We need more randomized trials in nutrition—preferably large, long-term, and with negative results

John PA Ioannidis*

Stanford Prevention Research Center, Meta-Research Innovation Center at Stanford (METRICS), and Departments of Medicine, Health Research and Policy, and Statistics, Stanford University, Stanford, CA

The key to optimal nutrition choices has been sought mostly in tens of thousands of observational analyses as well as in expert editorials (like mine) and recommendations. Epidemiology and opinion grossly outnumber randomized trials in this important field. Therefore, having 4 randomized trials of nutrition interventions published in the same issue of the Journal is an opportunity for celebration. It is even more rewarding that 3 of these trials have sample sizes >100 (and 1 trial has a sample size >1500) and 2 have long-term follow-up. The best news is that all 4 trials have “negative” results for some or all primary outcomes. DHA supplementation during pregnancy did not reduce BMI or fat mass in children at aged 5 y (1). A reduction in the n-6:n-3 long-chain PUFA ratio during pregnancy and lactation had no major impact on offspring body composition after 5 y (2). A trial of amino acid regimen and intravenous lipid composition in preterm parenteral nutrition showed largely nonsignificant differences between compared regimens (3). Finally, a small trial that compared fructose, sucralose, and sucrose acute effects after solid meals showed no differences in triglyceride concentrations (4). Admittedly, this last trial also showed significant differences in other outcomes. Although I would have preferred purely “negative” results, my disappointment is tempered, because the “positive” results pertained to physiologic rather than to major clinical outcomes.

Randomized trials have had a rough time establishing themselves in nutrition research. Influential epidemiologists have expressed skepticism about randomized trials for nutrition or lifestyle interventions in general (e.g., reference 5). Nonrandomized observational studies have been favored because people change their choices; thus, crossover, cross-in, withdrawals, and poor adherence are considered impediments for clinical trials, and the estimation of effect sizes by intention-to-treat may misrepresent the mechanistic effects.

These arguments are actually some of the reasons why randomized trials are not only preferable but also cannot be replaced by epidemiologic studies. Experiments are needed to study mechanisms and acute-change physiology, and the best way to perform well-controlled experiments is by randomization. Responses can be evaluated accurately with short-term data collection for physiologic changes. For studies that aim to assess long-term outcomes with clinical portend, epidemiologic studies are just too unreliable to capture subtle treatment effects whenever noise due to confounding and other biases exceeds plausible signals. For most clinical outcomes (as opposed to physiology or surrogate outcomes), the magnitude of the treatment effects of nutritional interventions is probably small or modest (6). Adherence, withdrawals, and treatment fidelity problems are not necessarily insurmountable and methods exist to improve retention (7). Moreover, any useful intervention should have mechanistic efficacy but should also be tolerable enough that people can adhere to it. If an intervention is mechanistically efficacious but cannot be used long term, it is practically the same as if it lacked efficacy to start with.

The combination of small effect sizes and unknown adherence means that these questions can only be addressed reliably with large, well-powered studies and long-term follow-up. Chasing small or modest effects with underpowered, small studies is an ideal recipe for getting both false-negative and even more false-positive results (8). Long-term follow-up is necessary both for increasing the power for patient-relevant clinical events as well as for making the results relevant for real life. People care mostly about what nutrition can achieve for them in the long term, not just whether they can improve some metabolic variables or their weight for a few months.

In this setting, “negative” results are much more desirable than “positive” ones. Let me explain this paradox. First, many “negative” results are not failures to show the superiority of 1 nutritional intervention compared with another but offer evidence of noninferiority or equivalence. They mean that different nutrition choices lead to equally acceptable outcomes. This is great news: it is better to have many equally good choices on what to eat than just one. Second, other “negative” results indeed show that nutrients or diets that we thought would be effective (e.g., on the basis of observational epidemiology) are not. The information gain (informativity, entropy change) from the results of any study depends on how much the study changes our previous
beliefs (9). In nutrition, there has been so much observational and mechanistic research that thousands of spuriously significant associations have already been produced and translated in heavily opinionated, debated recommendations. Getting another significant result in a field that is already saturated with so many significant results offers no information gain: we still (think we) know what (we thought) we knew. Conversely, “negative” results offer high information gain, because they change our probably false beliefs about potentially effective interventions. Looking at the list of large randomized trials that refuted epidemiologic associations of nutrients and diets (10), we should hope to get more “negative” results in the future.

Of some concern, the uncontrolled multiplicity that results in numerous false-positive results in epidemiologic studies can also affect nutrition randomized trials, when there are multiple outcomes analyzed in multiple different ways and selectively reported. To avoid this bias, we need complete preregistration of trial protocols with clear definitions of the outcomes and analyses to be performed.

Finally, cost is an undeniable issue for randomized trials. Nonrandomized studies are far cheaper, and analyses on existing databases can readily be streamlined for mass production by armies of nutrition doctoral and postdoctoral students. Nevertheless, it is difficult to fathom the cost of running zillions of such analyses or the induced cost of getting wrong conclusions that mislead research agendas to get even more wrong studies and mislead people, societies, and organizations to make wrong nutrition choices. Moreover, randomized trials on nutrition are not a total rarity. An evaluation of clinicaltrials.gov showed 4375 registered interventional trials on nutrition/diet as of August 2014 (3.1% of all registered intervention trials) (11) and many more trials are unregistered. However, the vast majority of these trials are small, without clear primary outcomes, and are selectively reported and published. Perhaps the same resources might suffice to conduct a smaller number of large-scale, long-term, fully registered, and transparent trials that would be fully and appropriately reported. Then, I wouldn’t be disappointed to also see some informative “positive” results.

The author had no conflicts of interest to declare.

REFERENCES
1. Muhlhausler BS, Yelland LN, McDermott R, Tapsell L, McPhee A, Gibson RA, Makrides M. DHA supplementation during pregnancy does not reduce BMI or body fat mass in children: follow-up of the DHA to Optimize Mother Infant Outcome randomized controlled trial. Am J Clin Nutr 2016;103:
3. Gallagher C, Keogh JB, Pedersen E, Clifton PM. Fructose acute effects on glucose, insulin, and triglyceride after a solid meal compared with sucralose and sucrose in a randomized crossover study. Am J Clin Nutr 2016;103: