PARTICIPANTS (POPULATION)

From the Therapy chapter for the 3rd edition of Clinical Epidemiology, by DL Sackett
17 April 2004 (day 108)

Participants Check List[1]:

1 ❍	Draft patient eligibility criteria that match your study question.
2 ❍	Make sure they exclude patients who cannot help you answer your study question.
3 ❍	Make sure they are objective and unambiguous by achieving "almost perfect" agreement when they are applied by different clinicians to the same patients.
4 ❍	Anticipate sample size requirements and begin to consider strategies for achieving them.
5 ❍	Generate clinician eligibility criteria for potential collaborators.
6 ❍	Create your patient-entry forms.
7 ❍	Decide how to handle "eligible but not randomized" patients.
8 ❍	Decide how to handle patients whose ineligibility is not discovered until after they are randomized.
9 ❍	Decide whether and how to carry out a "dry-run" of your forms and systems.
10 ❍	See whether what you've done so far is consistent with your study question.
11 ❍	Set up your "patient-flow" diagram.

Canvassing 26 centers across Canada, we recruited stroke-neurologists who were genuinely uncertain whether aspirin or sulfinpyrazone were efficacious for transient ischemic attacks (TIAs). TIA patients referred to them were eligible for our trial if they had experienced at least one transient cerebral or retinal Ischemic attack in the 3 months before entry (during the first year of the trial only patients with multiple attacks were admitted, and the protocol was then revised to include patients with single attacks). We developed definitions for symptoms of transient ischemic attacks that were agreed upon by all participating neurologists. Certain symptoms were sufficient for entry when they constituted the only manifestations of an attack, whereas others had to occur in predefined combinations (see below). Patients with residual symptoms beyond the 24-hour limit were eligible only if the symptoms were both stable and capable of subsequent observable further deterioration. Neurologically stable patients were nonetheless excluded if they had coexisting morbid conditions that could explain their symptoms, if they were likely to die from other illnesses within 12 months, or if they were unable to take the test drugs. Participating neurologists were asked to submit information on all patients excluded from the trial.

1. **Draft patient eligibility criteria that match your study question.**

2. **Make sure they exclude patients who cannot help you answer your study question.**

3. **Make sure they are objective and unambiguous by achieving "almost perfect" agreement when they are applied by different clinicians to the same patients.**

These first three steps describe the iterative process of successive attempts to describe, with ever greater precision, the sorts of patients who will permit you to answer your study question. To ever apply the results of your trial, clinicians must know how to identify exactly which patients should be offered the better treatment. In the RRPCE study we devoted considerable time and energy to hammering out exactly what, for purposes of our trial, a transient ischemic attack was. I interviewed several stroke-neurologists and pestered them until they specified which signs or

[1] Strategies for recruiting these participants are discussed in the section on sample size on page xx.

symptoms, singly or in combination, constituted transient ischemic attacks, and from what vascular territory (carotid or vertebrobasilar) they arose. For example, weakness of an extremity was sufficient all by itself. Diplopia had to be accompanied by either dysphagia, hearing loss, mental change, or vertigo. Drowsiness, headache, tinnitus, or vertigo were acceptable components of a transient ischemic attack, but insufficient for its diagnosis by themselves. As they discussed and debated these points, our neurologist colleagues began to take over the "ownership" of the trial.

During the first year of the trial, we documented large numbers of patients who were otherwise eligible for the study but had been referred to participating neurologists following just one, initial transient ischemic attack. We decided that, despite their slightly less firm diagnoses, they would be obvious candidates for any therapy found effective in patients with multiple attacks, and altered the eligibility criteria to admit them thereafter. We carefully considered this protocol change, as we needed to be sure that answers it generated wouldn't put us off-target. You should discuss any protocol changes in your trial with your trial monitor or Trial /Data Safety Monitoring Committee. Moreover, you'll need to describe and justify any protocol changes in study publications.

