

# Strategies to Optimize the Impact of Nutritional Surveys and Epidemiological Studies<sup>1,2</sup>

Densie Webb,<sup>3\*</sup> Margaret M. Leahy,<sup>4</sup> John A. Milner,<sup>5</sup> David B. Allison,<sup>6</sup> Kevin W. Dodd,<sup>7</sup> P. Courtney Gainey,<sup>8</sup> Robert A.J. Matthews,<sup>9</sup> Barbara O. Schneeman,<sup>10</sup> Katherine L. Tucker,<sup>11</sup> and S. Stanley Young<sup>12</sup>

<sup>3</sup>Austin, TX; <sup>4</sup>The Coca-Cola Company, Atlanta, GA; <sup>5</sup>USDA, Agricultural Research Service, Beltsville, MD; <sup>6</sup>University of Alabama, Birmingham, AL; <sup>7</sup>National Cancer Institute, Bethesda, MD; <sup>8</sup>North American Branch of the International Life Sciences Institute, Washington, DC; <sup>9</sup>Aston University, Birmingham, UK; <sup>10</sup>University of California, Davis, CA; <sup>11</sup>Northeastern University, Boston, MA; and <sup>12</sup>National Institute of Statistical Sciences, Research Triangle Park, NC

## ABSTRACT

The development of nutrition and health guidelines and policies requires reliable scientific information. Unfortunately, theoretical considerations and empirical evidence indicate that a large percentage of science-based claims rely on studies that fail to replicate. The session “Strategies to Optimize the Impact of Nutrition Surveys and Epidemiological Studies” focused on the elements of design, interpretation, and communication of nutritional surveys and epidemiological studies to enhance and encourage the production of reliable, objective evidence for use in developing dietary guidance for the public. The speakers called for more transparency of research, raw data, consistent data-staging techniques, and improved data analysis. New approaches to collecting data are urgently needed to increase the credibility and utility of findings from nutrition epidemiological studies. Such studies are critical for furthering our knowledge and understanding of the effects of diet on health. *Adv. Nutr.* 4: 545–547, 2013.

## Introduction

The development of nutrition and health guidelines and policies requires reliable scientific information. Unfortunately, theoretical considerations and empirical evidence indicate a large percentage of science-based claims rely on studies that fail to replicate. The session on “Strategies to Optimize the Impact of Nutrition Surveys and Epidemiological Studies” focused on the elements of design, interpretation, and communication of nutritional surveys and epidemiological studies to enhance and encourage the production of reliable, objective evidence for use in developing dietary guidance for the public.

The session addressed the question “How can we best integrate the information we have from nutrition studies and better understand it?” Dr. Leahy, co-chair, emphasized that the session goal is to foster dialogue to enhance production of reliable, objective evidence for use in developing dietary guidance.

The number of studies registered with ClinicalTrials.gov has grown 25-fold since 2000. Dr. Milner, co-chair, highlighted the urgent need to improve the quality and accuracy of designs used in the growing number of studies as well as in the quality of communication of study findings to the public.

Dr. Schneeman pointed to the importance of sound research to inform policy, and provided examples in the way of nutrition labeling regulations, development of the U.S. Dietary Guidelines for Americans, and guidelines and standards in international forums such as the WHO. Consistent, relevant data are essential for agencies to conduct evidence-based reviews that critically evaluate the totality of evidence. Ranking the level of the available scientific evidence is a vital part of that process. Dr. Schneeman posed a number of questions that must be asked to determine if scientific conclusions can be drawn from a human study, including the following: 1) Were the subjects healthy? 2) Was the disease or condition in question measured as the primary endpoint? 3) Was an appropriate control group included? 4) Were there

<sup>1</sup> This article is a summary of the symposium “Strategies to Optimize the Impact of Nutritional Surveys and Epidemiological Studies” held April 20, 2013 at the ASN Scientific Sessions and Annual Meeting at Experimental Biology 2013 in Boston, MA. The symposium was sponsored by the American Society for Nutrition and the North American Branch of the International Life Sciences Institute’s Committee on Carbohydrates. This is a free access article, distributed under terms (<http://www.nutrition.org/publications/guidelines-and-policies/license/>) that permit unrestricted noncommercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

