臨床試驗研究設計

國家衛生研究院
生物統計與生物資訊研究組
熊 昭 組主任

Planning a Trial

Considerations for Overall Clinical Development

Broad Aim of the clinical development of a new drug:

- To find out whether there is a dose range and schedule at which the drug can be shown to be simultaneously safe and effective, to the extent that the risk-benefit relationship is acceptable.

Also need to define:

- The particular subjects who may benefit from the drug
- The specific indications for its use

Development Plan

An ordered program of clinical trials

Each has its own specific objectives

A series of plans with appropriate decision points

- Marketing application
- Common standards:
  - Dictionaries of medical terms
  - Definition and timing of the main measurements
  - Handling of protocol deviations
- Statistical summary, meta analysis

Type of Trials

- Confirmatory Trial
- Exploratory Trial

Confirmatory Trial

- An adequately controlled trial in which the hypotheses (claims) are stated in advance and evaluated.
- To provide clear firm evidence of efficacy (clinical benefits) and/or safety
- The key hypothesis of interest follows from the trial's primary objective, always predefined.
- Equally important to estimate with precision the size of the effects attributable to the treatment of interest.

Confirmatory Trial

- Adherence to protocol and standard operating procedures.
- Unavoidable changes should be explained and documented.
- Justification of the design and statistical aspects.
- Each trial addresses only limited number of questions.
- The results of the confirmatory trial(s) should be robust.

Exploratory Trial

- A series of exploratory studies => rationale and design of confirmatory trials
- Should have clear and precise objectives, but may not lead to simple tests of predefined hypotheses
- Analyses may entail data exploration
- Can not be the basis of the formal proof of efficacy
- In confirmatory trials the data are also subjected to exploratory analyses which serve as a basis for suggesting further hypotheses for later research.

Scope of Trials

- Population
- Primary and Secondary Variables

Population

In earlier phases of drug development,

the choice of subjects for a trial may be influenced by the wish to maximize the chance of observing specific clinical effects of interest.

In confirmatory trials,

the subjects in the trials should closely mirror the target population

It is generally helpful to relax the inclusion and exclusion criteria as much as possible within the target population, while maintaining sufficient homogeneity to permit precise estimation of treatment effects.

Primary Variable

- Primary variable (“target” variable, primary endpoint) : the variable capable of providing the most clinically relevant and convincing evidence directly related to the primary objective of the trial.
- Generally only one primary variable, usually be an efficacy variable
- Safety/tolerability, quality of life and health economics are other possible primary endpoint.
- Should reflect the standards in the relevant field of research. Use of a validated variable.
- The primary variable should be the one used when estimating the sample size.

Primary variable

Primary variables used in the statistical analysis to measure the treatment effect

- Mortality:
  - Proportions alive at fixed points in time, e.g. 5 year survival rate.
  - Overall distributions of survival times.
- Recurring event:
  - A simple dichotomous variable: any occurrence of events during a specified interval.
  - Time to first occurrence
  - Rate of occurrence: events per time units of observation.

Primary variable

- Assessments of functional status over time:
  - In studying treatment for chronic disease
- Assessments done at the beginning and end of the interval of observations
- Proportions of subjects exceeding beyond a specified threshold
- Repeated measures data

Primary and Secondary Variables

- Redefinition of primary variables after unblinding is almost always unacceptable.
- Secondary variables are supportive measurements to the primary objective or measurements of effects related to the secondary objectives.
- For composite variables, a useful strategy is to integrate or combine the multiple measurements into a single or “composite” variable, using a predefined algorithm.

Primary and Secondary Variables

Other variables of interest:

- Rating scales used in arthritis, psychiatric disorders: Inter- and intra rater reliability.
- Global Assessment Variables
  - to measure the overall safety, overall efficacy
  - Usually a scale of ordered categorical ratings
- Categorized Variables
  - Dichotomization/categorization of continuous or ordinal variable: success/response, specify minimum percentage improvement relative to baseline

Design Techniques to Avoid Bias

- Blinding
- Randomization
- To use double blind approach in which treatments are prepacked in accordance with a suitable randomization schedule.
- No one in the conduct of the trial is aware of the treatment allocated to subject.