Your exclusion criteria must prevent the admission of three sorts of patients who will obscure the answer to your study question. First, you need to exclude patients with "mimicking" disorders which arise from, or respond to, other causes or cures. Therefore, the RRPCE study excluded patients with severe aortic stenosis, an unrelated illness that could cause identical symptoms but called for surgical, not medical, intervention.

Second, patients who succumb to other illnesses during the study period obscure your answer in two ways. Not only won't they survive long enough to display a benefit from experimental therapy. They also will add statistical "noise" to any analysis that includes death as an outcome event. Thus, the RRPCE study excluded patients likely to die from other illnesses within 12 months.

The third sorts of patients who cannot answer your question are those in whom it is already known, at entry, that they must or must not receive one of your study treatments. An example of the former would be when an otherwise eligible patient has to take your drug (or another drug from the same class) for an extraneous comorbid condition. An example of the latter is an otherwise eligible patient who has previously suffered an adverse reaction to your drug (or one that cross-reacts with it).

You can consume lots of time and energy carrying out the first three items on this check list. Moreover, your efforts can result in lengthy, complicated patient entry forms. But lengthy entry forms discourage patient accrual, so you may want to explore a radical approach to simplifying your eligibility criteria. For example, you could replace most of your eligibility criteria with an all-inclusive "uncertainty" decision between potential study patients and their clinicians[1]. When either or both of them feel certain, for any reason, that they know which treatment is better for them, they receive that treatment and do not enter the trial. Only when both the patient and their clinician are uncertain about which treatment is preferable is the patient invited to enter the trial. I discuss this "uncertainty principle" and its use in large, simple trials on pages xx and xx.

Back on page xx, I described a pre-trial "faintness-of-heart" period in which you would ask potential study patients to comply with tasks similar to those they'd face if they entered the trial. Because stratification for compliance in the analysis is valid only if it's determined before randomization, you need to decide whether low compliance could severely hamper your ability to answer your study question. If it will, you may need some pre-randomization filter for it.

Finally, at investigators' meetings during the planning of the RRPCE trial, we circulated neurological descriptions of hypothetical patients to identify and resolve disagreements about

their eligibility. We ultimately achieved "very good" (>80% greater than chance) agreement about eligibility (as measured by kappa[2]).

4. **Anticipate sample size requirements and begin to consider strategies for achieving them.**

During the discussions that led to specifying the exact sorts of patients you do and don't want in your trial, you should have begun to get a sense of how many such patients you'll need and where you might find them. There are several rules of thumb for estimating the availability of eligible patients for RCTs, and all of them are pessimistic (e.g., "The best way to make a disease disappear is to start an RCT about it."). Although I postpone discussing sample size determinations until page xx, you should start thinking realistically right now about how many, where, and how you will find patients for your trial. For example, some trialists ask potential collaborators to start keeping logs of potentially eligible patients they encounter as soon as they consider joining the trial.

5. **Generate clinician eligibility criteria for potential collaborators.**

Your early estimates of your sample size requirements often will reveal the necessity to recruit several clinical collaborators and perhaps even several centers. You need to involve them early on if they are to adopt the trial as "theirs" and make it succeed. Accordingly, you need to develop objective, unambiguous eligibility criteria for deciding whether a potential collaborator or center is eligible for joining your RCT. For one thing, they should be genuinely uncertain about the efficacy of the experimental treatment in the target illness. A point to remember is that initially "certain" clinicians may become "uncertain" (and eligible for the trial) when they realize how many of their colleagues are uncertain. A second major criterion concerns their level of expertise in applying the study treatments.

