Guidance on legislation

Clinical investigations of medical devices – statistical considerations

November 2013

Contents

This document replaces Guidance Note 17 'Guidance notes for manufacturers on statistical considerations for clinical investigations of medical devices'

Revision history

This version	Date published	Changes
V1.0	November 2013	n/a

Introduction

The collection and evaluation of sound clinical data are of importance in the affixing of a CE marking for many medical devices. The purpose of this document is to help manufacturers and others, by setting out the statistical elements required in the design, conduct and analysis of a clinical investigation of a medical device under the provisions of the Medical Devices Regulations 2002 (SI 2002 No 618) [1]. In these circumstances, the aim of a good clinical investigation is to provide objective evaluation of the safety and performance of the device in question, based on its intended purpose.

The principles and important aspects of carrying out clinical investigations of medical devices can be found in the European harmonised standard BS EN ISO 14155:2011: Clinical investigation of medical devices for human subjects. Good clinical practice [2].

More detailed information on the statistical principles for clinical trials can be obtained by reference to the International Conference on Harmonisation: Guidelines for Good Clinical Practice [3].

1   Clinical investigation design

Many of the principles of clinical investigation design are aimed at minimising known or suspected sources of bias which may compromise the ability to draw valid conclusions from clinical studies, whilst at the same time maximising precision.

1.1 Investigation objective

An effective and efficient design of a clinical investigation cannot be accomplished without a clear and concise objective. This must be formulated with great care and specificity. It is not adequate simply to state as an objective '… the safety and performance of a device …' The aims and objectives must be set out so as to evaluate accurately the particular use of the device in the target condition and with the appropriate population and must be properly established prior to any development of the protocol. Such aims and objectives will provide the essential focus for the investigation and should also provide the basis for labelling indications once the device is placed on the market.

In formulating the investigation objective(s), care needs to be taken in determining appropriate endpoints. These should be directly observable, objectively determined measures, subject to minimal bias and error and should be directly related to biological effects of the clinical conditions.

1.2 Sample size

The number of patients in a clinical investigation should balance the need for a reliable answer to the questions addressed whilst minimising the exposure of subjects to risk. The number is usually determined by the primary objective of the investigation. In certain circumstances, the objectives of the investigation may involve a comparative group which will influence not only the design but the appropriate sample size and subsequent investigation analysis. The usual method for determining the appropriate sample size for an investigation requires:
- specification of a primary variable
- the null hypothesis
- the alternative hypothesis
- the probability of erroneously rejecting the null hypothesis (the type 1 error, conventionally 5% or less); and
- the probability of erroneously accepting the null hypothesis (the type 2 error, conventionally 20% or less), as well as the approach to dealing with drop outs and other protocol deviations.

The method by which the sample size is calculated should be given in the protocol, together with the estimates of any quantities used in the calculation and the source of such estimates.

Sample size calculation should make appropriate allowance for drop-outs and any other anticipated protocol deviations. Not only should the sample size be increased to offset the loss of patients from the study or from their intended treatments but it should also allow for the fact that the size of effect observed may be less than expected.

1.3 Study population

The study population should be a relevant subset of the population targeted for the application of the medical device. The study population must be defined before the investigation by the development of strict, unambiguous inclusion and exclusion criteria. These criteria will characterise the study population and in this way help to define the intended use of the device. These criteria should also include an assessment of prognostic factors for the outcome variable(s) since one or more of these variables may influence the performance of the device, e.g. age, sex, stage of disease.

In the earlier stages of device development, the choice of patients for a clinical investigation may be influenced by the wish to maximise the chances of observing specific clinical effects of interest and hence they may come from a very narrow sub group of the total patient population for which the device may eventually be indicated. However, by the time the **pivotal** trials are undertaken, the patient populations should more closely mirror the intended treatment population.

1.4 Use and selection of controls

The objectives of a clinical investigation, as set out in the Medical Devices Regulations (SI 2002 No 618) are:
- to verify that, under normal conditions of use, the performance characteristics of the device are those intended by the manufacturer; and

- to determine any undesirable side effects and assess whether these constitute risks when weighed against the intended performance of the device.