<sup>2</sup> Author disclosures: M. M. Leahy is Vice-Chair of International Life Sciences Institute (ILSI) North America’s Carbohydrates Committee. J. A. Milner is an ILSI Global, IFIC, and McCormick Science Institute Board Member. D. B. Allison has received financial and other benefits from the Frontiers Foundation, Federal Trade Commission, and other government, nonprofit, and for-profit organizations. K. L. Tucker is a scientific advisor to ILSI North America and has received support from the Kraft Foods Group. S. S. Young is a member of the OmicSoft Scientific Advisory and the *AJCN* Editorial Boards. K. W. Dodd, P. C. Gainey, R. A. J. Matthews, and B. O. Schneeman, no conflicts of interest.

\* To whom correspondence should be addressed. E-mail: dwebb1@austin.rr.com.

relevant differences between control and treatment at baseline? and 5) In observational studies, was the substance a food or food component? Examples of fatal study flaws were presented, including no control, a lack of relevant statistics, key confounders or risks that were not controlled, and non-validated biomarkers used as endpoints.

Dr. Tucker discussed issues researchers face in assessing usual intakes in diverse populations. Assessment of dietary intakes is essential, but diets are complex and constantly changing. The methods currently used, mainly diet records, FFQs, and 24-h recalls, are subject to significant error and bias. Diet records require a certain amount of literacy and cooperation and may pose challenges for some populations. Twenty-four-hour recalls are a valid but only short-term measure of dietary intake. FFQs offer a long-term measure of usual intakes, but the use of different questionnaires may bias group comparisons. In addition, because diets are constantly changing, questionnaires may become obsolete as soon as they are developed. Furthermore, FFQs may not be specific enough for diverse groups and precision will vary among subgroups of the population. As a result, day-to-day variation cannot be assumed to be random. This can lead to systematic bias in subgroups. To obtain valid estimates for diverse populations, much more detailed dietary data are required. Existing statistical corrections are based on variance measurements and assumptions from the average population and do not consider major dietary pattern subsets. Using weighted average intakes removes important variation, but this differs with diverse cultural diets, which leads to confounding. A primary point made during the presentation was that improvements in dietary assessment methods are urgently needed to understand the effects of genetic factors on disease risk but that there are no shortcuts. Methods that combine FFQs with multiple 24-h recalls are currently the best, and most feasible, assessment techniques available for assessing diet composition.

Dr. Young pointed out that when tested rigorously, claims resulting from human medical observational studies often fail to replicate. Whereas randomized clinical trial findings replicate over 80% of the time, the findings of medical observational studies replicate only 10–20% of the time. However, in observational studies, if enough questions are asked and enough *P* tests run, a positive result will eventually emerge. For example, if 61 questions are asked of a given data set, there is a 95% chance of obtaining at least 1 positive result. Subjecting data to a large number of research questions is one way to obtain and publish a positive result. Dr. Young recommended that if a large number of claims are being tested in a paper and only a few yield small *P* values, skepticism is warranted.

There are many human health issues that can be examined only with observational data, but systematic problems in the way observational studies are conducted and analyzed have been identified and need to be corrected. Dr. Young called on funding agencies and journal editors to help fix a broken process and for consumers to be skeptical of observational study claims. He recommended that data generation and data

analysis be funded separately, that more replication studies be funded, and that funding be contingent on making the data publicly available. The following are among the technical problems he identified with observational studies: 1) the way in which data staging is performed; 2) the lack of an analysis protocol written in advance; 3) multiple testing; 4) multiple modeling (i.e., bringing covariates in and out of analysis); 5) not taking uncorrected bias, such as missing factors, unmeasured cofounders, and loss to follow-up, into account; and 6) self-serving paper writing and press releases. To better manage systematic problems in observational studies, Dr. Young suggested that funding be conditional upon publicly posting the protocol before the study has begun and the public posting of the data set. In addition, journal editors should be encouraged to look not only at the end result but at the entire analytical process behind a study. Data staging, the process by which raw data are included or excluded for analysis, based on such factors as gender, age, weight, health, and how missing values and outliers are handled was highlighted as a seldom-examined problem in research. Although staging can be done in myriad ways, it is rarely documented and is usually not reproducible. Yet, it can dramatically change the results of the research. Dr. Young suggested that data be split into initial analysis and verification sets. The number of questions under consideration should be clearly disclosed prior to the study and statistical methods that deal with multiple testing and multiple modeling should be employed.