In the RRPCE trial, most of the small family of Canadian stroke-neurologists had trained at the same institutions, were already known to be of high caliber, and were uncertain about the efficacy of aspirin and sulfinpyrazone in TIAs. Accordingly, we invited all of them to join. In other explanatory trials, however, you will want to go beyond pedigree, professed interest, reputation, pledges, and good will, and make sure they are expert clinicians. In the latter case your criteria may require potential collaborating clinicians to document not only their volume of eligible patients but also the quality of care provided to them. For example, in the North American Symptomatic Carotid Endarterectomy Trial (NASCET) we asked the "explanatory" or "efficacy" PICOT question: "Among patients with symptomatic carotid stenosis, does the performance of carotid endarterectomy by an expert neurosurgeon or vascular surgeon lower the overall risk of subsequent severe stroke or death (due both to their underlying diseases and to perioperative complications in patients randomized to surgery)?" We therefore set up a panel of surgical co-principal investigators, and they reviewed the past surgical performance (with special attention to complication rates) of potential surgical collaborators, offering participation to just the best of them.

6. **Create your patient-entry forms.**

Two issues are important here, substance and style, and each presents challenges to trialists. The substance challenge is how to ask for enough primary entry data to permit a thoughtful trial analysis, but not to ask for so much entry data of secondary importance (regardless of how "interesting" it might be) that the effort required for its collection discourages patients and clinicians from entering or continuing in the trial.

The long and short of patient-entry forms

There is no unanimity on this issue, and U.S. and U.K. trialists populate its extremes. Entry forms for the N.I.H.-sponsored North American Symptomatic Carotid Endarterectomy Trial (NASCET) of surgery vs. no surgery for symptomatic carotid stenosis were 33 pages long; for the simultaneous European Carotid Endarterectomy Trial that asked the same question and got the same answer, the entry forms were 2 pages long. The more extensive entry (and follow-up) data gathered in the NASCET trial are partially justified by their having permitted a number of ancillary studies and 44 additional publications.

The RRPCE entry forms were 17 pages long. The shortest entry form I've encountered is the ATLAS (Adjuvant Tamoxifen – Longer Against Shorter) trial that is randomizing women who appear to be disease-free following any type of curative surgery for their breast cancer, to stop or continue their adjuvant tamoxifen[3]. Its entry form is a single page and its investigators state the reason on their website: "To encourage wide participation, the **ATLAS** study involves virtually no extra work for collaborators, so that even the busiest clinicians can take part. The entry procedure is quick and easy, no examinations are required beyond those given as part of routine care and minimal, annual follow-up information is requested."

Vital and non-vital entry data

I suggest that you create two lists of entry data, one for the vital data and one for data that, although they might be interesting, are not vital, as shown in Table 5.1.

Table 5.1: Item-generation for RCT patient entry forms

Data that are <u>vital</u>: 1. To uniquely identify the patient (and how to keep track of them). 2. To confirm that the patient has the target disorder. 3. To confirm that the patient meets all the other eligibility criteria. 4. To permit stratification into the risk-response categories that will be used for stratifying patients both prior to allocation and during analysis. 5. To describe other baseline (pre-intervention) variables that will be used in the <u>primary</u> hypothesis-testing analysis (including socio-demographic information if relevant). 6. To permit combining the results with those of previous trials in systematic reviews.
Data that may be <u>interesting, but are not vital</u>: 1. To capture additional clinical, biochemical, physiological, or imaging data that might be interesting to explore in a hypothesis-forming analysis. 2. To document comorbid conditions not already known to affect responses to treatment.

The first, "vital" list begins with data required to uniquely identify the patient. In addition, if yours is a long-term trial you should consider getting contact information for someone "who will always know their whereabouts." This should be one or more of their younger relatives or close friends who do not live with them. These contacts can prove invaluable in preventing study patients from becoming lost to follow-up. The second set of vital entry data are those that confirm the patient's diagnosis, and the third, their eligibility for the trial.

