

Further guidance on the Statistical Information required by the Research Ethics Committee

Introduction

These guidelines are intended to help you to fill out the ethics committee submission form. We recommend researchers to seek statistical advice from a medical statistician at an early stage to ensure that the sample size and analysis plan are appropriate to the project. This is not, of course, compulsory; it may be that either your project is very straightforward or that you have the requisite skills yourself. In most cases, however, we have found that the input of a medical statistician speeds the process of gaining ethical approval for a study. These guidelines highlight the most common statistical reasons for submissions not being approved.

If you do not understand all the terms used in these guidelines then you should seek advice from a medical statistician.

Main outcome measures

We need to know what your main outcome measures are: the precise way you are measuring the outcome of your study. For example if you were looking at the death rates of patients using two different cancer treatments then *mortality* would not be a sufficient answer to this question, whereas *5-year mortality rate* might be.

Essentially, what we require here is the quantity (or quantities) that will be subjected to statistical analysis rather than a description of the variable. So, for example, in studying the effect of two toothpastes on dental health. In this case *dental health* would not be the main outcome measure but *dental health measured by the total number of decayed, missing and filled teeth after three months of treatment* could be.

The main outcome measure should normally feature in the sample size calculation - i.e. the thing you consider most important is normally the thing you base your sample size on. If there are several equally important outcome measures then the sample size should, ideally, be large enough to detect clinically important differences in *all* of them.

Sample size

- ❖ How many subjects will you be recruiting?

Make sure that you tell us how many people you intend to recruit! We prefer an exact number. *At least ... subjects* is not usually acceptable but *no more than ... subjects* may be reasonable.

- ❖ How did you decide on this number?

Have you done a formal sample size (power) calculation to determine your sample size? If you haven't done one you *must* explain why it wasn't necessary.

If you have done a sample size calculation you should tell us:

- ❖ the power of your study
- ❖ the statistical significance you will be using
- ❖ the difference you want to be able to detect (the 'minimum clinically significant difference')
- ❖ the expected variability of your measurements (e.g. the standard deviation)

Dropouts etc.

Have you made allowance, if necessary, for any potential reduction in your sample size due to, for example, patients withdrawing from the study whilst it is progressing or for non-responders to questionnaires?

Confounding

Have you made allowances for other problems such as confounding variables or multiple comparisons? Will this be dealt with by increasing the sample size or by using more sophisticated modelling techniques?

The statistical plan

What statistical tests or modelling methods are you going to use on which variables? For example:

- ❖ The difference between the matched subjects' systolic blood pressures will be analysed using a paired t test.
- ❖ The odds ratio for male vs. female patients being frequent attenders to GP surgeries will be modelled using multiple logistic regression, controlling for the confounding effects of age, social class and ethnicity.
- ❖ The effects of the two drugs on the length of survival of our subjects will be analysed using Cox regression, the variables age, cirrhosis status and serum albumin levels will be included in the model.

The analysis plan should be outlined in enough detail for it to be clear to a statistically informed reader that it is appropriate. If your data requires only the simplest of analysis plans (such as the first example above) then a single sentence may suffice. If the plan is complex or novel, more justification may be required and it is probably a good idea to speak to a statistician about this. (If the analysis plan is complex it may be explained in detail in the protocol rather than on the submission form.)

Non-analysed data

You may be collecting data that is not being analysed. You should, in general, justify this as ethical approval to collect data is only given if there is a good reason to collect it. It may be that some data has to be collected for safety purposes. Or you may feel that whilst collecting data for the main aim of your study you have an opportunity to collect data to generate hypotheses for future studies. Whatever the reasons, we would like a brief justification for the collection of this data. It would probably be better to explain this in the protocol rather than on the submission form.

We would not consider it sufficient justification to collect data on the grounds it might be analysed. It should be explicit from the outset of a project, which data are going to be subjected to statistical analysis according to the submitted analysis plan.

Who will carry out the statistical analysis?

This may be yourself or a member of your team, if the statistics are simple (t tests, one-way ANOVA) and you have the appropriate statistical skills. If the analysis is likely to be complex or time-consuming then we would like some indication of who will carry out this analysis and how it will be funded. (Again, it may be better to include this information in your protocol, rather than on the form.)

If you have discussed and agreed your analysis plan and power calculation with a statistician please let us know who it is. We may want to contact them to clarify any questions we have; this will help speed up the processing of your submission