

Perspective: Fundamental Limitations of the Randomized Controlled Trial Method in Nutritional Research: The Example of Probiotics

Dennis Zeilstra,¹ Jessica A Younes,² Robert J Brummer,³ and Michiel Kleerebezem⁴

¹Independent researcher, Nutriz, Enschede, The Netherlands; ²Winclove Probiotics, Amsterdam, The Netherlands; ³Nutrition-Gut-Brain Interactions Research Centre, Faculty of Health and Medicine, School of Medical Sciences, Örebro University, Örebro, Sweden; and ⁴Host Microbe Interactomics Group, Wageningen University, Wageningen, The Netherlands

ABSTRACT

Studies on the relation between health and nutrition are often inconclusive. There are concerns about the validity of many research findings, and methods that can deliver high-quality evidence—such as the randomized controlled trial (RCT) method—have been embraced by nutritional researchers. Unfortunately, many nutritional RCTs also yield ambiguous results. It has been argued that RCTs are ill-suited for certain settings, including nutritional research. In this perspective, we investigate whether there are fundamental limitations of the RCT method in nutritional research. To this end, and to limit the scope, we use probiotic studies as an example. We use an epistemological approach and evaluate the presuppositions that underlie the RCT method. Three general presuppositions are identified and discussed. We evaluate whether these presuppositions can be considered true in probiotic RCTs, which appears not always to be the case. This perspective concludes by exploring several alternative study methods that may be considered for future probiotic or nutritional intervention trials. *Adv Nutr* 2018;9:561–571.

Keywords: RCT limitations, probiotics, nutrition, study design

Introduction

For the last half-century, the relation between diet and health status has attracted ever-increasing attention from the scientific community. However, much of the research is observational in nature and concerns have been raised about the validity of the research findings based on these methods (1, 2). To improve confidence in nutritional study outcomes, investigators strive to improve the quality of evidence (3) and the evidence-based approach is being embraced by nutritional researchers (4). Each research methodology yields a certain level of evidence quality, and the “gold standard” of

clinical research—the randomized controlled trial (RCT)—is considered to offer medical and scientific evidence of the highest quality (5), because its design allows for elimination of multiple bias sources at the baseline. Only systematic reviews or meta-analyses of RCTs are considered to offer a higher grade of evidence (6) (Figure 1). To further improve confidence in the derived recommendations, several more refined evidence grading methodologies have been developed, such as GRADE (Grading of Recommendations Assessment, Development and Evaluation) (7), its counterpart for nutrition studies, NutriGrade (8), or the weight-of-evidence guideline by the European Safety Authority (9).

This grading of evidence plays a substantial role in how research studies are applied and used to valorize scientific results, and to provide evidence for health care products and, as Blumberg et al. (4) state, it seems that “there is often an almost exclusive reliance on the RCT as the only type of evidence worthy of such consideration.”

Unfortunately, the results of nutritional RCTs related to efficacy of disease-modulation often yield ambiguous results. As an example, a meta-analysis of RCTs testing glycemic-lowering effects of chromium supplementation in subjects with type 2 diabetes showed considerable variation in the

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.

DZ is partly funded by the Gut Flora Foundation (<http://www.gutflora.org/>).

Author disclosures: RJB and MK, no conflicts of interest. DZ is partly funded by the Gut Flora Foundation. JAY is an employee at Winclove Probiotics.

Address correspondence to DZ (e-mail: d.zeilstra@gmail.com).

Abbreviations used: NNT, numbers needed to treat; QC, quality criteria; RCT, randomized controlled trial.

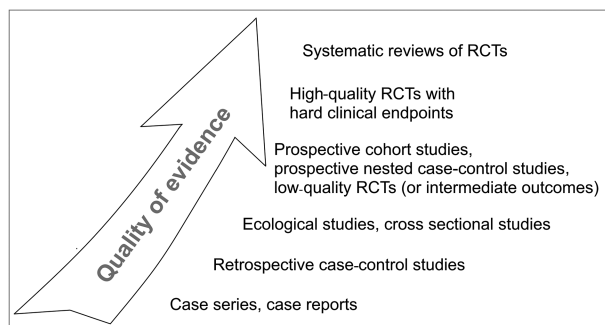


FIGURE 1 Levels of evidence as discussed in references 3, 5, and 6. RCT, randomized controlled trial.

obtained effect on fasting plasma glucose (10), even in studies that used exactly the same dietary supplement (600 µg Cr and 2 mg biotin) and very similar study populations [i.e., Singer and Geohas (11) compared with Albarracin et al. (12)]. In other words, very similar RCTs can yield different results. When even the gold standard method fails to provide convincing and consistent evidence, it is often concluded that either the RCT design or conduct is of insufficient quality, or that the alleged effect does not exist. Since RCTs became common clinical practice, however, scholars have expressed concerns about their applicability in various settings. Within the scope of nutritional research, several publications have argued that the RCT method might be ill-suited for this setting as well (13–15). Many of the arguments constitute largely practical considerations, ranging from inadequacy of the outcome measures and insufficient intervention duration, to high dropout rate, low adherence, variability of circumstances, or insufficient contrast between study groups.

In this Perspective article, we investigate whether there may be fundamental limitations to the application of the RCT method in nutritional research. To this end, and to limit the scope, we use studies of the microbiome and probiotics—a subarea of nutritional research—as an example. One reason to use this example is that probiotics are an example of nutrients that “can be packaged in a pill” (14) and are therefore seemingly suitable to study by means of RCTs. In addition, probiotics have attracted increasingly more interest from both researchers and clinicians. After the development of culture-independent microbial identification methods in the 1990s, interest in the microbiome has grown exponentially. The results are so promising that *Science* magazine selected the microbiome as a runner-up for Breakthrough of the Year in 2013 (16), and *Fortune Magazine* declared 2015 as “the year of the microbiome” (17). With this rapidly increasing interest in the microbiome and its modulation via probiotics, and the broad use of the RCT method in probiotic research, an analysis of the fundamental limitations of RCTs in this particular subarea of nutritional research seems warranted.

