Perspective: Limiting Dependence on Nonrandomized Studies and Improving Randomized Trials in Human Nutrition Research: Why and How

John F Trepanowski[1] and John PA Ioannidis[1,2,3,4,5,6]

[1]Stanford Prevention Research Center; [2]Meta-Research Innovation Center at Stanford (METRICS); and Departments of [3]Medicine, [4]Health Research and Policy, [5]Biomedical Data Science, and [6]Statistics, Stanford University, Stanford, CA

ABSTRACT

A large majority of human nutrition research uses nonrandomized observational designs, but this has led to little reliable progress. This is mostly due to many epistemologic problems, the most important of which are as follows: difficulty detecting small (or even tiny) effect sizes reliably for nutritional risk factors and nutrition-related interventions; difficulty properly accounting for massive confounding among many nutrients, clinical outcomes, and other variables; difficulty measuring diet accurately; and suboptimal research reporting. Tiny effect sizes and massive confounding are largely unfixable problems that narrowly confine the scenarios in which nonrandomized observational research is useful. Although nonrandomized studies and randomized trials have different priorities (assessment of long-term causality compared with assessment of treatment effects), the odds for obtaining reliable information with the former are limited. Randomized study designs should therefore largely replace nonrandomized studies in human nutrition research going forward. To achieve this, many of the limitations that have traditionally plagued most randomized trials in nutrition, such as small sample size, short length of follow-up, high cost, and selective reporting, among others, must be overcome. Pivotal mega-trials with tens of thousands of participants and lifelong follow-up are possible in nutrition science with proper streamlining of operational costs. Fixable problems that have undermined observational research, such as dietary measurement error and selective reporting, need to be addressed in randomized trials. For focused questions in which dietary adherence is important to maximize, trials with direct observation of participants in experimental in-house settings may offer clean answers on short-term metabolic outcomes. Other study designs of randomized trials to consider in nutrition include registry-based designs and "N-of-1" designs. Mendelian randomization designs may also offer some more reliable leads for testing interventions in trials. Collectively, an improved randomized agenda may clarify many things in nutrition science that might never be answered credibly with nonrandomized observational designs. *Adv Nutr* 2018;9:367–377.

Keywords: nutritional sciences, observational study, epidemiology, randomized controlled trial, research design

Introduction

In human nutrition science, nonrandomized observational studies outnumber randomized trials by a wide margin (**Text Box 1**). Here, we argue that this should no longer continue. A nonrandomized observational study of nutrition can produce positive value only when it probes large effect sizes (large in comparison to noise and biases) in a context in which randomization would be unethical or otherwise unfeasible. Severe nutritional deficiencies and other exceptional circumstances can therefore be examined reliably with epidemiologic research, but everyday questions about modest differences in dietary intake require random allocation of exposure for trustworthy answers. Overreliance on

This article is a review from the session "Observational Studies vs. RCTs in Nutrition: What Are the Tradeoffs?" presented at the 6th Annual Advances & Controversies in Clinical Nutrition Conference held 8–10 December 2016 at the Rosen Shingle Creek in Orlando, Florida. The conference was jointly provided by the American Society for Nutrition (ASN) and Tufts University School of Medicine.

Perspective articles allow authors to take a position on a topic of current major importance or controversy in the field of nutrition. As such, these articles could include statements based on author opinions or point of view. Opinions expressed in Perspective articles are those of the author and are not attributable to the funder(s) or the sponsor(s) or the publisher, Editor, or Editorial Board of *Advances in Nutrition*. Individuals with different positions of the topic of a Perspective are invited to submit their comments in the form of a Perspectives article or in a Letter to the Editor.

The Meta-Research Innovation Center at Stanford (METRICS) is funded by a grant from the Laura and John Arnold Foundation. The work of JPAI is also funded by an unrestricted gift from Sue and Bob O'Donnell and JFT is supported by the NIH (T32HL007034).

Author disclosures: JFT and JPAI, no conflicts of interest.

Address correspondence to JPAI (e-mail: jioannid@stanford.edu).

nonrandomized observational data has created widespread confusion about optimal nutrition (1, 2). Clearly, diet is important for health, and poor diet is a major contributor to the global burden of disease (3), but ambiguous and sometimes contradictory findings from nutritional epidemiology have made it difficult to identify best approaches for curtailing this burden. Many prominent epidemiologic associations (including highly cited studies on α-tocopherol, β-carotene, vitamin C, vitamin D, selenium, calcium, and low-fat diets) have not been corroborated by large randomized trials (4, 5). Discordant or even opposite results between nonrandomized observational studies and randomized trials have been summarized in meta-analyses as well (6, 7). Questionable nonrandomized data have led to dietary guidelines that did not curb the twin epidemics of obesity and type 2 diabetes, and it is unknown if the latest guidelines (8) will fare better. Progress in nutrition science may continue to be stunted until most observational research is replaced with randomized study designs.

Text Box 1. Ratio of Nonrandomized Observational Studies to Randomized Controlled Trials in Nutrition Science

On 7 April 2017, PubMed listed 511,648 papers with the keywords "diet OR nutrient OR nutrition" after filters for abstract availability and human species were both applied. We identified a random sample of 100 papers that reported results from either a nonrandomized observational study or a randomized controlled trial in the abstract. In this sample, 88 abstracts reported results from a nonrandomized observational study and 12 abstracts reported results from a randomized controlled trial, yielding a ratio of 7.3:1.

Nonrandomized observational research in nutrition is limited primarily by low signal and high noise, tiny effect sizes for nutritional risk factors and nutrition-related interventions (9), and massive confounding among densely correlated nutrients, clinical outcomes, and other variables of interest (10, 11). These largely unfixable problems pervade and cripple most epidemiologic analyses regardless of whether they use cohort, case-control, or cross-sectional designs. Dietary measurement error (12) and nontransparent research reporting (13) are 2 additional problems that undermine the credibility of nutritional epidemiology, although they are more fixable.

Random allocation of exposure can overcome some of these major problems, but there is also a need for revamping the randomized research agenda (14). The current agenda spreads resources too thinly over thousands of trials (15), with very few having the requisite size and duration for obtaining clear answers. It would be better to use these same resources instead to conduct a few dozen more-informative megatrials every decade, with many thousands of participants, long-term follow-up, and hard clinical endpoints. The megatrials would focus on pragmatic insights about nutrition that involve real-world (often low) adherence to dietary prescriptions. Although some of these trials may seem to have the disadvantage of requiring many years (or decades) of follow-up, observational nonrandomized studies have failed to give reliable answers for a century. For mechanistic problems and proof-of-concept questions in which compliance with the experimental protocol must be maximized, short-term randomized trials with in-house direct observation can be performed.

In this article, we review the inherent problems of nutritional epidemiology and the shortcomings of the current randomized research agenda and offer potential solutions for moving forward with more trustworthy nutrition science.

