Gut Microbes 1:4, 1-11; July/August 2010; © 2010 Landes Bioscience

Guide to designing, conducting, publishing and communicating results of clinical studies involving probiotic applications in human participants

Andi L. Shane,*[1] Michael D. Cabana,[2] Stéphane Vidry,[3] Dan Merenstein,[4] Ruben Hummelen,[5] Collin L. Ellis,[6] James T. Heimbach,[7] Susanne Hempel,[8] Susan V. Lynch,[9] Mary Ellen Sanders[10] and Daniel J. Tancredi[11]

[1]Department of Pediatrics; Emory University School of Medicine; Atlanta, GA USA; [2]Departments of Pediatrics; Epidemiology and Biostatistics and the Philip R. Lee Institute for Health Policy Studies; University of California; San Francisco (UCSF); San Francisco, CA USA; [3]International Life Sciences Institute (ILSI)-Europe; Brussels, Belgium; [4]Department of Family Medicine; Georgetown University; Washington, D.C USA; [5]Lawson Health Research Institute; London, Ontario, Canada and Erasmus MC; University Medical Centre; Rotterdam, The Netherlands; [6]Programs in Nutritional Biology, Biotechnology and Biophotonics; Department of Internal Medicine; Genome Facility; University of California; Davis Health System-School of Medicine; CA, USA; [7]JHeimbach; LLC; Port Royal, Virginia USA; [8]RAND Corporation; Santa Monica, CA USA; [9]Colitis and Crohn's Disease Center; Division of Gastroenterology; Department of Medicine; University of California; San Francisco (UCSF); San Francisco, CA USA; [10]Dairy & Food Culture Technologies; Centennial, CO USA; [11]Department of Pediatrics; University of California; Davis, CA USA

Key words: probiotics, clinical trials, human, guidelines, study design

Abbreviations: AFLP, Amplified Fragment Length Polymorphism; AHRQ, Agency for Healthcare Research and Quality; CONSORT, Consolidated Standards of Reporting Trials; DOT, Directly Observed Therapy; EFSA, European Food Safety Authority; ERIC, Enterobacterial Repetitive Intergenic Consensus; FDA, Food and Drug Administration; IBS-QOL, Irritable Bowel Syndrome-Quality of Life; IgE, Immunoglobulin E; ILSI, International Life Sciences Institute; IND, Investigational New Drug; ISAPP, International Scientific Association for Probiotics and Prebiotics; NCCAM, National Center for Complementary and Alternative Medicine; NEC, Necrotizing Enterocolitis; NIH, National Institutes of Health; ODS, Office of Dietary Supplements; OECD, Organisation for Economic Co-operation and Development; PCR, Polymerase Chain Reaction; RAPD, Random Amplified Polymorphic DNA; RCT, Randomized Controlled Trial; SF-36, Short Form 36

The heterogeneity of human clinical trials to assess the effectiveness of probiotics presents challenges regarding interpretation and comparison. Evidence obtained from clinical trials among a population with a disease or specific risk factors may not be generalizable to healthy individuals. The evaluation of interventions in healthy persons requires careful selection of outcomes due to the absence of health indicators and the low incidence of preventable conditions. Given the tremendous resources invested in such trials, development of consistent approaches to assessing the effectiveness of probiotics would be beneficial. Furthermore, the reporting, presentation and communication of results may also affect the validity of the scientific evidence obtained from a trial. This review outlines the challenges associated with the design, implementation, data analysis and interpretation of clinical trials in humans involving probiotics. Best practices related to their design are offered along with recommendations for enhanced collaboration to advance research in this emerging field.

*Correspondence to: Andi L. Shane; Email: ashane@emory.edu
Submitted: 03/31/10; Revised: 06/15/10; Accepted: 06/16/10
Previously published online:
www.landesbioscience.com/journals/gutmicrobes/article/12707

Introduction and Scope

The seventh Annual Conference of the International Scientific Association for Probiotics and Prebiotics (ISAPP) was held from November 4 through 5, 2009 in Irvine, California. ISAPP is an international collaboration of scientists and representatives from industry who exchange views and discuss topics related to scientific application of probiotics and prebiotics. The discussions of the 16 member work group titled, "Parameters for Designing, Publishing and Communicating Human Studies on Human Probiotics" formed the basis of this review.

The growing interest in the application of probiotics to the maintenance and improvement of human health is evidenced by a recent National Library of Medicine database search using the keywords "probiotic, human, clinical trial" spanning the dates 1992 to 2010. Over 1,186 peer-reviewed publications identified; of which 338 were review articles. Due to the interest in clinical research studies in humans involving probiotics, the ISAPP working group set out to identify the challenges of designing, conducting and interpreting results from human probiotic studies and to develop strategies to serve as standards for future endeavors.

This review will discuss study design, target populations, selection of placebo and probiotic microorganism(s), duration of follow up, outcome and endpoint measurements, safety assessments and regulatory considerations. The issues associated with each of

these topics and their contribution to clinical trials in probiotics in human participants will be detailed. Recommendations for optimizing clinical research involving probiotics will be addressed.

Study Design

Although other types of experimental and observational studies may contribute high quality scientific evidence to support claimed relationships between the use of probiotics and health and safety outcomes, well-performed randomized, controlled trials provide the strongest causal evidence. The latter are necessary to establish significant scientific agreement with the causal claims evaluated by government regulators. Considerations that influence key elements of the trial's design include the hypotheses and specific aims of the investigation, the resources available for the study, and human participant concerns including the protection, safety, privacy and autonomy of study participants. Trials should be designed and conducted using sound scientific principles and methods to minimize bias in the assessment of intervention effects. Main sources of bias can arise from systematic differences between comparison groups, how treatments are randomized and allocated to trial participants, how care is delivered, and how outcomes are assessed, as well as differences in study completion rates.[1] Well-designed, randomized-controlled trials use a variety of methods to minimize these biases, including randomization of subjects to study interventions, blinding of participants, caregivers and assessors and rigorous statistical comparisons between groups to estimate intervention effects.

Components that should be specified during the design phase of the study include: (1) the target population of interest, (2) the specific intervention(s) under study, (3) the control groups, (4) key safety and efficacy outcomes, including the designation of primary outcomes, (5) research hypotheses concerning the effects of experimental and comparison treatments on these outcomes and (6) the type and size of study for addressing hypotheses and producing evidence that generalizes to the target population.[2,3] Although on occasion, key safety and efficacy outcomes and research hypotheses are finalized before beginning the design phase, more often than not, these components are developed and refined iteratively during the design phase. A tentative specification of key study outcomes and hypotheses is followed by the development of a study design to address them.

Each of these components is then refined as necessary, given feasibility constraints. For example, if a researcher is interested in studying probiotics for preventing ulcerative colitis, a condition for which effective interventions exist, a placebo-controlled, superiority trial may not be ethical. However, an active-controlled non-inferiority trial that compares an investigational probiotic to previously recognized treatments would satisfy the principles of good clinical practices.[4]

Important caveats regarding study design pertain to the impact and validity of the outcome measurement(s). These should be selected based on their clinical relevance; if surrogate disease measures will be used, their relevance must be proven.[5] To justify health claims, a specific measurable disease or health-related condition must be identified and assessed by measuring the incidence, associated morbidity or scientifically validated surrogate endpoints for the specified disease or condition.

