECALL AS INSTRUMENTS FOR THE GENERATION OF SCIENTIFIC DATA

Overview

The use of M-BMs requires faith in the belief that human perception, memory, and recall are accurate and reliable instruments for the generation of scientific data. Nevertheless, more than 80 years of research demonstrates that this belief is patently false. Bernard et al58 concluded that “the results of all of these studies leads to one overwhelming conclusion: on average, about half of what informants report is probably incorrect in some way.” Bernard et al also provide a prescient commentary: “In sum, despite the evidence, the basic fact of informant inaccuracy seems not to have penetrated either graduate training or professional social science research. Informant inaccuracy remains both a fugitive problem and a well-kept open secret.” Given the substantial funding of M-BMs each year, it seems that this 30-year-old commentary also applies to nutrition and obesity research.

Furthermore, when events or behaviors are commonplace (eg, food and beverage consumption), previous experiences (eg, previous memories and mental schema of past meals) will determine what is encoded in memory and not the actual perception of behavior. For example, Freeman et al48 demonstrated a 52% error rate in recalling social interactions, with reports of social interactions shaped by typical past experiences. They explain their results by suggesting that when events are repeatedly experienced, each specific event will be minimally processed and the “actual memory of such elements will be poor,” and “attempts at recall result in a constructive process that taps into the general structure rather than the specific memory.”

Importantly, Bernard et al38 lamented 2 common problems with social scientific data: (1) the lack of an explicit formal theory of human behavior and (2) objective evidence from which to test the plausibility of self-reported data. Nevertheless, nutrition epidemiologists...
have both a formal theory (ie, human metabolism and the basic energy requirements of human life) and voluminous objective data by which to test the validity of M-BMs. Despite the availability of formal theory and overwhelming evidence that self-reported EI data are not accurate, “plausible,” or even “compatible with life,” self-reported EI continues to be assumed a valid measure of actual energy and nutrient consumption that can be used to inform public nutrition and dietary policy.

A detailed review of the social research literature is beyond the scope of this paper, and we direct our readers to Bernard et al’s review. Nevertheless, one more notable example is warranted. Immediately on leaving a restaurant, Kronenfeld et al had participants report on the attire of the waitstaff and the restaurants’ choice of music. Participants demonstrated much greater agreement on what the waiters were wearing compared with the waitresses’ attire. The interesting finding was that these restaurants had an all-female waitstaff (ie, there were no waiters in the restaurants). Participants provided much greater detail on the music from restaurants that were not playing music than from restaurants that were. These results raise the question: What is the possibility that self-reported food and beverage consumption in a restaurant setting will be a literal, accurate, or reliable representation of actual ingestive behavior?

Cognitive Neuroscience

The domain of cognitive neuroscience supports the hypothesis that human memory is an amalgam of dynamic constructive and reconstructive processes. For example, encoding is not a process that begins de novo with each perception. Encoding is the result of the limited amount of information available to perception at any given moment being “patched together to form memories with varying degrees of accuracy” (eg, the process of associative grouping via semantic relatedness and subject to “the distorting influences of present knowledge, beliefs, and... previous experience.” As such, the general knowledge and availability of mental schemas from previous eating occasions intrude on the encoding of current consumption to produce false and fuzzy (ie, gist) memories. Memory and recall are subject to a myriad of unintentional “sins,” including but not limited to distortions, misattribution, suggestibility, simple forgetting, falsehoods, and omissions. Because selective and elaborative processes operate on the perceptions that are encoded and recalled, “memory does not [and cannot] operate like a video recording.”

Recently, the process of reconsolidation (ie, the reconstruction and re-encoding of memories after recall) has been demonstrated in rodents, and the evidence in humans is supportive. Reconsolidation involves the same neural processes as the encoding of the original memory. Therefore, each time a memory is recalled, it is irretrievably changed such that the original memory no longer exists and a new memory of unquantifiable error replaces it. This fact has implications for the current state-of-the-art 24HR instrument the US Department of Agriculture (USDA) Automated Multiple-Pass Method. With each pass of the multipass procedure, the process of reconsolidation alters the original memory so that by the end of the data collection period, the result will be an amalgam of multiple new memories and reports with unquantifiable error. As such, neither the researchers nor the participants know the validity or reliability of the self-reported food and beverage consumption.