The fourth set of vital entry data are derived from prior evidence (or suspicion) that subsets of eligible patients differ to important degrees in their risk of an outcome event and/or their responsiveness to experimental therapy. Suppose that, by the play of chance, patients receiving your experimental treatment were at much higher risk of an outcome event than control patients at entry to your trial. They might make even a very effective treatment appear useless, mightn't they? The same misinterpretation would occur if, by chance, patients receiving your experimental

treatment were much less responsive to your experimental treatment. The effects of these differences in risk and responsiveness are described in detail in the "Principles" discussion of "physiological statistics" at the end of this section of the chapter.

If you decide you need to allocate different "risk-response" subgroups in precisely equal proportions to experimental and control treatments (rather than leave it to simple randomization), you must identify them at entry. In addition, because some study patients may not respond at all to (or might even be harmed by) an otherwise efficacious experimental treatment, you will also want to be able to identify any special features about them in your analysis. In the RRPCE study, we suspected that a number of subgroups might differ in their risk of an outcome event and/or their responsiveness to our study drugs. Accordingly, we gathered entry data on the site of their transient ischemic attacks, how many they had suffered, whether they displayed any permanent neurological damage, their age, sex, blood pressure, cholesterol level, and cigarette use, and whether they had diabetes or a history of myocardial infarction.

Items are far too often nominated to the second, "interesting but not vital" list for no better reason (especially in North America) than "it would be nice to have them." These "interesting but not vital" data create a three-edged sword. On the one hand, their analysis can generate important, exciting hypotheses for testing in other, independent investigations or in the next logical RCT. Second, however, their sheer volume can discourage busy investigators and patients. Similarly, documenting them can add considerable expense to the trial (and to patient, their insurer, or the institution where they are enrolled).

The third sword edge produced by collecting "interesting but not vital" data is the most damaging one. The potential for damage occurs when they become fodder for "exploratory sub-group analyses" (read "data dredging exercises"), some of which must generate statistically significant results by chance alone. The real damage occurs when their conclusions are used, not for hypothesis-generation, but for clinical pronouncements about efficacy. This is particularly so when the primary analysis is indeterminate or under-powered and fails to find the hypothesized benefit of the experimental treatment. The temptation to carry out extensive subgroup analysis in the hope of identifying at least one responsive subgroup can be overwhelming. I discuss the pitfalls of "looking for the pony" on page xx. For now I simply refer you to Richard Horton's Commentary on star signs and trial guidelines[4]. In it, he describes how he negotiated with the ISIS-2 investigators (who demonstrated the benefits of aspirin and streptokinase for suspected heart attacks) to include a nonsensical subgroup analysis, the patient's astrological birth sign, in their primary paper. Thus was it "revealed" that aspirin is ineffective in Geminis and Libras with heart attacks.

The RRPCE study is another case in point. We got into trouble even with our "vital" entry data on sex and comorbidity. Our subgroup analyses on these "vital" data led us to conclude that aspirin probably didn't benefit women, diabetics, or those with prior myocardial infarctions. Our publication of these cautions caused confusion among treating physicians and it took a few years and several more RCTs to refute our false-negative conclusions about the benefit of aspirin in these subgroups

Having considered the content of the patient entry form, you should turn your attention to its format and appearance. Both elements can have major impacts on the completeness and accuracy of both vital and merely interesting data. The strategies and tactics for generating effective forms have been well-described elsewhere, and we recommend these resources to you[5, 6, 7].

7. <u>Decide how to handle "eligible-but-not-randomized" patients</u>:

As you can see from the RRPCE patient-flow diagram on page xx, 141 patients were eligible for the trial but refused randomization. Whether they received neither, one or both of the study drugs was not determined by random allocation. Rather, it resulted from an unblind, joint decision with their clinicians. We took the view that they could not contribute to a valid efficacy analysis. Therefore, we did not engage in the expensive and labor-intensive task of keeping track of them and their outcomes.