Such clinical investigations should therefore be designed to provide firm evidence in support of these aims. Where the endpoints can be measured objectively, e.g. from radiological examination, the majority of clinical investigations of medical devices will not require a comparative group and a single arm study design will be sufficient to demonstrate the required objectives. However, the justification for the chosen design, particularly where there is omission of a control group, should be made explicitly in the documentation on the proposed investigation. In circumstances where the endpoints are subjective, e.g. improvement in pain, a control group will nearly always be necessary in order to validate the claims being made for the device in question.

Similarly, if a clinical investigation is intended to evaluate intervention with a device compared with an alternative/no intervention, then the design of the trial will need to include a control group. The safety and performance of the device is then evaluated through the comparison of differences in the diagnosis or outcome between the treated patients and the control group. A scientifically valid control population must be comparable to the study population in all important patient characteristic and prognostic factors.

There are four main types of control groups, namely:

- **Concurrent:** where the control group is under the direct care of the same clinical study investigator for the same condition.

- **Passive concurrent:** where the control group receives an alternative intervention, including no intervention, but is not under the direct care of the same clinical study investigator.

- **Self controls or cross over controls:** where following one intervention and after a prescribed period of time, the control group receives the alternative intervention. Here the alternative intervention usually follows a 'wash-out' period to allow time to elapse between the end of one experimental condition and the beginning of the next condition. There are instances where a patient may serve as his own control in circumstances where it can be clearly demonstrated that current clinical consensus has determined that there are no residual effects of the original device beyond the immediate treatment of the patient.

- **Historical controls:** where a non-concurrent group of patients with the same condition makes up the control group but is separated in time from the population under study.

Concurrent controls and, where applicable, self controls allow the largest degree of opportunity for comparability. The use of historical controls is the most difficult in assuring comparability with the study population since the practice of medicine in terms particularly of methods of diagnosis and criteria for treatment changes over time. There are often therefore differences in patient selections that may not be easily or adequately documented, and which lead to differences in outcome that are mistakenly attributed to the use of the new device.

Although in general terms the analysis of a comparative investigation and the interpretation of results are straightforward, biostatistical advice should be sought, since there may be additional features of the design which it is important or essential to take into account.

1.5 Techniques to avoid bias in controlled clinical investigations

- **Blinding:** there are a number of serious biases that may occur in a clinical investigation including investigator bias (which may arise due to knowledge by the investigator of treatment allocated to a particular patient), evaluator bias and placebo effects. To protect an investigation against these potential biases, blinding should be used if this is practical. Blinding is accomplished by coding the interventions and having an individual who is not in the patient care team to break the code. In practice, it has been demonstrated that blinding is often difficult or impossible in a clinical investigation of a medical device. Under the circumstances therefore, care must be exercised by the study staff to assure that these biases are minimised by assuring that the evaluator is blinded to the assignment of patients to a particular intervention or control group.

- **Randomisation:** this introduces a deliberate element of chance into the assignment of devices or no treatment to patients in order to produce device groups in which the distribution of prognostic factors, both known and unknown, are similar. Randomisation helps to avoid possible bias in the selection and allocation of patients arising from the predictability of device assignments.

The randomisation schedule of a clinical investigation documents the random allocation of devices to patients. In its simplest form it is a sequential list of devices, or corresponding codes, by patient number. Different study designs will require different procedures for generating randomisation schedules. The procedure should be capable of being reproduced (if the need arises) through the use of the same random number table, or the same computer routine. There are generally some advantages to be gained by randomising patients in blocks. This helps to increase the comparability of the device groups throughout the period of allocation. It also provides a better guarantee that the device groups will be of nearly equal size.

In multi-centre studies the randomisation procedure should always be organised centrally. It is often advisable to have a separate, random scheme for each centre. Details of the randomisation which facilitate predictability should not be contained in the main study protocol but should be set out in an

annex, which can be withheld from the study site. The randomisation schedule itself should be filed securely by the applicant in a manner which ensures that blindness is properly maintained throughout the trial. Access to the randomisation schedule during the investigation must take into account the possibility that, in an emergency, the blind may have to be broken for any patient, either partially or completely.

1.6 Outcome and prognostic variables

The observations in a clinical study involve two types of variables, namely outcome variables or endpoints and influencing variables.

- **Outcome variables** define an answer to the questions posed by the aims and objectives of the clinical investigation and should have a direct impact on the claims made for the device. These variables should be directly observable, as objectively determined as possible, subject to minimal bias and error and directly related to the biological effect of the clinical condition of the patient receiving the device intervention.

