4 Designing Trials for Testing the Efficacy of Pre- Pro- and Synbiotics

Stephen Lewis · Charlotte Atkinson

4.1 Introduction

Providing an evidence base for the rational delivery of medicines and treatments is the cornerstone of modern health care delivery. Much of this evidence base is gained through conducting clinical trials.

Superficially, designing a clinical trial seems straightforward. However, in practice many unforeseen difficulties arise with long setting up times, poor recruitment rates and patients or interventions not behaving in the way expected. Unfortunately, clinical trials examining the efficacy of pre-, pro- and synbiotics have developed a reputation for being published in low impact journals and reaching unconvincing conclusions. As a generalization, the reason for this poor reputation is that the trials have tended to be too small and have not used meaningful clinical endpoints. The level of evidence required to alter clinical practice is expected to be high and robust. Trials of drugs such as those used to treat hypertension are often very large with hundreds (if not thousands) of patients and have hard clinical end points, such as stroke, myocardial infarction or death. Many clinical trials involving pre-, pro- or synbiotics have less than 200 patients and often use surrogate markers of health benefit as main outcome measures. This chapter sets out to give an overview of how to design and run a clinical trial highlighting examples and problems related to studies using pre-, pro- and synbiotics.

4.2 Before You Begin

A thorough literature search is required to ensure that any prospective trialist has an in-depth knowledge of the subject area to be researched. Traditionally a trialist would start by looking in electronic databases (e.g., PUBMED, COCHRANE),

then searching through the references of relevant published papers. Writing to the manufacturers of a given pre-, pro- or synbiotic can be helpful as can writing to other trialists with experience in the field. In particular, it is worth checking to see if any other researchers are studying the chosen or a similar subject such that any new study can complement, rather than duplicate, existing trials. A number of web sites provide up-to-date information on currently registered clinical trials (e.g., www.controlled-trials.com, http://clinicaltrials.gov and http://eudract.emea.europa.eu/index.html).

Much attention is now paid to health claims for functional foods, especially those containing pre or probiotics. The Process for the Assessment of Scientific Support for Claims on Foods (PASSCLAIM) project provides a scheme by which health claims for functional foods could be justified in a scientific manner. The project was initiated by the International Life Science Institute (ILSI) Europe (http://europe.ilsi.org/), which is a European Union backed multiprofessional organization, principally funded by industry. The published criteria (Aggett et al., 2005) provide an excellent source of information for researchers designing both non-clinical and clinical trials. In particular the authors promote the importance of high scientific standards and the requirement to examine physiologically relevant end-points.

Collaboration with other researchers may be desirable. Indeed grant awarding committees are often impressed by a multidisciplinary team approach, which will draw on a breadth of expertise, e.g., pharmacologist, health economist, statistician as well as an experienced trialist.

The International Conference for Harmonization of Good Clinical Practice in Research (ICH-GCP in research) has produced a combined framework for research conducted in Europe, Japan and the United States of America. Researchers are obliged to be compliant with this framework and thus should ensure they are familiar with its requirements. It is also important to ensure that the appropriate resources and motivation are available, as most studies will take a minimum of 3–4 years from conception to publication.

4.3 Hypothesis

The hypothesis is the main focus of the study. It cannot be emphasized how important it is to define clearly the question you would like to answer. The hypothesis is usually based on previous observations or assumptions, and the goal of the study is to either prove or disprove the hypothesis. Keeping the question

simple and focused will greatly increase the study's success in proving or disproving the stated hypothesis. It is often desirable also to examine any underlying mechanisms that support the hypothesis, to explain why the trial either produced the desired result or to give an insight into the reasons for any unexpected results. Generally, researchers test the validity of the "null" hypothesis. The "null" hypothesis is where the assumption is that no clinically important difference of interest in the outcome exists between groups. For example, a trialist may set out to show that an interventional agent such as a probiotic yogurt when taken by patients does not alter the risk of developing endocarditis when compared with patients not taking the probiotic yogurt. An "alternative" hypothesis is that patients taking the probiotic yogurt will have a lower or higher risk of developing endocarditis when compared with patients not taking the probiotic yogurt.

4.4 Choosing an Interventional Agent, Placebo and Packaging

When choosing a pre-, pro- or synbiotic as an interventional agent, justification is required for the choice of agent and the proposed dose to be given. A trialist may be guided by factors such as faecal recovery rates as an indicator of intestinal viability, the demonstration of, such as: immune stimulation in previous studies, or other studies suggesting a benefit in terms of health outcomes for the target population of patients.

If a trialist chooses to use live microorganisms as the interventional agent, quality control and storage are important considerations as inappropriate bacteria may be present (e.g., through contamination) or viable counts of organisms may be less than desired if incorrectly stored (Gilliland and Speck, 1977; Hamilton-Miller et al., 1996). If a clinical trial requires multiple batches of microorganisms to be prepared or a long period of storage then repeated quality control is required.

In most countries pre-, pro- and synbiotic preparations are considered as foods or food additives. It is possible that some preparations – especially, if genetically modified to have specific characteristics (e.g., produce cytokines) would be classified as drugs. In such cases further approvals would be required (in the UK by the Medicines and Healthcare products Regulatory Agency (MHRA)) and considerably more stringent monitoring throughout the trial, especially for complications.

The power of suggestion should not be underestimated. It has been clearly shown that patients do better if they feel that they are receiving an active

intervention rather than placebo. Likewise, if the treatment arm is known to the trialist then there may be a temptation for them to interpret the data according to their personal views of the expected outcome for those taking the active intervention. As such it is desirable that the patient and the treating physician are both unaware of which intervention the patient is receiving (e.g., a "double-blind" trial). This can be achieved by using a placebo. The placebo, where possible, should be indistinguishable in sight (including packaging), smell and taste from any interventional agent. Trials that do not use a placebo are usually much diminished in terms of credibility.

