

Chapter 9

Setting the Ideal Statistical Methods

9.1 Introduction

The study design, sample size and statistical analysis must be able to properly evaluate the research hypothesis set forth by the clinical investigator. Otherwise, the consequences of a poorly developed statistical approach may result in a failure to obtain extramural funding and result in a flawed clinical study that cannot adequately test the desired hypotheses. Statisticians provide design advice and develop the statistical methods that best correspond to the research hypothesis [1].

9.2 Randomisation Plan

Random allocation of subjects to study groups is fundamental to the clinical trial design. Randomisation, which is a way to reduce bias, involves random allocation of the participants to the treatment groups. If investigators compare a new treatment

against a standard treatment, the study subjects are allocated to one of these treatments by a random process. A general description of the randomisation approach may be introduced in the clinical method section of the proposal; for example, ‘treatment assignment will be determined using stratified, blocked randomisation’. Specific randomisation details will need to be elaborated in the statistical method section, including how the allocation procedure will be implemented by, for example, computer programmes, a website, lists or sealed envelopes. If stratification is deemed necessary, include in the proposal a description of each stratification variable and the number of levels for each stratum, for instance, sex (male, female) or diabetes (type 1, type 2). However, keep the number of strata and stratum levels minimal [2].

9.3 Blinding

Knowledge of treatment assignment might influence how much of a dosage change is made to a study treatment or how an AE is assessed. Blinding or masking is another component of study design used to try to eliminate such bias [3]. In a double-blind randomised trial, neither the study subjects nor the clinical investigators know the treatment assignment. Describe the planned blinding scheme. For example, ‘this is a double-blind randomised study to investigate the effect of propranolol versus no propranolol on the incidences of total mortality and of total mortality plus nonfatal myocardial infarction in 158 older patients with congestive heart failure [CHF] and prior myocardial infarction’. Specify who is to be blinded and the steps that will be taken to maintain the blinding. It is important that evaluators such as radiologists, pathologists or laboratory personnel who have no direct contact with the study subjects remain blinded to treatment assignments.

9.4 Sample Selection/Allocation Procedures

1. *Matching*: When confounding cannot be controlled by randomisation, individual cases are matched with individual controls who have similar confounding factors, such as age, to reduce the effect of the confounding factors on the association being investigated in analytical studies. This is most commonly seen in case-control studies.
2. *Restriction (specification)*: Eligibility for entry into an analytical study is restricted to individuals within a certain range of values for a confounding factor, such as age, to reduce the effect of the confounding factor when it cannot be controlled by randomisation. Restriction limits the external validity (generalisability) to those with the same confounder values.
3. *Census*: A sample that includes every individual in a population or group (e.g. entire herd, all known cases). A census is not feasible when the group is large relative to the costs of obtaining information from individuals.
4. *Haphazard, convenience, volunteer, judgmental sampling*: Any sampling not involving a truly random mechanism. A hallmark of this form of sampling is that the probability that a given individual will be in the sample is unknown before sampling. The theoretical basis for statistical inference is lost and the result is inevitably biased in unknown ways. Despite their best intentions, humans cannot choose a sample in a random fashion without a formal randomising mechanism.
5. *Consecutive (quota) sampling*: Sampling individuals with a given characteristic as they are presented until enough with that characteristic are acquired. This method is possible for descriptive studies but unfortunately not much better than haphazard sampling for analytical observational studies.
6. *Random sampling*: Each individual in the group being sampled has a known probability of being included in the sample obtained from the group before the sampling occurs.

7. *Simple random sampling/allocation*: Sampling conducted such that each eligible individual in the population has the same chance of being selected or allocated to a group. This sampling procedure is the basis of the simpler statistical analysis procedures applied to sample data. Simple random sampling has the disadvantage of requiring a complete list of identified individuals making up the population (the list frame) before the sampling can be done.
8. *Stratified random sampling*: The group from which the sample is to be taken is first stratified on the basis of an important characteristic related to the problem at hand (e.g. age, parity, weight) into subgroups such that each individual in a subgroup has the same probability of being included in the sample, but the probabilities differ between the subgroups or strata. Stratified random sampling assures that the different categories of the characteristic that is the basis of the strata are sufficiently represented in the sample, but the resulting data must be analysed using more complicated statistical procedures (such as Mantel-Haenszel) in which the stratification is taken into account.
9. *Cluster sampling*: Staged sampling in which a random sample of natural groupings of individuals (houses, herds, kennels, households, stables) is selected and then all the individuals within the cluster are sampled. Cluster sampling requires special statistical methods for proper analysis of the data and is not advantageous if the individuals are highly correlated within a group (a strong herd effect).
10. *Systematic sampling*: From a random start in the first n individuals, sampling every n th subject/animal as they are presented at the sampling site (clinic, chute, etc.). Systematic sampling will not produce a random sample if a cyclical pattern is present in the important characteristics of the individuals as they are presented. Systematic sampling has the advantage of requiring only knowledge of the number of

subjects/animals in the population to establish n and that anyone presenting the subjects/animals is blind to the sequence so they cannot bias it [4].