FALSE REPORTING: AN INHERENT DESIGN FEATURE OF M-BMS

False Reporting and FFQs

False reports are the recollection of an event, or details of an event, that did not actually occur. False memories and recalls may be produced in multiple contexts (eg, during research, psychotherapy, and criminal investigatory interviews). Although research has demonstrated that false memories of ingestive behavior and subsequent false reporting of foods occur in laboratory settings, there is a larger literature base outside of nutrition. The Deese-Roediger and McDermott (DRM) paradigm is commonly used in research settings to elicit false reports. In this protocol, a list of semantically related words (eg, breakfast, bacon, sausage, orange juice, and cereal) are presented or read to participants. After a delay (minutes to days), participants are asked to report the words they remember. The mere presentation of lists of
Recent research has examined the effects of related but nonpresented words (eg, critical lures, eg, the word egg in the previous example). The DRM paradigm is so effective at inducing false reports that memory distortions occur even in the small percentage of individuals with highly superior memories. With the DRM paradigm, respondents are often more confident in their false reports than in the presented words.

Researchers familiar with FFQs will recognize that, by design, FFQs mimic the DRM protocol in that lists of semantically related words (ie, foods and beverages) are presented and respondents are expected to provide a response. Given that FFQs mimic the procedures designed to produce false recall, it is not surprising that FFQs with longer lists of semantically related words elicit more responses. Given the vast literature demonstrating misreporting with FFQs and the parallel literature on the extremely high level of false reports using the DRM paradigm, it is not a question of whether FFQs induce false reporting but to what extent. As stated previously, neither the researchers nor the participants know the validity or reliability of the reported food and beverage consumption, and neither can they quantify the error induced via false reporting. As we discuss in a later section, the inability of current nutrition epidemiologic research designs to independently falsify or confirm M-BM data renders the error due to false reports unquantifiable and, therefore, inadmissible as scientific data.

**False Reporting and the WWEIA-NHANES 24HR**

Recent research has examined the effects of creating “false memories for food preferences and choices.” Their work has established that it is relatively simple to “implant false beliefs and memories regarding a variety of early childhood food-related experiences.” We assert that false memories and reports are induced via the NHANES interview protocol itself, as has been demonstrated in other interviewing contexts. The factors that potentially induce false memories and reporting are well established. For example, the development of a rapport between an authority figure and respondents followed by the use of guided imagery, silence in responding, repetition, props, suggestive or repeated questioning, and encouragement to reminisce, imagine, or elaborate on past behaviors have all been shown to increase false recall. All of these factors are explicitly described in the training manual for the research personnel who conduct the NHANES 24HR. The use of rapport, silence, imagery, props, repeated questioning, eye contact, and “expectant looks,” to “motivate the respondent to answer more fully” are explicit and noteworthy in the training manual. For example, the following directive is an exemplar of the potentially false memory-inducing protocol: “If you sit quietly—but expectantly—your respondent will usually think of something. Silence and waiting are frequently your best probes for a ‘don’t know’ reply. Always try at least once to obtain a reply to a ‘don’t know’ response, before accepting it as the final answer.” The use of rapport combined with repeated questioning, silence, eye contact, and expectant looks is especially coercive when applied by an authority figure in a research context. In addition, NHANES personnel are directed to ask respondents to “imagine,” and “think” about their food intake and to “encourage” and ensure that the respondents are “convinced of the importance of the survey.” Throughout the manual there are examples of guided imagery and suggestive questioning, such as directing participants to begin “thinking about where you were, who you were with, or what you were doing, like working, eating out, or watching television,” and directives such as, “Your own state of mind—your conviction that the interview is important—will strongly influence the respondent’s cooperation. Your belief that the information you obtain will be significant and useful will help motivate the respondent to answer fully . . . .” Although the NHANES training manual states that “[t]his methodology is designed to maximize respondents’ opportunities for remembering and reporting foods they have eaten,” the scientific literature on false memories and recall strongly supports the contention that the NHANES M-BM generates significant false reporting. Given that imagination and coercive techniques (eg, the use of silence) are known to increase the probability of illusory (ie, false) recollections, it may be that
most 24HR data are false reports. If true, the NHANES 24HR is a mere exercise in number generation, and, therefore, by design, it does not provide proxy estimates of energy or nutrient consumption. This premise provides an empirically supported explanation why most M-BM data are implausible and have trivial relationships with reality (ie, actual EI and nutrient intake.) Nevertheless, without objective corroborations it is impossible to quantify what percentage of the recalled foods and beverages are completely false, grossly inaccurate, or somewhat congruent with actual consumption. Regardless, it is clear that people consistently “remember [and report] events that never happened.”

