

5.7: Basic statistical reporting for articles published in clinical medical journals: the Statistical Analyses and Methods in the Published Literature, or SAMPL guidelines

Thomas A. Lang

Tom Lang Communications and Training International, Kirkland, USA; tomlangcom@aol.com

Douglas G. Altman

Director, Centre for Statistics in Medicine, Oxford University, UK; doug.altman@csm.ox.ac.uk

Have they reflected that the sciences founded on observation can only be promoted by statistics?... If medicine had not neglected this instrument, this means of progress, it would possess a greater number of positive truths, and stand less liable to the accusation of being a science of unfixed principles, vague and conjectural.

Jean-Etienne Dominique Esquirol, an early French psychiatrist, quoted in *The Lancet*, 1838¹

Introduction

The first major study of the quality of statistical reporting in clinical journals was published in 1966². Since then, dozens of similar studies have been published, every one of which has found that large proportions of articles in these journals contain errors either in the application, analysis, interpretation, or reporting of statistics, or in the design or conduct of research. (See references 3–19 for examples.) Further, a large proportion of these errors is serious enough to call the authors' conclusions into question.^{5,18,19} The problem is made worse by the fact that many of these studies appear in the world's leading peer-reviewed general medical and specialty journals.

Although errors have been found in more complex statistical procedures,^{20,21,22} paradoxically, many errors are in basic, not advanced, statistical methods.²³ Perhaps advanced methods are suggested by consulting statisticians, who then competently perform the analyses, but many authors are far more likely to use only elementary statistical methods, if they use any at all, and may not apply these methods correctly.^{23–26} Yet articles with even major errors continue to pass editorial and peer review and be published in leading journals.

The truth is that the problem of poor statistical reporting is long-standing, widespread, potentially serious, concerns mostly basic statistics, and yet is largely unsuspected by most readers of the biomedical literature.²⁷

More than 30 years ago, O'Fallon and colleagues recommended that "standards governing the content and format of statistical aspects should be developed to guide authors in the preparation of manuscripts".²⁸ Despite the fact that this call has since been echoed by several others,^{17,18,29–32} most journals have still not included in their instructions for authors more than a paragraph or two about reporting

statistical methods.³³ However, given that many statistical errors concern basic statistics, a comprehensive – and comprehensible – set of reporting guidelines suitable for use by clinical journals might improve how statistical analyses are documented.

In light of these concerns, we present here a set of statistical reporting guidelines suitable for clinical journals to include in their instructions for authors. These guidelines tell authors, journal editors, and reviewers how to report basic statistical methods and results. Although these guidelines are limited to the most common statistical analyses used in clinical research, they are nevertheless sufficient to prevent most of the reporting deficiencies routinely found in articles reporting this research. They may also help to prevent some reporting errors by focusing attention on key points in the analyses.

The present guidelines were not developed by a formal consensus-building process, but they do draw considerably from published guidelines.^{27,32,34–37} In addition, a comprehensive review of the literature on statistical reporting errors in clinical journals reveals near universal agreement on how to report the most common methods.²⁷

Statistical analyses are closely related to the design and activities of the research itself. However, our guidelines do not address the issues related to the design and conduct of clinical research. Instead, we refer readers to the EQUATOR Network website (www.equator-network.org, see Chapter 5.6) where guidelines for reporting specific research designs can be found. (For example, see CONSORT,³⁸ TREND,³⁹ and STROBE.⁴⁰) These guidelines for reporting methodologies all include items on reporting statistics, but the guidelines presented here are more specific and complement, rather than duplicate, those in the methodology guidelines.

We welcome feedback and anticipate the need to update these guidelines in due course.

The 'Statistical Analyses and Methods in the Published Literature' or SAMPL Guidelines

Guiding principles for reporting statistical methods and results

Our first guiding principle for statistical reporting comes from the International Committee of Medical Journal Editors (ICMJE), whose Uniform Requirements for

Manuscripts Submitted to Biomedical Journals include the following excellent statement about reporting statistical analyses:

Describe statistical methods with enough detail to enable a knowledgeable reader with access to the original data to verify the reported results. When possible, quantify findings and present them with appropriate indicators of measurement error or uncertainty (such as confidence intervals). Avoid relying solely on statistical hypothesis testing, such as *P* values, which fail to convey important information about effect size. References for the design of the study and statistical methods should be to standard works when possible (with pages stated). Define statistical terms, abbreviations, and most symbols. Specify the computer software used.^{33,41} [Emphasis added.]