Inherent, Unfixable Problems in Nutrition Science

Despite all of the work that has been done in nutrition science, a healthy diet still cannot be defined professionally in a way that experts agree on—anything more specific than "eating with prudence" introduces controversy. Dietary guidelines are always argued over ferociously, and the relative merits of commonly consumed nutrients have been debated for decades (1, 2). This confusing state is not surprising when almost every single nutrient has been associated with almost any outcome in peer-reviewed publications (16). For example, not only have most nutrients been associated with cancer risk but most of the nutrients have published reports of increased risk in 1 study and decreased risk in another (17). Nutrition science has become an epidemic of questionable results. Meta-analyses of retrospectively compiled data suffering from biases and selective reporting do not necessarily make things better.

Two major epistemological problems stand out in nutrition science for being especially pervasive and difficult to fix. One problem is that many nutritional risk factors and nutrition-related interventions have tiny effect sizes for clinical outcomes, with RRs in the range of 0.95–1.05 or even 0.99–1.01 (9). For example, fruit consumption seems to have an HR for cancer risk of ~0.999/serving (100 g) (18). Tiny effects may be all that remain to be found in nutrition science, because most of the more conspicuous effects—which typically relate to either severe nutrient deficiency or excess (obesity)—have been found already (**Figure 1**). There will be many more claimed discoveries of tiny effects in the future with the emergence of "big data" (20).

Tiny effects create big controversies that cannot be settled easily. A recent example is the International Agency for Research on Cancer monograph that classified processed meat as fully proven to be carcinogenic (class 1) and red meat as a probable carcinogen (class 2A) (21). The HR for overall cancer risk may be ~1.01–1.02/serving (100 g) of red or processed meat. Even for colorectal cancer risk, where the observed effect is the strongest, a maximum HR of 1.18/50 g processed meat (22) is too small to avoid residual uncertainty given the other problems that we discuss below.

Researchers commonly try to sort out whether a newly found tiny effect is true or spurious by considering biological plausibility on the basis of external evidence

Outer ring — typical effect of smoking on lung cancer risk (RR ≥10)

Middle ring – minimum effect size needed to upgrade the strength of observational evidence using GRADE criteria (RR = 3)

Bullseye —— typical single nutrient association with myocardial infarction (protective or harmful; RR ≤ 1.05)

FIGURE 1 Epidemiologists may take aim at effects that differ greatly in size. One of the greatest achievements of observational epidemiology was the demonstration that smoking has a causal effect on lung cancer risk. In the case of smoking and lung cancer, the RR is very large (RR $\geq$ 10); this is akin to hitting the outer ring of the target. The middle ring must be hit when seeking the minimum effect size needed to upgrade the strength of observational evidence using the GRADE criteria (corresponding to RR = 3) (19). Few epidemiologic associations have such effect sizes. And the hit-to-miss ratio will be dismal when firing away at the small bullseye, which represents a typical association between a single nutrient and a clinical outcome (RR $\leq$ 1.05). GRADE, Grading of Recommendations Assessment, Development, and Evaluation.

(e.g., mechanistic data). If only highly relevant external evidence is invoked, this should greatly circumscribe the possible inferences that could be made with the primary data; however, too often, off-topic external findings are instead conscripted and forced to fight in support of whatever inference a researcher wishes to make (23). At its worst, a consideration of biological plausibility can unduly influence which primary results get published, because different results can be easily obtained (24) and selectively reported (13) depending on whatever the experts believe they should be (25).

Statistical significance has lost the authority to determine whether a tiny effect is real or illusory, now that almost all published articles report an analysis with P values <0.05 (26). Some researchers have proposed lowering the threshold to 0.005 or even 10^{-6} (27, 28). Alternatively, falsification endpoints could be used to calibrate thresholds of statistical significance (29, 30). This would involve prespecifying a number of effects known in advance to be null, and studying what P values they generate in each large database: for example, if every P value for a null effect were $>10^{-8}$, then 10^{-8} would become the cutoff for picking significant findings. These suggested approaches would reduce the risk of false positives, but they would also increase the risk of false negatives. Moreover, we still have limited experience about how these approaches could affect the sensitivity and specificity of findings in different settings

and data sets. With a massive data set, it could be easy to obtain P values $<10^{-100}$ for associations that are manifestly dubious (31).

When a field of research is saturated with tiny effects, even small errors or biases can result in innumerable misleading inferences, both by drowning out true effects and by generating spurious effects (32). This leads to the second major problem in nutrition science that is very difficult to fix. Confounding is a major problem generally, but it is an even bigger one in nutrition science, specifically as a result of very dense correlations among variables of interest. For practically every nutrient, amount of intake correlates (positively or negatively) with the intake of multiple other nutrients (10). It also correlates with many other environmental exposures (e.g., pollution), and with many variables that pertain to lifestyle, educational level, and socioeconomic status (11). Several of these exposures, including multiple nutrients, are associated with both clinical outcomes and intermediate outcomes (10, 11), but most of these significant associations probably do not reflect causal relations. Attempts to disentangle causes from spurious epiphenomena in these densely connected "correlation globes" rarely have good odds of success (10) (**Figure 2**).

For example, in an analysis with NHANES data of 317 exposures, serum *trans*-β-carotene was significantly associated with 68 other exposure variables, including 16 other nutrients (34). If any of the other exposure variables happen to have a genuine association with a major outcome (e.g., cancer), β-carotene will also seem to have that association whether it is real or not. This may explain why chemoprophylaxis with β-carotene and other antioxidants for cancer prevention was such a prevalent idea for years, and why it continues to be endorsed by some experts despite there being very strong evidence against it from multiple randomized trials (35).

Although epidemiologists may think carefully about how to deal with confounding, it is extremely difficult to precisely specify a regression model that could properly account for such a dense set of correlations among so many variables. Another example is the relation between tobacco and diet. Tobacco is known to be associated causally with multiple diseases. However, in another analysis with the same NHANES data, serum cotinine (a marker of tobacco exposure) had modest-to-strong associations with dozens of other environmental exposures, including 7 nutrients (34). If an association is found between one of these other exposures and a health outcome, one can never be sure how much of this should be attributed to the other exposure or to smoking. Simply adjusting for smoking in the regression model probably will not suffice, because several additional variables may need to be accounted for, and tobacco exposure and nutrient intake are both measured with considerable error (12, 36). Correlation globes will only become more jumbled over time due to rapid increases in the number of exposures that can be assessed through 'omics (37, 38), wearable technology (39), and other measurement tools.

Another Major Problem in Nutrition Science: Dietary Measurement

Tiny effects and dense correlation globes are problematic enough to disable most analyses in nutrition science by themselves. Another major problem, dietary measurement error (12), ruins most of what remains. Dietary measurement error affects nonrandomized studies, as well as randomized trials that attempt to measure adherence to an assigned diet (and almost all trials currently do this).

Most studies of nutrition do not record dietary consumption right when it occurs with a direct and objective method. Instead, inferences about past consumption are made by assessing participants' memories, typically with a 24-h dietary recall or an FFQ. Nutritional policy and dietary guidelines are largely informed by data obtained with these approaches (40), but they can be very inaccurate (12). Two out of 3 participants in the NHANES reported an amount of energy intake not compatible with life (41). Another analysis found that noise exceeds signal >9-fold for self-reported energy intake (12). Individual nutrients are also misreported differentially and unpredictably, so energy adjustment cannot salvage their analyses (42).