Target population for study. The results of a study should be generalizable to the consumers of the product. For study acceptance, The European Food Safety Authority (EFSA) requires that a study population be representative of the target population. Similarly, the United States Food and Drug Administration (FDA) requires that studies cited in support of health claims for foods be "scientifically appropriate" to extrapolate from the study sample to the population that is the subject of the health claim.[6] While institutions and clinics with academic affiliations and mass media advertisements are popular locations and methods of recruitment, unless these sources provide the population of interest, the external validity of the study may be greatly reduced. As the claims for probiotic products are applied to broader populations, it is important to ensure that a variety of subjects participate in such trials to ensure validity. To address these issues, researchers may take advantage of "grassroots efforts" including community-networking, targeted media advertisements, individual community practitioners and practice-based research networks.

Comparison group. The selection of a comparison group is a challenging issue in any study, especially when a probiotic is delivered in a food. When ethical and feasible, placebo agents provide an ideal comparison group. A placebo should be as similar to the active intervention as possible, with the exception of the presence of the probiotic under investigation. A report of a joint FAO/WHO expert consultation team recommends "For the testing of probiotic foods the placebo would be comprised of the food carrier devoid of the test probiotic".[7] However, food carriers may be comprised of multiple ingredients, blurring the differentiation between the study product and the food carrier. For example, when testing efficacy of a yogurt supplemented with a probiotic, several different placebos could be imagined. The placebo might be acidified milk with no active cultures, heat-treated probiotic yogurt containing fermentation end products and bacterial cell components or yogurt without the probiotic but retaining live, active starter bacteria. The choice depends on the research question being asked. However, regardless of the placebo chosen, it is imperative that the rationale for the decision is clearly described in the study protocol.

Length of follow up. There is no a priori preferred length of follow-up for subjects in clinical studies involving probiotics. The selection of an appropriate follow-up duration depends on the study question, outcome of interest and resources. For example, mortality outcomes require follow-up periods of many years, while morbidity outcomes such as duration of diarrhea may require follow-up for a few days. Typically, the power of a study for detecting important clinical effects on morbidity and mortality outcomes is enhanced by longer follow-up periods. If outcomes have seasonal patterns, (i.e., colds and influenza-like illness), follow-up for an entire year may be necessary. However, limited study resources and the urgency to complete studies as rapidly as possible argue for shorter follow-up periods, necessitating a suitable compromise that should be justified in study descriptions. One rationale is to extend follow up at least as

long as persistence of the probiotic in vivo—typically less than four weeks.

Sample size. A characteristic of a well-designed clinical trial is that it includes a sufficiently large sample size to provide a reliable assessment of study questions. Furthermore, it should have sufficient power to detect clinically significant effects *vis a vis* the primary outcome. A common misconception about sample size is that the larger the sample size, the better the study. It is more important that the study be appropriately powered at the design stage to detect realistic and meaningful effects on clinically relevant and reliably measured outcomes. The planned study design and analysis should account for the impact of loss to follow up and imperfect adherence on the effective sample size of the study. Two or three small studies with comparable conclusions may provide more robust data than one extremely large study, regardless of the statistical significance or p-value in the single large study. Additionally, effect size and thus sample size will be determined by the outcome of interest. For example, in a study of necrotizing enterocolitis (NEC), a low-cost and well-tolerated probiotic supplement may offer clinically meaningful effects, even if it only reduces the incidence of this serious condition by 4 or 5 percentage points. For detecting such small effects, larger sample sizes are required; however smaller sample sizes are often chosen due to financial constraints. This practice increases the chance that a smaller but clinically meaningful association will be missed, a Type II error.

The CONSORT standards of reporting trials. To standardize the reporting of results of randomized controlled trials (RCTs), with specific emphasis on describing key determinants of evidentiary quality, a standardized 22-item checklist and flowchart was devised by leading medical journal editors. The Consolidated Standards of Reporting Trials (CONSORT) Statement facilitates assessment of the design, analysis and interpretation of trials by readers. A companion flow diagram enables readers to track the progress of all participants through the trial. The CONSORT checklist has been translated into 10 languages, in addition to English, and has become a de facto standard for reporting of clinical trials.[8,9]

As CONSORT items reflect widely accepted markers of study quality, they serve as a helpful reference for investigators in the design, conduct, analysis and reporting of clinical trials. They further provide readers with information on patient flow, attrition and other sources of bias. For example, an enumerated item in the CONSORT statement concerns allocation concealment. Defined as the process of preventing foreknowledge of a potential enrollee's randomized treatment assignment, this blinding is crucial for minimizing bias in randomized trials. A clear statement that allocation concealment was used, is a widely used indicator of study quality for randomized controlled trials.[10,11] Similarly, intent-to-treat analyses offer protection against bias arising from selection effects and differential attrition. This approach, comparing study outcomes based on assigned treatment is preferred by many evaluators.[5,12,13]

Registration of clinical trials. The highest quality scientific evidence results when experiments are used to evaluate hypotheses that were declared prior to embarking on the experiment and when results of the study, whether positive or negative, are publically accessible. Concerns for the quality and availability of safety and efficacy data from industry-sponsored clinical trials led to the establishment of numerous online registries of clinical trials. Mandatory clinical trial registration is a standard of many regulatory bodies as well as journals and provides transparency for stakeholders. Additionally, these publically available databases of information may serve as a source for collaboration and partnership. To fulfill the objectives of good clinical practices, it is imperative that all trials be registered through a publicly accessible online clinical trial registry, accounting for all primary and secondary outcomes before subject enrollment.

Baseline Data

Randomization does not always guarantee equal distribution of patient characteristics across all groups. As a result, it is important to inspect data to assess for differences in the treatment groups on baseline (pretreatment) measures. After randomization, the groups can be compared to determine if there are important differences (e.g., randomization was not effective) between the groups. In general, this comparison is made by graphical and formal statistical methods to detect differences that may have occurred by chance. If major differences exist between the groups, specific techniques (e.g., stratification, multiple regression) can be used to mitigate these differences during the analysis stage. A clear statement of the intention to use such methods should be provided in the formal statistical analysis plan to protect against bias.[2]

Intervention

Identification and classification of therapeutic microorganisms. A number of probiotic species and strains used in trials involving human participants with a variety of conditions have demonstrated efficacy.[14-20] However, a key concern is the lack of both standardization and microbial characterization in the manufacturing of products being tested. To examine this issue, several recent studies have assessed commercially available probiotic supplements and probiotic-enriched foods to evaluate the accuracy of label claims. Discrepancies between reported and actual bacterial numbers present,[21] misidentification of starter cultures used for industrial-scale production,[22] absence of claimed probiotic strains and presence of non-reported strains[21,23] have been described. Inaccurate declaration of identity generally results from a lack of rigor in using newly available molecular tools as standard practice. Given the reported discordance between product content claims and actual content, it is reasonable to question the composition and potency of products used in some clinical trials. This highlights the importance of careful product characterization and potency assessment and reporting in clinical trials. Furthermore, this situation may account for some reported conflicting results.