In the next section, a brief overview is provided of the current status of knowledge of microbiome intervention studies and specifically probiotic trials. Many studies on probiotics are conducted with the use of a RCT design

and, as we will show, these too often yield equivocal and heterogeneous results. Consequently, the obvious question follows: are there fundamental limitations to the applicability of the RCT approach in probiotics research? In this manuscript, the apparent difficulty of demonstrating the efficacy of probiotic treatments through RCTs is reviewed by investigating the underlying presuppositions. After briefly outlining and explaining the meaning of presuppositions, we discuss those that are generally applicable for RCTs. Next, we discuss and review whether these presuppositions can be considered true if investigating probiotics, concluding that the RCT method may not always be suitable to demonstrate the effect of probiotic interventions. Consequently, alternative approaches are needed to either ensure that the RCT presuppositions can reasonably be considered to be true, or to demonstrate efficacy through the use of another method. We explore several alternative approaches that may be considered for future probiotic intervention trials. This perspective concludes with a general assessment of the impact of these conclusions regarding probiotic RCTs on nutritional research.

Current Status of Knowledge

Research focusing on the microbiome is well over a century old and began with the studies of Élie Metchnikoff, who was the first to suggest that supplementation of live bacteria could promote health (18). As mentioned previously, microbiome research has increased exponentially in the last few decades, in large part due to the development of culture-independent microbial identification methods which drastically improve interpretation and analysis of the microbiome composition and function. Although this research has uncovered many intriguing findings, studies on modulation of the microbiome to confer health benefits have not been very successful in providing solid evidence. It is known that diet can cause rapid and substantial changes in the composition of the microbiome (19). However, there are other means to improve microbiome-related health outcomes, such as probiotics (“live microorganisms that, when administered in adequate amounts, confer a health benefit on the host”) (20).

Because the characteristics of probiotics enable the use of high-quality placebo formulations, and due to the aforementioned evidence-grading, much probiotic research has been performed through RCTs. However, these RCTs often yield inconsistent results. Even RCTs that are designed to meticulously repeat each other can yield conflicting results. As an example, a double-blind RCT by Kalliomäki et al. (21), investigating the effect of perinatal administration of *Lactobacillus rhamnosus* GG (ATCC 53103) on the primary prevention of atopic eczema, described a 50% reduction of the disease frequency compared with a placebo, a conclusion that remained also after a 4-y follow-up (22). Conversely, an independent study by Kopp et al. (23) that used the same strain and a nearly identical study design showed no preventive effect.

It can be concluded that the RCT approach is considered the gold standard for probiotics intervention studies as

well, but that in many cases the results are ambiguous. Researchers are increasingly aware of the issues implied by such heterogeneous findings and are exploring means to overcome these difficulties, e.g., by the use of machine learning techniques (24). However, in order to ensure that different means of investigation do not encounter the same issues, it is important to understand the fundamental causes of the ambiguous results of many probiotics RCTs. To the best of our knowledge, no studies have been conducted to identify these fundamental causes. Here, we review these limitations by investigating the underlying presuppositions.

Presuppositions behind RCTs

Several aspects in the design and conduct of an RCT influence the internal validity of its conclusions. Internal validity is defined as the “extent to which systematic error (bias) is minimised in clinical trials” (25), and is influenced by methodological issues, such as inappropriate generation or concealment of patient randomization, additional treatment or detection rates which are unequally distributed between intervention arms, loss of follow-up, nonadherence, or incorrect definition or violation of eligibility criteria. In order to minimize the impact of these and other methodological considerations, several quality criteria (QC) guidelines have been formulated. Whereas methodological issues affect the internal validity of RCT results, several practical concerns limit the possibility to use the RCT method in dietary studies (e.g., time for follow-up, cost, ethics, inhibition of innovative research questions) (26, 27).

A detailed discussion of these methodological and practical considerations is beyond the scope of this text. Instead, we ask the question: which fundamental principles lead to the QC and ultimately determine the validity of the conclusions drawn from RCT results in probiotic studies? To answer this question, we use an epistemological approach and evaluate the underlying presuppositions.

Presuppositions are a fundamental aspect of science and are as important to the scientific method as evidence and logic. They are starting points of a hypothesis which cannot be proven and which are, often implicitly, assumed to be true, and can more formally be defined as: “the presuppositions of a question are things that must be true for the question to have an answer” (28). Some presuppositions are generic to the scientific procedure; e.g., as the American Association for the Advancement of Science states in a position paper, “science presumes that the things and events in the universe occur in consistent patterns that are comprehensible through careful, systematic study” (29). As shown in **Figure 2**, presuppositions form the frame within which the evidence, logic, and QC become meaningful.

Whereas QC can be checked, validated, and proven to be met, presuppositions cannot. The presupposition that events in the universe occur in consistent patterns cannot be proven, nor disproven. To put this into a more relevant context: researchers undertaking a clinical trial to investigate a treatment for a specific disease presuppose that this method of investigation can be used to find the most effective

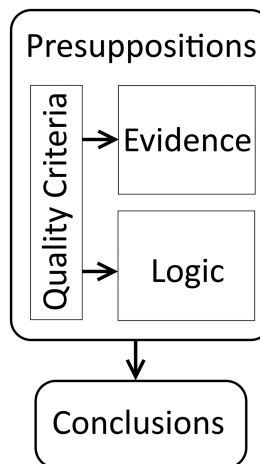


FIGURE 2 Relation between presuppositions, quality criteria, logic (and analyses), and evidence. Quality criteria are guidelines that aim to increase the quality of the evidence and logic in order to improve the validity of the conclusions. Presuppositions form the necessary frame within which research questions, approach, and assessment become meaningful and govern quality criteria, evidence, and logic. For the conclusions to be valid the presuppositions need to be true.

treatment. Again, it is not possible to prove nor to falsify this presupposition; for the trial to be meaningful, this presupposition has to be accepted by faith. This does not mean that presuppositions are arbitrary; they are legitimized within the context of a research topic through common sense. In order for the conclusions to be meaningful to the audience to whom they are disseminated, this common sense needs to be shared among the audience (28).

To better understand why the RCT method falls short within probiotic research, it is essential to explore the presuppositions underlying the RCT approach; a number of them are outlined in what follows. These presuppositions are made for any RCT, independent of the type of intervention, but their validity may differ from one study to another. As with any presupposition, the conclusions drawn from the RCT outcome are only valid if the presuppositions are true.

Uniformity

No single person is identical to another. Not only does individual constitution differ among people, there is also variation in lifestyle, genetics, diet, and health, among other factors. Any RCT has to cope with the reality of heterogeneity within the chosen population. It is presupposed that through the use of proper inclusion/exclusion criteria and randomization procedures, the participants in the different trial arms are uniform, which has also been referred to as “exchangeable at group level” (30). Conversely, presupposing uniformity in practice often also implies the idea that the eligibility criteria define the subpopulation to which the trial results apply. This is referred to as external validity (also termed applicability/generalizability), which is defined as the

relevance of the results to “a definable group of patients in a particular clinical setting” (31).