There are also a few theoretical reasons for disbelieving much of the data that come from memory-based dietary assessments. A memory is not a plain retelling of a past event but instead involves constructive and reconstructive processes (e.g., imagination) that are highly error prone (12). In addition, memory cannot be independently observed, quantified, or falsified, so recall data are pseudoscientific by definition (12). Furthermore, memory-based dietary assessments often interrogate in ways that have been shown to induce false recall in many other contexts (43–45): for example, when a participant claims to not know the answer to a particular question about intake, an interviewer might respond with silence to motivate a new answer (46). Some memory-based dietary assessments are better validated than others, but even the best validations are typically done against bronze standards.

The Current Research Agenda in Nutrition Science Cannot Handle These Problems

The aforementioned epistemologic problems cannot be overcome with the research designs that are currently used in nutrition science. Nonrandomized observational research simply involves too much confounding. It is also more vulnerable to publication and selective-outcome reporting biases (47) compared with randomized research, because it is more difficult to ascertain how many observational data sets are available worldwide that can address a given exposure-outcome relation (48). Although some epidemiologists register their analysis protocols, this is falsely reassuring because it can easily be done after peeking at the relevant data surreptitiously (49).

Random allocation of exposure is a necessary condition for overcoming the major epistemologic problems afflicting nutrition science. However, the current randomized research

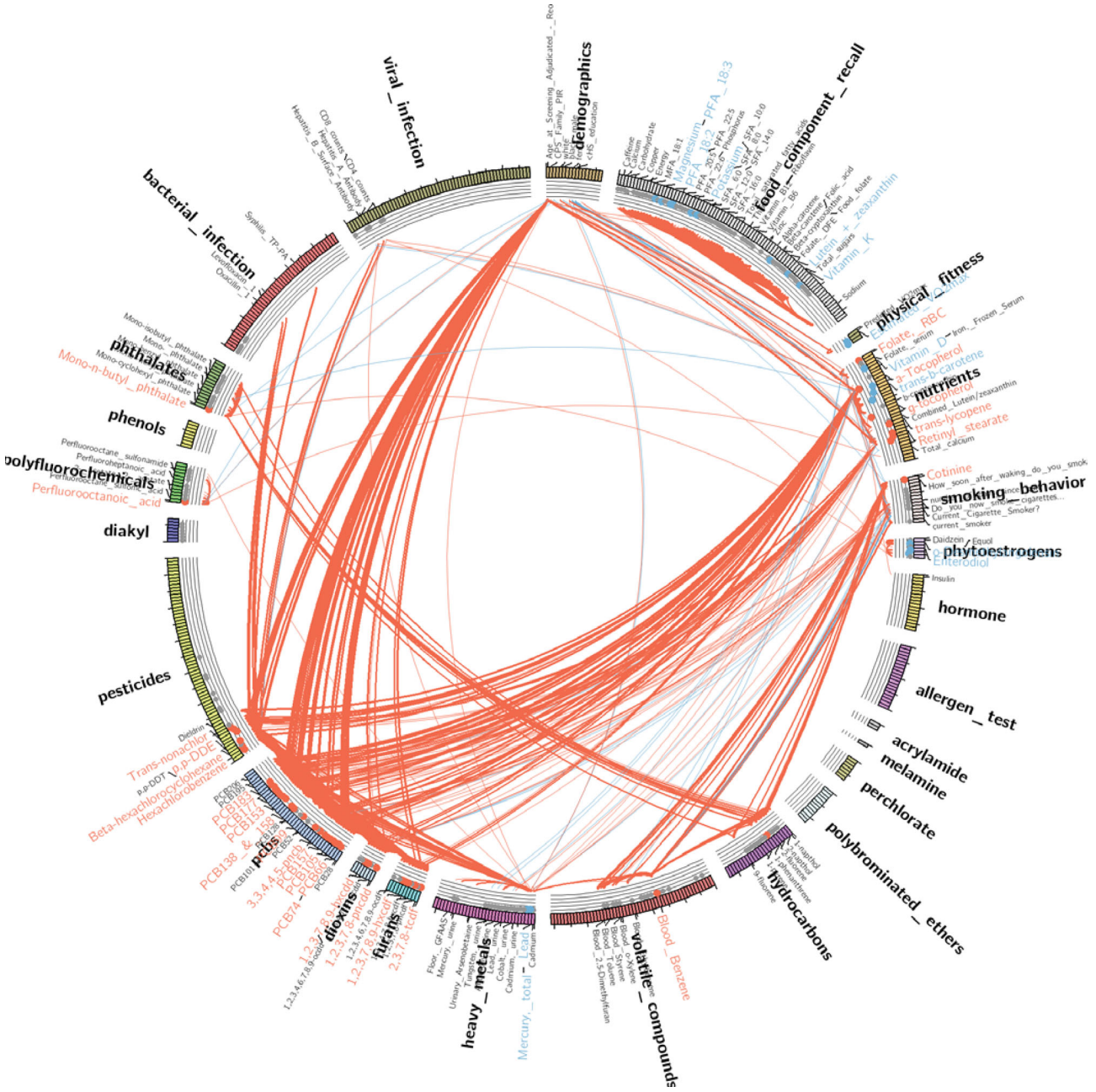

FIGURE 2 A beautiful, jumbled globe of correlations. This correlation globe depicts associations with fasting serum TGs and 575 exposures, including nutrients, food components, and other families of exposure variables. The strength of each association corresponds to line thickness, with red lines depicting positive associations and blue lines depicting negative associations. To examine whether any of these exposures causes fasting hypertriglyceridemia (rather than being merely correlated with it), the exposure of interest must first be disentangled from all the others, a daunting task. Data for this depiction derive from 4 individual survey periods, spanning the years 1999–2006, of the NHANES. Similar analyses have been presented in reference 33. Figure art courtesy of Chirag Patel.

agenda in nutrition science is far from optimal. It occasionally develops a pivotal, relatively large trial such as Prevención con Dieta Mediterránea (PREDIMED) (50), but it mostly produces thousands of small-sized, short-duration trials (15) that underdeliver for many reasons. To increase the likelihood of finding something publishable, these "microtrials" commonly examine only populations that have an exceptionally high risk of the main outcome of interest (51). Surrogate outcomes of questionable clinical relevance (52) are commonly selected for convenience. Adverse events are recorded haphazardly, if at all. Explanatory designs are necessarily favored over pragmatic designs (to be compatible with the small sample size), which introduces problems related to dietary measurement error as well as treatment nonadherence. A small sample size also provides limited statistical power to detect tiny effects, and it lowers the likelihood that a

significant result represents a true effect (53). Reformation is needed (14).

Megatrials Offer a Way to Answer a Small Number of the Most Important Nutritional Questions

Nutrition science needs to get rid of almost all nonrandomized observational research, as well as most of the microtrials, and conduct a few dozen megatrials at a time instead (14). Rather than try to answer a million different questions all at once without answering any one thing satisfactorily, as with the current approach, the main goal of our proposed megatrial approach should be to clearly answer a small number of the most important nutritional questions that we face. The megatrials should reflect pragmatic circumstances so that their results can be readily translated to recommendations for the general public (54). Hard outcomes (that preferably include death) should be evaluated, and the megatrials should not focus only on very-high-risk populations unless there is good reason. For example, a trial could use a low-fat compared with a low-carbohydrate diet assignment in a large, unselected population of participants, and measure death as the outcome instead of blood lipids or body weight.