The most frequently used bacterial probiotic species include members of the *Lactobacillus* and *Bifidobacterium*. Historically, species identification relied on phenotypic attributes, such as metabolite production and fermentation capability, to identify

and classify bacteria, particularly the Lactobacillaceae. The advent of genotypic characterization, largely based around sequence analysis of the 16S rRNA gene has led to reliable methods for species identification and for most species comprise the best approach to species determination.[24-27] This approach involves amplification of the target gene and phylogenetic assignment based on the sequence of specific variable regions within the gene. While many bacterial genes represent good targets for this approach, the 16S rRNA gene has been used most extensively and as a result the largest database of reference sequences are available for this genetic region. 16S rRNA sequence-based classification is a useful approach for identification of species for some organisms. In addition, probiotic efficacy is typically predicated on strain-specific characteristics and while a number of discriminatory loci have recently been described for Lactobacillus species-level differentiation.[28,29] Specific regions of the genome that distinguish bacteria at the strain level have not been determined. Therefore, identification of lactobacilli using 16S rRNA sequencing approaches exclusively likely will not provide the resolution necessary to discriminate specific probiotic strains. Polyphasic identification and classification[30] that utilizes a combination of genotypic and phenotypic approaches is currently the gold standard for accurate identification and classification of closely related probiotic species.[31]

More recently, PCR-based fingerprinting tools such as Random Amplified Polymorphic DNA (RAPD), Enterobacterial Repetitive Intergenic Consensus (ERIC) and Amplified Fragment Length Polymorphism (AFLP) have been successfully used to type members of the Lactobacillaceae.[32-34] A unique RAPD PCR product has recently been shown to be specific to *Lactobacillus rhamnosus* GG, and has been used to design quantitative PCR primers specific to this strain.[35]

To provide a more global analysis of strain capability, independent research groups have recently developed novel lactic acid bacteria-specific and bifidobacteria-specific microarrays to profile functional gene expression of these species using multiple genes involved in carbohydrate, pyruvate and amino acid metabolism, antibiotic resistance and bacteriocin production.[36-38] Such functional profiling under standardized conditions, may also provide a more comprehensive approach for identification and classification of probiotic strains. "Macroarray" approaches have also been designed to identify species using whole genome DNA-DNA hybridization. The advantage of this approach is that the genome from specific strains may be used as probes and that a large numbers of probes, generated from a variety of strains, may be used in parallel in a single assay.[39] These molecular approaches in development represent portable, rapid typing tools and provide the sensitivity and specificity necessary for accurate identification and classification of probiotic strain.

Probiotic viability. A unique aspect of probiotics is their viable nature. Maintaining probiotics in a live and robust state during the course of a study is essential. Assessment of the viable count of the probiotic product used at initiation and conclusion of a study is necessary to be certain that the potency is maintained. For a long-term study covering several months or more, assessment at one or more mid-points is advisable. Reporting of actual viable count and target viable count of the probiotic test article is recommended.

Probiotic viability is a function of numerous factors, summarized in **Figure 1**.[40] Factors that potentially influence probiotic viability and physiological state include production parameters, product parameters and features of the host. The extent to which these factors influence probiotic physiology should be expected to be strain-specific. The amount of active probiotic organisms recovered when isolated from different sites or products will depend on methods used for enumeration.[40-42]

Probiotic administration, delivery and host localization. Probiotic administration is typically via the oral,[43-45] rectal,[46,47] or vaginal[48,49] routes. Probiotic strains targeted at intestinal sites are often selected for their ability to withstand transit through the harsh environment of specific niches in the gastrointestinal tract. However, the materials chosen for microencapsulation of live microbial supplements can further enhance survival. Use of natural biodegradable polymers or those minimally engineered to enhance delivery of encapsulated microbial cells to the lower gastrointestinal tract, particularly the colon, are the focus of substantial current research.[50] While our understanding of the bacterial species that promote host health is dramatically increasing, parallel development of improved delivery systems represent a crucial component for efficacious probiotic therapeutics.

Successful transit of sufficient microbial numbers to target sites also appears to be a key aspect of probiotic efficacy. Several animal studies have demonstrated the localization of supplemented organisms at specific sites along the gastrointestinal tract, such as the Peyer's patches and in the mesenteric lymph nodes.[51,52] These studies also demonstrated that gastrointestinal colonization can be variable, even between isogenic animals,[53] a key finding given the added variability introduced when studying human populations.

Endpoint Selection

The characteristics of the trial endpoints will dictate to a large degree the design of a clinical trial. Several characteristics can be considered when selecting an endpoint. First, an endpoint should be *feasible to measure*, both technically and from a human participants' perspective. Second, an endpoint should be *sensitive* to the desired effect of the intervention being studied. Third, an endpoint should be a measure of *important benefit* to the population being studied and fourth, an endpoint should be an *objective* measurement.[54]

Surrogate endpoints. A surrogate endpoint is a marker that substitutes for a clinical endpoint. Often these markers are laboratory values or imaging results that are believed to be involved in the causal pathway of a particular condition. The use of surrogate endpoints is appealing as it may facilitate the conduct of studies with a smaller sample size and a shorter follow-up than would otherwise be possible. However, this approach has pitfalls. For example, interventions that decreased cardiac arrhythmias (the surrogate endpoint) were widely marketed under the assumption that they would also decrease mortality (clinical endpoint). However, post-marketing review showed that they actually increased mortality.[55] To avoid a situation such as this, a selected