When choosing or developing a placebo, it is important to avoid the use of any materials that could affect the outcome measures or potentially affect recruitment rates. For example, "carrier" substances such as lactose in capsules or milk products in yogurts may cause diarrhea in some people, and therefore must be provided in similar amounts (if at all) to participants taking either the active or placebo intervention. Using products like gelatin or additives (for coloring, taste etc) may exclude some potential volunteers to a study because of vegetarianism or allergy to certain additives. Such products should be avoided where possible.

Where an intervention is compared with a standard treatment, e.g., an anti-motility agent against a probiotic to treat diarrhea, the principles of blinding still apply and both products should be indistinguishable from one another. Alternatively placebos can be given for both products. However, care should be taken when considering such a trial, as it is not always wise to assume that a standard treatment is effective particularly for disorders such as irritable bowel syndrome. Any trial that involves the use of a licensed drug will have to comply with further sets of regulations (see above). Indeed there will be a requirement for the drug, placebo and packaging to be prepared by a licensed facility, which will increase the cost of the trial and also increase the administrative burden on the trial due to the requirement for increased monitoring.

4.5 Choosing the Primary Study End-Point

Choosing the principal outcome measure of any given study, e.g., pneumonia, diarrhea etc is critical. Above all the principal end-point being studied has to be clinically relevant and represent an improvement, which if achieved is likely to alter clinical practice. Generally clinical trials are designed to examine mortality, morbidity, quality of life and economic benefits. Ideally, a clinical trial should be able to isolate the effects of a treatment on a study outcome and provide results

that are free from bias. Surrogate markers of benefit, such as inflammatory or immune function markers are good for exploring mechanisms of action and for planning large, more definitive trials. However, surrogate markers of benefit, even if improved by the intervention, may not imply clinical benefit, especially if seen in healthy people and thus unlikely to alter clinical practice. A chosen primary end-point should be practical to measure and occur frequently enough in the target population to be statistically viable. If more than one principal end-point is chosen, then this usually will require the study to be larger, and thus increases its cost and duration and may reduce its chances of success. End-points can be binary, i.e., Yes/No, or continuous, e.g., blood pressure. A chosen endpoint must be clearly defined using a recognized and relevant definition. This can be difficult; for example what constitutes a wound infection? Trials examining the benefit of postoperative enteral nutrition revealed wound infection rates between 0–33% (Lewis et al., 2001), the large variation being due primarily to the different definitions used to describe a wound infection. A clinician may feel that a relevant definition of wound infection would include the need (or not) for treatment, increased cost, or patient discomfort, but the degree of erythema may be irrelevant, whereas another clinician may place more emphasis on the degree of erythema than other factors. It is also critical that any definition of an endpoint is reproducible; this subject has its own literature, which will not be covered in detail here but should be researched. For example what is a clinically meaningful and reproducible definition of diarrhea? Definitions range from one to three loose bowel motions over 1–3 days. Even trying to define what a loose bowel motion is can be fraught with difficulty, as different observers will interpret definitions such as semi-solid, loose and watery differently. Graphical representations of stool form have been developed such as the Bristol stool form scale, but even this scale has an intra and inter-observer error (Lewis and Heaton, 1997). Creating your own definitions of end-points will be open to criticism unless they have been substantiated. Again, any end point has to be clinically meaningful. If a clinical trial was to show that an intervention led to the resolution of diarrhea say one day earlier than the control group, would this change clinical practice? Probably not, but the answer to this would depend on a number of factors including the cost and ease of taking the intervention. It can be meaningful for trialists to adapt existing definitions or scales to complement their trial. Using the example of diarrhea, as well as collecting data on stool form it would also be relevant to know what a patients degree of urgency to defecate was or whether the presence of faecal incontinence was altered by an interventional agent.

In some clinical trials, assessing the economic benefits may be relevant, though this is difficult to do well. In order to do meaningful healthcare economic analysis it is usual to involve a heath economist early on within the trial design process. The most frequently used health economic measure is quality-adjusted life-years (QALYs) gained; this measurement takes into account the patient's duration of life, and heath related quality of life. QALYs can be used to calculate a cost-effectiveness ratio (the cost of gaining an extra year of good quality life). However, collecting data on costs can be problematic as identifying every health care cost may not be technically possible in many healthcare systems.

Many disease-specific as well as generic scoring systems are available to assess improvement in quality of life, or degree of dependence on carers, severity of illness and predicting outcome from illness. These can usually be easily incorporated into most trial designs. Indeed, it is becoming fashionable to use "composite" measures as the main end-point of clinical trials. Composite end-points may include several relevant outcome measures (e.g., for a study looking at outcomes of patients after a myocardial infarction the following measurements may be relevant: quality of life, heart rate, blood pressure and exercise tolerance tests) grouped together to give an overall score.

Secondary end-points may ask other relevant questions. When considering the study size, it may be appropriate to power the study to look at secondary end points if they are of sufficient interest to justify a larger study.

If possible, data on endpoints should be recorded prospectively rather than retrospectively. If the trial involves collecting data from a patient's notes going through their notes after the end of a study is inadequate. Data may have been either poorly recorded or not recorded at all, and it may not be possible to obtain the data at a later stage. Collecting data as the trial progresses will also allow interim analysis to be conducted if appropriate (see below). Furthermore, it could enable the identification of problems in data recording (e.g., patients may not understand a particular questionnaire and thus fill it in inadequately) or obtaining test results, which can be rectified quickly and with minimal loss of useable data.