This proposed scale-up is challenging but still quite doable. The cost of the average trial would go up considerably to pay for a larger sample size and lengthier follow-up, but the aggregate cost of all of the trials combined may stay the same or become even less because of their reduced number. Additional cost-savings will be achieved by streamlining the trial design (55) and, when possible, by building the trial on the platform of an already-existing health registry (56).

Our proposed reformation has several limitations—real ones as well as imagined ones—that are worth discussing. For instance, running only a few dozen megatrials at a time will leave many questions understudied or unaddressed altogether. Only a minority of all of the possible combinations of dietary exposures and clinical outcomes will be tested. Frustration will mount if a megatrial addresses a major question that becomes obsolete while the trial is ongoing (57).

These problems sound worse than they actually are. It is unrealistic to expect that every question about diet should be addressed in nutrition science. Many questions do not have answers that are valuable enough (i.e., translatable enough into improved human health) to justify the resources that would be needed to obtain them. For instance, trials of single nutrients may almost never be worth the trouble, because the expected effects are too tiny to be relevant (even to a huge population) or detectable reliably (9). In contrast, composite diets are much more likely to produce effects large enough to justify a trial (50). And although many highly important nutritional questions will not be addressed due to the small number of megatrials, this still represents a great improvement over the current research agenda, which publishes endless nominal answers but hardly any credible ones.

Similarly, it is unrealistic to expect that every minute variant of a research question should be addressed in its own trial. Much of the waste that has accumulated over the years in nutrition science has come from innumerable analyses and studies that were ever-so-slightly different from the previous ones. These incremental approaches usually add little or no value. We accept that a research question that is trialed only once will never be answered with a perfect definitiveness that convinces everybody: one can always look back and argue endlessly that a particular result could have been due to any number of factors, or that a slight change in the research question could have yielded a different result. However, most of these speculations lead nowhere, and we doubt that any amount of incremental research would put an end to them anyway.

New information can indeed make an important question obsolete. Therefore, major questions to be trialed must be evaluated beforehand to assess the likelihood that such information could soon emerge. This will not prevent every single megatrial from delivering a stillbirth, but the futility rate can be minimized.

Another potential limitation of our proposed reformation is that trials may not be pragmatic and may have poor representation of important populations among trial participants. For instance, some megatrials may exclude women, children, the elderly, or those with common medical conditions (58). In addition, trial participants may rarely be an ethnic minority and may rarely live in geographically remote areas (59). Of special concern, trial participants' responses to dietary intervention might be seriously affected by volunteer bias (60).

Although the megatrials should try to include traditionally underrepresented populations whenever possible, there is little evidence to suggest that average dietary intervention effects often differ meaningfully across broad populations. Rarely are formal interaction tests performed to evaluate this specifically (61), and even more rarely are claims of interaction shown to be credible (62, 63).

Interventions that could be seriously affected by volunteer bias can be examined in randomized trials that are nested within larger observational cohorts (64). This allows for data to be collected on treatment refusers (because they are still in the observational cohort despite refusing the intervention randomly assigned to them) so that comparisons can be made with treatment acceptors.

Randomized trials are often regarded as being ill-equipped to deal with unintended participant behavior, which is commonplace in nutrition studies. Participants assigned to the intervention group may adhere poorly because of study fatigue or because the intervention is genuinely difficult. They could also make changes to ancillary behaviors that are not directly targeted by the intervention but affect the main outcome of interest nonetheless, often in unpredictable ways. Participants in the control group may decide to adopt the intervention for themselves or they may withdraw from the trial because of disappointment. Nevertheless, most of these unintended behaviors are not introduced by the trial itself (apart from delivery of an intervention), so they are not especially problematic for pragmatic trials that aim to evaluate interventions in real-world settings.

Current methods for measuring diet are not accurate enough for randomized trials (12), and newer methods based on biochemical, Web, camera, mobile, or sensor tools have yet to establish suitable validity (65, 66). Until this is done, the megatrials will evaluate only the effects of prescribing dietary interventions, and adherence will not be measured (but will be strongly encouraged).

Study assignment cannot be blinded in trials of whole diets (apart from outcome assessment, which can and should be blinded), so dietary preferences and expectations will factor into the results. This is not very problematic. The interventions are being trialed because there is no definitive pretrial evidence of superiority, and this message will be reiterated to the participants throughout the study. In addition, preferences and expectations are parts of real-world medicine, so it is better for these "noises" to be included in the results.

Dietary interventions are not randomly assigned in real life. Although randomization removes allocation bias, it also removes preference effects (for the participants who are assigned an intervention that is other than their favorite) that are important to examine for results to be maximally generalizable. We think that the overall benefits of our approach are well worth this limitation. In addition, randomized preference designs can partially overcome this problem (67).

Traditional parallel-arm randomized trials in nutrition test dietary interventions that are fixed throughout the study, which does not allow for investigation of the effectiveness of continuing, modifying, or stopping an intervention altogether depending on the previous response. Investigations of this sort can first be evaluated in Sequential Multiple Assignment Randomized Trials to develop an adaptive intervention, which can then be evaluated in a randomized confirmatory trial against an appropriate alternative (68).

For selected questions in which adherence needs to be maximized (e.g., to get mechanistic metabolic insights rather than pragmatic insights on clinical outcomes), we propose that short-term randomized trials with direct observation of participants should be considered, as we discuss in the next section.

Moving Forward with Traditional and Novel Randomized Trial Designs

Having cataloged the many problems with traditionally performed randomized trials and some potential solutions, we now present how we can make progress with novel study designs or improvements to existing traditional randomized trial designs.

Large, simple trials

Large, simple trials (LSTs) can overcome many of the limitations that relate to pragmatism with the current randomized agenda in nutrition. LSTs aim to maximize benefits and minimize cost. Each trial can compare $\geq$2 substantially different dietary prescriptions with intention-to-treat methodology.

LSTs try to maximize real-world relevance and generalizability (69). Eligibility criteria are as inclusive as possible. Data collection focuses on objective measurements of the most relevant clinical outcomes for efficacy and harms. Follow-up continues until ascertainment of the primary outcome or death for all participants regardless of adherence. Ideally, participants and researchers agree in advance to allow for passive follow-up for major trial outcomes and vital status in the event of dropout or nonadherence (55).

LSTs are designed to minimize cost and complexity. Already available resources are used whenever possible. Study enrollment and data capture can be done in part or entirely online. Data collection is streamlined to capture only key information and outcomes that provide a level of scientific benefit that exceeds added costs. Monitoring and adjudication processes are similarly streamlined.

Rather than verifying all of the data, random samples are chosen for verification instead. An error rate regarded as acceptable is chosen in advance, and targeted on-site monitoring strategies are used when key indicators are triggered (55).

