Some nutrition scientists and much of the public often consider epidemiologic associations of nutritional factors to represent causal effects that can inform public health policy and guidelines. However, the emerging picture of nutritional epidemiology is difficult to reconcile with good scientific principles. The field needs radical reform.

In recent updated meta-analyses of prospective cohort studies, almost all foods revealed statistically significant associations with mortality risk.1 Substantial deficiencies of key nutrients (eg, vitamins), extreme over-consumption of food, and obesity from excessive caloric intake may indeed increase mortality risk. However, small intake differences of specific nutrients, foods, or diet patterns with similar calories causally, markedly, and almost ubiquitously affect survival?

Assuming the meta-analyzed evidence from cohort studies represents life span–long causal associations, for a baseline life expectancy of 80 years, eating 12 hazelnuts daily (1 oz) would prolong life by 12 years (ie, 1 year per hazelnut).1 Drinking 3 cups of coffee daily would achieve a similar gain of 12 extra years,2 and eating a single mandarin orange daily (80 g) would add 5 years of life.1 Conversely, consuming 1 egg daily would reduce life expectancy by 6 years, and eating 2 slices of bacon (30 g) daily would shorten life by a decade, an effect worse than smoking.1 Could these results possibly be true? Authors often use causal language when reporting the findings from these studies (eg, “optimal consumption of risk-reducing foods results in a 56% reduction of all-cause mortality”).1 Burden-of-disease studies and guidelines endorse these estimates. Even when authors add caveats, results are still often presented by the media as causal.

These implausible estimates of benefits or risks associated with diet probably reflect almost exclusively the magnitude of the cumulative biases in this type of research, with extensive residual confounding and selective reporting.3 Almost all nutritional variables are correlated with one another; thus, if one variable is causally related to health outcomes, many other variables will also yield significant associations in large enough data sets. With more research involving big data, almost all nutritional variables will be associated with almost all outcomes. Moreover, given the complicated associations of eating behaviors and patterns with many time-varying social and behavioral factors that also affect health, no currently available cohort includes sufficient information to address confounding in nutritional associations.

Furthermore, the literature is shaped by investigators who report nonprespecified results that are possible to analyze in very different ways.4 Consequently, meta-analyses become weighted averages of expert opinions. In an inverse sequence, instead of carefully conducted primary studies informing guidelines, expert-driven guidelines shaped by advocates dictate what primary studies should report. Not surprisingly, an independent assessment by the National Academies of Sciences, Engineering, and Medicine of the national dietary guidelines suggested major redesign of the development process for these guidelines: improving transparency, promoting diversity of expertise and experience, supporting a more deliberative process, managing biases and conflicts, and adopting state-of-the-art processes.5

Proponents of the status quo may maintain that the true associations are even larger than what are reported because of attenuation from nondifferential misclassification. Indeed, self-reported data have error,6 but there is no guarantee it is nondifferential. Nevertheless, if error is nondifferential and estimated effects are attenuated, reported results become even more implausible: eating 12 hazelnuts daily would increase life expectancy by 20 to 30 years, not just 12 years.

Individuals consume thousands of chemicals in millions of possible daily combinations. For example, there are more than 250 000 different foods and even more potentially edible items, with 300 000 edible plants alone. Seemingly similar foods vary in exact chemical signatures (eg, more than 500 different polyphenols). Much of the literature silently assumes disease risk is modulated by the most abundant substances; for example, carbohydrates or fats. However, relatively uncommon chemicals within food, circumstantial contaminants, serendipitous toxicants, or components that appear only under specific conditions or food preparation methods (eg, red meat cooking) may be influential. Risk-conferring nutritional combinations may vary by an individual’s genetic background, metabolic profile, age, or environmental exposures. Disentangling the potential influence on health outcomes of a single dietary component from these other variables is challenging. If not impossible.

To use an analogy from genetics, studying associations of specific foods is like studying whether large chromosomal regions increase mortality risk. For decades, genome linkage scans struggled to link large chromosomal areas to disease risk. According to current knowledge, these previous efforts were doomed: each chromosomal area contains thousands of genetic variants. Linkage scans resulted in numerous articles, but limited useful information. Retrospectively, using a few hundred microsatellite markers to study an entire genome with many million polymorphisms seems naive. Similarly, limited self-reported nutrition data ascertained with a handful of questions and self-reported items fail to acknowledge or accurately measure a system that matches or exceeds the genome in complexity.

Beyond food studies, results of single-nutrient studies have largely failed to be corroborated in randomized trials. False-positive associations are common in the literature. For example, updated meta-analyses of published...
In recent years, many studies have been conducted to investigate the role of nutrition in health outcomes. For example, the PREDIMED study, a trial of the Mediterranean diet, has shown a benefit on mortality compared to a control group.

However, there are several challenges and limitations in nutritional epidemiology. One of the biggest challenges is the difficulty in establishing causality. Epidemiological findings often lack the precision to be considered causal, especially after multiple large trials have yielded conflicting results.

Moreover, there are several other factors that contribute to the difficulty of nutritional epidemiology. For example, the public is often interested in diet and nutrition, but these interests can be misleading. The media often portrays extreme diets that are not sustainable, which can confuse the public and detract from the agenda of preventing and treating obesity.

There are also several challenges in nutritional research that contribute to the difficulty of establishing causality. For example, the use of small trials that are not powered for the outcome of interest can lead to misleading conclusions. A recent example is the retraction of the results of the PREDIMED study after it was realized that there were multiple subversions of randomized trials.

In conclusion, nutritional epidemiology is a complex field that requires careful study and analysis. Researchers should be transparent about their methods and results, and they should strive to conduct high-quality trials that are designed to answer specific research questions. The nutritional epidemiology community should also work to improve research practices and reduce the potential for bias in these studies.

REFERENCES

© 2018 American Medical Association. All rights reserved.