In any clinical investigation of a device, there will be primary and sometimes secondary outcome variables. The primary variables should be the variables capable of providing the most relevant and convincing evidence directly related to the primary objective of the investigation. There should generally be only one primary outcome variable which should be specified in the protocol. This is the most appropriate one for estimating the same size. It may, however, sometimes be desirable to use more than one primary variable to cover the range of effects of the device. Secondary outcome variables either support measurements related to the primary objective or are measurements of effect related to the secondary objective (if any). Their pre-definition in the protocol is also important.

- **Influencing variables (confounding factors) or prognostic factors** are any aspect of the study that can affect the end point or the relationship between treatment and outcome. In a controlled study, imbalance in prognostic factors between comparative groups, e.g. patient age, can lead to false conclusions by improperly attributing an effect observed in the outcome variable to an intervention when it was mainly due to imbalance. Therefore, in the development of a clinical investigation of a device, care must be taken to identify influencing variables that are likely to affect outcome.

1.7 Study sites and investigators

In the majority of device clinical investigations, pooling of data across study sites and investigators will be necessary in order to attain the required sample size. The selection of such sites and investigators is critical in planning a clinical investigation.

The sites selected must have sufficient numbers of eligible patients who are representative of the target population for the device. Each centre must have facilities that are capable of processing patients in the manner prescribed by the protocol and must have staff who are qualified to conduct the trial. The principal investigator at each site must be able to recruit eligible patients to the trial and must be willing to abide by the procedures established by the protocol.

2   Pre-specified data analysis

When designing a clinical investigation it is essential to consider the eventual analysis of the data and document the analysis plan, since such factors may well affect the choice of what variables to collect and possibly other aspects of the study design.

2.1 Study population

The analysis plan should first determine the population of patients whose data are to be included in the main analysis. As a minimum, documentation is required for all patients for whom study procedures were initiated and who have given their informed consent. The content of this patient documentation depends on detailed features of the particular investigation, but at least demographic and baseline data on disease status should be collected whenever possible.

Single arm studies: it is particularly important that for a single arm study the outcome is described for all patients who are identified as recipients of the device under investigation. Thus, even if for some reason the device is not used (say for administrative reasons totally unrelated to the device rather than for medical reasons) full details are required.

Intention to treat: the principle of 'intention to treat' implies that all randomised patients should be included in the analysis. In most clinical investigations it provides a conservative approach and also gives estimates of treatment effects which are more likely to mirror those observed when the device in question is put into service. Wherever possible, the protocol should also define prospectively how any foreseeable problems will be addressed. Any objective entry criteria, measured before randomisation, which will be used to exclude patients from analysis, should be pre-specified and justified. In the case of a randomised trial where patients are withdrawn prior to randomisation, such withdrawals should be reported to allow assessment of the degree to which the patients who are included in the trial are a select subgroup of those who might have been included.

Per protocol: the per protocol population is characterised by the following criteria:

- the completion by the patient of a certain pre-specified minimal exposure to the device in question

- the availability of measurements of the primary outcome variable(s) at relevant and pre-specified time(s)

- the absence of any major protocol violation including the violation of entry criteria.

This population generally maximises the opportunity for a new device to fully meet the objectives of the investigation.

2.2 Missing values and outliers

Missing values represent a potential source of bias in a clinical trial. Hence, every effort should be undertaken to fulfil all the requirements of the protocol concerning the collection of data.

Care should be adopted when exploring the influence of outliers or influential observation. Clear identification of a particular value as an outlier is most convincing when justified medically as well as statistically.

If no procedures for dealing with outliers was included in the investigation protocol, one analysis with the actual values and at least one other analysis eliminating or reducing the outlier effect should be performed and differences between their results discussed.

2.3 Estimation, confidence intervals and hypothesis testing

The statistical analysis plan should specify the hypotheses which are to be tested and/or the device performance characteristics which are to be estimated in order to satisfy the objectives of the clinical investigation. The statistical methods to be used to accomplish these tasks should be described for the primary (and preferably the secondary) variables. Estimates of device characteristics should be accompanied by confidence intervals, wherever possible, and the way in which these will be calculated should be identified.