Blinding

- Limit the conscious and unconscious bias in the conduct of a trial
- Include double-blind, single-blind, open-label
- Assessment should better be made by medical staff who are not involved in treating the subjects and who remain blind to treatment.
- Minimize the various sources of bias
- Set up adequate standard operating procedures to unblind.

Randomization

- Provide an element of chance into the assignment of treatments to subjects and provide a statistical basis in evaluation effects.
- Randomize subjects in blocks
  - Investigators/staff should be blind to the block length. (not stated in the protocol)
  - use two or more block lengths, randomly selected for each block.
- In multicenter trials, the randomization should be done centrally.

Randomization

- In order to promote balanced allocation within strata
  - To stratify by center
  - To stratify by important prognostic factors measured at baseline (e.g. severity of disease, age, sex)
  - The use of more than two or three stratification factors is rarely necessary.
  - Stratified factors should be accounted for in the analysis.

Trial Design

- Parallel Group Design
- Crossover Design

Parallel Group Design

- The most common design:

  Subjects are randomized to one of two or more arms.

- Investigational product at one or more doses versus one or more control treatments (placebo/active comparator)
- Statistics methods for two sample or *K*-sample problems

Crossover Design

- Each subject is randomized to a sequence of two or more treatments and acts as his/her own control for treatment comparisons.
- It reduces the number of subjects
- 2-2 crossover design:

  A (washout) B

  B (washout) A

  randomize

- Main difficulty: carryover

Crossover Design

- The disease under study should be chronic and stable
- The relevant effects of the medication should develop fully within the treatment period.
- The washout period should be sufficiently long
- A common use of the 2-2 crossover design is to demonstrate the bioequivalence of 2 formulations of the same medication in healthy volunteers.

Type of Comparison

- Superiority
- Equivalence or Noninferiority
- Dose-Response Relationship

Superiority Trial

- Efficacy is most convincingly established by demonstrating superiority to placebo in a placebo-controlled trial, by showing superiority to an active control treatment, or by demonstrating a dose-response relationship.
- For serious illness, when an active therapeutic treatment exists, a placebo-controlled trial may be unethical.

Equivalence or Noninferiority Trial

- Equivalence trial:

  Bioequivalence trial

  Clinical equivalence trials

  to demonstrate the clinical equivalence of a generic product to the marketed product when the compound is not absorbed.

- Noninferiority trial:

  Many active control trials are designed to show that the efficacy of an investigational product is no worse than that of the active comparator.

Equivalence or Noninferiority Trial

- A suitable active comparator should be a widely used therapy whose efficacy has been clearly established.
- The new trial should have the same important design features as the previous superiority trials in primary variables, the dose of the active comparator, eligibility criteria etc.
- An equivalence margin should be specified, this margin is the largest difference that is clinically acceptable, should be smaller than differences observed in superiority trials of the active comparator.

Equivalence or Noninferiority Trial

- For the active control equivalence trial, both the upper and the lower equivalence margins are needed.
- For the active control noninferiority trial, only the lower margin is needed.
- Statistical analysis is based on the use of confidence intervals.(CI)

Equivalence or Noninferiority Trial

- For equivalence trial, two-sided C.Is are used.

  When the entire C.I. falls within the equivalence margins→Equivalence

  Equivalent to using two simultaneous one-sided tests to test

  H_0: the treatment difference is outside the equivalence margins.

  vs H_1: the treatment difference is within the margins.

- For noninferiority trials, one-sided intervals are used.

  H_0: the tx difference (investigational product-control) is equal to the lower margin.

  vs H_1: the tx difference is $>$ the lower margin.

Equivalence or Noninferiority Trial

- Note: The following is inappropriate:

  Concluding equivalence based on observing a nonsignificant test result to test

  H_0: there is no difference between the investigational product & the active comparator

Group Sequential Designs

- To facilitate the conduct of interim analysis.
- The statistical methods and early stopping rules should be fully specified in advance.
- The sequential design often used in large, long term trials of mortality.
- An independent data monitoring committee review the interim analysis.
- Safety must be monitored in all trials.