Our second guiding principle for statistical reporting is to **provide enough detail so that the results can be incorporated into other analyses, specifically, systematic reviews and meta-analyses.** In general, this principle requires reporting the descriptive statistics from which other statistics are derived, such as the numerators and denominators of percentages, especially in risk, odds, and hazards ratios. Likewise, *P* values are not sufficient for re-analysis. Needed instead are descriptive statistics for the variables being compared, including sample size of the groups involved, the estimate (or “effect size”) associated with the *P* value, and a measure of precision for the estimate, usually a 95% confidence interval.

General principles for reporting statistical methods

Preliminary analyses

- Identify any statistical procedures used to modify raw data before analysis. Examples include mathematically transforming continuous measurements to make distributions closer to the normal distribution, creating ratios or other derived variables, and collapsing continuous data into categorical data or combining categories.

Primary analyses

- Describe the purpose of the analysis.
- Identify the variables used in the analysis and summarize each variable using descriptive statistics.
- When possible, identify the smallest difference considered to be clinically important.
- Describe fully the main methods for analysing the primary objectives of the study.
- Make clear which method was used for which analysis, rather than just listing in one place all the statistical methods used.
- Verify that the data conformed to assumptions of the test used to analyse them. In particular, specify that 1) skewed data were analysed with non-parametric tests; 2) paired data were analysed with paired tests, and 3) the underlying relationship analysed with linear regression models was linear.

- Indicate whether and how any allowance or adjustments were made for multiple comparisons (performing multiple hypothesis tests on the same data).
- If relevant, report how any outlying data were treated in the analysis.
- Say whether tests were one or two-tailed and justify the use of one-tailed tests.
- Report the alpha level (e.g., 0.05) that defines statistical significance.
- Name the statistical package or program used in the analysis.

Supplementary analyses

- Describe methods used for any ancillary analyses, such as sensitivity analyses, imputation of missing values, or testing of assumptions underlying methods of analysis.
- Identify post-hoc analyses, including unplanned subgroup analyses, as exploratory.

General principles for reporting statistical results

Reporting numbers and descriptive statistics

- Report numbers – especially measurements – with an appropriate degree of precision. For ease of comprehension and simplicity, round to a reasonable extent. For example, ‘mean age’ can often be rounded to the nearest year without compromising either the clinical or the statistical analysis. If the smallest meaningful difference on a scale is, say, 5 points, scores can be reported as whole numbers; decimals are not necessary.
- Report total sample and group sizes for each analysis.
- Report numerators and denominators for all percentages.
- Summarize data that are approximately normally distributed with means and standard deviations (SD). Use the form: mean (SD), not mean \pm SD.
- Summarize data that are not normally distributed with medians and inter-percentile ranges, ranges, or both. Report the upper and lower boundaries of inter-percentile ranges and the minimum and maximum values of ranges, not just the size of the range.
- Do NOT use the standard error of the mean (SE) to indicate the variability of a data set. Use standard deviations, inter-percentile ranges, or ranges instead. (The SE is an inferential statistic, not a descriptive one.)
- Display data in tables or figures. Tables present exact values, and figures provide an overall assessment of the data.^{42,43}

Reporting risk, rates, and ratios

- Identify the type of rate (e.g., incidence rates, survival rates), ratio (e.g., odds ratios, hazards ratios), or risk (e.g., absolute risks, relative risk differences), being reported.
- Identify the quantities represented in the numerator and denominator (e.g., the number of men with prostate cancer divided by the number of men in whom prostate cancer can occur).

- Identify the time period over which each rate applies.
- Identify any unit of population (that is, the unit multiplier: e.g., x 100 persons, x 10,000 persons) associated with the rate.
- Report a measure of precision (a confidence interval) for estimated risks, rates, and ratios.

Reporting hypothesis tests

- State the hypothesis being tested.
- Identify the variables in the analysis and summarize the data for each variable with the appropriate descriptive statistics.
- If possible, identify the minimum difference considered to be clinically important.
- For equivalence and non-inferiority studies, report the largest difference between groups that will still be accepted as indicating biological equivalence (the equivalence margin).
- Identify the name of the test used in the analysis. Report whether the test was one or two-tailed, and for paired or independent samples. Justify the use of one-tailed tests.
- Confirm that the assumptions of the test were met by the data.
- Report the alpha level (e.g., 0.05) that defines statistical significance.
- At least for primary outcomes (e.g., differences or agreement between groups, diagnostic sensitivity, and slopes of regression lines) report a measure of precision (such as the 95% confidence interval). Do NOT use the standard error of the mean (SE) to indicate the precision of an estimate. (The SE is essentially a 68% confidence coefficient: use the 95% confidence coefficient instead.)
- Although not preferred to confidence intervals, if desired, *P* values should be reported as equalities, when possible, and to one or two decimal places (e.g., $P = 0.03$ or 0.22 , not as inequalities: e.g., $P < 0.05$). Do NOT report ‘NS’; give the actual *P* value. The smallest *P* value that need be reported is $P < 0.001$, except in studies of genetic associations in which the values can be much, much smaller.
- Report whether and how any adjustments were made for multiple statistical comparisons.
- Name the statistical software package used in the analysis.