3   The conduct and monitoring phase

3.1 Changes in inclusion and exclusion criteria

Inclusion and exclusion criteria should, if possible, remain constant throughout the period of patient recruitment. However, this may not always prove possible particularly in longer term studies. Changes may also result from the discovery by monitoring staff that regular violations of the entry criteria are occurring, or that recruitment rates are seriously low due to over-restrictive criteria. The protocol amendment implementing such changes should cover any statistical consequences, such as sample size adjustments, arising from different event rates, or modifications to the analysis plan.

The MHRA must be informed of any change in these criteria. They must not be implemented until agreement has been obtained in writing (see the MHRA's document 'Clinical investigations of medical devices – guidance for manufacturers' [4]).

3.2 Checking the design assumption

In larger trials there will usually be an opportunity to check the assumptions which underlie the original design and sample size calculation. This may be particularly important if the trial specifications have been made on preliminary or uncertain information. An interim check conducted on the blinded data may reveal that overall response, event rates or survival experience are not as was anticipated. A revised sample size may then be calculated using suitably modified assumptions and should be justified and documented in a protocol amendment and in the final report.

If such analysis leads to the need for a revised sample size, the MHRA must be informed prior to any changes being made to the numbers of patients to be included within the clinical investigation and justification provided for the proposed increase.

4   Follow-up and reporting

All patients and devices that enter the clinical investigation should be accounted for in the final report. All reasons for exclusion from analysis must be carefully documented. Similarly, for all patients and devices included in an analysis population, the measurements of all important variables must be accounted for at all relevant time points. Additional information that is available on patients screened for entry but not randomised should also be summarised.

While it is sometimes impossible to locate all patients included within the clinical investigation, the sponsor must demonstrate that everything possible was done in an attempt to find patients lost to follow-up.

The effect of all losses of patients or of data, withdrawals from treatment and major protocol violation on the main analysis of the primary variables should be considered carefully. Patients lost to follow-up or withdrawn from device use should be identified and a descriptive analysis of them provided, including the reasons for their loss and its relationship to treatment and outcome.

Suitable tables and/or graphical representations should illustrate clearly the important features of the primary and secondary variables and of key prognostic and demographic variables. The results of the main analysis relating to the objectives of the investigation should be the subject of particularly careful descriptive presentation. Although the primary goal of the analysis of a clinical investigation should be to answer the questions posed by its main objectives, new questions based on the observed data may well emerge during the analysis. This additional work should be strictly distinguished in the report from the work which was planned in the protocol. In general, sparing use should be made of unplanned subsidiary analysis.

Statistical judgement should always be brought to bear on the analysis, interpretation and presentation of results of the clinical investigation. To this end, the investigation statistician should be a member of the team responsible for the study report and should be a signatory to it.

5   Safety issues

Clinical investigations under the provision of the Medical Devices Directives are rarely sufficiently large to detect infrequent adverse effects of devices. Nevertheless, it is important to monitor clinical investigations for those adverse effects which do become apparent. Safety variables should subsequently be collected as comprehensively as possible from these patients.

All safety variables will require attention during analysis and the broad approach should be indicated in the protocol. All adverse events should be carefully recorded and notified to the MHRA (see Regulation 16(10) of the Medical Devices Regulations 2002 (SI 2002 No 618) [1] and Annex X of the Medical Devices Directive 93/42 EEC [5]). Such notification should include incidents indirectly related to the device, in addition to those directly attributable to the device or use of the device in question. In particular adverse incidents should be notified if they have led to hospitalisation, prolongation of hospitalisation, additional surgical or medical intervention or death.

6   Further information

Enquiries regarding this document or the clinical investigation procedure should be addressed to: Mrs Daniella Smolenska, European and Regulatory Affairs, Medicines & Healthcare products Regulatory Agency Tel: 020 3080 7363
Email: daniella.smolenska@mhra.gsi.gov.uk or mb-mda-era@mhra.gsi.gov.uk

7   References and bibliography

1 The Medical Devices Regulations 2002 (SI 2002 No 618)

2 European Standard BS EN ISO 14155:2011: Clinical investigation of medical devices for human subjects. Good clinical practice. Available from BSI.