The goal of nutritional epidemiology is to relate dietary intakes to health outcomes. Dr. Dodd discussed how measurement error in dietary assessment can sabotage nutritional epidemiology findings. Measuring an individual's average intake over a long period of time is challenging. Any one of several types of measurement errors can arise from the interaction between the instrument used and the population under study. If these factors are ignored, it can affect the analysis in numerous ways. Small observed effect sizes, coupled with incorrect scaling of exposure, complicate interpretation of even significant results. The pros and cons of FFQs compared with 24-h dietary recalls were discussed. Also discussed was the false economy of selecting imperfect, but less costly dietary measurement instruments that may require 2–11 times the sample size to obtain a significant result compared with using a better, but more expensive instrument. It was also pointed out that although statistical techniques used to correct for measurement error theoretically require a true reference instrument in conjunction with an error prone measure, in many cases adjusting with a better, but still imperfect instrument is preferable to ignoring the problem of measurement error. Research is ongoing to develop better statistical methods for observational studies. Dr. Dodd offered suggestions for using better dietary intake instruments and statistical methods in the future.

Robert Matthews highlighted long-standing concerns about the very concept of significance and the validity of the traditional practice of using *P* values in the analysis of dietary assessment studies. Despite widespread belief, they do not give the probability that the null hypothesis is false. Furthermore,

they do not allow the new findings to be combined with extant (prior) evidence, the key process in assessing the plausibility of a new finding. Matthews emphasized his belief that failure to include prior evidence leads to unreliable inferences; although the results of a study may be “statistically significant,” they may not be plausible. Matthews showed how Bayesian statistical methods would allow analysis to move on from a pass/fail dichotomy, assessing both significance and plausibility of new findings. Matthews then summarized his presentation with examples illustrating the benefits of Bayesian statistical analysis for observational studies.

Dr. Allison presented simple steps that could improve the use, reporting, and interpretation of epidemiologic research findings. Dr. Allison pointed out that problems with research integrity have to do with human interests that create conflicts. Too much time and too many resources and journal pages are devoted to research that increases belief rather than knowledge. A positive step would be to make mandatory the publication of all studies performed at nonprofit institutions, with human subjects or funded with government or philanthropic funds, whether the findings are positive or negative. Another suggestion was to maintain consistency between trial registration and published papers. Like Dr. Young, Dr. Allison strongly advocated that raw data be made publicly accessible. That said, he acknowledged the substantial challenges in doing so, but rather than throwing up our hands in defeat, he suggested that the scientific community begin moving toward making nearly all raw data publicly available. As a first step toward that goal, journal editors could, in appropriate circumstances, require data sets to be deposited to

a repository prior to publication of a paper. Once they are available, raw data should be made publicly accessible. Journal editors should require data sets to accompany the publication of a paper.

Currently, press releases from media offices of institutions and journals often include distortions of findings that are then passed on to the public by journalists who use the press release as their primary source. To avoid this, scientific journals could be provided with resources to issue peer-reviewed press releases.

In summary, reliable scientific information is essential for nutrition policy development, yet a considerable number of observational studies cannot be replicated and their findings are distorted when presented to the public. Speakers at the session called for more transparency of research raw data, consistent data-staging techniques, and improved data analysis. Diet measurement problems represent another major issue. The field of nutrition is at a turning point, interacting with other fields such as genetics. However, it may not be possible to effectively study gene/diet interactions and the findings could have limited interpretability if the diet cannot be measured with reasonable validity. New approaches to collecting data are urgently needed to increase the credibility and utility of findings from nutrition epidemiological studies. Such studies are critical resources for furthering our knowledge and understanding of the effects of diet on health.

### **Acknowledgments**

All authors read and approved the final manuscript.