

Statistical Guidelines for Authors

Guidelines are intended to aid authors in reporting statistical aspects of their work in ways that are clear and helpful to readers. Statistical aspects should be described in sufficient detail to allow a knowledgeable reader with access to the original data to verify the reported results.

Statistical guidelines are provided below along with a brief explanation; for a more detailed description see Bailar III and Mosteller (1).

Manuscript Section

Abstract

- Provide a brief description of the study design (e.g. randomized, double-blind, placebo-controlled trial or retrospective cohort study).
- Describe the number of observations with a clear statement as to the unit of analysis (e.g. total number of patients, or total number of biopsies).

Statistical Methods

- Describe the study design. Distinguish between an observational study (e.g. retrospective or prospective cohort study) and a randomized controlled trial.
- Report what statistical methods were used and why. Common techniques (e.g. t-test) do not need to be described in detail, but more complex methods (e.g. mixed effects linear regression) require explanation with appropriate references.
- Report reasons for and method of selecting subjects or other observational units.
- Provide sample size justification for a clinical trial if applicable (e.g. a Phase III study including all assumptions for calculation).
- Provide details about randomization and methods for masking of treatments, if appropriate.
- Describe the basic observational units (e.g. total number of patients or total number of biopsies) and specify which observational units are included in denominators. Include description of the number and reasons for exclusions of observational units from the study population.
- Distinguish between prevalence, a proportion of subjects or observations, and incidence or rate, a quantity expressed per unit of time.
- Distinguish between multivariable, referring to several predictors or explanatory variables or risk factors, and multivariate, referring to simultaneous analysis of several outcome variables.
- Provide details regarding methods used to impute missing data (e.g. pattern-mixture (mixed) models or multiple imputation).
- Specify whether tests are two-tailed versus one-tailed.
- Describe any multiple comparison techniques used to control the overall type I error rate when performing multiple tests (e.g. pairwise comparison among several groups).
- Describe any transformations of the data used in the analysis.
- Specify any general software used to analyze the data (e.g. SAS version 9.2 [SAS Institute, Cary, NC]).
- For multivariable models in which any variable selection models were used (e.g. backward selection), clearly state which variables were considered for the model and what probability threshold was utilized for inclusion/exclusion of variables. Otherwise state the rationale for selecting the variables that are utilized in the final multivariable model.

Notation and Terminology for Methods

- Explain meaning of notations such as SE, SD, CL or CI in abstract, methods, results, tables and figures when they first appear.
- Mean (SD) is preferred over mean \pm SD, for example.

Results

- Report exact p values rather than statements like “p < 0.05” or “NS” and no more than 3 significant digits to the right of the decimal point (e.g. for a p = 0.12346, report p=0.123). For highly significant p values, e.g. p = 0.00001, report p < 0.001.
- Units should always be specified in text, tables and figures.
- Report the number of observational units including the number of observations excluded and reasons why. For clinical trials and complex observational studies use a CONSORT diagram (2).
- CONSORT checklists are required at submission for all studies reporting randomized, controlled trials (3).
- Report losses to follow-up (such as drop-outs in a clinical trial).
- Report the frequency of missing variables and outcomes; reasons and patterns for the missing data. Method to handle missing data should be described in detail in the methods section.
- Continuous variables (e.g. age) should be summarized using the mean and standard deviation (SD). If the distribution of measurements is not normally distributed, the median and a percentile range (e.g. inter-quartile range) provide a more appropriate summary. For ordinal data (e.g. data on an ordered scale such as acute rejection grade), the use of mean and SD is incorrect; instead proportions should be employed. Confidence intervals should be presented as an indication of the uncertainty of sample means, proportions and other summary statistics.
- Reporting comparative results in terms of confidence intervals in addition to p values is recommended (e.g. 95% confidence interval for the mean difference between groups or 95% confidence for the hazard ratio for graft failure).
- When presenting means, standard deviations and other statistics the authors should bear in mind the precision of the original data. Means should not be given to one decimal place more than the raw data, but standard deviations and standard errors may need to be quoted to one extra decimal place.
- For time-related events, results reported at periodic intervals (e.g. one-year posttransplant) should include point-estimates using the product-limit (Kaplan-Meier) or Life-Table method and confidence intervals.
- For survival curves, report the number of events and the number of subjects at risk at periodic intervals along the x-axis (e.g. at one, three and five years posttransplantation).

References

1. Bailar III JC, Mosteller F. Guidelines for statistical reporting in articles for medical journals: amplifications and explanations. *Ann Intern Med* 1988; 108: 266-273.
2. <http://www.consort-statement.org/consort-statement/flow-diagram0/>
3. <http://www.consort-statement.org/consort-statement/overview0/#checklist>