Reporting association analyses for categorical variables

- Describe the association of interest.
- Identify the variables used and summarize each with descriptive statistics.
- Identify the test of association used.
- Indicate whether the test was one or two-tailed. Justify the use of one-tailed tests.
- For **tests** of association (e.g., a *chi*-square test), report the *P* value of the test (because association is usually defined as a statistically significant result).
- For **measures** of association (e.g., the *phi* coefficient), report the value of the coefficient and a confidence

interval. Do not describe the association as low, moderate, or high unless the ranges for these categories have been defined. Even then, consider the wisdom of using these categories, given their biological implications or realities.

- For primary comparisons, consider including the full contingency table for the analysis.
- Name the statistical package or program used in the analysis.

Reporting correlation analyses for continuous variables

- Describe the purpose of the analysis.
- Summarize each variable with the appropriate descriptive statistics.
- Identify the correlation coefficient used in the analysis (e.g., Pearson, Spearman).
- Confirm that the assumptions of the analysis were met.
- Report the alpha level (e.g., 0.05) that indicates whether the correlation coefficient is statistically significant.
- Report the value of the correlation coefficient. Do not describe correlation as low, moderate, or high unless the ranges for these categories have been defined. Even then, consider the wisdom of using these categories given their biological implications or realities.
- For primary comparisons, report the (95%) confidence interval for the correlation coefficient, whether or not it is statistically significant.
- For primary comparisons, consider reporting the results as a scatter plot. The sample size, correlation coefficient (with its confidence interval), and *P* value can be included in the data field.
- Name the statistical package or program used in the analysis.

Reporting regression analyses

- Describe the purpose of the analysis.
- Identify the variables used in the analysis and summarize each with descriptive statistics.
- Confirm that the assumptions of the analysis were met. For example, in linear regression indicate whether an analysis of residuals confirmed the assumption of linearity.
- If relevant, report how any outlying values were treated in the analysis.
- Report how any missing data were treated in the analyses.
- For either simple or multiple (multivariable) linear regression analyses, report the regression equation.
- For multiple regression analyses: 1) report the alpha level used in the univariate analysis; 2) report whether the variables were assessed for a) collinearity and b) interaction, and 3) describe any variable selection process by which the final model was developed (e.g., forward-stepwise; best subset).
- Report the regression coefficients (beta weights) of each explanatory variable, and the associated confidence intervals and *P* values, preferably in a table.
- Provide a measure of the model’s “goodness-of-fit” to the data (e.g., the coefficient of determination, r^2 ,

for simple regression and the coefficient of multiple determination, R^2 , for multiple regression).

- Specify whether and how the model was validated.
- For primary comparisons analyzed with simple linear regression analysis, consider reporting the results graphically in a scatter plot showing the regression line and its confidence bounds. Do not extend the regression line (or the interpretation of the analysis) beyond the minimum and maximum values of the data.
- Name the statistical package or program used in the analysis.

Reporting analyses of variance (ANOVA) or of covariance (ANCOVA)

- Describe the purpose of the analysis.
- Identify the variables used in the analysis and summarize each with descriptive statistics.
- Confirm that the assumptions of the analysis were met. For example, indicate whether an analysis of residuals confirmed the assumptions of linearity.
- If relevant, report how any outlying data were treated in the analysis.
- Report how any missing data were treated in the analyses.
- Specify whether the explanatory variables were tested for interaction, and if so, how these interactions were treated.
- If appropriate, report, in a table, the P value for each explanatory variable, the test statistics and, where applicable, the degrees of freedom for the analysis.
- Provide an assessment of the goodness-of-fit of the model to the data, such as R^2 .
- Specify whether and how the model was validated.
- Name the statistical package or program used in the analysis.