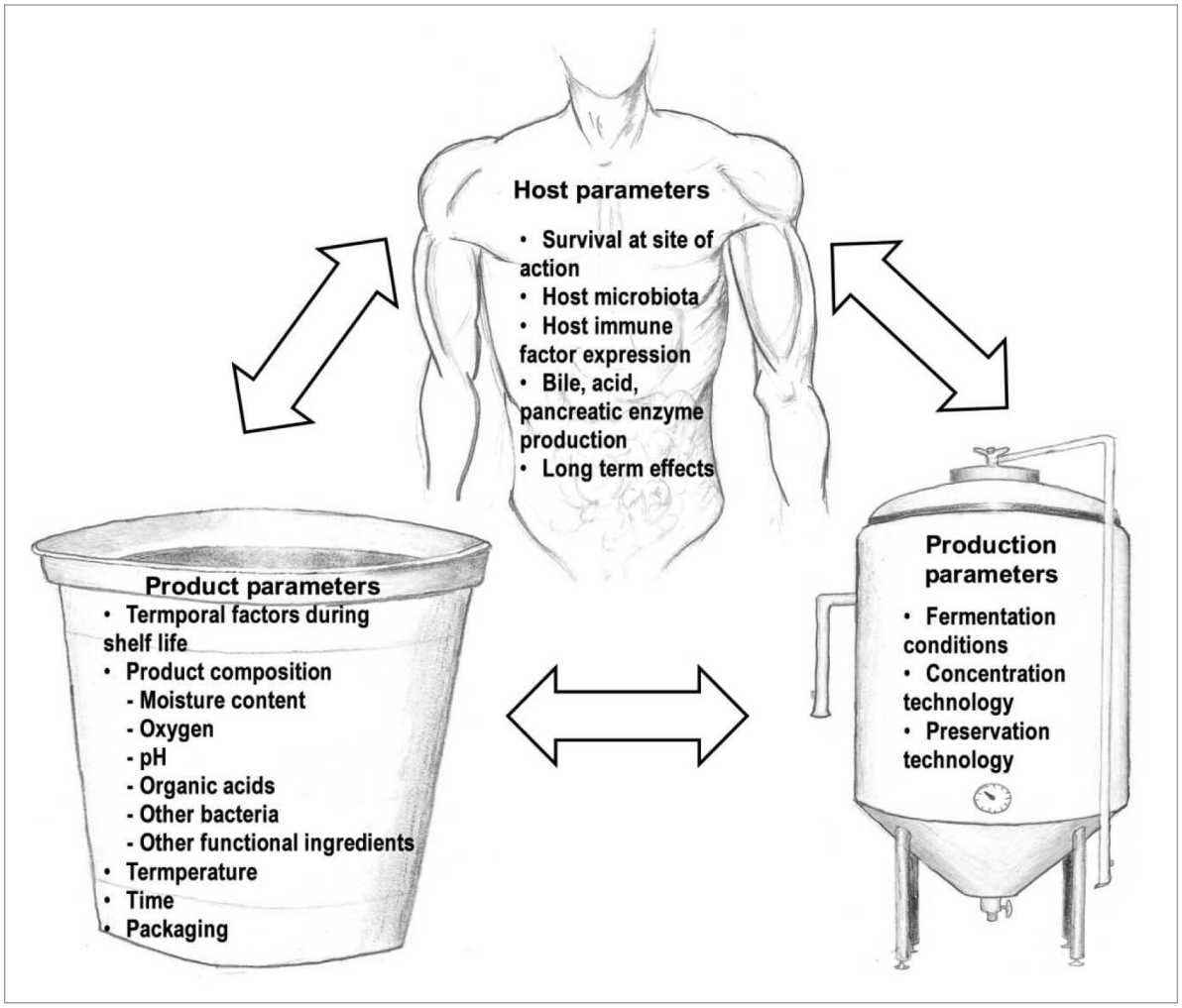

Figure 1. Factors that potentially impact probiotic viability and physiological state. Factors include production parameters, product parameters and features of the host. The extent to which these factors influence probiotic physiology should be expected to be strain-specific. The amount of active probiotic organisms recovered when isolated from different sites or products will depend on methods used for enumeration. Reprinted with permission from Sanders and Marco, 2010.

surrogate marker should fulfill stringent criteria. Firstly, a surrogate marker should have a good positive predictive value. For example, if a surrogate marker predicts that there is benefit to an intervention, it will have the same effect on a clinical endpoint. Conversely, if a surrogate marker predicts that an intervention is not beneficial, the intervention should also not have an effect on the clinical endpoint. Secondly, there should be an understanding of the role of the surrogate marker in the causal chain of events and this should be a validated measure in both the surrogate endpoint and the clinical endpoint.[56] The use of surrogate endpoints that do not meet these criteria in clinical trials should be avoided, as the scientific evidence generated might not be sufficiently correlated with the clinical endpoint.[57]

Laboratory endpoints. Endpoints measured in the laboratory can be of value for understanding physiological processes or for documenting potential safety parameters that may otherwise remain undetected. As alterations in laboratory values often precede clinical symptom onset, the former may serve as an early indicator of dysfunction.[58] The same considerations should be applied to the selection of laboratory and clinical endpoints; they should be feasible, sensitive, a measue of an important benefit, and objective. Other important considerations are the variability and timing of laboratory measurements. For example, since allergen-specific immunoglobulin E (IgE) has a half-life of 6 months to 2 years, it would not be a useful surrogate marker in a study with a follow up period of a shorter duration.

Participant reported endpoints. Endpoints directly reported by the participant are a valuable contribution to a clinical trial as these subjective assessments may provide a measurement of the impact of an intervention on lifestyle. Participant-reported endpoints can be designed to target a relatively simple endpoint (e.g., headache), a more complex endpoint (e.g., ability to perform daily activities) or even an extremely complex endpoint (e.g., Quality of life). Quality of life assessments may consist of multiple domains that consider the effects of the physical, emotional and social components, however quality of life should not

Table 1. Comparison of methods for measuring adherence in probiotic protocols

Method	Advantages	Disadvantages
Directly observed therapy (DOT)	Objective; tends to be the most accurate May be more feasible in an inpatient setting for short-term protocols	Expensive Easier to apply in an inpatient setting
Measurement of probiotic in stool	Objective Allows assessment probiotic exposure in 'control' subjects	May unblind investigator Only measures recent adherence Probiotic may still be detected with infrequent adherence
Electronic medication monitor	Objective Precise	Expensive; requires downloading data and specific equipment Not applicable for patients who use separate pill boxes Data can be altered
Pill count	Objective	Can be easily altered
Patient questionnaire	Simple to use	Error secondary to recall or distortion by patient
Patient diary	Can assist patient with recall of adherence	Poor adherence Can be manipulated

(Modified from L. Osterberg et al. 2005)[62]

be limited to solely an assessment of impairment or simply the presence of symptoms.[59] Several standardized and validated endpoints have been compiled into condition-specific instruments such as the IBS-QOL, specifically designed for a population with irritable bowel syndrome,[60] and more generic quality of life instruments such as the SF-36.[59] Instruments should be evaluated for reliability (consistent results with repeated testing), validity (ability to measure stated concepts) and ability to detect change (does score change when condition changes) before being applied to the clinical trial setting.[61]

Adherence

Subject adherence may be defined as the extent to which persons comply with or follow a prescribed regimen. Although clinical trials volunteers with underlying health conditions may be seen as motivated and compliant, data from clinical trials suggest that patient adherence rates typically range between 43 to 78%.[62]

Trials that examine the efficacy or effectiveness of probiotic supplements generally stipulate routine consumption of the probiotic supplement by the study subject. The longer or more complicated the study regimen, the lower the likelihood of subject adherence.[63] The level of subject adherence may affect outcomes, as well as how clinical trial results are interpreted. Therefore, the assessment of adherence is an important component in the conduct of probiotic clinical trials in human participants, especially if the intended use of a probiotic is on a daily basis.

Clinicians may tend to over-estimate patient adherence and are unable to consistently identify which specific patients are non-adherent.[64] To address this issue, several methods have been developed to measure adherence. Each method has advantages and disadvantages as listed in **Table 1**.

Methods to assess adherence to probiotic protocols. Directly observed therapy (DOT) involves witnessing a participant ingest a study product. Although this technique is the most accurate, its expense may make DOT impractical, especially in the outpatient setting. Detection of the probiotic supplement in the stool of clinical trial subjects has the advantage of being an objective measure of ingestion (for probiotic strains with this capability), as well as serving as a surrogate assessment for probiotic exposure of subjects who receive active product versus those who receive placebo. However, the measurement of probiotic organisms in the stool is a marker of probiotic organism survival, and does not imply activity or benefit. In addition to the cost, additional steps may need to be taken to ensure that assessment of survival in the stool does not unblind the investigators during the trial. One option might be to store stool samples and assess for probiotic organism content after unblinding has occurred to avoid this risk.