For many drugs, presupposing uniformity is seemingly acceptable, because the signal-to-noise ratio (i.e., the effect of the intervention compared with the effect of interpersonal variations) is sufficiently large in many cases. Sometimes there are known and recorded variations between the included participants (e.g., age or gender) that result in response differences within subgroups, which can be corrected for through stratification. In that case, the uniformity presupposition is deferred to the stratified subgroups.

There are circumstances in which the presupposition of uniformity may not hold. For example, even if the trial arms are truly exchangeable at baseline (i.e., at randomization), factors unrelated to the intervention may impair the exchangeability during the trial, also called postbaseline exchangeability (32). As an example, in a trial with a disease that has cyclic characteristics, all participants may be in the same disease phase at baseline, guaranteeing trial arm exchangeability at baseline. However, a change in disease phase may occur unevenly among trial arms during the trial, causing the trial arms to be no longer exchangeable (Table 1). In that case, the presupposition of uniformity (between trial arms) cannot be considered to be true and consequently no valid study-internal conclusions can be drawn.

In addition, effect-modifying factors can impair the validity of the group-definition, or external validity (30, 34). Effect-modifying factors are variables for which the effect of the intervention varies across different levels of the variable (30). As an example, in a study with nonsmokers, the treatment effect may differ between ex-smokers and never-smokers (Table 1). Proper randomization will ensure that the observed group-level results are internally valid for the particular study population enrolled in the trial. However, the conclusion may not be very usable because, even if a very narrowly defined group is enrolled, this group definition may not be sufficient to ensure a uniform group. A repeat of the trial that uses exactly the same stringent eligibility criteria may yield different results, due to a different level of effect-modifying factors in the new trial. Although gender, age, or ethnic differences for example are well-known effect modifiers and can often be corrected for through stratification, there can be other less obvious reasons for nonuniformity that may be unknown and which therefore cannot be corrected for. In that case, the often implied study-external conclusion, that the observed results apply to populations that meet the eligibility criteria, is invalid.

Independence of effects

Although interpersonal variations do result in variations of the obtained effects, in an RCT, the intervention is considered to be the sole cause of the observed results. This presupposes that there are no major interactions between the active component and other factors, allowing the inference of a causal relation. Causal inferences would not be valid if the effect size would strongly depend on the presence of additional, uncontrolled variables. Effect modification

can be one reason that this presupposition cannot be considered true, as exemplified by the impact of uncontrolled use of grapefruit juice on drug metabolism (Table 1). As previously mentioned, the uniformity presupposition can also be affected by effect modification. Though related, these 2 presuppositions are intrinsically different, yet are both fundamental to the RCT method: whereas uniformity is about the validity of group-definition, independence of effects focuses on the intervention and causal inference thereof.

A second related, but different, cause of invalidity of the independence of effects presupposition is the presence of interactions. Contrary to effect modification, interactions are characterized by a joint exposure of ≥ 2 factors that each affect the outcome measure (and not just modifies the effect of the intervention), and which can act either in synergy or antagonistic (35). One example of such interaction is the effect of magnesium and pharmaceutical treatment on pain, either alone or combined (36). As a simplified example of a situation where this affects the independence of effects presupposition, uncontrolled magnesium use may adjust the measured efficacy of an analgesic treatment, rendering the conclusion about the effect size of the tested treatment invalid (Table 1).

In general, researchers are forced to rely on and account for known effect modifiers (such as grapefruit juice) or interactions (such as magnesium) when designing an RCT. Hypothetical reasoning can be of use beforehand based on other studies, but even in well-designed and well-conducted RCTs, it is not possible to know every potential effect-modifying factor or interaction, especially when they include >2 variables. Consequently, there may remain effect modifications or interactions that are uncontrolled for, disallowing any valid study-internal conclusions because the presupposition that the response to a treatment is independent from other factors is not true.

The intervention and placebo are well-defined

As mentioned previously, causal inference is considered valid for most drug trials because the drug is considered the sole cause of the observed effect when compared with the placebo. This presupposes that the tested substance is well-defined, enabling the conclusion that administration of *A* yields effect *B*. However, as illustrated in Table 1, if the treatment is ill-defined, which aspect of it resulted in the observed overall effect?

Although the use of a well-defined intervention product or procedure may seem obvious, and is one of the reasons for stringent quality control in pharmaceutical products, this is not always possible with the use of nutritional products. As an example, pharmacokinetics studies on curcumin (a compound of *Curcuma longa*) suggest that addition of just 1% by weight of piperine (a compound of *Piper nigrum*) can result in a 2000% increase in the bioavailability (37). If this is the case, even tiny changes in product composition that may be well within manufacturing quality standards may yield very different effects and it may be difficult to properly define

TABLE 1 Presuppositions behind RCTs and simplified examples in which the RCT-presuppositions are not valid¹

| Presupposition | Simplified example within RCT framework | Invalid conclusion | Reason for invalidity of the conclusion | General implication, no valid conclusion can be drawn when: | Comments |
|---|---|---|--|---|--|
| Uniformity (effect modification) | Suppose that a very narrowly defined group of participants is enrolled in a trial testing a treatment against headache (e.g., Caucasian, male, age 55–60 y, nonsmoker, BMI 20–25 kg/m ² , not using drugs, >20 d headache/mo). Suppose that 60% are ex-smokers and they have a different response (+4 d headache/mo) than never-smokers (–6 d headache/mo). Ideal randomization is obtained. | The net result is no change in the number of days per month with headache, thus the treatment is ineffective for Caucasian males, aged 55–60 y, who are nonsmokers, etc., etc. (as per inclusion/exclusion criteria). | Even though a very narrow set of inclusion/exclusion criteria was used in this example, the observed result was due to a nonuniform group. When the trial is repeated with a different ratio of ex-smokers to nonsmokers, a different result will be obtained. | The definition of the included group is insufficient to obtain a uniform response and it is not possible to correct for effect-modifying factors. | Although smoking status (such as ex-smoker and never-smoker) is a known potential effect-modifying factor which can relatively easily be corrected for through stratification, other effect modifiers causing nonuniformity might be unknown (and not recorded), but could be a cause of unrepeatability of trial results. |
| Uniformity (postbaseline exchangeability) | Suppose that a trial is conducted to test the efficacy of a treatment for a disease with cyclic characteristics. Periods of physical decline are alternated with periods of stability or even slight improvement. Patients that are included have a similar disease-state and are all in a stable period when randomization is performed. One day into the 4-wk trial, the disease phase has changed to progressive in 40% of the placebo group and in 10% of the active group. | After the trial the treated group had less disease-progression, thus the treatment effectively slows disease progression. | Although the treatment and placebo arms were exchangeable at baseline (during randomization), they were no longer uniform at trial start. | There are relevant and substantial (postbaseline) differences between the different (treatment or placebo) groups within a trial. | Individuals are different by definition, thus groups of individuals are different as well. Whether or not these intergroup differences are relevant for a specific trial cannot always be known, because not all aspects that are relevant may be known. Even when differences are known and controlled for in the randomization procedure, these might change during the trial owing to aspects unrelated to the treatment. This risk is increased in diseases with rapid progression or with cyclic characteristics, or in trials with long follow-up periods. |
| Independence of effects (effect modification) | Suppose that a comparison study is conducted between drugs A and B and that the metabolism of drug A (but not drug B) to its active metabolite is slowed by grapefruit juice. The intake of grapefruit juice is uncontrolled for during the trial, only the intake of fruit juice in general is recorded. The group taking drug B has less disease progression than the group taking drug A. | The group taking drug B had a favorable outcome compared with the group taking drug A, thus drug B has a better efficacy than drug A. | In this example, the interaction of drug A with the grapefruit juice effectively decreased the blood concentrations of drug A's active metabolite, which resulted in an unfavorable observed efficacy of drug A. Had the intake of grapefruit (juice) not been permitted, the outcome would have been different. | There is ≥1 effect-modifying factors that are uncontrolled for. | The impact of grapefruit juice on the metabolism of certain drugs, such as statins and benzodiazepines, is well-known (33). However, even in well-designed and well-conducted RCTs it is not possible to know every potential effect-modifying factor. |