Registry-based designs

Some LSTs may be conducted with the use of existing registries. Registry-based randomized trials (RRTs) are pragmatic trials that utilize a health registry as a platform for case records, data collection, randomization, and follow-up (70). The data can originate from reports by patients or physicians, medical chart abstraction, electronic health records, administrative databases, institutional or organizational databases, and other sources (70).

RRTs allow for enrollment of potentially thousands of participants in very little time. A quick rate of enrollment is facilitated by identifying eligible trial participants with already-existing clinical data (70). The use of a health registry also allows for near-complete follow-up, collection of data, and recording of outcomes of the reference population (all participants who were eligible for the trial), as well as for the participants who were not eligible (71). Typically, eligibility criteria are not as stringent, and monitoring and follow-up are more similar to everyday medical practice than with traditional randomized trials (70).

Registries can provide main outcome data long after termination of an intervention. For instance, the Alpha-Tocopherol, Beta-Carotene Cancer Prevention (ATBC) Study had a median of 6.1 y of intervention (72) and then utilized national registries to obtain follow-up data for 18 additional years (73, 74). The trial showed that, contrary to earlier claims from nonrandomized studies, both β-carotene and α-tocopherol did not offer any benefit for survival or overall cancer risk during the main intervention period. β-Carotene was actually associated with increased overall mortality risk. The follow-up postintervention period showed no further differences between the compared arms: once the interventions were stopped, the mortality disadvantage with β-carotene shrunk and became undiscernible after 8 y. The same shrinking was seen for occasional signals of increased or decreased risk of specific cancer types (which were not primary endpoints) that had been seen in the original intervention period (i.e., increased lung cancer risk

TABLE 1 ATBC study (enrolled in 1985–1988): initial and postintervention-period results[1]

	β-Carotene		α-Tocopherol	
	All deaths	Lung cancer	All deaths	Prostate cancer
Intervention to April 1993 (72)	1.08 (1.01, 1.16)	1.18 (1.03, 1.36)	1.02 (0.95, 1.09)	0.68 (0.53, 0.88)
Postintervention				
To April 1999 (73)	—	1.06 (0.94, 1.20)	—	0.88 (0.76, 1.03)
To April 2001 (73)	1.07 (1.02, 1.12)	—	1.01 (0.96, 1.05)	—
To December 2009 (74)	1.02 (0.99, 1.05)	1.04 (0.96, 1.01)	1.02 (0.98, 1.05)	0.97 (0.89, 1.05)

[1] Values are relative risk (95% CIs). ATBC, Alpha-Tocopherol, Beta-Carotene Cancer Prevention.

with β-carotene and decreased prostate cancer risk with α-tocopherol) (**Table 1**).

Meta-analysis of multiple long-term trials

Most intervention effects in nutrition are small (9), so meta-analysis of multiple (preferably large and long-term) trials may offer the best opportunity for reliable answers. The evaluation of antioxidants once again provides a useful example. Several meta-analyses (75–78) of multiple trials have shown convincingly that, contrary to earlier epidemiologic expectations, antioxidant vitamins such as β-carotene and α-tocopherol do not offer an overall mortality advantage or preventive benefit for cancer or cardiovascular disease. Very high doses may even be associated with excess mortality (76, 78). Of course, meta-analyses have their own strengths, weaknesses, and caveats, and their discussion goes beyond the scope of this article.

Embedding multiple trials in the same study population

Multiple nutrition- or diet-related questions may be addressed concurrently in the same study population. Factorial randomization may allow for optimal use of resources and also maximizes power for assessing interactions (79). It has been proposed that a very large number of research questions can be studied concurrently in the same population. The proposed design, Multiple Lifestyle Factorial Experimental (multi-LIFE) trials (80), can be thought of as specialized RRTs with some important distinctions: participants can choose from a long list of simple lifestyle randomization options, and several interventions can be tested concurrently with factorial randomization. Health-conscious, motivated individuals will likely be attracted to this study design, and they have been shown to show good adherence to lifestyle interventions (81–84). Adherence is also fostered by deliberately assigning participants to interventions that they feel neutral toward.

N-of-1 trials

N-of-1 trials are multiple crossover trials. Although each N-of-1 trial examines a single individual, they are often conducted in a series, and their results can be aggregated or even combined with results from parallel-arm trials (85). However, N-of-1 trials have some potential limitations for applications in nutrition science. First, they have a low throughput. All of the N-of-1 trials combined have examined only ∼2000 participants to date (86). Despite having the theoretical capability to examine many different treatment options in each individual, N-of-1 trials assess only 2 interventions in 93% of cases, and the median period length is only 10 d (86). This may not be a long enough duration to capture most onset and offset effects for diet. Second, generalizability is highly questionable in practice. Third, assumptions of performed statistical tests are frequently violated as a result of small sample size, limited data, nonnormal distribution, carryover effects, priming from exposure to previous interventions, dropouts, and other concerns (86).

Trials in experimental settings with direct observation of participants

Many focused, mechanistic metabolic questions and proof-of-concept questions require high treatment fidelity to be answerable, and one cannot afford for nonadherence, crossover, or dropout to be substantial. These questions also require the measurement of outcomes that can respond in a short time frame, typically metabolic or laboratory markers. For these questions, trials can be performed that involve continuous direct observation of participants in in-house settings (87, 88). These trials are typically very expensive, but if done sparingly for crucial questions, they may be worth the investment.

Mendelian randomization studies

Mendelian randomization studies (89, 90) are possible to build within nonrandomized observational cohorts and data sets, in which the availability of genetic instruments allow for the creation of a randomized trial equivalent. The Mendelian randomization approach has major advantages in that it allows a randomized trial equivalent to be set up without extra cost and without the need of new follow-up, with the use of the available observational data. However, they can have some shortcomings: for example, only weak genetic instruments may be available and the assumptions about the specificity of genetic instruments may not hold. Although Mendelian randomizations may have better validity than traditional observational designs, it is unknown whether they can be used for policy decisions, entirely replacing formal randomized trials. Interestingly, most well-conducted Mendelian randomization studies show negative results (91, 92), which aligns with the argument that most observational claims about causal associations are spurious.

Compared with traditional observational analyses that have such a poor record of identifying causality in nutrition, well-done Mendelian randomization studies may be a good investment for analyzing observational data toward identifying factors that may have a higher chance of being causal. Interventions that affect these factors can then be selectively prioritized for evaluation in randomized trials, whenever feasible.

Final Comments

Our perspective—on the relative merits of nonrandomized and randomized studies in nutrition science, on the fixability of certain epistemologic problems, and on replacing a current highly prolific agenda of nonrandomized studies and microtrials with a small number of megatrials, as well as a few fit-for-purpose designs—may not meet with agreement by some nutrition experts. We encourage debate and we recommend the reader to also examine other perspectives on these issues (93–98). Regardless, we think that we have reached a saturation point in nutrition science, with limited or no further progress being made, and thus some reformed research agenda is necessary.