Pill counts have traditionally been used in clinical trials to measure adherence to dosing; however these assessments can be manipulated by subjects. Electronic medication monitors can offer additional data regarding how often patients open their medication container.[65] Subject questionnaires or symptom diaries can also be used to assess adherence. The most commonly used method includes patient self-report through surveys. Specific techniques to improve self-reported adherence includes 'normalizing' non-adherence, asking about adherence over a period of time, and using multiple questions to create an index (e.g., difficulty taking medication on time, frequency of missed doses and time since most recent missed dose).[65]

Safety Assessments

When safety assessments are detailed, many studies neglect to report harms in a standardized and comprehensive way.[66] The extension of the CONSORT statement for harms provides guidance for reporting adverse events for RCT evaluations.[67] There are also established systems that allow a classification and grading of harms.[68] Despite the system used, adverse events should be classified according to operational definitions specified before the initiation of the trial.[69] Commonly, these include a scale of severity and a likelihood of association with the intervention based on time correlation and other potential causes. As specified in the CONSORT guidelines for the reporting of clinical trials, authors should provide estimates of the main adverse events, reasons for

treatment discontinuation for each intervention group, and the methods and timing for assessing and recording safety information, including operational definitions.[70]

The Agency for Healthcare Research and Quality (AHRQ) in conjunction with the Office of Dietary Supplements (ODS), the National Center for Complementary and Alternative Medicine (NCCAM) and the FDA has commissioned an evidence report on the safety of probiotics summarizing the existing evidence on health for human participants.[71] Priority outcomes for these reviews are symptomatic, rather than intermediate outcomes; for example, harms that are important to decision makers and users of interventions.[72] This report will include an evaluation of the overall body of evidence available for the key research questions pertaining the safety of probiotics. The greater need for reporting of safety information may be accomplished as journal editors provide accessible online space for reporting of safety assessments that are unable to be included in the body of a manuscript. The publication of the AHRQ report will greatly enhance the evidence base for safety, in concert with effectiveness evidence, for evaluation of human probiotic applications.

Using human trials to assess the safety of probiotics. For most chemical substances, most of the burden of evaluating safety falls on tests performed on well-understood animal models. For the safety-related endpoints important in the assessment of probiotics, validated animal models do not exist and, as a result, the determination of safety rests primarily on human studies. Consequently, it would be of value to incorporate safety-related design elements and endpoints into the planning, the execution, and the reporting of human studies. Guidance is available from the guidelines developed by the Organisation for Economic Co-operation and Development (OECD) for oral toxicity studies on animals.

One key design element is the use of multiple doses—normally three levels in addition to at least one 'zero-dose' control group. There are numerous reasons why single-dose studies are rarely acceptable, but two key reasons are listed below. First, multiple doses are needed in order to elucidate the presence of dose-dependence and to characterize the dose/response relationship. Secondly, multiple doses assist in distinguishing real effects from random variability. If there is only one dose group, and the subjects in this group have a statistically significantly higher incidence of diarrhea, for example, is this an adverse effect of the probiotic or just chance? With multiple doses, it is possible to discern dose-dependence for adverse effects. If it is not present— that is, if the mid-dose group rather than the high-dose group has the highest incidence of diarrhea—then one can more confidently ascribe the effect to chance.

A second design element we can acquire from clinical studies is the need for complete characterization of the test article and the vehicle, including repeated testing throughout the study to assure that concentration, purity and potency remain "as advertised."

Safety-related endpoints should be considered during the trial design stage. These might include hematology, clinical chemistry and immune parameters as well as clinical observations and self-reporting. Clinical observations should begin prior to dosing, should be frequent and regular, systematic and not haphazard, and that they should be included in appropriate detail in published reports of the study. Many human studies of probiotics do not mention adverse effects at all or they contain a meaningless observation that "the treatment was generally well tolerated" or perhaps that "no adverse effects were noted." In order to be of any real value in evaluating safety, the efforts to observe any adverse effects must be determined during trial planning. Measures of the primary endpoint and the published report must describe these methods along with the results. In short, the safety component needs to be as thoroughly thought out, executed, analyzed and reported as the efficacy component.

Regulatory Considerations

Recently, regulatory issues are a focus for those involved in the development and marketing of probiotic products. The European Food Safety Authority (EFSA) recently determined that none of the claims for specific probiotic strains submitted to date were adequately substantiated by the scientific data that were provided as evidence of support.[73] Health Canada recently provided industry guidance on probiotic-containing foods[74] and is developing guidance for probiotic natural health products. The Indian Council of Medical Research is developing guidelines for India that would require probiotic strains to be backed by clinical trials—preferably conducted in local populations—if they are to be marketed.[75] Other activities are underway in countries across the globe to address regulatory issues pertaining to probiotics. Regulatory issues relevant to probiotics have been reviewed previously[76-78] and many of the details from these papers regarding allowable claims, target populations for claims, and safety standards are summarized in **Table 2**.

Although the focus of this review is on the design and conduct of clinical trials on probiotics, mindfulness of the regulatory framework is important at the outset and throughout development of a probiotic product—from concept to market. Regulatory issues prior to commercialization pertain to product characterization, safety and efficacy. Once commercialized, regulatory impact expands to include issues such as good manufacturing practices, product labeling, advertising and post-market surveillance. This discussion will be limited to regulatory impact on design of human studies for safety and efficacy.

Proper and thorough characterization of the probiotic to be commercialized is essential. It is worth noting that the vast majority of probiotic claims evaluated by EFSA were rejected based on improper product characterization. With the availability of inexpensive genomic sequencing services, that complete sequencing and annotation of commercial strains is a reasonable and advisable expectation prior to probiotic commercialization.

Perhaps least appreciated is that the design of human studies involving probiotics should reflect the intended path to commercialization of the product. The intended use of the probiotic (drug, food, dietary supplement) will impose different requirements for filings with the FDA, study populations and study outcomes. Determination of the regulatory category is dependent on the intent for use, according to the FDA. Furthermore, use as indicated in the stated outcomes for a study dictates the product category. A product studied for its ability to treat, prevent, mitigate or cure a disease is a drug, even if the study product is in the form of a food.

Table 2. Distinctions among products to be used as drugs, conventional foods or dietary supplements[1]

Product parameter	Drug	Conventional food	Dietary supplement
Investigational new drug application filed with FDA	Required	Not required	Not required
Study outcomes	Treatment, prevention, mitigation or cure of disease	Impact on normal structure or function of the human body; or reduction of the risk of disease[1]	Impact on normal structure or function of the human body; or reduction of the risk of disease[2]
Study populations	Healthy at-risk or patients	Healthy or healthy at-risk	Healthy or healthy at-risk
Clinical trial registration	Advisable	Advisable	Advisable
Acceptable claims for products	Drug claim (FDA approval required)	Health Claim (FDA approval required), which characterizes the relationship of the food to a disease or health-related condition; Structure Function Claim (no FDA approval required), which describes the mechanism of or the effect on the structure or function of the human body or general well-being	Health Claim (FDA approval required) Structure Function Claim (no FDA approval required)
Safety standard	Risk/benefit analysis	Reasonable certainty of no harm	Unsafe if significant or unreasonable risk of illness or injury under the conditions of the recommended use
Burden of proof of safety	Rests with manufacturer	FDA must prove product is unsafe	FDA must prove product is unsafe
Approach to safety assessment	Manufacturer must submit evidence of safety to FDA	GRAS process (with or without FDA notification) commonly used	GRAS or New Dietary Ingredient process
Route of administration	No restrictions	Orally consumed	Orally consumed

[1]Medical foods are not considered here as these products are specially formulated and intended for the specific dietary management of a disease condition for which distinctive nutritional requirements are medically established and intended for use under medical supervision. This is not the case for probiotics. Medical foods are not intended to cure, treat, mitigate or prevent the underlying disease condition. [2]FDA currently views the appropriate target to be chronic, diet-related disease, not acute disease.