4.6 Independent Variables

Independent variables are factors that may influence the main study outcome measure. A thorough review of previous literature to identify independent variables is important, and similar to primary end-points they need to be robustly defined. With regards to studies looking at *Clostridium difficile* related diarrhea, patient age, degree of disability, immune suppression, taking of proton pump inhibitors and

recent antibiotic use need to be recorded, as they can all influence the development of the disease. It is likely that any study will have to be larger to be suitably powered to examine the influence of independent variables on the main outcome end-point, e.g., the use of immune suppressing drugs. The most commonly used approach to remove the potential effects of an independent variable on the outcome of interest is to use a randomized trial design where patients are randomly allocated to an intervention. The independent variables are recorded and if recruited numbers are large enough, then these independent variables will be equally distributed between study groups and their influence on the final outcome (e.g., *C difficile* related diarrhea) should also be equally distributed. If there is concern over the influence of a variable on the main trial outcome measure, e.g., immunosuppressive drugs on the development of diarrhea due to *C difficile*, then the patient's random allocation to different groups can be stratified by the variable. Where stratification is used, this may result in the requirement for greater numbers of patients in the trial to ensure reasonable numbers in each subgroup.

4.7 Clinical Trials

Any new medical drug has to go through various levels of clinical trials, often described as first, second, third and fourth phase trials. Pre-, pro- and synbiotics have not gone through these hurdles because in most countries they are considered to be nutritional supplements and are not subject to the same rules and regulations as drugs. Trialists are, however, interested in substantiating health claims, including benefit and lack of harm, for a given pre-, pro- or synbiotic. Whilst most pre-, pro- or synbiotics are considered safe, it must not be assumed that they are free from potential side effects. Indeed, there is a considerable literature on septicemias caused by many probiotic bacteria and yeasts. Furthermore, prebiotics may cause diarrhea. Thus, it would be wise to pay considerable attention to possible detrimental effects of an intervention when given to certain groups of patients such as those who are immunosuppressed or who are prone to diarrhea (e.g., those with ulcerative colitis).

4.7.1 Phase I Trials (Clinical Pharmacology and Toxicity, Typically 20–80 People)

These are the first experiments in humans (usually healthy volunteers or patients in whom usual treatments have failed), and are primarily concerned with drug

safety and drug pharmacokinetics. Trials often involve small numbers of people and dose escalation studies using predetermined criteria can be conducted.

4.7.2 Phase II Trials (Initial Clinical Investigation of Treatment Effect, Typically 40–100 People)

These are small-scale investigations (although usually larger than phase I trials) into the effectiveness and safety of a drug, and require close monitoring of patients. Usually, but not always, there is no randomization process.

4.7.3 Phase III Trials (Evaluation of Intervention, Typically Greater than 200 People)

After a drug has been shown to be potentially effective in a phase II trial, it is essential to compare this drug against the current standard treatment or a placebo within a real life environment in order to provide definitive data on the effectiveness of an intervention. Phase III trials may require relatively large sample sizes and lengthy follow-up of study participants.

4.7.4 Phase IV Trials (Post Marketing Surveillance)

Large, long-term follow-up studies looking at morbidity and mortality are termed phase IV trials, and may detect uncommon problems, which were not picked up in phase III trials.

4.8 Trial Design

This chapter is principally concerned with phase III trials that are designed to show clinical benefit. The simplest and most widely used design involves an intervention and placebo randomly allocated to subjects or group of subjects, these are known as randomized placebo controlled trials. Where both the trialist and trial subject are blinded to the intervention these trials are called, "double blinded, placebo controlled" (DBPC) and are considered the "gold standard" of clinical trial design. Other ways of designing a clinical trial include studying cohorts of subjects or using a crossover design so that all subjects receive both

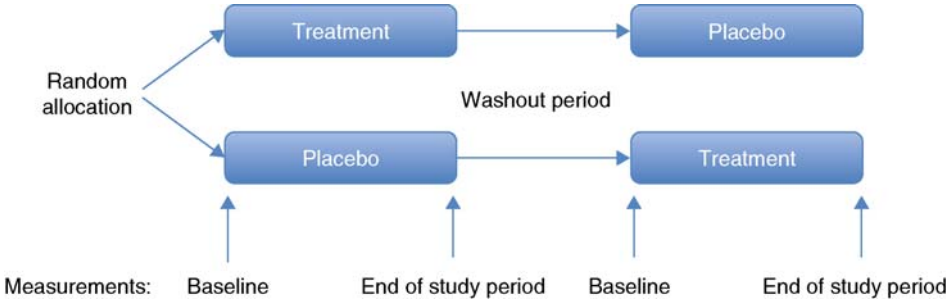

Figure 4.1
Basic cross-over design for a clinical trial Patients are randomly allocated to one of two groups (treatment and placebo groups), they are assessed at the beginning and end of a study period. After a washout period the patients undergo another study period with beginning and end assessment. If there has been no "carry over of effect" then the two baseline assessments for each individual should be similar.

the interventional agent and placebo. The best trial design will depend on the study question, the study population and the nature of the end-points being studied.