Wider adoption of optimal research practices would also greatly benefit nutrition science (99). These include, but are not limited to, preregistration for randomized trials and other prespecified hypothesis-testing and validation studies, availability of protocols and raw data, complete reporting of all results, and provision of proper rewards and incentives for reproducible research. As we discussed above, randomized trials have their own limitations as well, and they are not immune to many of the problems encountered in nonrandomized studies. For example, in the absence of detailed preregistration of outcomes and analyses (15), selective reporting can be as severe an issue in randomized trials as in observational studies. Unaccounted multiplicity of analyses can also become a major problem in randomized trials. In extreme cases, a single trial can generate dozens or even hundreds of secondary articles (100).

But despite the many drawbacks of randomized trials, they alone can possibly overcome the epistemological problems of tiny effects, dense correlation globes, and high dietary measurement error. Our proposed megatrial approach—augmented with novel study designs, such as direct, continuous observation of participants—will form a sounder basis for informing dietary recommendations, both at the population level and, in select circumstances, at the individual level as part of precision nutritional therapy.

Acknowledgments

Both authors read and approved the final manuscript.

References

1. Trinquart L, Johns DM, Galea S. Why do we think we know what we know? A metaknowledge analysis of the salt controversy. Int J Epidemiol 2016;45(1):251–60.
2. Kanter MM, Kris-Etherton PM, Fernandez ML, Vickers KC, Katz DL. Exploring the factors that affect blood cholesterol and heart disease risk: is dietary cholesterol as bad for you as history leads us to believe? Adv Nutr 2012;3(5):711–7.
3. Benziger CP, Roth GA, Moran AE. The Global Burden of Disease Study and the preventable burden of NCD. Glob Heart 2016;11(4):393–7.
4. Ioannidis JP. Contradicted and initially stronger effects in highly cited clinical research. JAMA 2005;294(2):218–28.
5. Young SS, Karr A. Deming, data and observational studies. Significance 2011;8(3):116–20.
6. Mente A, de Koning L, Shannon HS, Anand SS. A systematic review of the evidence supporting a causal link between dietary factors and coronary heart disease. Arch Intern Med 2009;169(7):659–69.
7. Miller PE, Perez V. Low-calorie sweeteners and body weight and composition: a meta-analysis of randomized controlled trials and prospective cohort studies. Am J Clin Nutr 2014;100(3):765–77.
8. US Department of Health and Human Services; USDA. 2015–2020 Dietary guidelines for Americans. 8th ed. December 2015. Available from : https://health.gov/dietaryguidelines/2015/guidelines/, Last accessed April 2, 2018.
9. Siontis GC, Ioannidis JP. Risk factors and interventions with statistically significant tiny effects. Int J Epidemiol 2011;40(5):1292–307.
10. Ioannidis JP, Loy EY, Poulton R, Chia KS. Researching genetic versus nongenetic determinants of disease: a comparison and proposed unification. Sci Transl Med 2009;1(7):7ps8.
11. Patel CJ, Cullen MR, Ioannidis JP, Butte AJ. Systematic evaluation of environmental factors: persistent pollutants and nutrients correlated with serum lipid levels. Int J Epidemiol 2012;41(3):828–43.
12. Archer E, Pavela G, Lavie CJ. 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. Mayo Clin Proc 2015;90(7):911–26.
13. Patel CJ, Ioannidis JP. Placing epidemiological results in the context of multiplicity and typical correlations of exposures. J Epidemiol Community Health 2014;68(11):1096–100.
14. Ioannidis JP. We need more randomized trials in nutrition—preferably large, long-term, and with negative results. Am J Clin Nutr 2016;103(6):1385–6.
15. Dal-Ré R, Bracken MB, Ioannidis JP. Call to improve transparency of trials of non-regulated interventions. BMJ 2015;350:h1323.
16. Ioannidis JP. Implausible results in human nutrition research. BMJ 2013;347:f6698.
17. Schoenfeld JD, Ioannidis JP. Is everything we eat associated with cancer? A systematic cookbook review. Am J Clin Nutr 2013;97(1):127–34.
18. Ioannidis JP, Siontis GC. Re: Fruit and vegetable intake and overall cancer risk in the European Prospective Investigation into Cancer and Nutrition. J Natl Cancer Inst 2011;103(3):279–80.
19. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schünemann HJ. Rating quality of evidence and strength of recommendations: GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ 2008;336(7650):924.
20. Khoury MJ, Ioannidis JP. Big data meets public health. Science 2014;346(6213):1054–5.
21. Bouvard V, Loomis D, Guyton KZ, Grosse Y, El Ghissassi F, Benbrahim-Tallaa L, Guha N, Mattock H, Straif K. Carcinogenicity of consumption of red and processed meat. Lancet Oncol 2015;16(16):1599–600.
22. Chan DS, Lau R, Aune D, Vieira R, Greenwood DC, Kampman E, Norat T. Red and processed meat and colorectal cancer incidence: meta-analysis of prospective studies. PLoS One 2011;6(6):e20456.
23. Ioannidis JP, Polyzos NP, Trikalinos TA. Selective discussion and transparency in microarray research findings for cancer outcomes. Eur J Cancer 2007;43(13):1999–2010.
24. Patel CJ, Burford B, Ioannidis JP. Assessment of vibration of effects due to model specification can demonstrate the instability of observational associations. J Clin Epidemiol 2015;68(9):1046–58.