Reporting survival (time-to-event) analyses

- Describe the purpose of the analysis.
- Identify the dates or events that mark the beginning and the end of the time period analysed.
- Specify the circumstances under which data were censored.
- Specify the statistical methods used to estimate the survival rate.
- Confirm that the assumptions of survival analysis were met.
- For each group, give the estimated survival probability at appropriate follow-up times, with confidence intervals, and the number of participants at risk of death at each time. It is often more helpful to plot the cumulative probability of not surviving, especially when events are not common.
- Reporting median survival times, with confidence intervals, is often useful to allow the results to be compared with those of other studies.
- Consider presenting the full results in a graph (e.g., a Kaplan-Meier plot) or life table.
- Specify the statistical methods used to compare two or more survival curves.

- Report the P value of the comparison when comparing two or more survival curves with hypothesis tests.
- Report the regression model used to assess the associations between the explanatory variables and survival or time-to-event.
- Report a measure of risk (e.g., a hazard ratio) for each explanatory variable with a confidence interval.

Reporting Bayesian analyses

- Specify the pre-trial probabilities (“priors”).
- Explain how the priors were selected.
- Describe the statistical model used.
- Describe the techniques used in the analysis.
- Identify the statistical software program used in the analysis.
- Summarize the posterior distribution with a measure of central tendency and a credibility interval.
- Assess the sensitivity of the analysis to different priors.

References

- 1 Esquirol J.E.D. [Passage from Lancet, 1838]. In: Pearl R. *Introduction to Medical Biometry and Statistics*. Philadelphia: W.B. Saunders, 1941.
- 2 Schor S. and Karten I. Statistical evaluation of medical journal manuscripts. *Journal of the American Medical Association* 1966; 195: 1123–1128.
- 3 Nagele P. Misuse of standard error of the mean (SEM) when reporting variability of a sample. A critical evaluation of four anaesthesia journals. *British Journal of Anaesthesia* 2003; 90: 514–516.
- 4 Neville J.A., Lang W. and Fleischer A.B. Jr. Errors in the Archives of Dermatology and the Journal of the American Academy of Dermatology from January through December 2003. *Archives of Dermatology* 2006; 142: 737–740.
- 5 Glantz S.A. Biostatistics: how to detect, correct and prevent errors in the medical literature. *Circulation* 1980; 61: 1–7.
- 6 Lionel N.D. and Herxheimer A. Assessing reports of therapeutic trials. *British Medical Journal* 1970; 3: 637–640.
- 7 Altman D.G. Statistics in medical journals: developments in the 1980s. *Statistics in Medicine* 1991; 10: 1897–1913.
- 8 White S.J. Statistical errors in papers in the British Journal of Psychiatry. *British Journal of Psychiatry* 1979; 135: 336–342.
- 9 Gore S.M., Jones I.G. and Rytter E.C. Misuse of statistical methods: critical assessment of articles in BMJ from January to March 1976. *British Medical Journal* 1977; 1: 85–87.
- 10 Scales C.D. Jr, Norris R.D., Peterson B.L. et al. Clinical research and statistical methods in the urology literature. *Journal of Urology* 2005; 174: 1374–1379.
- 11 Kurichi J.E. and Sonnad S.S. Statistical methods in the surgical literature. *Journal of the American College of Surgery* 2006; 202: 476–484.
- 12 Gardner M.J., Altman D.G., Jones D.R., et al. Is the statistical assessment of papers submitted to the British Medical Journal effective? *British Medical Journal* 1983; 286: 1485–1488.
- 13 Bakker M. and Wicherts J.M. The (mis)reporting of statistical results in psychology journals. *Behavior Research* 2011; 43: 666–678.
- 14 Avram M.J., Shanks C.A., Dykes M.H. et al. Statistical methods in anesthesia articles: an evaluation of two American journals during two six-month periods. *Anesthesia and Analgesia* 1985; 64: 607–611.
- 15 Godfrey K. Comparing the means of several groups. *New England Journal of Medicine* 1985; 313: 1450–1456.