(Continued)

TABLE 1 (Continued)

| Presupposition | Simplified example within RCT framework | Invalid conclusion | Reason for invalidity of the conclusion | General implication, no valid conclusion can be drawn when: | Comments |
|---|--|--|---|--|--|
| Independence of effects (interaction) | Suppose that a treatment against pain is tested via an RCT. Unknown to the researchers, 40% of the participants take an over-the-counter magnesium supplement that acts synergistically to the treatment. Suppose that the pain-score improvement with magnesium alone is 3, with treatment alone 4, and combined 10. Randomization is ideal. The pain score improvement is 1.2 in the placebo group and 6.4 in the treated group. | The pain-score improvement was 5.2 points better for treatment than for placebo, thus the treatment efficacy is a 5.2-point improvement in the pain score. | The observed response in the treatment arm was partly due to the (uncontrolled) synergistic effect of magnesium. | There are interactions that have a substantial influence on the outcome, but which are not or cannot be corrected for. | |
| Intervention and placebo are well-defined | Suppose that a multicenter trial is conducted with a treatment consisting of 3 different substances and each center has to prepare the cocktail on-site, but the protocol does not properly define the ratios between the substances. In 2 out of 5 centers the outcome in the treatment group outperforms the placebo, in 1 center there is no difference, and in the other 2 centers the outcome in the placebo group was better than in the active group. | On average the treatment did not perform better than the placebo, thus the treatment consisting of substances A, B, and C is ineffective. | The result depends on the composition of the treatment cocktail. In this example, the composition varied between the centers because it was not well-defined. Effectively the 5 centers used 5 different treatments, which does not allow 1 overall conclusion. | The treatment varies in composition from participant to participant or over time; in other words, the treatment is not well-defined. | This is widely acknowledged in medicine and is, together with safety concerns, an important reason for rigorous production process control, quality assurance, and preclinical testing. This is also the reason that neither RCT QC nor ethical standards would allow a trial with such a poorly defined product. Moreover, the fact that production of medicines complies with stringent QC in turn means that is considered common sense to accept this presupposition to be true for pharmaceuticals. |

¹ Assuming a well-powered, well-controlled, and properly blinded trial. QC, quality criteria; RCT, randomized controlled trial.

the tested intervention. In addition, studies investigating natural materials should not only acknowledge but also account for compositional differences depending on variety, cultivation method, place of origin, or time of harvest. For example, one batch of blackcurrant (*Ribes nigrum*) may be very different from the next (38), further complicating proper definition of the tested intervention.

Thus, although a well-defined intervention is key to be able to draw valid conclusions from RCT results, whether or not this is always achieved remains a presupposition.

These 3 presuppositions are made for every RCT and if any of them is untrue, the trial will not deliver valid conclusions. However, this strict validity is not binary

and the impact on the study conclusions ultimately depends on the magnitude of nonconformance. Assessment of the validity of these presuppositions for specific RCT designs can help researchers to understand why some interventions find unreliable, mixed, negative, or positive outcomes.

Implications of Presuppositions on Probiotics RCTs

Many RCT trials with nutritional interventions do not yield convincing results for numerous reasons, including methodological and outcome-related factors. However, one may also question whether or not the underlying presuppositions can be considered true in such trials.

Contrary to pharmaceutical products, dietary supplements, or food products, probiotics exhibit one fundamental difference: they consist of living organisms. This results in several characteristics that affect the validity of the RCT-presuppositions. In the following sections, we will focus in on the gastrointestinal tract, but the arguments apply to other microbiome niches as well.

After administration, the living organisms arrive in the gut, an environment containing a vast number of other microbes as well as chyme of variable composition. Each bacterial strain in this complex ecosystem, whether probiotic or commensal, competes or cooperates with its neighbors and the hosts' systems. These interactions occur through the production and metabolism of a range of bioactive compounds. In fact, these compounds play a critical role in the physiologic or therapeutic effect observed with probiotics (39, 40). In addition, in situ abundance is in part determined by the gastric survival rate (which is affected by the presence of other foodstuff), by competition success, and by the interaction between the probiotic organisms and the host, and vice versa. With the effect of probiotics being so highly dependent on the multitude of interactions and effect modifiers, the validity of the presupposition of independence may be questionable.

