Statistical analyses and methods in the published literature: The SAMPL guidelines

Thomas A. Lang[1] and **Douglas G. Altman**[2]

1 Tom Lang Communications and Training International, Kirkland, WA, USA

2 Centre for Statistics in Medicine, Oxford University, Oxford, UK

Correspondence to:

Tom Lang
10003 NE 115th Lane
Kirkland, WA 98933
USA
tomlangcom@aol.com
+1 425 242 1370

Abstract

Despite calls for guidelines on reporting statistical aspects of studies, most journals have still not included in their instructions for authors more than a paragraph or two about reporting statistical methods and results. However, given that many statistical errors concern basic statistics, a comprehensive – and comprehensible – set of reporting guidelines might improve how statistical analyses are documented. The SAMPL guidelines are designed to be included in a journal's 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, they are nevertheless sufficient to prevent most of the reporting deficiencies routinely found in scientific articles.

This article is an update of an article originally published in the *Science Editors' Handbook* (Lang T, Altman D. Basic statistical reporting for articles published in clinical medical journals: the SAMPL Guidelines. In: Smart P, Maisonneuve H, Polderman A (eds). Science Editors' Handbook. European Association of Science Editors; 2013). References documenting the incidence of statistical errors have been updated in this revision.

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*, 1838[1]

Introduction

The first major study of the quality of statistical reporting in the biomedical literature was published in 1966.[2] Since then, dozens of similar studies have been published, every one of which has found that large proportions of articles contain errors in the application, analysis, interpretation, or reporting of statistics or in the design or conduct of research. (See, for example, references 3 through 19.) Further, large proportions of these errors are serious enough to call the authors' conclusions into

question.[5,18,19] The problem is made worse by the fact that most of these studies are of the world's leading peer-reviewed general medical and specialty journals.

Although errors have been reported for more complex statistical procedures,[19-22] paradoxically, many errors are in basic, not advanced, statistical methods.[23] Perhaps advanced methods are suggested by consulting statisticians, who perform the analyses competently, but it is also true that authors are far more likely to use only elementary statistical methods, if they use any at all.[23-26] Still, articles with even major errors continue to pass editorial and peer review and to 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,[29-32] most journals have still not included in their Instructions for Authors more than a paragraph or two about reporting statistical methods and results.[33] However, given that many statistical errors concern basic statistics, a comprehensive – and comprehensible – set of reporting guidelines might improve how statistical analyses are documented.

The SAMPL guidelines are designed to be included in a journal's 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, they are nevertheless sufficient to prevent most of the reporting deficiencies routinely found in scientific articles.

Unlike most of the other guidelines in this book, the SAMPL guidelines were not developed by a formal consensus-building process, but they do draw considerably from published guidelines.[27,34-37] In addition, a comprehensive review of the literature on statistical reporting errors 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, we do not address these issues here. Instead, we refer readers to the EQUATOR Network website (www.equator-network.org) 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, not duplicate, those in the methodology guidelines.

We welcome feedback and anticipate the need to update this guidance in due course.

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.

Guiding principles for reporting statistical methods and results

Our first guiding principle for statistical reporting comes from The International Committee of Medical Journal Editors, 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. [Emphasis added.] 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]

Our second guiding principle for statistical reporting is to **provide enough detail that the results can be incorporated into other 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 with 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 each analysis, rather than just listing in one place all the statistical methods used.
- Verify that that data conformed to the 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 programme 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 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.
- Summarise data that are approximately normally distributed with means and standard deviations (SD). Use the form: mean (SD), not mean ± SD.
- Summarise 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 – it is about a 68% confidence interval – not a descriptive statistic.)
- 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 with each rate applies.
- Identify any unit of population (that is, the unit multiplier: e.g., x 100; x 10,000) associated with the rate.
- Consider reporting 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 (justify the use of one-tailed tests) and for paired or independent samples.
- 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, such as 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 $P < 0.001$, save in studies of genetic associations.