25. Brown AW, Ioannidis JP, Cope MB, Bier DM, Allison DB. Unscientific beliefs about scientific topics in nutrition. Adv Nutr 2014;5(5):563–5.
26. Chavalarias D, Wallach JD, Li AHT, Ioannidis JP. Evolution of reporting P values in the biomedical literature, 1990–2015. JAMA 2016;315(11):1141–8.
27. Johnson VE. Revised standards for statistical evidence. Proc Natl Acad Sci USA 2013;110(48):19313–7.
28. Belbasis L, Bellou V, Evangelou E, Ioannidis JP, Tzoulaki I. Environmental risk factors and multiple sclerosis: an umbrella review of systematic reviews and meta-analyses. Lancet Neurol 2015;14(3):263–73.
29. Schuemie MJ, Hripcsak G, Ryan PB, Madigan D, Suchard MA. Robust empirical calibration of p-values using observational data. Stat Med 2016;35(22):3883–8.
30. Prasad V, Jena AB. Prespecified falsification end points: can they validate true observational associations? JAMA 2013;309(3):241–2.
31. Patel CJ, Ji J, Sundquist J, Ioannidis JP, Sundquist K. Systematic assessment of pharmaceutical prescriptions in association with cancer risk: a method to conduct a population-wide medication-wide longitudinal study. Sci Rep 2016;6:31308.
32. Fewell Z, Smith GD, Sterne JA. The impact of residual and unmeasured confounding in epidemiologic studies: a simulation study. Am J Epidemiol 2007;166(6):646–55.
33. Patel CJ, Manrai AK. Development of exposome correlation globes to map out environment-wide associations. Pac Symp Biocomput 2015;20:231–42.
34. Patel CJ, Ioannidis JP. Studying the elusive environment in large scale. JAMA 2014;311(21):2173–4.
35. Tatsioni A, Bonitsis NG, Ioannidis JP. Persistence of contradicted claims in the literature. JAMA 2007;298(21):2517–26.
36. Gorber SC, Schofield-Hurwitz S, Hardt J, Levasseur G, Tremblay M. The accuracy of self-reported smoking: a systematic review of the relationship between self-reported and cotinine-assessed smoking status. Nicotine Tob Res 2009;11(1):12–24.
37. Athersuch TJ. The role of metabolomics in characterizing the human exposome. Bioanalysis 2012;4(18):2207–12.
38. Mischak H, Allmaier G, Apweiler R, Attwood T, Baumann M, Benigni A, Bennett SE, Bischoff R, Bongcam-Rudloff E, Capasso G et al. Recommendations for biomarker identification and qualification in clinical proteomics. Sci Transl Med 2010;2(46):46ps2.
39. Naci H, Ioannidis JP. Evaluation of wellness determinants and interventions by citizen scientists. JAMA 2015;314(2):121–2.
40. Dietary Guidelines Advisory Committee. Scientific report of the 2015 Dietary Guidelines Advisory Committee. Washington (DC): USDA, US Department of Health and Human Services; 2015.
41. Archer E, Hand GA, Blair SN. Validity of US nutritional surveillance: National Health and Nutrition Examination Survey caloric energy intake data, 1971–2010. PLoS One 2013;8(10):e76632.
42. Bellach B, Kohlmeier L. Energy adjustment does not control for differential recall bias in nutritional epidemiology. J Clin Epidemiol 1998;51(5):393–8.
43. Bernstein DM, Loftus EF. The consequences of false memories for food preferences and choices. Perspect Psychol Sci 2009;4(2):135–9.
44. Straube B. An overview of the neuro-cognitive processes involved in the encoding, consolidation, and retrieval of true and false memories. Behav Brain Funct 2012;8(1):35.
45. Johnson MK. Memory and reality. Am Psychol 2006;61(8):760–71.
46. Centers for Disease Control and Prevention. NAHNES ES MEC in-person dietary interviewers procedures manual. Atlanta, GA: CDC; 2008.
47. Dwan K, Gamble C, Williamson PR, Kirkham JJ. Systematic review of the empirical evidence of study publication bias and outcome reporting bias—an updated review. PLoS One 2013;8(7):e66844.
48. Ioannidis JP. The importance of potential studies that have not existed and registration of observational data sets. JAMA 2012;308(6):575–6.
49. Dal-Ré R, Ioannidis JP, Bracken MB, Buffler PA, Chan A-W, Franco EL, La Vecchia C, Weiderpass E. Making prospective registration of observational research a reality. Sci Transl Med 2014;6(224):224cm1.
50. Estruch R, Ros E, Salas-Salvadó J, Covas M-I, Corella D, Arós F, Gómez-Gracia E, Ruiz-Gutiérrez V, Fiol M, Lapetra J. Primary prevention of cardiovascular disease with a Mediterranean diet. N Engl J Med 2013;368(14):1279–90.
51. Ioannidis JP, Lau J. The impact of high-risk patients on the results of clinical trials. J Clin Epidemiol 1997;50(10):1089–98.
52. Fleming TR, DeMets DL. Surrogate end points in clinical trials: are we being misled? Ann Intern Med 1996;125(7):605–13.
53. Button KS, Ioannidis JP, Mokrysz C, Nosek BA, Flint J, Robinson ES, Munafò MR. Power failure: why small sample size undermines the reliability of neuroscience. Nat Rev Neurosci 2013;14(5):365–76.
54. Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. BMJ 2015;350:h2147.
55. Calvo G, McMurray JJ, Granger CB, Alonso-García Á, Armstrong P, Flather M, Gómez-Outes A, Pocock S, Stockbridge N, Svensson A. Large streamlined trials in cardiovascular disease. Eur Heart J 2014;35(9):544–8.
56. Lauer MS, D'Agostino RB Sr. The randomized registry trial—the next disruptive technology in clinical research? N Engl J Med 2013;369(17):1579–81.
57. Lissner L, Majem LS, de Almeida MDV, Berg C, Hughes R, Cannon G, Thorsdottir I, Kearney J, Gustafsson J-Å, Rafter J. The Women's Health Initiative: what is on trial: nutrition and chronic disease? Or misinterpreted science, media havoc and the sound of silence from peers? Public Health Nutr 2006;9(02):269–72.
58. Van Spall HG, Toren A, Kiss A, Fowler RA. Eligibility criteria of randomized controlled trials published in high-impact general medical journals: a systematic sampling review. JAMA 2007;297(11):1233–40.
59. Murthy VH, Krumholz HM, Gross CP. Participation in cancer clinical trials: race-, sex-, and age-based disparities. JAMA 2004;291(22):2720–6.
60. Rosenthal R, Rosnow RL. The volunteer subject. Artifacts in behavioral research, Oxford University Press, Oxford, 2009:48–92.
61. Sainani K. Misleading comparisons: the fallacy of comparing statistical significance. PM R 2010;2(6):559–62.
62. Hernández AV, Boersma E, Murray GD, Habbema JDF, Steyerberg EW. Subgroup analyses in therapeutic cardiovascular clinical trials: are most of them misleading? Am Heart J 2006;151(2):257–64.
63. Sun X, Briel M, Busse JW, You JJ, Akl EA, Mejza F, Bala MM, Bassler D, Mertz D, Diaz-Granados N. Credibility of claims of subgroup effects in randomised controlled trials: systematic review. BMJ 2012;344:e1553.
64. Relton C, Torgerson D, O'Cathain A, Nicholl J. Rethinking pragmatic randomised controlled trials: introducing the "cohort multiple randomised controlled trial" design. BMJ 2010;340:c1066.
65. Illner AK, Freisling H, Boeing H, Huybrechts I, Crispim SP, Slimani N. Review and evaluation of innovative technologies for measuring diet in nutritional epidemiology. Int J Epidemiol 2012;41(4):1187–203.
66. Ioannidis JPA, Bossuyt PMM. Waste, leaks, and failures in the biomarker pipeline. Clin Chem 2017;63(5):963–72.
67. Walter SD, Turner R, Macaskill P, McCaffery KJ, Irwig L. Beyond the treatment effect: evaluating the effects of patient preferences in randomised trials. Stat Methods Med Res 2017;26(1):489–507.
68. Lei H, Nahum-Shani I, Lynch K, Oslin D, Murphy SA. A "SMART" design for building individualized treatment sequences. Annu Rev Clin Psychol 2012;8:21–48.
69. Vickers AJ, Scardino PT. The clinically-integrated randomized trial: proposed novel method for conducting large trials at low cost. Trials 2009;10(1):14.
70. Li G, Sajobi TT, Menon BK, Korngut L, Lowerison M, James M, Wilton SB, Williamson T, Gill S, Drogos LL. Registry-based randomized controlled trials—what are the advantages, challenges, and areas for future research? J Clin Epidemiol 2016;80:16–24.