**THE INADMISSIBILITY OF M-BM DATA**

**Criteria for Scientific Research: Observable, Measurable, and Falsifiable**

Although the terms science and research are used interchangeably, they are not synonymous. Science is more than mere data collection; it is an attempt to discover order, a potentially self-correcting, explanatory, and predictive process that demonstrates lawful relations (eg, diets high in vitamin C prevent scurvy). In contrast, research is simply the process of collecting information, and many forms of research do not meet the rigor necessary for the results to be scientific. There is a long history of efforts to formally demarcate scientific from nonscientific and pseudoscientific data, the most famous of which may be Popper’s falsifiability criterion. For example, in US jurisprudence, the Daubert standard provides the rules of evidence for the admissibility of expert testimony. The criterion of falsifiability is central to expert scientific testimony and was used by Judge William Overton in ruling in McLean v. Arkansas Board of Education. This case determined that creation science was not a science because it was not falsifiable and, therefore, could not be taught as science in Arkansas public schools. As we detail in later sections, we assert that M-BM data are akin to creation science in that they fail to meet the basic requirements of scientific research.

Although philosophers continue to debate demarcation criteria, practicing scientists must set forth principles from which to judge the admissibility of data in scientific research. We extend Popper’s criterion and proffer the following widely accepted principles of scientific inquiry. First, for results to be scientific, the study’s protocols must produce outcomes that are subject to replication. To accomplish this goal, the data must be (1) independently observable (ie, accessible by others), (2) measurable, (3) falsifiable, (4) valid, and (5) reliable. These nonmetaphysical criteria were first suggested by Roger Bacon in the 13th century and later were elaborated on by the “father of empiricism,” Sir Francis Bacon, in the late 16th century. They were again reiterated by Sir Isaac Newton in the 17th century and have been subsequently clarified and defined. The skepticism and empirical rigor inherent in these criteria are of such importance to science that The Royal Society of London, the oldest scientific society in the modern world, succinctly summarized them in its motto, Nullius in Verba. This phrase, derived from Horace’s Epistles, is translated as “on the word of no one” or “take no one’s word for it” and suggests that scientific knowledge should be based not on authority, rhetoric, or mere words but on objective evidence.