3 International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. http://www.ich.org/

4 Medicines & Healthcare Products Regulatory Agency. Clinical investigations of medical devices – guidance for manufacturers. July 2013

5 The Medical Devices Directive 93/42/EEC

Altman D G: Better Reporting of Randomised Controlled Trials: the Consort Statement. British Medical Journal (1996), 313.570 - 571

CPMP Working Party on the Efficacy of Medicinal Products (1995): Biostatistical Methodology in Clinical Trials in Applications for Marketing Authorisations for Medicinal Products. Statistics in Medicine, 14.1659 - 1682

Hill A B: Principles of Medical Statistics (1967): Oxford University Press, New York

Pocock S J: Clinical Trials: A Practical Approach (1983): John Wiley & Sons, New York

8   Glossary

Alternative hypothesis: The hypothesis that there is a difference or association of some specified degree against which the null hypothesis is tested (see null hypothesis).

Bias: This results when any characteristic of the investigator, study population, or study conduct interferes in a systematic way with the ability to measure a variable accurately.

Blind study: One in which the patient or the investigator (or both) are unaware of the treatment to be received by a patient (see double-blind study and single-blind study).

Clinical investigation: Any systematic study in human subjects undertaken to verify the safety and performance of a specific medical device under normal conditions of use (Standard ISO 14155).

Comparative study: One in which the investigative product is compared against another product, either device or placebo.

Confidence limits: An evaluation of a parameter of a population, an interval or range of values, within which the parameter is most likely to be. The likelihood is specified by a probability.

Control group: The group of subjects in a controlled study that receives no treatment, or placebo, or a standard treatment.

Confidence interval: A range of values within which it is believed (with a particular probability-often 95%) that the true value of the parameter will lie.

Controlled study: The study in which a test article is compared with a treatment that has known effect. The control group may receive no treatment, standard treatment, or placebo.

Crossover trial: A trial in which each subject receives both treatments being compared or the treatment and control. Such trials are used for patients who have a stable, usually chronic condition, during both treatment periods.

Double-blind study: A study in which neither the subject nor the investigators know what treatment a subject is receiving.

Endpoint: An indicator measured in a subject or biological sample to assess the safety, performance, or other objective of a clinical investigation.

Equivalence: Having the same effect as one another (for example, two interventions are equivalent if they result in the same efficacy and safety profiles as one another).

Error: This is the result of inability to accurately measure a variable.

Exclusion criteria: A list of criteria, any one of which excludes the potential subject from participation in the study.

Final report: Complete, comprehensive description of a completed clinical investigation that describes the experimental materials and statistical designs. It also presents and evaluates the trial results and statistical analysis.

Inclusion criteria: The criteria that prospective subjects must meet to be eligible for participation in a study.

Informed consent: Voluntary confirmation of the subject's agreement to take part in a particular investigation, documented in accordance with national guidance or regulation.

Investigator: The person(s) responsible for conducting a clinical investigation and for the health and welfare of subjects during the investigation.

Multi-centred study: A trial conducted under a single protocol but at more than one investigational site and by more than one investigator.

Null hypothesis: A null hypothesis is the assumption to be tested, usually in a significance test, against an alternative hypothesis, that there is no difference between the groups being compared (e.g. no difference in efficacy between groups receiving old and new devices) or no association between two variables being studied (for example no increase in risk with changing patient characteristics).

Open study: A trial in which subjects and investigators know which product each subject is receiving; opposite of double-blind study.

Parallel trial: Volunteers are randomised to one of two differing treatment groups and usually receive the assigned treatment during the entire investigation.

Placebo: A product that takes no active part in diagnosis or treatment. In blinded studies, it is generally made to look like the active product.

Protocol: A document which states the rationale, objectives, statistical design, and methodology of the investigation and the conditions under which it is to be performed and managed.

Randomisation: A process that aims to prevent bias by secretly and arbitrarily assigning subjects to treatment or control groups.

Risk: The probable rate of occurrence of a hazard causing harm and the degree of severity of the harm.

Single arm design: This is an investigation in which there is no parallel comparative group.

Standard deviation: Indicator of the relative variability within a group. The square root of the variance.

Single-blind study: One in which subjects do not know whether they are receiving the treatment or a placebo.

Statistical significance: Level at which an investigator can conclude that observed differences are not due to chance alone.

Type one error: Error made when a correct null hypothesis is rejected.

Type two error: Error made when an incorrect null hypothesis is not rejected.

Validity: The extent to which what is measured is what it was intended to measure.