If the end-points of a trial can be measured repeatedly over time (e.g., serum cholesterol) then using more complex trial designs such as crossover trial (*Figure 4.1*) can be more economic in terms of numbers of subjects required for adequate statistical power. Any carry-over of effect between different phases of a trial can be assessed by collecting further "baseline" data after the washout period (but prior to the start of the next period) for comparison with the initial baseline data collected at the start of the trial. Using a crossover design any number of study periods can occur, though increasing the length of the time subjects are in the trial may increase the likelihood that they may withdraw from the trial due to the additional burden. If a subject withdraws from a crossover trial the loss of statistical power may be greater than if a simpler trial design had been used, because all of that subjects data from each of the study periods may then be unavailable for analysis.

There are many other design options and the exact format of any trial will depend exactly on what is being studied.

4.9 Protocol and Other Study Documents

Often, trial protocols go through many versions prior to being finalized, which enables ideas to be developed and refined (*Table 4.1*). It is important to discuss

Table 4.1

Ideal contents of a protocol

A clear statement of the clinical question being asked, then how the trial will answer the question
Scientific background, and why the trial results will be meaningful
Trial design and methodology compliant with Good Clinical Practice (GCP) guidelines in research
Study population and how they will be recruited
How the trial will be analyzed then presented

trial design with people who may have done similar trials, in order to identify potential problems - many of which may not be initially obvious.

Protocols are commonly structured along the lines of: Title, Hypothesis, Background/Rationale, Objective/Value of research, Design, Methods, Timetable, Statistics, References, Costings. Some medical journals, e.g., The Lancet require that clinical trials conform to the CONSORT Statement (www.consort-statement.org, Altman et al., 2001; Moher et al., 2001), which is a flow diagram of how a clinical trial should be structured and is worth studying (*Table 4.2*).

The protocol is used as the base document for submission of the study to any "ethics" committee for approval (see below) and, in many centres, for peer review by local research and development (R&D) committees. In most countries the conduct of clinical trials is regulated and audited. Within the European Union (EU), clinical trials are governed by the EU Clinical Trials Directive (http://www.wctn.org.uk/downloads/EU_Directive/Directive.pdf). Particular attention must be paid to data protection and reporting of adverse events. A protocol will be required to state how these requirements will be met.

The final version of the protocol needs to be well written, clear and be in compliance with ethical and regulatory requirements. Many trialists involve patients or members of the public in the design of protocols, as obtaining the view of potential participants may highlight difficulties not appreciated by the trialists; for example, procedures that cause patient discomfort (e.g., colonoscopy), may be off-putting to potential volunteers but may seem perfectly acceptable to a gastroenterologist!

Patient information leaflets and consent forms need to be prepared, which highlight any potential risks or inconveniences to the patient. The text needs to be easily understandable to a lay person. The more involved and complex the study,

Table 4.2

CONSORT Statement 2001 – Checklist Items to include when reporting a randomized trial

(*Cont'd* p. 122)

Paper section and topic	Item	Descriptor	Reported on page #
Title & Abstract	1	How participants were allocated to interventions (e.g., "random allocation," "randomised," or "randomly assigned")	
Introduction	2	Scientific background and explanation of rationale	
Background			
Methods	3	Eligibility criteria for participants and the settings and locations where the data were collected	
Participants			
Interventions	4	Precise details of the interventions intended for each group and how and when they were actually administered	
Objectives	5	Specific objectives and hypotheses	
Outcomes	6	Clearly defined primary and secondary outcome measures and, when applicable, any methods used to enhance the quality of measurements (e.g., multiple observations, training of assessors)	
Sample size	7	How sample size was determined and, when applicable, explanation of any interim analyses and stopping rules	
Randomization – sequence generation	8	Method used to generate the random allocation sequence, including details of any restrictions (e.g., blocking, stratification)	
Randomization – allocation concealment	9	Method used to implement the random allocation sequence (e.g., numbered containers or central telephone), clarifying whether the sequence was concealed until interventions were assigned	
Randomization – implementation	10	Who generated the allocation sequence, who enrolled participants, and who assigned participants to their groups	
Blinding (masking)	11	Whether or not participants, those administering the interventions, and those assessing the outcomes were blinded to group assignment. If done, how the success of blinding was evaluated	
Statistical methods	12	Statistical methods used to compare groups for primary outcome(s); Methods for additional analyses, such as subgroup analyses and adjusted analyses	

Table 4.2

Paper section and topic	Item	Descriptor	Reported on page #
Results Participant flow	13	Flow of participants through each stage (a diagram is strongly recommended). Specifically, for each group report the numbers of participants randomly assigned, receiving intended treatment, completing the study protocol, and analyzed for the primary outcome. Describe protocol deviations from study as planned, together with reasons	
Recruitment	14	Dates defining the periods of recruitment and follow-up	
Baseline data	15	Baseline demographic and clinical characteristics of each group	
Numbers analysed	16	Number of participants (denominator) in each group included in each analysis and whether the analysis was by "intention-to-treat." State the results in absolute numbers when feasible (e.g., 10/20, not 50%)	
Outcomes and estimation	17	For each primary and secondary outcome, a summary of results for each group, and the estimated effect size and its precision (e.g., 95% confidence interval)	
Ancillary analyses	18	Address multiplicity by reporting any other analyses performed, including subgroup analyses and adjusted analyses, indicating those pre-specified and those exploratory	
Adverse events	19	All important adverse events or side effects in each intervention group	
Discussion Interpretation	20	Interpretation of the results, taking into account study hypotheses, sources of potential bias or imprecision and the dangers associated with multiplicity of analyses and outcomes	
Generalisability	21	Generalisability (external validity) of the trial findings	
Overall evidence	22	General interpretation of the results in the context of current evidence	

the less likely patients will volunteer to be part of it. However, the information leaflet and consent form must provide a complete and balanced view of the study to enable the potential study participant to make an informed choice as to whether or not to participate.