The first 3 criteria (ie, independently observable, measurable, and falsifiable) define the phenomena that are in the domain of science (ie, able to be examined via the scientific method), and the final 2 criteria (ie, validity and reliability) refer to the concordance between a measurement and its referent as well as the error associated with the measurement protocols used to collect the data. Together, the 5 basic tenets distinguish scientific research from mere data collection and pseudoscience. For example, if someone is eating an apple, his or her behavior can be independently observed, measured, and verified or refuted. Yet, if he or she reports eating an apple at some point in the past (eg, with an FFQ or 24HR), neither the past behavior nor the neural correlates of the memory of that behavior are independently observable or quantifiable, and without additional information, his or her statement cannot be falsified or confirmed. It is a rather obvious fact that the respondent is the only person who has access to the raw data of M-BMs (ie, his or her memories of consumption). As such, researchers cannot examine the validity of the
memory and base M-BM research results on their faith in the verbal report (ie, the belief that the participant is telling the truth). Nevertheless, faith and belief are basic tenets of religion, not science. The unwavering credulity of nutrition epidemiologists with respect to verbal reports is literally in direct opposition to Nullius in Verba (ie, take no one’s word for it) and skeptical, rigorous science. The confluence of these simple facts and the well-documented failure of self-reported EI to accurately correspond to reality, demonstrate that the memory and subsequent recall of ingestive behavior are not within the realm of the scientific investigation of nutrition and obesity. As the philosopher Karl Popper stated, “all the statements of empirical science must be capable of being finally decided, with respect to their truth and falsity, and it is wholly impossible to verify or refute something that cannot be directly or indirectly independently observed and measured (eg, memories).

The Pseudoscience of Nutrition Epidemiology

The term pseudoscience describes data or results that are presented as scientific but lack plausibility because they cannot be reliably, accurately, and independently observed, quantified, and confirmed or refuted. When M-BMs are examined from the perspective of the basic tenets of science, the reason for the explanatory and predictive failure of epidemiologic nutrition research becomes obvious. First and foremost, scientific conclusions cannot result from nonempirical (ie, unobserved) or subjective (ie, private, not publically accessible) data that are not subject to independent observation, quantification, and falsification. When a person provides a dietary report, the data collected are not actual food or beverage consumption but rather an error-prone and highly edited anecdote regarding memories of food and beverage consumption. As such, M-BMs do not meet the basic requirements of the scientific method and, by definition, are pseudoscientific when presented as actual estimates of energy or nutrient consumption. Two famous physicists of the 20th century, Wolfgang Pauli and Arthur Schuster, summed up the problem with pseudoscientific data eloquently when they stated, respectively, that a pseudoscientific conclusion “is not only not right, it is not even wrong.”

It is difficult to determine the empirical consequences of M-BMs because the primary data (ie, memories: private information to which the respondents have privileged access) do not meet the basic tenets of scientific methods (eg, independent observation of data, falsifiability, and accuracy). If neither the researchers nor the participants are able to quantify what percentage of the recalled foods and beverages are completely false reports, grossly inaccurate, or reports that are somewhat congruent with actual consumption, it is impossible to know the validity and the error associated with each report. As Dhurandhar et al recently suggested, the use of M-BM–based data is a context in which “…something is not better than nothing.” Given the forgoing, M-BM–derived data are inadmissible and constitute a substantial ongoing threat to nutrition and obesity research and national dietary guidelines because the greatest obstacle to scientific progress is not ignorance but the illusion of knowledge created by pseudoscientific data that are neither right nor wrong.

Nevertheless, performing rigorous science is a skill that can be learned, but only if mentors understand and practice rigorous science. Given the ubiquitous use of M-BMs over many decades, it seems that nutritional epidemiologists have eschewed the inherent rigor and skepticism of Nullius in Verba (ie, take no one’s word for it) and literally replaced it with Totius in Verba (ie, take everyone’s word for it). As a result, skeptical rigorous science is not practiced or taught in nutrition and obesity epidemiologic research.