A second issue arises because of the dependence on the aforementioned interactions: there is often substantial interpersonal variation of the endogenous microbiome as well as the host characteristics (41–43), which strongly influence each other (44, 45). Neither the commensal community nor the host tissue activity themselves are stable, because they both adjust to the ever-changing conditions in the intestinal tract. This implies that it is not possible to define a stable baseline. Consequently, one can legitimately question whether this can be regarded as “uniform”. In addition, the fact that the intestinal tract conditions are ever-changing implies that differences between trial arms can occur post-baseline, affecting the uniformity, or exchangeability, of trial arms and thus the validity of the study-internal conclusions. Moreover, the effect size of nutritional interventions is usually within the normal biological variability (4). This also applies to probiotics, as exemplified by the study by Van Baarlen et al. (42) and is captured by the “bandwidth of health” concept (43). That is, the ratio between the effective probiotic-induced “signal” and the “noise” of the multitude of signals induced (e.g., by the commensal microbiome, chyme constituents, and the host tissue itself) is relatively small. Thus, the relative influence of baseline differences between individuals on the treatment effect is more pronounced, and only certain baseline statuses may yield a physiologically relevant response. Contrary to many pharmaceutical treatments, the interindividual background variation is relatively large compared with the intraindividual effect of the intervention. Consequently, even small interpersonal differences may have a distinct impact, making trial population uniformity, and thus group definition, more questionable. In addition, probiotic intervention trials often involve a relatively long intervention duration to establish treatment effects, which

increases the risk of protocol deviations due to intercurrent illness, lifestyle changes, or noncompliance, which also affect the uniformity, because it impairs postbaseline exchangeability. Even if a trial is properly powered, these issues with uniformity can result in a situation where the effects cancel each other out, leading to a conclusion that the average effect is only marginally positive, absent, or even negative. Such a conclusion may be invalid if it is unreasonable to consider the trial population to be uniform.

Because probiotics are living microorganisms producing bioactive compounds which contribute to their therapeutic effects, the term “well-defined” can be confusing. Most probiotic formulations used in clinical trials or by healthcare providers are very well characterized in terms of the used strains, because stringent quality assurance requirements are followed for their production (46, 47). Thus, the probiotics that are administered to the patient can be considered to be well-defined in terms of ingredients. However, with regard to the bioactive compounds or downstream effects generated, the story is very different. Because these processes are highly dependent on, and variable due to, the aforementioned interactions, the physiologic exposure to these compounds may be regarded as similar to administration of composite drugs of which the composition varies from day to day and participant to participant. In other words, the same product does not imply the same treatment, and as such, the validity of the presupposition that the treatment is well-defined may be questionable.

Given the unique characteristics of probiotics, it seems reasonable to conclude that there will be many cases where the presuppositions that underlie the RCT method are not valid, implying that no valid conclusions can be drawn from the results. In these cases it is impossible to conclude that the intervention is either effective or ineffective, or to define a population for which it may be effective. A wide distribution of the response among the participants of an RCT, or a large heterogeneity of outcomes between high-quality RCTs, may be indications for a situation where the aforementioned mechanisms have a substantial influence and, consequently, where the presuppositions cannot be considered to be true.

How to Increase Validity?

If we accept that the presuppositions behind the RCT method are not fully met, it should be questioned whether this approach yields valid conclusions. It follows that we should either seek ways to meet the presupposed conditions, or explore different methods of investigation.

To identify potential methods to obtain a greater level of confidence in the internal validity of the conclusions, an important question is which goals to pursue when undertaking studies to investigate the efficacy of probiotic treatments. The decisive role that RCTs play in the economics of treatment development is in sharp contrast with the external validity of many RCTs. External validity is of vital importance to practitioners and regulatory agencies, because they seek the best and safest treatment for a particular patient. In the end, the fundamental goal of investigating the efficacy

of an intervention should be to aid the patient and support the patient care decision-making process. In other words, the conclusions of the total body of research should be applicable to individuals (strong external validity) and the results should help to advance the scientific understanding of said intervention (strong internal validity).

Given the reason that the presuppositions underlying RCTs are potentially invalid for probiotics, one way forward may be to identify participants who exhibit a similar gastrointestinal environment. With a greater level of similarity one may expect that the results of the interactions between the probiotic product and the host and commensal microbiome will be more uniform among the participants. One strategy might be to stratify patients before enrolment, or to use pretrial screening tests to differentiate between responders and nonresponders. In the latter approach, the probiotic could be briefly supplemented while monitoring certain temporary biomarkers (predefined in the study protocol). Only responders would be subsequently enrolled into a placebo-controlled trial. One potential issue would be a lasting pilot response effect, because this may confound the outcomes of the subsequent trial. As argued by Hanekamp et al. (48) in the case of bioactive components in foods, screening for small alterations of biomarkers within the normal homeostatic response width may be needed, because large responses to these types of compounds are rare. Therefore, combining the small response of a number of relevant biomarkers may enhance the validity of the screening results.

An important practical limitation of stratification is that to date, few reliable biomarkers are available (49). Additionally, it may be difficult to apply this approach in an everyday clinical setting, because many academically proposed biomarkers are not easily assessable within regular clinical practice. A pragmatic way to overcome this and to increase external validity might be to differentiate between responders and nonresponders based on a clinical assessment of absence or presence of a favorable effect. If this same selection process is clinically employed, the translation from trial results to the individual patient becomes more straightforward. An alternative approach, which is based on similar considerations, is an adaptive intervention; this is where the intervention is tailored to the individual by means of continuously monitoring the response and adapting the treatment (e.g., the dose) (15). When subjects are not enrolled at the same time, another implementation of an adaptive intervention is to adjust the probability of assigning subjects to a treatment arm based on the response obtained with previously enrolled subjects, reducing the probability of assigning subjects to an inferior intervention (50).

The best-case scenario for external validity would consist of a trial performed on participants that are identical to the patient in every sense. Although nonexistent at the group level, there are techniques that employ the patient as their own control, called n-of-1 trials (51). N-of-1 trials are randomized crossover studies with a single subject, and preferably use a double-blinded design. Because guidelines for conducting n-of-1 trials are readily available (52), any

clinician can perform this kind of research in their own practice (53). However, an aforementioned potential problem with the crossover design is the lasting effect, such that even a long washout period cannot prevent carryover effects. Consequently, testing a placebo after an active treatment may result in the absence of differences between interventions. On the other hand, such lasting effect is desirable for any treatment and when aware of this, the practitioner can take it into account when evaluating the results, possibly discontinuing the n-of-1 trial earlier.

When equivalent n-of-1 trials are conducted among a large number of individuals, meta-analyses may be used to derive more generally applicable conclusions (54). In light of today's ever-increasing computational power and data storage abilities, big data techniques have attracted the interest of the scientific community as a powerful method to enhance the level of differentiation (55, 56). Big data methods employ the exponentially growing and extremely versatile amount of data, can work with unstructured data (57), and can predict responsiveness to an intervention through identification of stratifying determinants. Various machine-learning techniques are being developed to extract medical concepts from unstructured data, such as natural language processing of free-text documents. Even the unstructured consultation notes of a practitioner could be used as a data source to guide informed treatment decisions based on the practitioner's own patient population. Thus, it is no longer necessary to conduct and report trials in an identical manner in order to extract generally applicable data. An example of employing big data-based machine-learning techniques is a study by Zeevi et al. (24), who used semicontinuous measurement of blood glucose concentrations, microbiome data from stool samples, and food intake to track and optimize the glycemic response of 800 study participants.