- Report whether and how any adjustments were made for multiple statistical comparisons.
- Name the statistical software package used in the analysis.

Reporting association analyses

- Describe the association of interest.
- Identify the variables used and summarise 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 defined as a statistically significant result).
- For *measures* of association (i.e., 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

- Describe the purpose of the analysis.
- Summarise 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 programme 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 assumptions 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) 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. co-linearity and b. interaction; and 3. describe the 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 (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 analysed 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 programme 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, in a table, report 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 programme 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 for 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 table.
- Specify the statistical methods used to compare two or more survival curves.
- When comparing two or more survival curves with hypothesis tests, report the P value of the comparison
- 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.
- Summarise 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 JED. Cited in: Pearl, R. Introduction to Medical Biometry and Statistics. WB Saunders, Philadelphia; 1941.
2. Schor S, Karten I. Statistical evaluation of medical journal manuscripts. J Am Med Assn. 1966;195:1123-8.
3. Prescott RJ, Civil I. Lies, damn lies and statistics: errors and omission in papers submitted to Injury 2010-2012. Injury. 2013;44:6-11.
4. Fernandes-Taylor S, Hyun JK, Reeder RN, Harris AH. Common statistical and research design problems in manuscripts submitted to high-impact medical journals. BMC Res Notes. 2011;4:304.
5. Bosker T, Mudge JF, Munkittrick KR. Statistical reporting deficiencies in environmental toxicology. Environ Tox Chem. 2013;32(8):1737-9.
6. Vesterinen HM, Egan K, Deister A, Schlattmann P, Macleod MR, Dirnagl U. Systematic survey of the design, statistical analysis, and reporting of studies published in the 2008 volume of the Journal of Cerebral Blood Flow and Metabolism. J Cerebr Blood Flow Metab. 2011;31:1064-72.
7. Kim JS, Kim DK, Hong SJ. Assessment of errors and misused statistics in dental research. Int Dent J. 2011;61:163-7.
8. Lee HJ, Jung SK. The use of statistical methodology in articles in medical journals and suggestions for the quality improvement of the Pediatric Allergy and Respiratory Disease. Pediatr Allergy Respir Dis. 2011;21:144-55.
9. Yim KH, Nahm FS. Han KA, Park SY Analysis of statistical methods and errors in the articles published in the Korean Journal of Pain. Kor J Pain. 2010;23:35-41.
10. Al-Benna S, Al-Ajam Y, Way B, Steinstraesser L. Descriptive and inferential statistical methods used in burns research. Burns. 2010;36(3):343-6.
11. Robinson PM, Menakuru S, Reed MW, Balasubramanian SP. Description and reporting of surgical data – scope for improvement? Surgeon. 2009;7(1);6-9.
12. Afshar K, Jafari S, Seth A, Lee JK, MacNeily AE. Publications by the American Academy of Pediatrics Section on Urology: the quality of research design and statistical methodology. J Urol. 2009;182(4 Suppl):1906-10.
13. Barbosa FT, de Souza DA. Frequency of the adequate use of statistical tests of hypothesis in original articles published in the Revista Brasileira de Anestesiologia between January 2008 and December 2009. Rev Brasil Anestesiol. 2010;60:528-36.
14. Neville JA, Lang W, Fleischer AB. Errors in the Archives of Dermatology and the Journal of the American Academy of Dermatology from January through December 2003. Arch Dermatol. 2006;142:737-40.
15. Kurichi JE, Sonnad SS. Statistical methods in the surgical literature. J A Coll Surg. 2006;202:476-84.
16. Scales CD, Norris RD, Preminger GM, Vieweg J, Peterson BL, Dahm P. Evaluating the evidence: statistical methods in randomized controlled trials in the urological literature. J Urol. 208;180 1463-7.
17. Gaskin CJ, Happell B. Power, effects, confidence, and significance: an investigation of statistical practices in nursing research. Int J Nurs Stud. 2014;51(5):795-806.
18. Jaykaran YP. Quality of reporting statistics in two Indian pharmacology journals. J Pharmacol Pharmacother. 2011;2(2):85-9.