NATIONAL NUTRITION SURVEILLANCE: M-BM DATA AND USDA FOOD AVAILABILITY ECONOMIC DATA

If the 2 major components of US national nutritional surveillance are valid (ie, NHANES M-BM data and USDA Food Availability economic data), estimates from these surveillance tools should track together and independently provide population-level approximations of trends in food consumption or use. Nevertheless, history demonstrates that this is not the case. Trends in estimates of macronutrient consumption from population-level epidemiologic
surveys (ie, M-BMs) exhibited statistically significant trends that were in opposition to those of USDA economic data for fat, carbohydrates, protein, and energy (ie, kilocalories per day) from the 1960s to the late 1980s.\textsuperscript{116} It should be apparent that US residents could not be simultaneously consuming more and less fat, protein, carbohydrates, and energy over time. The contradictory patterns and striking lack of correspondence between the 2 primary US nutrition surveillance tools suggest that 1 or more likely both protocols are invalid. As with the severe misreporting demonstrated across the globe,\textsuperscript{45,80} these contradictory patterns are not limited to the United States; many countries exhibit considerable disparity between national surveillance via M-BMs and economic/food supply data.\textsuperscript{117-120} This fact is further evidence that M-BMs are fatally flawed and diet-health inferences from M-BM\textemdash derived data are meaningless.

**PA AND CRF: ESSENTIAL ELEMENTS IN NUTRITION, OBESITY, AND HEALTH RESEARCH**

The lack of explanatory and predictive power of epidemiologic nutrition research may also be explained by the limited acknowledgement of nonnutritional determinants of health and disease, such as nongenetic evolution,\textsuperscript{6-8} PA,\textsuperscript{121,122} CRF,\textsuperscript{123} and other components of nutrient partitioning and energy balance.\textsuperscript{124-130} For example, more than 50 years ago the Food and Agriculture Organization of the United Nations and the World Health Organization determined that human food energy requirements should be estimated using TEE and that PA and basal energy expenditure were the primary determinants.\textsuperscript{131,132} Yet, most nutrition research does not measure any form of energy expenditure or objectively quantify PA. Currently, there is only 1 manuscript of which we are aware that uses the NHANES objectively measured PA data to directly assess nutrition-related outcomes\textsuperscript{133} and no nutrition-related publications that include the NHANES treadmill CRF data in analyses.\textsuperscript{134} The lack of publications may be due to the fact that only 2 waves in the more than 40-year history of the NHANES include objective measures of PA, and despite the widespread acknowledgment of the necessity of daily PA for health and well-being, it is routinely discounted by governmental public health funding agencies. For example, PA, CRF, and exercise are not even listed on the National Institutes of Health’s spreadsheet of categorical spending of nearly 250 classifications through 2016.\textsuperscript{9} This is unfortunate given that 80% of Americans are not at risk for most nutritional deficiencies,\textsuperscript{3} but 95% of Americans are at risk for PA deficiency (ie, inactivity or high sedentary behavior) and do not meet the federal recommendations of 30 minutes per day of moderate to vigorous PA.\textsuperscript{134}

Given that PA and CRF are major determinants of health\textsuperscript{122,123,133,135-137} and that PA is the only major modifiable determinant of TEE and nutrient-energy partitioning (ie, the metabolic fate of the foods we consume),\textsuperscript{6,124-130,133} it is clear that PA and CRF must be objectively measured and controlled for in analyses if the health effects of any dietary intervention are to be examined accurately. Yet, because PA questionnaires are susceptible to many of the same systematic biases\textsuperscript{75,138,139} and inadmissibility issues as M-BMs, the failure to objectively measure PA and control for it in analyses renders health inferences from previous nutrition epidemiologic studies moot. Fortunately, for the science of health and disease, there are objective tools for the measurement of PA (eg, pedometers and accelerometry-based PA monitors),\textsuperscript{140} and despite limitations,\textsuperscript{141} these should be used in place of surveys and questionnaires to quantify PA in future examinations of health and disease.

**SUMMARY AND FUTURE DIRECTIONS**

A wise man proportions his belief to the evidence.

David Hume\textsuperscript{142, p87}

This critical review provides empirical and analytic evidence to support the position that (1) M-BM estimates of EI and nutrient intake have trivial relationships with actual EI and nutrient intake; (2) the assumption that human memory and recall provide literal, accurate, or precise reproductions of past ingestive behavior is indisputably false; (3) M-BMs require participants to submit to protocols that mimic procedures known to induce false recall; (4) the subjective (ie, private, not publicly accessible) mental phenomena (ie, memories) from which M-BM data are derived are not subject to independent observation, quantification,
falsification; therefore, these data are pseudoscientific and inadmissible in scientific research; and (5) the failure to objectively measure and control for PA and CRF in analyses renders inferences regarding most diet-health relationships moot.