One consequence of this situation is that products marketed for the treatment, prevention, mitigation or cure of a disease (drug use) may trigger a request for an Investigational New Drug (IND) application to be filed with the FDA prior to commencing with human research. In fact, although the probiotic products currently on the United States market are in the form of foods or dietary supplements, many of the studies documenting health effects have evaluated endpoints that would classify them as drugs. For example, treatment of symptoms of irritable bowel syndrome, prevention of antibiotic associated diarrhea, treatment of acute diarrhea in children and extension of remission of pouchitis would all be considered drug trials by the FDA. In addition, a researcher must consider the value of research on a drug endpoint in an effort to substantiate an efficacy claim for a food or dietary supplement. For example, research that targets a population with a medical condition will likely not be relevant to a healthy population. To complicate matters further, in the United States, if the probiotic under study has not previously been sold as a food or dietary supplement ingredient, conducting clinical trials under an IND may permanently bar the studied probiotic from ever being used in foods or as a dietary supplement. Evaluation of a probiotic under an IND would label it as a biotherapeutic, and limit other intended uses.

Additional considerations regarding regulatory impact on the design of clinical trials include the fact that regulatory authorities will evaluate the totality of evidence, published and unpublished. It is worth considering that an underpowered study that fails to show a relationship between the probiotic and a health endpoint will contribute to the conclusion that the evidence is inconsistent and therefore not convincing. It may be better not to conduct an underpowered study. Furthermore, it is preferable to match the product dose and delivery format to what is anticipated to be marketed. Proof-of-concept studies early in the efficacy assessment may use higher doses or laboratory-scale product formulations. If so, confirmatory studies should be conducted validating the efficacious dose and delivery method intending to be marketed.

In addition to realistic doses, dosing schedules and carriers, evidence to substantiate an efficacy claim must be obtained using endpoints and target populations stipulated or implied by the claim. Requirements will differ for different categories of claims. To substantiate a drug claim, evidence must causally support the value of the drug to cure, treat, prevent or mitigate disease. The target population might be currently healthy, at-risk or diseased. FDA interprets a health claim for a food or dietary supplement to be limited to reduction in the risk of incurring a diet-related, chronic disease or condition by the currently healthy population. The target population must be the currently healthy or at-risk population or a subgroup from which results can be extrapolated. The endpoint must be the target disease itself or a biological endpoint accepted as a valid marker of development or manifestation of the disease.[79] Many chronic diseases take years or decades

to develop. Further, even common chronic diseases are relatively rare events requiring extremely large study populations in order to have a high probability of accruing adequate numbers of individuals with the disease. This combination of factors may place the costs of an adequately powered trial beyond reach. Trials with a relatively quick-response, valid biomarker are more plausible. Unfortunately, there are few biomarkers universally accepted as valid indicators, and those that exist (e.g., LDL cholesterol as a marker for coronary heart disease, bone density for osteoporosis) are not likely endpoints for probiotic effects. An efficient application of probiotic research funds would be a concerted effort to validate markers of disease endpoints that would be amenable to probiotic risk reduction.

To substantiate a structure/function claim, the evidence must support the role of a nutrient or dietary ingredient to affect the structure or function of some aspect of human physiology or metabolism, characterize a documented mechanism by which a nutrient or dietary ingredient acts to maintain such structure or function or describe general well-being from consumption of a nutrient or dietary ingredient. For example, the FDA has stated that use of a probiotic to help maintain intestinal microbial balance during antibiotic therapy is intended to produce a drug effect, not to affect the structure or function. The FDA has provided guidance on types of evidence it considers appropriate for structure/function claims[80] and health claims.[6]

Conclusions

This review highlights some of the challenges associated with the design, implementation, data analysis and interpretation of clinical trials in humans involving probiotics. From these challenges, best practices to address these challenges are suggested. Consideration of outcome measurements and assessment of safety issues are paramount to conducting reproducible assessments. Ensuring identity and potency of probiotic products utilizing molecular techniques is vital to ensure that the product(s) being evaluated are the intended ones. Even if sophisticated molecular techniques are not available to researchers, detailed description(s) of the probiotic genus, species and strain designation(s) provides key information. Issues of effectiveness including adherence and generalizability should be accounted for in the trial planning stage. Finally, the concerns of regulatory agencies should be considered in the trial design and conception stage. Enhanced collaboration among clinicians, microbiologists and representatives from industry and regulatory agencies is vital to the conduct of quality clinical trials, to advance research in the field, and to improve human health. The use of international standards for probiotic trials on human health may facilitate the comparison of results from different probiotic products. In addition, common standards can allow meta-analyses and systematic reviews to augment the power of such studies. Given the costs of large, well-done clinical trials, it is important to leverage the resources used for clinical trials to ensure their widest application to current therapy. In addition, the establishment of an international human probiotic clinical trials network would facilitate such discussion as well as bring together a group of individuals who could collaborate to design and conduct quality human clinical trials with probiotics.

Acknowledgements

Travel expenses for all non-industry participants in this meeting were paid by ISAPP. No honoraria were paid for time involved in preparation for or participation in the ISAPP meeting. Collin L. Ellis received support from the UC Davis-NIH T32 Predoctoral Clinical Research Training Program Scholar Award (NIH-NCRR & NIH-Roadmap for Medical Research Grant. ULI RR024146) and the U.S. National Academy of Sciences Sackler Colloquia on Microbes & Health/International Scientific Association for Probiotics and Prebiotics-Trainee Travel Awards.

References

1. Higgins J, Green S, eds. Cochrane Handbook for Systematic Reviews of Interventions. Chicester, UK: John Wiley & Sons, Ltd 2006.
2. European Medicines Agency E. General Considerations for Clincial Trials. ICH Topic E8 1998.
3. European Medicines Agency E. Notice for Guidance on Choice of Control Group in Clinical Trials. ICH Topic E 10. London 2001.
4. Kruis W, Fric P, Pokrotnieks J, Lukás M, Fixa B, Kascák M, et al. Maintaining remission of ulcerative colitis with the probiotic *Escherichia coli* Nissle 1917 is as effective as with standard mesalazine. Gut 2004; 53:1617-23.
5. Ebell M, JS, Weiss B, Woolf S, Susman J, Ewigman B, et al. Strength of recommendation taxonomy (SORT): a patient-centered approach to grading evidence in the medical literature. Am Fam Physician 2004; 69:548-56.
6. FDA. Guidance for Industry: Evidence-Based Review System for the Scientific Evaluation of Health Claims-Final. In: Nutrition CfFSaA, ed., 2009.
7. FAO WHO. Health and Nutritional Properties of Probiotics in Food including Powder Milk with Live lactic Acid Bacteria. Cordoba, Argentina: WHO/FAO 2001.
8. Moher D, Hopewell S, Schulz K, Montori V, Gøtzsche P, Devereaux P, et al. CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials. J Clin Epidemiol 2010.
9. Schulz K, Altman D, Moher D. CONSORT 2010 Statement: Updated Guidelines for Reporting Parallel Group Randomised Trials. PLoS Med 2010; 7.
10. Haynes R. Incorporating allocation concealment and blinding in randomised controlled trials. Evid Based Med 2000; 5:38.
11. Jüni P, Altman D, Egger M. Systematic reviews in health care: Assessing the quality of controlled clinical trials. BMJ 2001; 323:42-6.
12. Canada H. Guidance Document for Preparing a Submission for Food Health Claims. In: Bureau of Nutritional Sciences Food Directorate HPaFB, ed., 2009.
13. Ebell M, Siwek J, Weiss B, Woolf S, Susman J, Ewigman B, et al. Strength of recommendation taxonomy (SORT): a patient-centered approach to grading evidence in the medical literature. J Am Board Fam Pract 2004; 17:59-67.
14. Nishitani Y, Tanoue T, Yamada K, Ishida T, Yoshida M, Azuma T, et al. *Lactococcus lactis* subsp. cremoris FC alleviates symptoms of colitis induced by dextran sulfate sodium in mice. Int Immunopharmacol 2009; 9:1444-51.
15. Prisciandaro L, Geier M, Butler R, Cummins A, Howarth G. Probiotics and their derivatives as treatments for inflammatory bowel disease. Inflamm Bowel Dis 2009; 15:1906-14.
16. Miele E, Pascarella F, Giannetti E, Quaglietta L, Baldassano R, Staiano A. Effect of a probiotic preparation (VSL#3) on induction and maintenance of remission in children with ulcerative colitis. Am J Gastroenterol 2009; 104:437-43.
17. Llopis M, Antolin M, Carol M, Borruel N, Casellas F, Martinez C, et al. *Lactobacillus casei* downregulates commensals' inflammatory signals in Crohn's disease mucosa. Inflamm Bowel Dis 2009; 15:275-83.
18. Maldonado J, Lara-Villoslada F, Sierra S, Sempere L, Gómez M, Rodriguez J, et al. Safety and tolerance of the human milk probiotic strain *Lactobacillus salivarius* CECT5713 in 6-month-old children. Nutrition 2009.
19. Senok A, Verstraelen H, Temmerman M, Botta G. Probiotics for the treatment of bacterial vaginosis. Cochrane Database Syst Rev 2009; 006289.
20. Giamarellos-Bourboulis E, Bengmark S, Kanellakopoulou K, Kotzampassi K. Pro- and synbiotics to control inflammation and infection in patients with multiple injuries. J Trauma 2009; 67:815-21.
21. Coeuret V, Gueguen M, Vernoux J. Numbers and strains of lactobacilli in some probiotic products. Int J Food Microbiol 2004; 97:147-56.
22. Huys G, Vancanneyt M, D'Haene K, Vankerckhoven V, Goossens H, Swings J. Accuracy of species identity of commercial bacterial cultures intended for probiotic or nutritional use. Res Microbiol 2006; 157:803-10.