4.10 Selection of Target Study Population

For clinical trials to be conducted for the prevention of a given problem, such as *C. difficile* associated diarrhea, then clearly it is clearly important to identify a target population where this is relevant. It may not be appropriate to exclude high-risk populations such as those receiving immunosuppressive drugs or suffering memory impairment, as these are the patients who suffer the majority of the morbidity and mortality associated with this disease. Excluding these patients will require the study to be much larger, be more likely to produce a non-significant result and would not be clinically relevant. Conversely, it may be harder for a clinical trial to show benefit in patients with more severe or end stage disease where there may be less ability to influence the disease process. As such, it may be more appropriate to study patients with milder disease in order to assess any benefits of an intervention. Thus, although it may be more convenient to study for example patients attending hospital clinics, for many conditions such as irritable bowel or asthma, these patients may have more severe problems than patients in the community and demonstration of benefit may be more difficult, and the results of the study may be not be generalisable. However, involving patients in the community rather than for example in a hospital setting, may bring its own set of difficulties that need to be addressed, e.g., low attendance at clinics if they live far away or have poor access to transport.

For the study to be clinically relevant, then the recruited patients should be as similar as possible to those on which the intervention will ultimately be used. Exclusion criteria should be as narrow as possible as the more people that are excluded the harder it will be to recruit and the less generally applicable the results will be. Furthermore, there may be ethical implications of excluding certain populations or groups of people. Thought should be given as to how volunteers will be identified and recruited. This is not always easy, especially if disease-specific databases are not available, and even then potential volunteers may be difficult to recruit if they are not prepared to come to a hospital or clinic, if that is necessary for recruitment purposes. The method of approach may also influence the characteristics of volunteers, e.g., advertisements in magazines or newspapers will only target the readership of those publications. For example, if trying to test whether a probiotic preparation will prevent travellers diarrhea, recruiting volunteers via high street travel agents may produce results that are not relevant to an intended "high-risk" target population (i.e., young backpackers roughing it in hostels with poor sanitation and food hygiene), as the recruited population may

consist primarily of individuals who will be staying in expensive hotels (with hopefully good sanitation and food hygiene).

Clearly considerable thought needs to be spent on identifying a relevant target population of subjects to study, who are easy to recruit and who represent a population to whom the results of the study will be clinically relevant.

4.11 Pilot Studies

Obtaining local data pertinent to the area(s) in which the trial will be run may be informative (e.g., the prevalence of a disease within the local population) as local situations may not reflect national or previously published data. Such information may be useful in terms of assessing the feasibility of running a trial in a particular area.

Where there is little previous data on which to base a trial design a pilot trial will usually be very helpful. The results can provide information that can be used in the calculation of the sample size needed for a definitive study (see below) and the experience gained in running a pilot study may help to iron out any major study design problems prior to starting the main trial. Furthermore, the pilot study may identify any unforeseen ethical problems. A pilot study can quantify likely recruitment rates and enables examination of inclusion and exclusion criteria. They can also give an indication of likely compliance with the protocol. However, whilst doing a pilot study is generally a valuable exercise, many trialists decide not to do them because of the increased effort and expense of doing two studies.

4.12 Statistical Considerations: Power and Sample Size

Obtaining statistical advice when designing a trial is almost an obligatory prerequisite and usually well worth while. Determining the sample size of a study is not an exact science, as one has to make realistic assumptions before doing calculations to decide on an appropriate sample size. It is essential that trials be designed to recruit sufficient numbers of volunteers to avoid Type 2 errors (i.e., where there is a difference between treatments, but the study has failed to detect it), but clearly avoid recruiting too many patients. Type 2 errors occur where the natural variation for the outcome being measured is wider than expected, and the sample size was insufficient to detect any difference. When doing sample size

calculations, consideration needs to be given to the likely "drop-out" rates and potential for poor compliance. Many studies using pre-, pro- or synbiotics are too small to detect benefit even if it occurs. Determining the number of patients required for recruitment will depend on the frequency of occurrence (and standard deviation) of the primary end-point in the study population, and the predicted degree of improvement likely to be seen with the interventional agent. This data may be available from previously published trials or pilot studies, but if not, educated guesses are required. Often trialists overestimate the degree of benefit likely to be seen which may result in an unrealistically small sample size estimate. The smaller the expected difference between treatment groups, the more people will be needed for the trial to be definitive. Another approach is to assume the degree of improvement of the primary outcome measure required from an intervention to be of clinical significance (likely to change clinical practice), e.g., a 20% reduction in length of hospital stay. When planning a large or expensive study, if there are no robust data on which to base sample size calculations it may be desirable to do a pilot study or even schedule an interim analysis to provide information, which will help determine the final size of the study.

Given a fixed sample size, it is nearly always true that simple one to one random allocation (i.e., one person assigned to the intervention for every one person assigned to the placebo/comparison treatment) is statistically the most efficient approach to the randomization process. Placing more participants in one group relative to the other reduces the chance of observing a difference if the sample size is fixed, although the power of the statistical test does not greatly decline unless the ratio exceeds 3–1. If the sample sizes can be increased then unequal distribution of subjects between groups may be beneficial if there are resource constraints or costs (i.e., if the intervention is very expensive or labor intensive), or if a high dropout rate is expected from an intervention because of poor tolerability.