We had three reasons not to follow the "eligible-but-not-randomized" patients (besides the huge amounts of time and money required to do that). First and foremost was the proposition that following cohorts of non-randomized patients couldn't tell us whether sulfinpyrazone and aspirin were efficacious, useless, or harmful. Indeed, if non-randomized patients could have answered that question, we wouldn't have needed the RRPCE or any other RCTs! Jumping ahead 30 years, this claim was most recently validated by Regina Kunz, Gunn Vist and Andrew Oxman. They carried out a Cochrane systematic review that compared the estimates of efficacy that were found in non-randomized and randomized studies[8]. Their bottom line: "On average, non-randomised trials and randomised trials with inadequate concealment of allocation tend to result in larger estimates of effect than randomised trials with adequately concealed allocation. However, it is not generally possible to predict the magnitude, or even the direction, of possible selection biases and consequent distortions of treatment effects."

Second, we realized that the <u>proportion</u> of eligible patients who are randomized into a trial is irrelevant to decisions about efficacy, unless they have importantly different risks or responsiveness to the experimental treatment. For example, only a small minority of children with leukemia were joining RCTs back then, but the results among that minority were already shown to save the lives of the majority of leukemic children. This fact has been repeatedly demonstrated since.

The third reason is in press and will be added here when it is published.

Taking all this into consideration, you'd have to have pretty convincing evidence that the low-reliability data you could obtain from following "eligible-but-not-randomized" patients in your RCT could ever justify the expense or effort required to generate it.

8. <u>Decide how to handle patients whose ineligibility is not discovered until after they are randomized</u>:

Early in an RCT, collaborating clinicians and centers often err in applying the eligibility criteria. The result is a few ineligible patients in your trial. Accordingly, you need to decide in the trial's planning stage how to handle such patients. You'll learn in the "Principles" part of this section (page xx) that including patients with the wrong diagnosis, other terminal illnesses, and a prior allergy to a study drug will add noise to your analysis and decrease its power[10,11]. So, you'd better establish a bias-free means for removing them right now. I recommend reviewing the eligibility of <u>all</u> randomized patients, <u>blind</u> to their allocated treatment. For example, early In the RRPCE study the routine, blind review of all randomized patients discovered that 64 of them were actually ineligible for the trial. They didn't have transient ischemic attacks, they had brain tumors, aortic stenosis, migraine, and the like. Or, they had a second disease likely to kill them within a few months, or already were known to be allergic to aspirin. We didn't know their treatment group when we decided their eligibility, so it was unbiased. Note, however, that this strategy is valid only when there is equally intensive application and follow-up of the eligibility criteria in all treatment groups.

A large number of such post-randomization exclusions can detract from the credibility, if not the validity, of your trial. The best way to deal with them is to reject them prior to randomization. You can accomplish this by training your research staff (for example, with test cases) and by giving

rapid feedback to centers when they enter a patient who proves to be ineligible. You can use one of the automated data systems that "reads" and rejects entry forms of ineligible patients. Finally, you can send bulletins to all study investigators, outlining common mistakes. This sort of "quality assurance" program is exceptionally important at the beginning of your trial, so that everyone learns quickly from their mistakes and those of others.

The foregoing strategies apply when patient eligibility can be unambiguously determined from data obtained before they are invited to join the trial. In other trials, it may take weeks to process the eligibility evidence (e.g., when you need to create and review special pathological preparations). In this latter situation, your proper course of action is, as usual, determined by the question you posed. Suppose, for example, that you are conducting a management trial to test a treatment policy among patients whose eligibility cannot be ascertained prior to their treatment, both inside the trial and in routine practice. In that case, it makes sense to include the "late" ineligible patients in your intention-to-treat analysis. This policy is especially sensible when your treatment is "permanent", such as an operation, rather than a medication with negligible adverse effects. Either way, when you report your results you'll need to include the numbers, types, and justifications for excluding every patient you remove from your trial. Finally, your RCT's results will be the most credible when the reintroduction of all ineligible patients makes no difference to its conclusion.