- 16 Kim J.S., Kim D.K. and Hong S.J. Assessment of errors and misused statistics in dental research. *International Dental Journal* 2011; 61: 163–167.
- 17 Pocock S.J., Hughes M.D. and Lee R.J. Statistical problems in the reporting of clinical trials. A survey of three medical journals. *New England Journal of Medicine* 1987; 317(7): 426–432.
- 18 Murray G.D. Statistical aspects of research methodology. *British Journal of Surgery* 1991; 78: 777–781.
- 19 Yancy J.M. Ten rules for reading clinical research reports [Editorial]. *American Journal of Surgery* 1990; 159: 553–559.
- 20 Burton A. and Altman D.G. Missing covariate data within cancer prognostic studies: a review of current reporting and proposed guidelines. *British Journal of Cancer* 2004; 91: 4–8.
- 21 Mackinnon A. The use and reporting of multiple imputations in medical research – a review. *Journal of Internal Medicine* 2010; 268: 586–593.
- 22 Schwarzer G., Vach W. and Schumacher M. On the misuses of artificial neural networks for prognostic and diagnostic classification in oncology. *Statistics in Medicine* 2000; 19: 541–561.
- 23 George S.L. Statistics in medical journals: a survey of current policies and proposals for editors. *Medical and Pediatric Oncology* 1985; 13: 109–112.
- 24 Colditz G.A. and Emerson J.D. The statistical content of published medical research: some implications for biomedical education. *Medical Education* 1985; 19(3): 248–255. DOI: <http://dx.doi.org/10.1111/j.1365-2923.1985.tb01315.x>.
- 25 Golden J., Zhu W. and Sayre J.W. A review of the statistical analysis used in papers published in Clinical Radiology and British Journal of Radiology. *Clinical Radiology* 1996; 51(1): 47–50.
- 26 Lee C.M., Soin H.K. and Einarson T.R. Statistics in the Pharmacy Literature. *Annals of Pharmacotherapy* 2004; 38(9): 1412–1418. DOI: <http://dx.doi.org/10.1345/aph.1D493>.
- 27 Lang T. and Secic M. *How to Report Statistics in Medicine: Annotated Guidelines for Authors, Editors, and Reviewers*, Second edition. Philadelphia: American College of Physicians, 2006.
- 28 O’Fallon J.R., Duby S.D., Salsburg D.S. et al. Should there be statistical guidelines for medical research papers? *Biometrics* 1978; 34: 687–695.
- 29 Shott S. Statistics in veterinary research. *Journal of the American Veterinary Medical Association* 1985; 187: 138–141.
- 30 Hayden G.F. Biostatistical trends in Pediatrics: implications for the future. *Pediatrics* 1983; 72: 84–87.
- 31 Altman D.G. and Bland J.M. Improving doctors’ understanding of statistics. *Journal of the Royal Statistical Society, series A* 1991; 154: 223–267.
- 32 Altman D.G., Gore S.M., Gardner M.J., et al. Statistical guidelines for contributors to medical journals. *British Medical Journal* 1983; 286: 1489–1493.
- 33 Bailar J.C. 3rd and Mosteller F. Guidelines for statistical reporting in articles for medical journals. Amplifications and explanations. *Annals of Internal Medicine* 1988; 108(2): 266–273.
- 34 Bond G.R., Mintz J. and McHugo G.J. Statistical guidelines for the Archives of PM&R. *Archives of Physical Medicine & Rehabilitation* 1995; 76: 784–787.
- 35 Wilkinson L. and the Task Force on Statistical Inference. Statistical methods in psychology journals. Guidelines and explanations. *American Psychologist* 1999; 54: 594–604.
- 36 Curran-Everett D. and Benos D.J. American Physiological Society. Guidelines for reporting statistics in journals published by the American Physiological Society. *American Journal of Physiology*

— *Endocrinology and Metabolism* 2004; 287: E189–91 (plus other journals).

- 37 Curran-Everett D. and Benos D.J. Guidelines for reporting statistics in journals published by the American Physiological Society: the sequel. *Advances in Physiology Education* 2007; 31: 295–298.
- 38 Moher D., Schulz K. and Altman D.G. for the CONSORT Group. CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *Annals of Internal Medicine* 2001; 134: 657–662.
- 39 Des Jarlais D.C., Lyles C., Crepaz N. et al. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. *American Journal of Public Health* 2004; 94(3): 361–366. PMID: 14998794
- 40 von Elm E., Altman D.G., Egger M. et al. The Strengthening of Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies. *Annals of Internal Medicine* 2007; 147(8): 573–577. PMID: 17938396
- 41 International Committee of Medical Journal Editors. Uniform requirements for manuscripts submitted to biomedical journals: writing and editing for biomedical publication, <http://www.icmje.org>. (2011, accessed 12 December 2012).
- 42 Schriger D.L., Arora S. and Altman D.G. The content of medical journal instructions for authors. *Annals of Emergency Medicine* 2006; 48: 743–749, 749.e1–4.
- 43 Lang T. *How to Write, Publish, and Present in the Health Sciences: A Guide for Clinicians and Laboratory Researchers*. Philadelphia: American College of Physicians, 2010.