To determine the sample size required for a clinical trial the "power" of a trial needs to be chosen. A trials "power," is the ability to show a significant difference between groups if it exists. The power of a study is calculated as 100%- β, where β is the Type 2 error (chance of arriving at a false negative result). Traditionally β values of 20–10% are used which equates to a power of 80 or 90%. The higher the study's "power" the increased chance the study has of detecting any difference, if a difference between groups exists. The increased confidence in detecting difference between groups is gained with increased study power and requires a larger sample size. Sample size calculations also require assumptions on the Type 1 error rate (detection of a false positive result); this is usually taken as a

1 in 20 chance (5%) and is denoted by α. The motivation behind most clinical trials is to show superiority for an interventional agent over a placebo or other treatment option. Occasionally, trials deliberately set out to show that two interventions are equally effective (called non-inferiority for clinical studies and equivalence for pharmacokinetic studies). In such cases a larger number of patients are usually required than for trials that are designed to show superiority of one agent over another. Calculating the sample size is best done with the aid of a statistician and suitable computer software (e.g., SPSS or STATA).

After obtaining a sample size estimate for the trial size the next step is to assess the likely recruitment rate of patients into a trial. Commonly, trialists overestimate how many patients would be eligible and willing to consent to participate. Once the likely recruitment rate has been established, an estimated time frame for the trial can be calculated. Long accrual periods are associated with failure to complete the trial, as the trialists may lose motivation. The best way to get around this is by getting more investigators and centers to participate. However, the organization of multicentre trials is considerably more complicated and expensive than a single centre trial, which perhaps explains why most studies of pre, pro and synbiotics are done in single centres.

4.13 Randomization Process and Labeling of Packaging

If the trial involves random allocation of one group of patients to a particular intervention (treatment) and another group to a placebo or current/standard treatment, computers are usually used to generate random allocations so that the investigators cannot influence who is given which treatment. Ideally, to avoid bias, the allocations need to be kept secret from the people running the trial and also from the trial subjects themselves (i.e., double-blind). The best way to achieve this is through the use of a separate randomization coordinator or service that can be contacted by telephone or email each time a new randomization code is required. The use of sealed envelopes, randomization lists, etc, are more open to abuse. A master code is often kept locally to enable decoding should there be a need due to, for example, a complication from a trial intervention.

Labeling of the packaging of trial products individually, i.e., 1,2,3,4,5. . ., rather than as "A" or "B" makes it harder for the trialist to predict the content. If all the placebos are labeled as "A" and the active interventions as "B," it is possible that the trialists could eventually work out which was which and potentially introduce bias. Ideally, a person unconnected with data collection and

analysis should supervise the administration of the trial intervention to the study participant. However, this would not completely exclude the possibility of bias if they are not blinded to which treatment the study volunteer is getting. As an estimate of the quality of blinding, when each study participant has completed the study, they and the trialists can be asked whether they thought that they had received the active or placebo treatment.

4.14 Interim Analysis

Interim analysis is generally not encouraged as it may increase the opportunity for bias within the trial, and there are several instances in which early results were published which were later not substantiated by the completion of the whole trial. However, in some cases interim analysis may be appropriate (e.g., where there was little pre-existing data on which to base power calculations prior to the start of the study or if it is important to know if the intervention is providing a beneficial effect or if it is increasing morbidity). If an interim analysis is needed, it should be conducted by people independent from the trialists and if possible, the trialists should be kept blinded from the results if it does not impact on the subsequent running of the trial. Strict criteria for any interim analysis (end-points, safety data, patient accrual rates, quality issues or complaints) should be established in the design stage. The original protocol should contain details of how the trial will be analysed and have predefined criteria for either stopping the study early (either because the effect being studied is too small and extending the study will not detect it, or the effect is so obvious that a larger study is not needed), or allowing it to continue and perhaps increasing the sample size. Interim analysis may also look at recruitment rates as well as safety and compliance issues. Often, after the interim analysis is complete protocol changes are made after discussion with the trialists, interim monitors, trial sponsor and trial financer. Any substantial changes to the protocol would require ethical committee approval.

4.15 Data Analysis

The exact statistical analysis used will depend on the trial design and type of data collected (e.g., quantitative or qualitative). Collecting accurate data and entering it into a spreadsheet in a timely fashion is essential to avoid any potential for later

confusion (e.g., from incorrectly recorded data or if original hard copies of data are lost). Analysis of data from a randomized trial is usually done on an intention to treat basis, where the analysis is based on the initial assigned treatment, not the treatment received. In real life, patients do not take all their medications for a variety of reasons (e.g., side effects), thus it is important to know how well an intervention would be expected to perform and be tolerated if used in a clinical situation. Indeed, one of the "problems" in conducting a clinical trial is that participation is voluntary; thus, the subject may be more or less likely to be compliant with an intervention, than if given to a patient as part of their standard clinical care.

Baseline data particularly on covariates can be used to demonstrate that the random allocation of groups was successful and that both treatment groups were similar at the start of the trial.

4.16 Ethical Considerations

There is a responsibility of the trialist to ensure that the trial design is of the highest quality and that the question being asked is meaningful. It is important for the researcher to be familiar with the ICH-GCP guidelines and any local guidelines that may be in place.

It is vitally important that serious consideration is given before subjecting patients to any risk or inconvenience such that the potential benefits are (1) proportionate to the value of the clinical question being asked, and (2) essential to the correct interpretation of the trial outcome. All clinical research requires approval by a research and ethics committee. These committees will demand that due attention has been paid to the ethical involvement of patients/volunteers, and especially to the involvement of vulnerable patients, such as those with learning difficulties or psychiatric problems. The research trial should be compliant with "The declaration of Helsinki" and subsequent amendments (http://www.wma.net/e/policy/b3.htm). Participants should be informed of what the trial involves and be fully aware of their rights whilst they are taking part. Patients should be given a suitable amount of time to think about their involvement and to ask questions before consenting to take part. They should be aware of the right for them to withdraw from a study at any time, without having to give a reason (and without it affecting their usual care). The involvement of minors or patients who are unable to comprehend the trial design for whatever reason is clearly even more complex. Informed consent needs to be obtained and if dealing with

a patient population who are unable to give this then an ability to obtain appropriate consent from carers should be part of the trial design.