Given the overwhelming evidence in support of our hypotheses, we conclude that M-BM data cannot be used to inform national dietary guidelines and that continued funding of M-BMs constitutes an unscientific and major misuse of research resources. In addition, given that there are objective data on the nutrition-related health status of Americans,2 we find the DGAC’s reliance on M-BMs to be without scientific support or merit. We think that skepticism and rigor are essential requirements in scientific investigations, and we fault the overly credulous nature of nutrition epidemiology for the obvious and well-demonstrated failures of the scientific community to properly inform previous federal dietary guidelines (eg, cholesterol consumption).30,46,74 We think that our nation’s dietary guidelines should not be based on the pseudoscientific and highly edited anecdotes of M-BMs, and although others may disagree, we ask that they do as we have done and provide empirical evidence rather than rhetoric to support their positions. Without valid evidence, the dogmatic defense of illusory knowledge and the status quo in nutrition and obesity research (eg, see previous commentaries and guidelines30,46,74) is an impediment to scientific progress and empirically supported public nutrition and obesity policy.

We began this critical review with evidence that our nation’s food supply and the nutritional status of Americans have improved to a level unparalleled in human history.2,3,5 Given this reality and recent work on the intergenerational transmission of obesity and type 2 diabetes mellitus,6-8 we posit that the American diet is no longer a significant risk factor for disease for most individuals. This hypothesis is supported by multiple lines of evidence, such as a 40% decline in the age-adjusted mortality rate from 1969 to 2010,144 a progressive decades long reduction in age-adjusted cardiovascular disease incidence and mortality,145,146 and a 1.5% per annum reduction in age-adjusted mortality rates from all major cancers as well as significant reductions in lung cancer incidence in men and women between 2001 and 2010.147 Given the foregoing and the evidence presented herein demonstrating the pseudoscientific nature of M-BMs, we assert that research efforts and funding of M-BMs and diet-health research are misdirected and argue that those resources would be better targeted to the most prevalent disease of deficiency of the 21st century: inactivity (ie, a lack of PA and exercise and high levels of sedentary behavior).121,134

CONCLUSION
In this critical review, we argued that the essence of science is the ability to discern fact from fiction, and we presented evidence from multiple fields to support the position that the data generated by nutrition epidemiologic surveys and questionnaires are not falsifiable. As such, these data are pseudoscientific and inadmissible in scientific research. Therefore, these protocols and the resultant data should not be used to inform national dietary guidelines or public health policy, and the continued funding of these methods constitutes an unscientific and major misuse of research resources.

ACKNOWLEDGMENTS
The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Abbreviations and Acronyms: BMR = basal metabolic rate; CRF = cardiorespiratory fitness; DRM = Deese-Roediger and McDermott; DGAC = Dietary Guidelines Advisory Committee; EI = energy intake; FFQ = food frequency questionnaire; M-BM = memory-based dietary assessment method; NHANES = National Health and Nutrition Examination Survey; PA = physical activity; rEI = reported energy intake; TEE = total energy expenditure; USDA = US Department of Agriculture; WWEIA = What We Eat in America; 24HR = 24-hour dietary recall

Grant Support. This work was supported by grant T32DK062710 (E.A., G.P.) from the National Institute of Diabetes and Digestive and Kidney Diseases of the National Institutes of Health.

Potential Competing Interests. Dr Archer has received honoraria from the International Life Sciences Institute and The Coca-Cola Company. Dr Lavie reports receiving consulting fees and speaking fees from The Coca-Cola Company and writing a book on the obesity paradox with potential royalties.

Correspondence. Address to Edward Archer, PhD, Office of Energetics, Nutrition Obesity Research Center, University
REFERENCES


43. Heinlitz BL, Milman N, Hansen GL. Relationship between dietary iron intake, corrected for diet reporting error, and...