9. **Decide whether and how to carry out a "dry-run" of your forms and systems.**

If this is your first trial, or it is the initial trial for a particular clinical condition, most or all of your forms and systems for recruiting, investigating, and determining the eligibility of your study patients will be new. It would be a shame (in terms of both validity and credibility) to discover important flaws in them only after you've started the trial. Accordingly, you would be wise to perform a "dry-run[2]" of your forms and systems before you start the formal trial. You can use the dry-run to correct errors of omission, commission, and ambiguity in your data forms. It will also let you identify and solve problems in the flow of study patients and study data. In our trials, we test, revise and retest draft forms until our study patients, interviewers, clinicians, and data managers are satisfied that the data coming from the field are accurate.

What sorts of study patients should you use in "dry-runs?" It would be vital not to "use up" eligible patients for this. We usually put members of our study staff through first, since they are already tuned into potential problems and aren't shy about pointing them out. After fixing those problems, we often perform a second "dry-run" with consenting patients who, although they have the target condition, are already known to be ineligible for our trial.

10. **See whether what you've done so far is consistent with your study question.**

In executing items 1-9 on the checklist, you may have made decisions that unintentionally damage your ability to answer your PICOT question. For example, if you're asking a pragmatic question (can the treatment work among typical patients?), did you wind up with eligibility criteria that excluded otherwise typical patients because they have comorbid conditions, have a track record of low compliance, or are older than your arbitrary age cut-off? If so, you're going to have to change either your question or your eligibility criteria. Reiterate, reiterate, reiterate.

The sorts of patients you do and don't admit to your trial, the extent to which they comply with your study treatment, how well you keep track of them, and how accurately you ascertain their

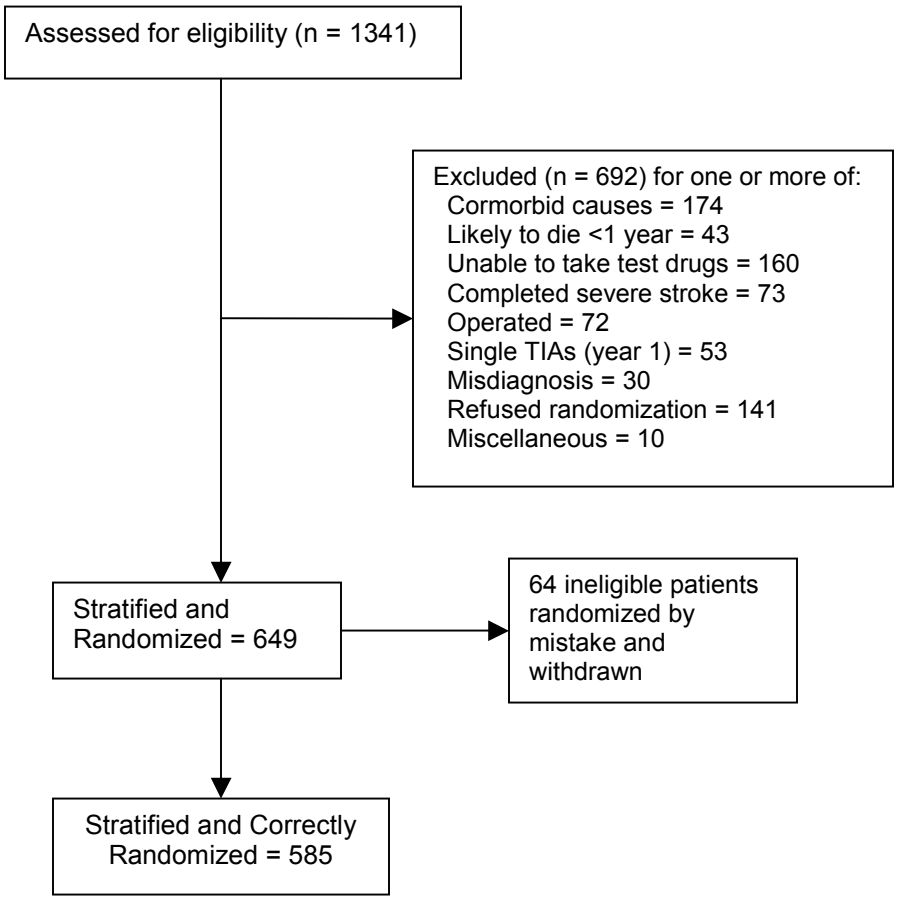
[2] I prefer "dry-run" to the more ambiguous "pilot" or "feasibility" terms that address other issues such as the availability of study patients or the skills of a trial's clinicians. I discuss these other issues in the section on sample size on page xx.