Many trials rely on the motivation of a few individuals and the goodwill of patients. The larger and more complex the study, the more difficult it will be to manage any trial to the required standard without providing incentives for trialists and patients. Financial incentives can be beneficial with respect to improved recruitment rates and motivation of trialists, but ethically these can be problematic (Lexchin et al., 2003). If financial incentives are available to trialists, it is possible that they may also be detrimental, as a trialist may ignore recruitment criteria, or recruit patients without the appropriate motivation, for financial gain. Similarly, if financial incentives are provided to study participants, it is possible that they may be interested only in financial gain and not the collection of reliable data.

All studies should have adequate indemnity insurance. Often the trial sponsor, whether a hospital, university or commercial company will be responsible for arranging this.

Any conflicts of interest must be declared when publishing the findings of a trial and many journals will require a declaration of not only where the monies for the trial came from, but also the sources of any relevant past contributions the trialists may have received.

4.17 Misconduct

Trialists have an ethical and legal responsibility to conduct research and present the results honestly. ICH-GCP framework highlights that all clinical trials have to be conducted to a high standard and that all data are made available for audit, thus ensuring compliance with the approved protocol and honesty in data collection. It is essential to have someone knowledgeable and up to date with this framework who is responsible for collecting trial related data in the appropriate way. Commercially sponsored trials often include continuous audit to ensure completeness in data collection and to highlight any potential problems early. It is important to ensure that data is clear and not open to miss-interpretation.

Data should be kept in a secure environment and if stored on a computer, it should be at least password protected. It is a usual requirement that data is not personalised and is linked only to a named patient/volunteer via a trial specific identification number, with the master key being held elsewhere.

4.18 Recruitment, Consent and Data Collection

Trialists should aim not only gain a legal knowledge of the consenting procedure but also improve their practical skills in this important aspect of a trial. How potential volunteers are approached and encouraged to participate in a trial can have a major influence on the way they are recruited and their subsequent compliance with the trial protocol (Donovan et al., 2002). Potential volunteers should be given background information, a trial specific information sheet, and be encouraged to ask questions. As noted above, potential volunteers should be given sufficient time to make up their minds and ask friends or relatives for advice before signing a consent form and should be aware that they can withdrawal from a trial without having to give a reason. Whilst a trialist should be enthusiastic about recruiting to a trial, a degree of realism is required when enrolling volunteers who are not completely comfortable with being in the study: if a volunteer later withdrawals as a result of, for example not having had a full understanding of the requirements of the trial from the outset, then this is counter productive. Conversely, being too negative about the demands of the trial will result in poor recruitment rates.

It is usual to keep a log not only of those who agree to participate in the trial, but also of those who do not and their given reason(s) for declining participation. Potentially this information can be used to help future recruitment.

Patients in clinical trials are often managed more closely than patients not in a trial. Occasionally this could lead to an overall better outcome in these patients. Conversely the improved data collection and patient management may identify higher than predicted complication rates. In order to reduce bias, individuals blinded to which intervention the patient received should collect and interpret the data.

Compliance with a trial should be recorded if practical (e.g., by recording the amount of remaining tablets at the end of the study, concentrations of an agent in stool or urine, or by direct observation). Information on side effects should also be recorded along with comments on the likelihood that the side effect is due to the interventional agent or not.

Trial data should be collected accurately and anonymously whether on paper or computer, and should always be compliant with any data protection laws. There is a requirement for the original data generated to be stored securely and be available for audit. The exact length of time data is required to be stored will depend on the type of trial and local advice should be taken.

4.19 Monitoring Trial Progress and Protocol Deviations

There is a requirement for trials to have a monitoring committee. It is clearly important to identify and deal with any unpredicted problems that may arise during a trial. This ensures compliance with the protocol and that adverse effects are identified and data is processed appropriately. There is a requirement for adverse event reporting to be done throughout a trial. If serious adverse events occur frequently or unexpectedly they can be investigated in a timely fashion and the trial stopped if it meets the pre-defined criteria for stoppage due to such events.

Maintaining the enthusiasm of the trialists, especially over long recruitment periods can be a challenge, especially if the trial is being conducted in more than one centre. Regular contact, meetings, newsletters and feedback on recruitment rates and problems are essential.

There are numerous reasons why trials do not run according to plan. Planning how protocol deviations, non-compliance, and withdrawal of patients are handled, is ideally done in the trial planning stage. Amendments to the trial protocols are often made after the trial has begun based on unforeseen problems that may have been encountered, newly available information from other published studies, or the results of interim analyses. Clearly identifying any issues with design or recruitment early on in the trial is imperative.

4.20 Dissemination of Research Findings

Even before starting a research project the trialist should ideally have given thought as to how the results will be written up and distributed. If the intention is to publish the trial results in a peer-reviewed journal, then the design and implementation will have to conform to the journal's requirements. Thought should also be given as to how publishable the trial will be if the results were not as expected or "null." Although null results are scientifically important, it is often more difficult to publish the results of a trial if it shows no difference between treatments. Anything that can be done to make a study more attractive even if null should be considered, e.g., producing data on the natural history of a disease process, looking at the underlying mechanism of the disease process, or the mode of action of the intervention. Any trial will be considerably enhanced if it not only answers the question as to whether or not the intervention provided benefit, but also if it sheds light on the mechanism of benefit.