23. Drisko J, Bischoff B, Giles C, Adelson M, Rao R, McCallum R. Evaluation of five probiotic products for label claims by DNA extraction and polymerase chain reaction analysis. Dig Dis Sci 2005; 50:1113-7.
24. Turroni F, Marchesi J, Foroni E, Gueimonde M, Shanahan F, Margolles A, et al. Microbiomic analysis of the bifidobacterial population in the human distal gut. ISME J 2009; 3:745-51.
25. Zoetendal E, Rajilic-Stojanovic M, de Vos W. High-throughput diversity and functionality analysis of the gastrointestinal tract microbiota. Gut 2008; 57:1605-15.
26. Costello E, Lauber C, Hamady M, Fierer N, Gordon J, Knight R. Bacterial community variation in human body habitats across space and time. Science 2009; 326:1694-7.
27. Turnbaugh P, Gordon J. The core gut microbiome, energy balance and obesity. J Physiol 2009; 587:4153-8.
28. Baele M, Vaneechoutte M, Verhelst R, Vancanneyt M, Devriese L, Haesebrouck F. Identification of Lactobacillus species using tDNA-PCR. J Microbiol Methods 2002; 50:263-71.
29. Desai A, Shah N, Powell I. Discrimination of dairy industry isolates of the *Lactobacillus casei* group. J Dairy Sci 2006; 89:3345-51.
30. Vandamme P, Pot B, Gillis M, de Vos P, Kersters K, Swings J. Polyphasic taxonomy, a consensus approach to bacterial systematics. Microbiol Rev 1996; 60:407-38.
31. Gancheva A, Pot B, Vanhonacker K, Hoste B, Kersters K. A polyphasic approach towards the identification of strains belonging to *Lactobacillus acidophilus* and related species. Syst Appl Microbiol 1999; 22:573-85.
32. Huang C, Lee F, Liou J. Rapid discrimination and classification of the *Lactobacillus plantarum* group based on a partial dnaK sequence and DNA fingerprinting techniques. Antonie Van Leeuwenhoek 2010; 97:289-96.
33. Stephenson D, Moore R, Allison G. Comparison and utilization of repetitive-element PCR techniques for typing Lactobacillus isolates from the chicken gastrointestinal tract. Appl Environ Microbiol 2009; 75:6764-76.
34. Masco L, Huys G, Gevers D, Verbrugghen L, Swings J. Identification of Bifidobacterium species using rep-PCR fingerprinting. Syst Appl Microbiol 2003; 26:557-63.
35. Ahlroos T, Tynkkynen S. Quantitative strain-specific detection of *Lactobacillus rhamnosus* GG in human faecal samples by real-time PCR. J Appl Microbiol 2009; 106:506-14.
36. Weckx S, Allemeersch J, Van der Meulen R, Vrancken G, Huys G, Vandamme P, et al. Development and validation of a species-independent functional gene microarray that targets lactic acid bacteria. Appl Environ Microbiol 2009; 75:6488-95.
37. Klaassens E, Boesten R, Haarman M, Knol J, Schuren F, Vaughan E, et al. Mixed-species genomic microarray analysis of fecal samples reveals differential transcriptional responses of bifidobacteria in breast- and formula-fed infants. Appl Environ Microbiol 2009; 75:2668-76.
38. González R, Klaassens E, Malinen E, de Vos W, Vaughan E. Differential transcriptional response of *Bifidobacterium longum* to human milk, formula milk and galactooligosaccharide. Appl Environ Microbiol 2008; 74:4686-94.
39. Weber D, Sahm K, Polen T, Wendisch V, Antranikian G. Oligonucleotide microarrays for the detection and identification of viable beer spoilage bacteria. J Appl Microbiol 2008; 105:951-62.
40. Sanders ME, Marco M. Food formats for effective delivery of probiotics. Ann Rev Food Sci Technol 2010; 1:65-85.
41. Saarela M, Virkajärvi I, Alakomi H, Mattila-Sandholm T, Vaari A, Suomalainen T, et al. Influence of fermentation time, cryoprotectant and neutralization of cell concentrate on freeze-drying survival, storage stability and acid and bile exposure of *Bifidobacterium animalis ssp. lactis* cells produced without milk-based ingredients. J Appl Microbiol 2005; 99:1330-9.
42. Saarela M, Alakomi H, Puhakka A, Mättö J. Effect of the fermentation pH on the storage stability of *Lactobacillus rhamnosus* preparations and suitability of in vitro analyses of cell physiological functions to predict it. J Appl Microbiol 2009; 106:1204-12.
43. Arribas B, Rodríguez-Cabezas M, Camuesco D, Comalada M, Bailón E, Utrilla P, et al. A probiotic strain of *Escherichia coli*, Nissle 1917, given orally exerts local and systemic anti-inflammatory effects in lipopolysaccharide-induced sepsis in mice. Br J Pharmacol 2009; 157:1024-33.
44. Selvam R, Maheswari P, Kavitha P, Ravichandran M, Sas B, Ramchand C. Effect of *Bacillus subtilis* PB6, a natural probiotic on colon mucosal inflammation and plasma cytokines levels in inflammatory bowel disease. Indian J Biochem Biophys 2009; 46:79-85.
45. Kokesová A, Frolová L, Kverka M, Sokol D, Rossmann P, Bártová J, et al. Oral administration of probiotic bacteria (*E. coli* Nissle *E. coli* O83, *Lactobacillus casei*) influences the severity of dextran sodium sulfate-induced colitis in BALB/c mice. Folia Microbiol (Praha) 2006; 51:478-84.
46. Sleator R, Hill C. Improving probiotic function using a patho-biotechnology approach. Gene Therapy and Molecular Biology 2007; 11:269-74.
47. Kim J, Kwon J, Ahn S, Lee S, Han Y, Choi Y, et al. Effect of probiotic mix (*Bifidobacterium bifidum*, *Bifidobacterium lactis*, *Lactobacillus acidophilus*) in the primary prevention of eczema: a double-blind, randomized, placebo-controlled trial. Pediatr Allergy Immunol 2009.
48. Czaja C, Stapleton A, Yarova-Yarovaya Y, Stamm W. Phase I trial of a Lactobacillus crispatus vaginal suppository for prevention of recurrent urinary tract infection in women. Infect Dis Obstet Gynecol 2007; 2007:35387.
