International Journal of Nursing Studies 52 (2015) 5–9

Contents lists available at ScienceDirect

International Journal of Nursing Studies

journal homepage: www.elsevier.com/ijns

Basic statistical reporting for articles published in Biomedical Journals: The "Statistical Analyses and Methods in the Published Literature" or the SAMPL Guidelines☆

Thomas A. Lang [a], Douglas G. Altman [b]

[a] Tom Lang Communications and Training International, United States
[b] Centre for Statistics in Medicine, Oxford University, United Kingdom

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 Esquirol (1941).

Introduction

The first major study of the quality of statistical reporting in the biomedical literature was published in 1966 (Schor and Karten, 1966). 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, Altman, 1991; Avram et al., 1985; Bakker and Wicherts, 2011; Gardner et al., 1983; Glantz, 1980; Godfrey, 1985; Gore et al., 1977; Kurichi and Sonnad, 2006; Lionel and Herxheimer, 1970; Murray, 1991; Nagele, 2003; Neville et al., 2006; Pocock et al., 1987; Scales et al., 2005; White, 1979; Yancy, 1990). Further, large proportions of these errors are serious enough to call the authors' conclusions into question (Glantz, 1980; Murray, 1991; Yancy, 1990). 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 found in more complex statistical procedures (Burton and Altman, 2004; Mackinnon, 2010; Schwarzer et al., 2000), paradoxically, many errors are in basic, not advanced, statistical methods (George, 1985). Perhaps advanced methods are suggested by consulting statisticians, who then competently perform the analyses, but it is also true that authors are far more likely to use only elementary statistical methods, if they use any at all (Emerson and Colditz, 1985; George, 1985; Golden et al., 1996; Lee et al., 2004). 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 (Lang and Secic, 2006).

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" (O'Fallon et al., 1978). Despite the fact that this call has since been echoed by several others (Altman and Bland, 1991; Altman et al., 1983; Hayden, 1983; Murray, 1991; Pocock et al., 1987; Shott, 1985), most journals have still not included in their Instructions for Authors more than a paragraph or two about reporting statistical methods (Bailar and Mosteller, 1988). However, given that many statistical errors concern basic statistics, a comprehensive — and comprehensible — set of reporting guidelines might improve how statistical analyses are documented.

In light of the above, we present here a set of statistical reporting guidelines suitable for medical 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

☆ This paper was originally published in: Smart P, Maisonneuve H, Polderman A (eds). Science Editors' Handbook, European Association of Science Editors, 2013. Reproduced with kind permission as part of a series of Classic Methods papers. An introductory Commentary is available as http://dx.doi.org/10.1016/j.ijnurstu.2014.09.007.

http://dx.doi.org/10.1016/j.ijnurstu.2014.09.006
0020-7489/© 2014 Published by Elsevier Ltd.

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; they may also help to prevent some reporting errors by focusing attention on key points in the analyses.

Unlike many of other guidelines, the SAMPL guidelines were not developed by a formal consensus-building process, but they do draw considerably from published guidelines (Bond et al., 1995; Curran-Everett and Benos, 2004, 2007; Lang and Secic, 2006; Wilkinson and Task Force on Statistical Inference, 1999).

In addition, a comprehensive review of the literature on statistical reporting errors reveals near universal agreement on how to report the most common methods (Lang and Secic, 2006).

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 research. 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 the CONSORT (Moher et al., 2001), TREND (Des Jarlais et al., 2004), STROBE (von Elm et al., 2007)). 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.

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" (Bailar and Mosteller, 1988; International Committee of Medical Journal Editors, 2011).

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 analyzing 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 analyze them. In particular, specify that (1) skewed data were analyzed with non-parametric tests, (2) paired data were analyzed with paired tests, and (3) the underlying relationship analyzed 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 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 ± SD.
- Summarize data that are not normally distributed with medians and interpercentile ranges, ranges, or both. Report the upper and lower boundaries of interpercentile 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 (Lang, 2010; Schriger et al., 2006).

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., ×100; ×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 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 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.
- 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 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) colinearity 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 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, 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 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 analyzed.
- 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.
- Summarize the posterior distribution with a measure of central tendency and a credibility interval.
- Assess the sensitivity of the analysis to different priors.

References

Altman, D.G., 1991. Statistics in medical journals: developments in the 1980s. Stat. Med. 10, 1897–1913.
Altman, D.G., Bland, J.M., 1991. Improving doctors' understanding of statistics. J. R. Stat. Soc. A 154, 223–267.
Altman, D.G., Gore, S.M., Gardner, M.J., Pocock, S.J., 1983. Statistical guidelines for contributors to medical journals. BMJ 286, 1489–1493.

Avram, M.J., Shanks, C.A., Dykes, M.H., et al., 1985. Statistical methods in anesthesia articles: an evaluation of two American journals during two six-month periods. Anesth. Analg. 64, 607–611.

Bailar 3rd, J.C., Mosteller, F., 1988. Guidelines for statistical reporting in articles for medical journals. Amplifications and explanations. Ann. Intern. Med. 108 (2), 266–273.

Bakker, M., Wicherts, J.M., 2011. The (mis)reporting of statistical results in psychology journals. Behav. Res. 43, 666–678.

Bond, G.R., Mintz, J., McHugo, G.J., 1995. Statistical guidelines for the Archives of PM&R. Arch. Phys. Med. Rehabil. 76, 784–787.

Burton, A., Altman, D.G., 2004. Missing covariate data within cancer prognostic studies: a review of current reporting and proposed guidelines. Br. J. Cancer 91, 4–8.