71. James S, Rao SV, Granger CB. Registry-based randomized clinical trials—a new clinical trial paradigm. Nat Rev Cardiol 2015;12(5):312–6.
72. Alpha-Tocopherol, Beta Carotene Cancer Prevention Study Group. The effect of vitamin E and beta carotene on the incidence of lung cancer and other cancers in male smokers. N Engl J Med 1994;330(15):1029–35.
73. Virtamo J, Pietinen P, Huttunen JK, Korhonen P, Malila N, Virtanen MJ, Albanes D, Taylor PR, Albert P; ATBC Study Group. Incidence of cancer and mortality following alpha-tocopherol and beta-carotene supplementation: a post-intervention follow-up. JAMA 2003;290(4):476–85.
74. Virtamo J, Taylor PR, Kontto J, Männistö S, Utriainen M, Weinstein SJ, Huttunen J, Albanes D. Effects of α-tocopherol and β-carotene supplementation on cancer incidence and mortality: 18-year postintervention follow-up of the Alpha-tocopherol, Beta-carotene Cancer Prevention Study. Int J Cancer. 2014;135(1):178–85.
75. Jeon YJ, Myung SK, Lee EH, Kim Y, Chang YJ, Ju W, Cho HJ, Seo HG, Huh BY. Effects of beta-carotene supplements on cancer prevention: meta-analysis of randomized controlled trials. Nutr Cancer 2011;63(8):1196–207.
76. Bjelakovic G, Nikolova D, Gluud LL, Simonetti RG, Gluud C. Antioxidant supplements for prevention of mortality in healthy participants and patients with various diseases. Cochrane Database Syst Rev 2012;3:CD007176.
77. Ye Y, Li J, Yuan Z. Effect of antioxidant vitamin supplementation on cardiovascular outcomes: a meta-analysis of randomized controlled trials. PLoS One 2013;8(2):e56803.
78. Bjelakovic G, Nikolova D, Gluud C. Meta-regression analyses, meta-analyses, and trial sequential analyses of the effects of supplementation with beta-carotene, vitamin A, and vitamin E singly or in different combinations on all-cause mortality: do we have evidence for lack of harm? PLoS One 2013;8(9):e74558.
79. Piantadosi S. Clinical trials: a methodological approach. Hoboken, NJ: Wiley; 1997.
80. Ioannidis JP, Adami H-O. Nested randomized trials in large cohorts and biobanks: studying the health effects of lifestyle factors. Epidemiology 2008;19(1):75–82.
81. Kriska AM, Bayles C, Cauley JA, LaPorte RE, Sandler RB, Pambianco G. A randomized exercise trial in older women: increased activity over two years and the factors associated with compliance. Med Sci Sports Exerc 1986;18(5):557–62.
82. Irwin ML, Tworoger SS, Yasui Y, Rajan B, McVarish L, LaCroix K, Ulrich CM, Bowen D, Schwartz RS, Potter JD. Influence of demographic, physiologic, and psychosocial variables on adherence to a yearlong moderate-intensity exercise trial in postmenopausal women. Prev Med 2004;39(6):1080–6.
83. Courneya KS, Segal RJ, Reid RD, Jones LW, Malone SC, Venner PM, Parliament MB, Scott CG, Quinney HA, Wells GA. Three independent factors predicted adherence in a randomized controlled trial of resistance exercise training among prostate cancer survivors. J Clin Epidemiol 2004;57(6):571–9.
84. Women's Health Initiative Study Group. Dietary adherence in the Women's Health Initiative Dietary Modification Trial. J Am Diet Assoc 2004;104(4):654.
85. Punja S, Xu D, Schmid CH, Hartling L, Urichuk L, Nikles CJ, Vohra S. N-of-1 trials can be aggregated to generate group mean treatment effects: a systematic review and meta-analysis. J Clin Epidemiol 2016;76:65–75.
86. Punja S, Bukutu C, Shamseer L, Sampson M, Hartling L, Urichuk L, Vohra S. N-of-1 trials are a tapestry of heterogeneity. J Clin Epidemiol 2016;76:47–56.
87. Hall KD, Bemis T, Brychta R, Chen KY, Courville A, Crayner EJ, Goodwin S, Guo J, Howard L, Knuth ND. Calorie for calorie, dietary fat restriction results in more body fat loss than carbohydrate restriction in people with obesity. Cell Metab 2015;22(3):427–36.
88. Hall KD, Chen KY, Guo J, Lam YY, Leibel RL, Mayer LE, Reitman ML, Rosenbaum M, Smith SR, Walsh BT. Energy expenditure and body composition changes after an isocaloric ketogenic diet in overweight and obese men. Am J Clin Nutr 2016;104(2):324–33.
89. Smith GD, Ebrahim S. Mendelian randomization: prospects, potentials, and limitations. Int J Epidemiol 2004;33(1):30–42.
90. Smith GD, Timpson N, Ebrahim S. Strengthening causal inference in cardiovascular epidemiology through Mendelian randomization. Ann Med 2008;40(7):524–41.
91. Theodoratou E, Timofeeva M, Li X, Meng X, Ioannidis JPA. Nature, nurture, and cancer risks: genetic and nutritional contributions to cancer. Annu Rev Nutr 2017;37:293–320.
92. Li X, Meng X, Timofeeva M, Tzoulaki I, Tsilidis KK, Ioannidis JP, Campbell H, Theodoratou E. Serum uric acid levels and multiple health outcomes: umbrella review of evidence from observational studies, randomised controlled trials, and Mendelian randomisation studies. BMJ. 2017;357:j2376.
93. Maki KC, Slavin JL, Rains TM, Kris-Etherton PM. Limitations of observational evidence: implications for evidence-based dietary recommendations. Adv Nutr 2014;5:7–15.
94. Schwingshackl L, Knüppel S, Schwedhelm C, Hoffmann G, Missbach B, Stelmach-Mardas M, Dietrich S, Eichelmann F, Kontopantelis E, Iqbal K et al. NutriGrade: a scoring system to assess and judge the meta-evidence of randomized controlled trials and cohort studies in nutrition research [perspective]. Adv Nutr 2016;7(6):994–1004.
95. Yang C, Pinart M, Kolsteren P, Van Camp J, De Cock N, Nimptsch K, Pischon T, Laird E, Perozzi G, Canali R et al. Essential study quality descriptors for data from nutritional epidemiologic research [perspective]. Adv Nutr 2017;8(5):639–51.
96. Magni P, Bier DM, Pecorelli S, Agostoni C, Astrup A, Brighenti F, Cook R, Folco E, Fontana L, Gibson RA et al. Improving nutritional guidelines for sustainable health policies: current status and perspectives [perspective]. Adv Nutr 2017;8(4):532–45.
97. Temple NJ. How reliable are randomised controlled trials for studying the relationship between diet and disease? A narrative review. Br J Nutr 2016; 116(3):381–9.
98. Ankarfeldt MZ. Comment on "Limitations of observational evidence: implications for evidence-based dietary recommendations". Adv Nutr 2014;5(3):293.
99. Munafò MR, Nosek BA, Bishop DV, Button KS, Chambers CD, du Sert NP, Simonsohn U, Wagenmakers E-J, Ware JJ, Ioannidis JP. A manifesto for reproducible science. Nat Hum Behav 2017; 1:0021.
100. Ebrahim S, Montoya L, el Din MK, Sohani ZN, Agarwal A, Bance S, Saquib J, Saquib N, Ioannidis JP. Randomized trials are frequently fragmented in multiple secondary publications. J Clin Epidemiol 2016;79:130–9.