The various options outlined here are not fully defined methods, but merely starting points for approaches that may lead to a greater level of confidence in both the internal and external validity of conclusions. It is recommendable to employ these and other alternative approaches, in order to find ways to advance probiotic research beyond the limitations imposed by the RCT method.

Discussion

In the previous sections we have shown that there are legitimate reasons to question whether the aforementioned presuppositions can be considered true for probiotics RCTs in many cases, and that this may be the cause of heterogeneous findings. However, although inconsistent RCT results are found for some probiotic interventions, others succeed to repeatedly and consistently lead to the same overall conclusions. One example is the treatment of antibiotic-associated diarrhea with *Saccharomyces boulardii* probiotics. For this intervention, a recent Cochrane review of RCTs reported an average risk reduction of 53% and, more importantly, no significant heterogeneity among the included RCTs (58). This indicates that in certain applications the presuppositions may reasonably be regarded to be true.

Although it is not exactly known under which circumstances this may be the case, we propose that this may be due to a combination of 3 related aspects. First, a higher level of homogeneity may be found when the treatment depends on a local effect (e.g., within the gastrointestinal tract) rather than a systemic effect. Second, the mechanism of action of certain probiotic strains may be relatively independent of the microbiome composition and host's molecular expression, eliciting more or less the same effect in all individuals. Third, the signal (effect of probiotics)-to-noise (individual variation) ratio may be relatively high. Whatever the exact reason(s), it is reasonable to accept that if high-quality RCTs consistently deliver the same results, the underlying presuppositions are likely to be valid. The fact that the particularities of probiotics *could* affect the validity of the presuppositions underlying RCTs, is no reason to dismiss all probiotic RCTs. It is reason, however, to critically evaluate the validity of these implicit premises if the RCTs do yield inconsistent results, which is unfortunately not seldom the case.

In addition to the 3 presuppositions explored here, medical and nutritional research often is based on another presupposition as well: the idea that diseases and interventions can be studied through the use of a reductionist approach. This presupposition fuels the idea that a disease can be treated with the intervention by targeting a single pathway or set of pathways. This does not necessarily mean that the pathophysiology of the disease has to be fully understood or that the targets are known in detail, but only that the implicit premise is made that there is a pathway or set of pathways that are involved in the disease, which can be targeted with the intervention, and which is effective in treating the disease. However, this premise becomes problematic if the impact of the targeted pathways varies during the course of the disease, between patients, or from cell to cell, or if the effectiveness of modulation of the targeted pathways varies between cell types, tissues, or patients, and this is the root of the responder/nonresponder phenomenon. Cancer is a well-known example of a disease where many different pathways are involved in very similar types of cancer and consequently personalized treatment is a major research theme (59). Consequently, if trials are conducted within the frame of the specific target pathway premise, they are bound to either be underpowered, finding no treatment effect, or result in high numbers needed to treat (NNTs). The NNT expresses how many patients have to be treated in order to benefit one, and this can be as high as 100 for adjuvant therapies of resected tumors (60).

In the case of probiotics, the idea of targeting a specific set of pathways may be problematic. Even if an exact target pathway is known (influenced by the bacterium itself or a metabolite), it may depend on the baseline physiochemical and microbiome makeup of the individual whether or not this pathway can actually be relevantly modulated by the probiotic. This potentially leads to high NNTs, many nonresponders, or trials that appear to be underpowered.

Conclusions

Although the RCT method is considered the gold standard of clinical research, this status is not based on logic alone; presuppositions form the frame within which this approach makes sense. However, to paraphrase the previously given definition, the presuppositions of a scientific inquiry must be true for the inquiry to have an answer. Here, we have shown that within probiotic research it is in many cases reasonable to doubt the validity of the 3 RCT-presuppositions, which means that the RCT method may not be able to provide valid conclusions in this particular research area.

This conclusion may apply to investigations of the health outcomes of diet and nutrients in general. In order to determine whether or not this is the case for a particular research topic, the analysis of the plausibility of the presuppositions needs to be repeated. Some general outlines can be drawn, however. For example, a change in diet is likely to affect the entire local and systemic ecosystem. In addition, it is known that there are interactions between different nutrients, as exemplified by the competition for intestinal uptake that can occur between amino acids (61). With such a variety of interactions, it seems reasonable to conclude that in many cases the presupposition of independence is not valid. Moreover, in light of the highly diverse and personalized character of the complex intestinal ecosystem, the validity of the uniformity presupposition is likely also affected. Finally, the validity of the presupposition of a well-defined intervention and placebo may be questionable for several reasons (Box 1). These issues, together with the polyvalent character of nutrients (13), question the validity of the presupposition that an intervention can target a specific pathway.

Box 1. Reasons why a particular dietary intervention may not be well-defined

- The composition of foodstuffs can vary substantially.
- The food matrix or its preparation can affect the physiologic response (62), potentially making the usual characterization of dietary interventions (i.e., diet composition and nutritional value) insufficiently well-defined.
- The exact composition of the diet and the foodstuff quantity may be difficult to control and quantify.
- Complex interactions influence the uptake of nutrients and thus the physiologic exposure.

It can be concluded that there are several reasons to question the validity of the RCT-presuppositions for probiotic and nutritional research in general. Because no valid conclusions can be drawn if these presuppositions are not valid, different means of investigation should be explored. However, even when exploring other research methods, the critical question remains: which presuppositions are involved and can they be considered valid? Although the alternative research approaches discussed here may potentially provide means to consider the RCT-presuppositions to be true, they may also introduce new

presuppositions. The most important step to verify whether or not any presupposition can be considered true, is to make it explicit. This is especially important when a new method is introduced or when an existing method is adopted from one field to another, which was the case for the adoption of the RCT method from pharmaceutical research into microbiome studies and the greater arena of nutritional research.

Acknowledgments

We thank FM Rombouts, LMA Akkermans, S van Hemert, P Pekelharing, and D Veerman for fruitful discussions and their valuable input and feedback. DZ is partly funded by the Gut Flora Foundation and thanks them for enabling this project. The authors' contributions were as follows—DZ: created the conceptual idea; all authors contributed equally on the analysis of the concept and writing of the manuscript; and all authors have read and approved the final manuscript.