19. Mikolajczyk RT, DiSilvestro A, Zhang J. Evaluation of logistic regression reporting in current obstetrics and gynecology literature. Obst Gynecol. 2008;111(2 Pt 1):413-9.
20. Burton A, Altman DG. Missing covariate data within cancer prognostic studies: a review of current reporting and proposed guidelines. Br J Cancer. 2004;91:4-8.
21. Mackinnon, A. The use and reporting of multiple imputation in medical research – a review. J Intern Med. 2010;268:586-93.
22. Abraira V, Muriel A, Emparanza JI, Pijoan JI, Royuela A, Plana MN. Cano A, Urreta I, Zamora J. Reporting quality of survival analyses in medical journals still needs improvement. A minimal requirements proposal. J Clin Epidemiol. 2013;66:1340-6.
23. Kim M. Statistical methods in Arthritis & Rheumatism – current trends. Arthr Rheum. 2006;54:3741-9.
24. Reed JF, Salen P, Bagher P. Methodological and statistical techniques: what do residents really need to know about statistics? J Med Syst. 2003;27(3):233-8.
25. Aljoudi AS. Study designs and statistical methods in the Journal of Family and Community Medicine: 1994-2010. J Fam Commun Med. 2013;20(1):8-11
26. Lee CM, Soin HK, Einarson TR. Statistics in the pharmacy literature. Ann Pharmacother. 2004;38(9): 1412–8
27. Lang T, Secic M. How to Report Statistics in Medicine: Annotated Guidelines for Authors, Editors, and Reviewers, 2nd edn. American College of Physicians, Philadelphia; 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;34:687–95.
29. Shott S. Statistics in veterinary research. J Am Vet Med Association 1985; 187:138–41.
30. Hayden GF. Biostatistical trends in Pediatrics: implications for the future. Pediatrics. 1983;72:84–7.
31. Altman DG, Bland JM. Improving doctors' understanding of statistics. J Royal Stat Soc SerA. 1991;154, 223–67.
32. Altman DG, Gore SM, Gardner MJ, Pocock SJ. Statistical guidelines for contributors to medical journals. BMJ. 1983;286, 1489–93.
33. Bailar JC, Mosteller F. Guidelines for statistical reporting in articles for medical journals. Amplifications and explanations. Ann Intern Med. 1988;108 (2), 266–73.
34. Bond GR. Mintz J, McHugo GJ. Statistical guidelines for the Archives of PM&R. Arch Phys Med Rehab. 1995;76:784–7.
35. Wilkinson L, Task Force on Statistical Inference. Statistical methods in psychology journals. Guidelines and explanations. Am Psych. 1999;54: 594–604.
36. Curran-Everett D, Benos DJ. American Physiological Society. Guidelines for reporting statistics in journals published by the American Physiological Society. Am J Physiol Endocrinol Metab. 2004;287: E189–91.
37. Curran-Everett D, Benos DJ. Guidelines for reporting statistics in journals published by the American Physiological Society: the sequel. Adv Physiol Ed. 2007;31:295–8.
38. Moher D, Schulz K, Altman DG, for the CONSORT Group. CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. Ann Intern Med. 2001;134:657–62.
39. DesJarlais DC, Lyles C, Crepaz N, Trend Group. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. Am J Publ Health. 2004;94(3):361–6.
40. von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies. Ann Inter Med. 2007;147 (8), 573–7.
41. International Committee of Medical Journal Editors. Uniform requirements for manuscripts submitted to biomedical journals: writing and editing for biomedical publication. 2011 [cited 12 Dec 2012]. Available from: http://www.icmje.org.
42. Schriger DL, Arora S, Altman, DG. The content of medical journal instructions for authors. Ann Emerg Med. 2006;48:743–9.
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.