The order in which the names of the trialist are presented on reports or publications should reflect their degree of contribution to the study. It is well to agree this order upfront to avoid later disagreements. It is inappropriate to include the names of people who have not contributed significantly to a trial (Hwang et al., 2003; Slone, 1996). At the time of submission for publication some journals require authors to state their contributions to the trial.

Most trials are written up using a standard format of Abstract, Introduction, Methods, Results and Discussion. If you are submitting your findings to a journal for publication, it is advisable to check the "information for authors" section as individual journals may have slightly different requirements. Generally, the best approach is to keep things as clear and succinct as possible. The title should indicate what the trial is about and possibly its results. The abstract should be brief and give the reader a quick overview of what was done, the main results and their implications. Trialists need to be realistic about their findings and it is advisable to avoid extravagant claims.

Introduction: The introduction should give the reader background information on the topic area, the experimental hypothesis and the importance of the research question being asked.

Methods: This section should include what was done in the trial in enough detail so that others could replicate the study if they so wished. Results should not be presented in this section. This section can be subdivided as relevant, e.g., Design, Participants, Apparatus, Procedures, Laboratory analysis and Data analysis.

Results: Traditionally, this section begins with descriptive statistics such as the age of participants in each group and distributions of other baseline variables, such as sex. The text then goes on to list the main findings of the study in an organized manner. The results should be intelligible to most readers. If results are presented in tables and/or as graphs, they should be understandable without reference to the text. Data should be presented without speculating.

Discussion: This should start with a brief statement of the main results. It is then usual to discuss the implications of the results and how they relate to previous studies and the provisional hypothesis. Comment should be made on any problems that may have occurred during the study, the strength and weaknesses of the study, and other possible interpretations of the results. Comment can then be made on the other findings and ideas for further work.

4.21 Summary

- The results of well-conducted clinical trials provide the evidence base for present day medical practice.
- Setting up and seeing a well-designed clinical trial through to completion and publication requires a considerable amount of energy and enthusiasm from both trialists and study participants.
- Trialists should ensure that they are asking a clear and useful question and seek appropriate advice and support where needed.
- The strict requirements to comply with ICH-GCP guidelines can place a significant burden on the trialists, and may require the employment of persons to help run the trials and ensure compliance with the guidelines.
- The choice of interventional agent and outcome measures needs careful thought.
- Obtaining appropriate advice from other trialists and statisticians is invaluable.
- A multidisciplinary team approach is often required for the successful completion of a clinical trial.
- Ensure an appropriate placebo is used and both trialists and subjects are blinded as to the interventional groups.
- Seeing a clinical trial through to completion and publication is a rewarding experience for those involved and the results may directly improve patient care.
- For further reading consult Field and Hole, 2006; Institute of Clinical Research, 2008; Torgerson and Torgerson, 2008; Wang and Bakhai, 2006.

List of Abbreviations

DBPC	Double Blinded, Placebo Controlled
ICH-GCP	International Conference for Harmonization of Good Clinical Practice

References

Aggett PJ, Antoine JM, Asp NG, Bellisle F, Contor L, Cummings JH, Howlett J, Muller DJ, Persin C, Pijls LT, Rechkemmer G, Tuijtelaars S, verhagen H (2005) PASSCLAIM: consensus on criteria. Eur J Nutr 44(Suppl 1):i5–i30

Donovan J, Mills N, Smith M, Brindle L, Jacoby A, Peters T, Frankel S, Neal D,

Hamdy F (2002) Quality improvement report: improving design and conduct of randomised trials by embedding them in qualitative research: Protect (prostate testing for cancer and treatment) study. Commentary: presenting unbiased information to patients can be difficult. Br Med J 325:766–770

Field A, Hole G (2006) How to design and report experiments. Sage Publications, London.

Gilliland SE, Speck ML (1977) Enumeration and identity of lactobacilli in dietary products. J Food Prot 40:760–762

Hamilton-Miller JMT, Shah S, Smith CT (1996) "Probiotic" remedies are not what they seem. Br Med J 312:55–56

Hwang SS, Song HH, Baik JH, Jung SL, Park SH, Choi KH, Park YH (2003) Researcher contributions and fulfillment of ICMJE authorship criteria: analysis of author contribution lists in research articles with multiple authors published in radiology. International Committee of Medical Journal Editors. Radiology 226:16–23

Institute of Clinical Research (2008) ICH harmonised tripartite guidelines for good clinical practice. Institute of clinical research, Marlow

Lewis SJ, Egger M, Sylvester PA, Thomas S (2001) Early enteral feeding versus "nil by mouth" after gastrointestinal surgery: systematic review and meta-analysis of controlled trials. Br Med J 323:773–776

Lewis SJ, Heaton KW (1997) Stool form scale as a useful guide to intestinal transit-time. Scand J Gastroenterol 32:920–924

Lexchin J, Bero LA, Djulbegovic B, Clark O (2003) Pharmaceutical industry sponsorship and research outcome and quality: systematic review. BMJ 326:1167–1170

Slone RM (1996) Coauthors' contributions to major papers published in the AJR: frequency of undeserved coauthorship. AJR Am J Roentgenol 167:571–579

Torgerson DJ, Torgerson CJ (2008) Designing randomised trials in health, education and the social sciences Palgrave macmillan, Basingstoke

Wang D, Bakhai A (2006) Clinical trials: practical guide to design, analysis, and reporting Remedica, London