Sample Size

- The sample size should be large enough to provide a reliable answer.
- The number is determined by the primary objective; a primary variable, the test statistics, H_0, H_1(working hypothesis (minimal effect which has clinical relevance)), type I error probability ($\leq$ 5%), type II error probability (10%-20%)
- Assumptions should be based on published data or on the results of earlier trials.
- The method of calculating sample size should be given in protocol.

Multicenter Trial

- To accrue sufficient subjects within a reasonable timeframe.
- To provide better basis for the subsequent generalization.
- The protocol should be clear and common at all centers.
- Procedures should be standardized.

Multicenter Trial

- Variations of evaluation criteria and schemes can be reduced by investigational meetings, personnel training in advance, and careful monitoring during the trial.
- In statistical analysis, the main treatment effect may be investigated first using a model that allows for center effects. In some trials with very few subjects per center, it is impractical to include the center effects in the model.

Statistical Section of the Protocol

- Confirmatory analysis of the primary variables
- Exploratory trials

  General principles and directions
- In order to satisfy the primary objectives of the trial, should specify the hypotheses that are to be tested and/or the treatment effects that are to be estimated
- The statistical methods to be used should be described for the primary variable.

Statistical Analysis Plan

- A separate document
- A more technical and detailed elaboration
- Detailed procedures for executing the statistical analysis of the primary and secondary variables
- Finalized before breaking the codes
- If the blind review suggests changes to the principal features in the protocol, should be documented in an amendment.
- Only results from analyses stated in the protocol (including amendments) can be regarded as confirmatory.

Analysis Sets

- The set of subjects whose data are to be included in the main analyses should be defined in the statistical section
- Full analysis set (Intention- to treat)

  per protocol set
- In confirmatory trials, usually appropriate to plan to conduct both an intention- to treat analysis and a per protocol analysis

Decisions Concerning the Analysis Set

Guided by

1. To minimize bias
2. To avoid inflation of Type I error
3. To provide a secure foundation for statistical tests

Full Analysis Set

Intention-to-treat principle:

- The primary analysis should include all randomized subjects .
- May include subjects with errors in treatment assignment, the use of excluded medication, poor compliance, loss to follow up and missing data.
- Provide estimates of treatment effects that are more likely to mirror those observed in subsequent practice.

Full Analysis Set

Intention-to-treat principle: (con't)

- Might exclude (from the set of all randomized subjects) subject with
  - the failure to satisfy major entry criteria
  - the failure to take at least one dose of trial medication
  - the lack of any data post randomization

Per Protocol Set

Evaluable subjects who are more compliant with the protocol

- The completion of a certain prespecified minimal exposure to the treatment regimen
- The availability of measurements of the primary variables
- The absence of any major protocol violations

Imputation for Missing Data

- The carrying forward of the last observation
- The use of complex mathematical models

Data Transformation

- The decision to transform key variables prior to analysis is made during the design of the trial on the basis of similar data from earlier clinical trials.
- Transformations (e.g. square root, logarithm) should be specified in the protocol and a rationale provided.
- To ensure that the assumptions underlying the statistical methods are met.
- The decision on whether and how to transform a variable should be influenced by the preference for a scale that facilitates clinical interpretation.

Types of Variables

- Continuous, proportions, categorical, ordinal, time to event, repeat measures.
- Derived variables
  - The use of change from baseline
  - Percentage change from baseline
  - The “area under the curve” of repeated measures
  - The ratio of two different variables

Estimation, Confidence Intervals and Hypothesis Testing

- Estimates of treatment effects should be accompanied by confidence intervals, and the way in which these will be calculated.
- Use baseline data to improve precision or to adjust for potential baseline differences, for example, by means of analysis of covariance.
- All effects (covariates) to be fitted in the analysis (for example, in analysis of variance (covariance) models) should be fully specified.

Hypothesis Testing

- Important to clarify whether one or two-sided tests of statistical significance will be used.
- Setting Type I error (2.5%) for one-sided tests at half the conventional Type I error used in two-sided tests is preferable.

Scientific merit

- Quality

Speed