events all can have huge effects on how confident you (and your readers) will be about the answer you get. The relationships between these factors are complex, and are the bread and butter of sophisticated biostatisticians. Many clinician trialists are bewildered and intimidated by the formulas and statistics used to determine and control for these factors. However, if clinicians contemplate these relationships in physiological, rather than mathematical, terms, they can not only understand them, but also manipulate them. Accordingly, a discussion of "physiological statistics" follows shortly. But before you delve into that, you need to start your patient-flow diagram.

11. <u>Set up your patient-flow diagram</u>:

We agree with the revised CONSORT statement[12] that the flow of participants through each stage of an RCT ought to be described, preferably in the form of a diagram. Accordingly, we'll provide an updated patient-flow diagram for the RRPCE study at the end of each section of this chapter. Figure 02-03-05-1 shows what the RRPCE Study looked like at this stage.

Figure 02-03-05-1: A flow diagram of the recruitment phase of the RRPCE trial



REFERENCES

[1] Baigent C. The need for large-scale randomized evidence. Br J Clin Pharmacol. 1997;43:349-53.

[2] Altman DG. Practical Statistics for Medical Research. London: Chapman & Hall, 1991. Pp 404-9.

[3] http://www.ctsu.ox.ac.uk/~atlas/ (from the Clinical Trial Service Unit of Oxford University).

[4] Horton R. From star signs to trial guidelines. Lancet 2000;355:1033-4.

[5] Meinert CL. Clinical Trials; Design, Conduct and Analysis. Oxford: Oxford University Press, 1986. Chapter 12: Data collection considerations, Pp 119-37; and Appendix F: Data items and forms illustrations, Pp 379-416.

[6] Pocock SJ. Clinical Trials; A Practical Approach. Chichester: John Wiley & Sons, 1983. Chapter 11: Forms and Data Management. Pp160-6.

[7] Spilker B. Guide to Clinical Trials. Philadelphia: Lippincott Williams & Wilkins, 1991. Chapter 36: Preparing Data Collection Forms. Pp 262-71.

[8] Kunz R, Vist G, Oxman AD. Randomisation to protect against selection bias in healthcare trials (Cochrane Methodology Review). In: *The Cochrane Library,* Issue 3, 2002. Oxford: Update Software.

[9] Vist GE, Hagen KB, Devereaux PJ, Dianne Jackowski, Oxman AD. Outcomes of patients who participate in randomised controlled trials versus those of similar patients who do not participate (Protocol for a Cochrane Methodology Review). In: *The Cochrane Library,* Issue 3, 2002. Oxford: Update Software.

[10] Sackett DL, Gent M: Controversy in counting and attributing events in clinical trials. N Engl J Med. 1979;301:1410-2.

[11] Ferguson D, Aaron SD, Guyatt G, Hebert P. Post-randomisation exclusions: the intention to treat principle and excluding patients from analysis. BMJ 2002;325:652-4.

[12] Moher D, Schulz KF, Altman D for the CONSORT Group. The CONSORT statement: Revised recommendations for improving the quality of reports of parallel-group randomized trials. JAMA 2001;285:1987-91.