References

- Young SS, Karr A. Deming, data and observational studies. Significance [Internet] 2011;8:116–20.
- Ioannidis JPA. Implausible results in human nutrition research. BMJ [Internet] 2013;347:f6698.
- Maki KC, Slavin JL, Rains TM, Kris-Etherton PM. Limitations of observational evidence: implications for evidence-based dietary recommendations. Adv Nutr [Internet] 2014;5:7–15.
- Blumberg J, Heaney RP, Huncharek M, Scholl T, Stampfer M, Vieth R, Weaver CM, Zeisel SH. Evidence-based criteria in the nutritional context. Nutr Rev [Internet] 2010;68:478–84.
- Pan A, Lin X, Hemler E, Hu FB. Diet and cardiovascular disease: advances and challenges in population-based studies. Cell Metab [Internet] 2018;27:489–96.
- Burns PB, Rohrich RJ, Chung KC. The levels of evidence and their role in evidence-based medicine. Plast Reconstr Surg [Internet] 2011;128:305–10.
- Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S, Guyatt GH, Harbour RT, Haugh MC, Henry D et al. Grading quality of evidence and strength of recommendations. BMJ [Internet] 2004;328:1490.
- Schwingshackl L, Knüppel S, Schwedhelm C, Hoffmann G, Missbach B, Stelmach-Mardas M, Dietrich S, Eichelmann F, Kontopanteis E, Iqbal K et al. Perspective: NutriGrade: a scoring system to assess and judge the meta-evidence of randomized controlled trials and cohort studies in nutrition research. Adv Nutr [Internet] 2016;7:994–1004.
- Hardy A, Benford D, Halldorsson T, Jeger Michael J, Knutsen Helle K, More S, Naegeli H, Noteborn H, Ockleford C, Ricci A et al. Guidance on the use of the weight of evidence approach in scientific assessments. EFSA J [Internet] 2017;15:e04971.
- San Mauro-Martin I, Ruiz-León AM, Camina-Martín MA, Garicano-Vilar E, Collado-Yurrita L, Mateo-Silleras B de, Redondo Del Río MDP. Chromium supplementation in patients with type 2 diabetes and high risk of type 2 diabetes: a meta-analysis of randomized controlled trials. Nutr Hosp 2016;33:27.
- Singer GM, Geohas J. The effect of chromium picolinate and biotin supplementation on glycemic control in poorly controlled patients with type 2 diabetes mellitus: a placebo-controlled, double-blinded, randomized trial. Diabetes Technol Ther [Internet] 2006;8:636–43.
- Albarracín CA, Fuqua BC, Evans JL, Goldfine ID. Chromium picolinate and biotin combination improves glucose metabolism in treated, uncontrolled overweight to obese patients with type 2 diabetes. Diabetes Metab Res Rev [Internet] 2008;24:41–51.
- Heaney RP. Nutrients, endpoints, and the problem of proof. J Nutr [Internet] 2008;138:1591–5.
- Satija A, Yu E, Willett WC, Hu FB. Understanding nutritional epidemiology and its role in policy. Adv Nutr [Internet] 2015;6:5–18.
- Hébert JR, Frongillo EA, Adams SA, Turner-McGrievy GM, Hurley TG, Miller DR, Ockene IS. Perspective: randomized controlled trials are not a panacea for diet-related research. Adv Nutr [Internet] 2016;7:423–32.
- Your microbes, your health. Science [Internet] 2013;342:1440–1.
- FORTUNE Editors. 29 predictions for 2015: Fortune's crystal ball [Internet]. Fortune [cited 2017 Jan 10]. Available from: <http://fortune.com/2014/12/04/predictions-2015-future-business/>.
- Podolsky SH. Metchnikoff and the microbiome. The Lancet [Internet] 2012;380:1810–11.
- David LA, Maurice CF, Carmody RN, Gootenberg DB, Button JE, Wolfe BE, Ling AV, Devlin AS, Varma Y, Fischbach MA et al. Diet rapidly and reproducibly alters the human gut microbiome. Nature [Internet] 2013;505(7484):559–63.
- Hill C, Guarner F, Reid G, Gibson GR, Merenstein DJ, Pot B, Morelli L, Canani RB, Flint HJ, Salminen S et al. Expert consensus document: the International Scientific Association for Probiotics and Prebiotics consensus statement on the scope and appropriate use of the term probiotic. Nat Rev Gastroenterol Hepatol [Internet] 2014;11:506–14.
- Kalliomäki M, Salminen S, Arvilommi H, Kero P, Koskinen P, Isolauri E. Probiotics in primary prevention of atopic disease: a randomised placebo-controlled trial. The Lancet [Internet] 2001;357:1076–9.
- Kalliomäki M, Salminen S, Poussa T, Arvilommi H, Isolauri E. Probiotics and prevention of atopic disease: 4-year follow-up of a randomised placebo-controlled trial. The Lancet [Internet] 2003;361:1869–71.
- Kopp MV, Hennemuth I, Heinzmann A, Urbanek R. Randomized, double-blind, placebo-controlled trial of probiotics for primary prevention: no clinical effects of Lactobacillus GG supplementation. Pediatrics [Internet] 2008;121:e850–6.
- Zeevi D, Korem T, Zmora N, Israeli D, Rothschild D, Weinberger A, Ben-Yacov O, Lador D, Avnit-Sagi T, Lotan-Pompan M et al. Personalized nutrition by prediction of glycemic responses. Cell [Internet] 2015;163:1079–94.
- Jüni P, Altman DG, Egger M. Assessing the quality of controlled clinical trials. BMJ [Internet] 2001;323:42–6.
- Sanson-Fisher RW, Bonevski B, Green LW, D'Este C. Limitations of the randomized controlled trial in evaluating population-based health interventions. Am J Prev Med [Internet] 2007;33:155–61.
- Black N. Why we need observational studies to evaluate the effectiveness of health care. BMJ [Internet] 1996;312:1215–18.
- Gauch HG. Scientific method in practice. New York: Cambridge University Press; 2003.
- Rutherford FJ, Ahlgren A. Science for all Americans. New York: Oxford University Press; 1990.
- Hernán MA, Robins JM. Causal inference [Internet]. Boca Raton, FL: Chapman & Hall/CRC; forthcoming [cited 2018 Apr 25]. Available from: <https://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/>.
- Rothwell PM. External validity of randomised controlled trials: “to whom do the results of this trial apply?” The Lancet [Internet] 2005;365:82–93.
- Toh S, Hernán MA. Causal inference from longitudinal studies with baseline randomization. Int J Biostat [Internet] 2008;4(1):22.
- Kane GC, Lipsky JJ. Drug–grapefruit juice interactions. Mayo Clin Proc [Internet] 2000;75:933–42.
- Dekkers OM, von Elm E, Algra A, Romijn JA, Vandenbroucke JP. How to assess the external validity of therapeutic trials: a conceptual approach. Int J Epidemiol [Internet] 2010;39:89–94.
- Corraini P, Olsen M, Pedersen L, Dekkers OM, Vandenbroucke JP. Effect modification, interaction and mediation: an overview of theoretical insights for clinical investigators. Clin Epidemiol [Internet] 2017;9:331–8.