Curran-Everett, D., Benos, D.J., 2004. American Physiological Society Guidelines for reporting statistics in journals published by the American Physiological Society. Am. J. Physiol. Endocrinol. Metab. 287, E189–E191 (plus other journals).

Curran-Everett, D., Benos, D.J., 2007. Guidelines for reporting statistics in journals published by the American Physiological Society: the sequel. Adv. Physiol. Educ. 31, 295–298.

Des Jarlais, D.C., Lyles, C., Crepaz, N., Trend Group, 2004. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. Am. J. Public Health 94 (3), 361–366, PMID: 14998794.

Emerson, J.D., Colditz, G.A., 1985. The statistical content of published medical research: some implications for biomedical education. Med. Educ. 19 (3), 248–255, http://dx.doi.org/10.1111/j.1365-2923.1985.tb01315.x.

Esquirol, J.E.D., 1941. Cited in: Pearl R. Introduction to Medical Biometry and Statistics. WB Saunders, Philadelphia.

Gardner, M.J., Altman, D.G., Jones, D.R., Machin, D., 1983. Is the statistical assessment of papers submitted to the British Medical Journal effective? BMJ 286, 1485–1488.

George, S.L., 1985. Statistics in medical journals: a survey of current policies and proposals for editors. Med. Pediatr. Oncol. 13, 109–112.

Glantz, S.A., 1980. Biostatistics: how to detect, correct and prevent errors in the medical literature. Circulation 61, 1–7.

Godfrey, K., 1985. Comparing the means of several groups. N. Engl. J. Med. 313, 1450–1456.

Golden, J., Zhu, W., Sayre, J.W., 1996. A review of the statistical analysis used in papers published in Clinical Radiology and British Journal of Radiology. Clin. Radiol. 51 (1), 47–50.

Gore, S.M., Jones, I.G., Rytter, E.C., 1977. Misuse of statistical methods: critical assessment of articles in BMJ from January to March 1976. BMJ 1, 85–87.

Hayden, G.F., 1983. Biostatistical trends in Pediatrics: implications for the future. Pediatrics 72, 84–87.

International Committee of Medical Journal Editors, 2011. Uniform requirements for manuscripts submitted to biomedical journals: writing and editing for biomedical publication www.icmje.org (accessed 12.12.12).

Kurichi, J.E., Sonnad, S.S., 2006. Statistical methods in the surgical literature. J. Am. Coll. Surg. 202, 476–484.

Lang, T., 2010. How to Write, Publish, and Present in the Health Sciences: A Guide for Clinicians and Laboratory Researchers. American College of Physicians, Philadelphia.

Lang, T., Secic, M., 2006. How to Report Statistics in Medicine: Annotated Guidelines for Authors, Editors, and Reviewers, 2nd edition. American College of Physicians, Philadelphia.

Lee, C.M., Soin, H.K., Einarson, T.R., 2004. Statistics in the pharmacy literature. Ann. Pharmacother. 38 (9), 1412–1418, http://dx.doi.org/10.1345/aph.1D493.

Lionel, N.D., Herxheimer, A., 1970. Assessing reports of therapeutic trials. BMJ 3, 637–640.

Mackinnon, A., 2010. The use and reporting of multiple imputation in medical research – a review. J. Intern. Med. 268, 586–593.

Moher, D., Schulz, K., Altman, D.G., 2001. for the CONSORT Group CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. Ann. Intern. Med. 134, 657–662.

Murray, G.D., 1991. Statistical aspects of research methodology. Br. J. Surg. 78, 777–781.

Nagele, P., 2003. Misuse of standard error of the mean (SEM) when reporting variability of a sample. A critical evaluation of four anaesthesia journals. Br. J. Anaesth. 90, 514–516.

Neville, J.A., Lang, W., Fleischer Jr., A.B., 2006. Errors in the Archives of Dermatology and the Journal of the American Academy of Dermatology from January through December 2003. Arch. Dermatol. 142, 737–740.

O'Fallon, J.R., Duby, S.D., Salsburg, D.S., et al., 1978. Should there be statistical guidelines for medical research papers? Biometrics 34, 687–695.

Pocock, S.J., Hughes, M.D., Lee, R.J., 1987. Statistical problems in the reporting of clinical trials. A survey of three medical journals. N. Engl. J. Med. 317 (August (7)), 426–432.

Scales Jr., C.D., Norris, R.D., Peterson, B.L., et al., 2005. Clinical research and statistical methods in the urology literature. J. Urol. 174, 1374–1419.

Schor, S., Karten, I., 1966. Statistical evaluation of medical journal manuscripts. JAMA 195, 1123–1128.

Schriger, D.L., Arora, S., Altman, D.G., 2006. The content of medical journal instructions for authors. Ann. Emerg. Med. 48, 743–749, 749.e1–749.e4.

Schwarzer, G., Vach, W., Schumacher, M., 2000. On the misuses of artificial neural networks for prognostic and diagnostic classification in oncology. Stat. Med. 19, 541–561.

Shott, S., 1985. Statistics in veterinary research. J. Am. Vet. Med. Assoc. 187, 138–141.

von Elm, E., Altman, D.G., Egger, M., Pocock, S.J., Gotzsche, P.C., Vandenbroucke, J.P., 2007. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies. Ann. Intern. Med. 147 (8), 573–577, PMID: 17938396.

White, S.J., 1979. Statistical errors in papers in the British Journal of Psychiatry. Br. J. Psychiatry 135, 336–342.

Wilkinson, L., Task Force on Statistical Inference, 1999. Statistical methods in psychology journals. Guidelines and explanations. Am. Psychol. 54, 594–604.

Yancy, J.M., 1990. Ten rules for reading clinical research reports. Am. J. Surg. 159, 553–559 (editorial).