HAD 5303H
Controlled Clinical Trials
Statistical Analysis

Prof. Kevin E. Thorpe

Dept. of Public Health Sciences
University of Toronto

Outline

The Analysis Section

Introduction

- The analysis section of a protocol attempts to describe the planned statistical analyses that will be carried out to answer the research questions.
- This is planned in advance so that the choice of statistical analysis is not *inappropriately* influenced by the data.

The Analysis Section

Things to Include

- A description of the primary analysis of the primary outcome.
- A description of the primary analyses of any secondary outcomes.
- A description of secondary analyses of the primary and secondary outcomes.
- Other secondary analyses.
- Which analyses will follow the intention-to-treat principle.
- How missing data will be handled.

The Analysis Section

Things Which Might be Included

- Whether or not statistical tests will be one or two sided and the significance level for testing, if not specified elsewhere.
- Dummy tables.

Primary and Secondary Analyses

Defining the Concepts

- The terms *primary analysis* and *secondary analysis* often get confounded with the terms *primary outcome* and *secondary outcome*.
- The *primary analysis* of an *outcome* is, generally, the simple comparison between the treatment groups.
- The *secondary analysis* of an *outcome* is, generally, the more complex analysis where you might obtain adjusted treatment effects, perform subgroup analyses, etc.
- The *primary analysis* for a *trial* is the primary analysis for the primary outcome. All other analyses are considered *secondary* with respect to the trial.

Primary and Secondary Analyses

Primary Analysis

- An outcome, whether it be primary or secondary, has a single primary analysis which can usually be described in 1–2 sentences.

Example

The primary research question will be addressed by the comparison of 2 proportions using a chi-square test. Specifically, the 2 high dose groups (650 and 1300 mg/day) will be combined and compared with the combined low dose groups (80 and 325 mg/day), on the proportion of patients suffering stroke (of any type) or death (from any cause) in the 30 day perioperative period.

- The primary analysis of the primary outcome should mirror the sample size calculation for the trial.

Primary and Secondary Analyses

Secondary Analyses

- The secondary analyses describe adjusted analyses or exploratory or hypothesis generating analyses.

Example

Additional analyses, employing logistic regression, will examine, and control for, the influence of patient and surgical factors which might be associated with the risk of perioperative stroke and death including: degree of stenosis, presence/absence of ulceration, medical conditions (e.g. diabetes, prior myocardial infarction, hypertension, etc.), recency and type of cerebrovascular events, ASA regimen prior to randomization, use of heparin during clamping, type of cerebral monitoring, type of arteriotomy closure, and appearance of the plaque during surgery.

Intention-to-Treat

From the *Trial Protocol Tool* glossary:

Definition (Intention-to-Treat)

An intention-to-treat analysis is one in which all the participants in a trial are analysed according to the intervention to which they were allocated, whether they received it or not.

Intention-to-Treat

Operational Characteristics

- All randomized patients are included, even those found to be ineligible post-randomization.
- Participants are analysed according to the treatment they were allocated to receive irrespective of compliance to the allocation.

Examples

- A patient randomized to receive a surgical procedure refuses. The patient would still be included in the surgical group for analysis.
- A patient taking only half of the drug allocated to them would be included in the arm they were randomized to ignoring the non-compliance.

- Participants who are lost to follow-up are not removed from the analysis.

Intention-to-Treat

The Value

- Non-compliance is a reality in clinical practice. Intention-to-treat analysis implies that the treatment is being compared in a more "real world" setting.
- An analysis that includes only compliant participants may result in a biased estimate of the treatment effect.

Common Statistical Techniques

Introduction

- Although there exist many complex statistical methods, for the primary analyses typical of RCTs, there are a few commonly used methods.
- The choice of method depends on the type of outcome (eg. measured, binary, etc.).
- The slides that follow will highlight the common outcome types and statistical methods used in clinical trials.

Common Statistical Techniques

Continuous/Measured Data

- Continuous data is typically something measured with an instrument (eg. blood pressure, viral load)
- The data are summarized by means and standard deviations and treatment effects are computed as differences in the mean response between groups.
- A primary analysis for continuous data is typically a t-test (two group comparisons) or Analysis of Variance (more than two groups).
- Multiple linear regression is the usual tool for adjusted analyses (often the secondary analyses).
- For cluster randomized trials, there are cluster-aware versions.

Common Statistical Techniques

Binary Data

- Binary outcomes have two states that may by thought of as success or failure (eg. alive/dead, cured/not cured) where the actual status can be determined for each individual.
- The data are typically summarized as a contingency table and treatment effects are usually odds ratios.
- The primary analysis usually involves a χ^2 test or a Fisher's exact test.
- Multiple logistic regression is the usual tool for adjusted analyses (often the secondary analyses).
- For cluster randomized trials, there are again, cluster-aware versions.

Common Statistical Techniques

Survival or Time to Event Data

- This data arises when you are interested in the time to some event that may or may not happen to all patients (eg. time to stroke, time to relapse) or an event that you cannot observe for all who will have it because the trial concludes (eg. death, stroke).
- Survival is typically summarized graphically by the Kaplan-Meier curve and treatment effect is often given as a hazard ratio.
- Survival is typically compared by the log-rank test.
- Adjusted analyses are most often performed by the (Cox) proportional hazard model.
- Once again, there are similar methods available for cluster randomized trials.

Dummy Tables

- Dummy tables are table shells that show how you plan to summarize the trial data.
- If included in a protocol, they should go in an appendix.
- If not in a protocol, they should be prepared prior to beginning analysis as a part of a *Statistical Analysis Plan (SAP)*.
- The purpose is to ensure that the choice of analyses are not driven by the results.