36. Kim EM, Kim M-S, Han S-J, Moon BK, Choi EM, Kim EH, Lee J-R. Magnesium as an adjuvant for caudal analgesia in children. *Paediatr Anaesth* 2014;24:1231–8.
37. Sharma RA, Steward WP, Gescher AJ. Pharmacokinetics and pharmacodynamics of curcumin. In: BB Aggarwal, Y-J Surh, S Shishodia, editors. *The molecular targets and therapeutic uses of curcumin in health and disease*. Boston, MA: Springer; 2007. p. 453–70.
38. Khoo GM, Clausen MR, Pedersen HL, Larsen E. Bioactivity and chemical composition of blackcurrant (*Ribes nigrum*) cultivars with and without pesticide treatment. *Food Chem* 2012;132:1214–20.
39. O’Shea EF, Cotter PD, Stanton C, Ross RP, Hill C. Production of bioactive substances by intestinal bacteria as a basis for explaining probiotic mechanisms: bacteriocins and conjugated linoleic acid. *Int J Food Microbiol* 2012;152:189–205.
40. Arpaia N, Campbell C, Fan X, Dikiy S, van der Veeken J, deRoos P, Liu H, Cross JR, Pfeffer K, Coffey PJ et al. Metabolites produced by commensal bacteria promote peripheral regulatory T cell generation. *Nature* [Internet] 2013;504:451–5.
41. Consortium THMP. Structure, function and diversity of the healthy human microbiome. *Nature* [Internet] 2012;486:207–14.
42. van Baarlen P, Troost F, van der Meer C, Hooiveld G, Boekschoten M, Brummer RJM, Kleerebezem M. Human mucosal in vivo transcriptome responses to three lactobacilli indicate how probiotics may modulate human cellular pathways. *Proc Natl Acad Sci USA* [Internet] 2011;108:4562–9.
43. Bron PA, van Baarlen P, Kleerebezem M. Emerging molecular insights into the interaction between probiotics and the host intestinal mucosa. *Nat Rev Microbiol* [Internet] 2012;10:66–78.
44. Shulzhenko N, Morgun A, Hsiao W, Battle M, Yao M, Gavrilova O, Orandle M, Mayer L, Macpherson AJ, McCoy KD et al. Crosstalk between B lymphocytes, microbiota and the intestinal epithelium governs immunity versus metabolism in the gut. *Nat Med* [Internet] 2011;17:1585–93.
45. Hooper LV, Littman DR, Macpherson AJ. Interactions between the microbiota and the immune system. *Science* [Internet] 2012;336:1268–73.
46. Sanders ME, Merenstein DJ, Ouwehand AC, Reid G, Salminen S, Cabana MD, Paraskevacos G, Leyer G. Probiotic use in at-risk populations. *J Am Pharm Assoc* [Internet] 2016;56:680–6.
47. International Probiotics Association. 2017 best practices guidelines [Internet]. [cited 2017 Mar 5]. Available from: <http://internationalprobiotics.org/resources/guidelines/2017-best-practices-guidelines/>.
48. Hanekamp JC, Bast A, Calabrese EJ. Nutrition and health – transforming research traditions. *Crit Rev Food Sci Nutr* [Internet] 2015;55:1074–80.
49. Reid G, Gaudier E, Guarner F, Huffnagle GB, Macklaim JM, Munoz AM, Martini M, Ringel-Kulka T, Sartor BR, Unal RR et al. Responders and non-responders to probiotic interventions. *Gut Microbes* [Internet] 2010;1:200–4.
50. Ghiglietti A, Scarale MG, Miceli R, Ieva F, Mariani L, Gavazzi C, Paganoni AM, Edefonti V. Urn models for response-adaptive randomized designs: a simulation study based on a non-adaptive randomized trial. *J Biopharm Stat* 2018;1–13.
51. Schork NJ. Personalized medicine: time for one-person trials. *Nature* [Internet] 2015;520:609–11.
52. Kravitz R, Duan N, editors. *Design and implementation of n-of-1 trials: a user’s guide*. Rockville, MD: Agency for Healthcare Research and Quality, US Department for Health and Human Services; 2014. Available from: <https://effectivehealthcare.ahrq.gov/topics/n-1-trials/research-2014-5/>.
53. Tsapas A, Matthews DR. N of 1 trials in diabetes: making individual therapeutic decisions. *Diabetologia* [Internet] 2008;51:921–5.
54. Lillie EO, Patay B, Diamant J, Issell B, Topol EJ, Schork NJ. The n-of-1 clinical trial: the ultimate strategy for individualizing medicine? *Pers Med* [Internet] 2011;8:161–73.
55. Murdoch TB, Detsky AS. The inevitable application of big data to health care. *JAMA* [Internet] 2013;309:1351–2.
56. Marx V. Biology: the big challenges of big data. *Nature* [Internet] 2013;498:255–60.
57. Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. *Health Inf Sci Syst* [Internet] 2014;2:3.
58. Szajewska H, Kolodziej M. Systematic review with meta-analysis: *Saccharomyces boulardii* in the prevention of antibiotic-associated diarrhoea. *Aliment Pharmacol Ther* [Internet] 2015;42:793–801.
59. Simon R, Roychowdhury S. Implementing personalized cancer genomics in clinical trials. *Nat Rev Drug Discov* [Internet] 2013;12:358–69.
60. Kruse M, Parthan A, Coombs J, Sasane M, Taylor D. Comparison of different adjuvant therapies for 9 resectable cancer types. *Postgrad Med* [Internet] 2013;125:83–91.
61. Gropper SAS. *Advanced nutrition and human metabolism*. 5th ed. Belmont, CA: Wadsworth/Cengage Learning; 2009.
62. Flood-Obbagy JE, Rolls BJ. The effect of fruit in different forms on energy intake and satiety at a meal. *Appetite* [Internet] 2009;52: 416–22.