49. Uehara S, Monden K, Nomoto K, Seno Y, Kariyama R, Kumon H. A pilot study evaluating the safety and effectiveness of Lactobacillus vaginal suppositories in patients with recurrent urinary tract infection. Int J Antimicrob Agents 2006; 28:30-4.
50. Kosaraju S. Colon targeted delivery systems: review of polysaccharides for encapsulation and delivery. Crit Rev Food Sci Nutr 2005; 45:251-8.
51. Barth S, Duncker S, Hempe J, Breves G, Baljer G, Bauerfeind R. *Escherichia coli* Nissle 1917 for probiotic use in piglets: evidence for intestinal colonization. J Appl Microbiol 2009; 107:1697-710.
52. Schultz M, Watzl S, Oelschlaeger T, Rath H, Göttl C, Lehn N, et al. Green fluorescent protein for detection of the probiotic microorganism *Escherichia coli* strain Nissle 1917 (EcN) in vivo. J Microbiol Methods 2005; 61:389-98.
53. Kleta S, Steinrück H, Breves G, Duncker S, Laturnus C, Wieler L, et al. Detection and distribution of probiotic *Escherichia coli* Nissle 1917 clones in swine herds in Germany. J Appl Microbiol 2006; 101:1357-66.
54. Weiss K, Siegel J. Issues in endpoint selection. In: Cato A, Suton L, Cato A, eds. Clinical Drug Trials and Tribulations. New York: Marcel Dekker Inc. 2002; 79-92.
55. Echt D, Liebson P, Mitchell L, Peters R, Obias-Manno D, Barker A, et al. Mortality and morbidity in patients receiving encainide, flecainide or placebo. The Cardiac Arrhythmia Suppression Trial. N Engl J Med 1991; 324:781-8.
56. Chin R, Lee B. Endpoints. Principals and Practice of Clinical Trial Medicine: Elsevier 1996; 121-44.
57. Grimes D, Schulz K. Surrogate end points in clinical research: hazardous to your health. Obstet Gynecol 2005; 105:1114-8.
58. Chuang-Stein C. Laboratory data in clinical trials: a a staistician's perspective. Controlled Clinical Trials 1998; 19:167-77.
59. Garratt A, Schmidt L, Mackintosh A, Fitzpatrick R. Quality of life measurement: bibliographic study of patient assessed health outcome measures. BMJ 2002; 324:1417.
60. Drossman D, Patrick D, Whitehead W, Toner B, Diamant N, Hu Y, et al. Further validation of the IBS-QOL: a disease-specific quality-of-life questionnaire. Am J Gastroenterol 2000; 95:999-1007.
61. FDA. Guidance for Industry: Patient Reported Outcome Measurements, Use in Medical Product Development to Support Labeling Claims., 2006; http://wwwfdagov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282pdf.
62. Osterberg L, Blaschke T. Adherence to Medication. N Engl J Med 2005; 353:487-97.
63. Claxton A, Cramer J, Pierce C. A systematic review of the associations between dose regimens and medication compliance. Clin Ther 2001; 23:1296-310.
64. Gilbert J, Evans C, Haynes R, Tugwell P. Predicting compliance with a regimen of digoxin therapy in family practice. Can Med Assoc J 1980; 123:119-22.
65. Miller L, Hays R. Measuring adherence to antiretroviral medications in clinical trials. HIV Clin Trials 2000; 1:36-46.
66. McIntosh H, Woolacott N, Bagnall A. Assessing harmful effects in systematic reviews. BMC Med Res Methodol 2004; 4:19.
67. Ioannidis J, Evans S, Gøtzsche P, O'Neill R, Altman D, Schulz K, et al. Better reporting of harms in randomized trials: an extension of the CONSORT statement. Ann Intern Med 2004; 141:781-8.
68. National Institutes of Health. Common Terminology Criteria for Adverse Events (CTCAE). In: Services United States Department of Health and Human Services, 2009.
69. Ioannidis J, Lau J. Completeness of safety reporting in randomized trials: an evaluation of 7 medical areas. JAMA 2001; 285:437-43.
70. Altman D, Schulz K, Moher D, Egger M, Davidoff F, Elbourne D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. Ann Intern Med 2001; 134:663-94.
71. AHRQ. Safety of Probiotics Used to Reduce Risk and Prevent or Treat Disease. Rockville: Agency for healthcare Research and Quality:http://www.ahrq.gov/clinic/tp/probiotictp.htm.
72. Chou R, Aronson N, Atkins D, Ismaila A, Santaguida P, Smith D, et al. Assessing harms when comparing medical interventions: AHRQ and the Effective Health-Care Program. J Clin Epidemiol 2008.
73. European Food Safety Authority. Briefing document for Member States and European Commission on the evaluation of Article 13.1 health claims. Parma, Italy, 11 November 2009; EFSA Journal 2009; 7:1386.
74. Health Canada. Guidance Document—The Use of Probiotic Microorganisms in Food 2009, Food Directorate, Health Products and Food Branch, Health Canada, April 2009. Accessible via http://www.hc-sc.gc.ca/fn-an/alt_formats/hpfb-dgpsa/pdf/legislation/probiotics_guidance-orientation_probiotiques-eng.pdf.
75. Gokhale G, Nadkarni A. Probiotic foods market in India soon to be regulated. India Law Journal, 2007, Accessible via http://www.indialawjournal.com/volume3/issue_1/article_by_aditi.html.
76. Sanders M, Tompkins T, Heimbach J, Kolida S. Weight of evidence needed to substantiate a health effect for probiotics and prebiotics: regulatory considerations in Canada EU and U.S. Eur J Nutr 2005; 44:303-10.
77. Sanders M, Hamilton J, Reid G, Gibson G. A nonviable preparation of *Lactobacillus acidophilus* is not a probiotic. Clin Infect Dis 2007; 44:886.
78. Pineiro M, Stanton C. Probiotic bacteria: legislative framework—requirements to evidence basis. J Nutr 2007; 137:850-3.

79. Katz R. Biomarkers and surrogate markers: an FDA perspective. NeuroRx 2004; 1:189-95.
80. FDA. Guidance for Industry: Substantiation for Dietary Supplement Claims Made Under Section 403 (r) (6) of the Federal Food, Drug and Cosmetic Act. In: Office of Nutrition L and Dietary Supplements, ed. 2008.