9.5 Statistical Analysis Methodology

The statistical analysis methods for analysing study outcomes must be carefully detailed. Specifying these methods in advance is another way to minimise bias and maintain the integrity of the analysis. Any changes to the statistical methods must be justified and decided on before the blind is broken [5]. In the statistical analysis plan, not only must the statistical hypotheses to be tested be described and justified but which subjects and observations will be included or excluded in each analysis must also be detailed. The statistical analysis plan is driven by the research questions, the study design and the type of the outcome measurements. The analysis plan includes a detailed description of statistical testing for each of the variables in the specific aim(s). If several specific aims are proposed, an analysis plan should be written for each specific aim. Plan descriptive analyses for each group or planned subgroup. If subjects were randomly assigned to groups, there should be a description of subject characteristics that includes demographic information as well as baseline measurements or comorbid conditions. Specify anticipated data transformations that may be needed to meet analysis assumptions, and describe derived variables to be created such as area under the curve. Incorporate confidence intervals in the analysis plan for reporting treatment effects. Confidence limits are much more informative to the reader than are P values alone [6].

Statistical details and terminologies are not intended to be an obstacle for a young investigator. Instead, this is where a statistical expert can be a valuable resource to help the investigators use the

appropriate statistical methods and language that address the research hypotheses. Brief statistical analysis descriptions are written below.

9.5.1 Statistical Analysis Example for a Randomised Study

Statistical analysis. The full analysis set will include patients who have received at least one dose of medication or had one or more post randomisation, follow-up evaluations. Descriptive statistics will be computed for each treatment group; medians and percentiles will be reported for skewed continuous variables. For primary and secondary outcomes, descriptive statistics and 95 % confidence intervals will be used to summarise the differences between groups. The primary outcome of systolic blood pressure and other continuous variables will be assessed with a repeated-measures analysis using a mixed linear model approach. Because many of the inflammatory markers are positively skewed, interleukin 6 and C-reactive protein levels will be log transformed before analysis. The Wilcoxon rank sum test will be used to compare pill counts between groups. Hypothesis tests will be two sided using the 0.05 significance level. Bonferroni-type adjustments for multiple testing will be implemented to control type I errors. Statistical analysis will be performed using SAS software (SAS Institute, Cary, NC) [1].

9.5.2 Statistical Analysis Example for a Longitudinal Cohort Study

Descriptive/comparative statistics defines the biomarker levels in the different disease activity classes. We will compute and compare the mean/median and interquartile range of urine biomarker levels in different disease activity groups, after partitioning

patients in various ways: patients who attain any of the primary disease outcomes, i.e. World Health Organization class III or IV glomerulonephritis, patients with nephritic or nephrotic flares, or patients with end-stage renal disease. In addition, we will define the biomarker levels in patients with the following disease features: anaemia, leucopenia or thrombocytopenia. To compare multiple patient groups, analysis of variance (ANOVA) or the Kruskal-Wallis test will be used, depending on whether the biomarker values are normally distributed. Data transformations will be performed if necessary. If the omnibus ANOVA or Kruskal-Wallis test yields $P < 0.05$, we will conduct pairwise group comparisons using either t tests or Wilcoxon rank sum tests with Bonferroni corrections. The generalised estimating equation approach will be used to evaluate whether urinary biomarkers vary significantly over time among different disease activity classes [1].

References

1. Adams-Huet B, Ahn C. Bridging clinical investigators and statisticians: writing the statistical methodology for a research proposal. *J Investig Med*. 2009;57(8):818–24. doi:[10.2310/JIM.0b013e3181c2996c](https://doi.org/10.2310/JIM.0b013e3181c2996c).
2. Pocock S. *Clinical trials: a practical approach*. New York: Wiley; 1983.
3. Day SJ, Altman DG. Statistics notes: blinding in clinical trials and other studies. *BMJ*. 2000;321:504.
4. Clinical Study Design and Methods Terminology. Clinical epidemiology & evidence-based medicine glossary. <http://people.vetmed.wsu.edu/jmgay/courses/GlossClinStudy.htm>. Accessed online at 16 Oct 2015.
5. International Conference on Harmonisation (ICH). *Guidance for industry: E9 statistical principles for clinical trials*. Rockville: Food and Drug Administration; 1998.
6. Altman D. *Practical statistics for medical research*. New York: Chapman & Hall/CRC; 1991.