Anesthesiology 2007; 107:193–4

Statistical Reporting for Current and Future Readers

THE communication of research findings through statistical reporting is a crucial element of the scientific writing process. The success of this aspect of manuscript preparation can directly affect the impact of a study because easily comprehensible results can make an article much more accessible and, consequently, much more read. Despite this fact, scientific articles are becoming harder to read.[1] This is the case because of both the emergence of subdisciplines within a field, each with their own unique technical jargon, and advancing statistical analyses that are increasingly not familiar to many readers. Thoughtful statistical reporting can partially address this issue by placing readers closer to the actual data, allowing them to focus more on the implications of the study and less on deciphering the methods.

There is not one right way to present a data analysis, and there are certainly many elegant ways of doing so. It is for that reason that ANESTHESIOLOGY presents only guiding principles to authors and refrains from mandating a strictly uniform approach. The current editorial examines some of the issues involved in optimal statistical reporting. The discussion that follows is organized by two of the most common features of inadequate reporting and is accompanied by suggestions for improving the clarity of presentation, with the goal of making our research easier to read and more easily incorporated into our knowledge base.

Insufficient Information for Replication

Perhaps the most common omission of reporting is not including enough information in the methods section to allow a reader to follow the conducted statistical methods. In the Uniform Requirements for Manuscripts Submitted to Biomedical Journals,* from which the instructions to ANESTHESIOLOGY authors is based, a guiding principle is proposed:

> Describe statistical methods with enough detail to enable a knowledgeable reader with access to the original data to verify the reported results.

This principle is usually achieved by creating a "Statistical Analyses" subsection in the context of the procedural methods. Although unique reporting issues exist for randomized controlled trials,† diagnostic studies,‡ and meta-analyses,[2] several pieces of information are crucial to replication of all studies and must be reported. This description involves more than a simple list of conducted statistical tests and includes which variables were analyzed by which procedures, any conducted *post hoc* comparisons, the nature of the hypothesis tests (*i.e.*, two-tailed tests are the standard unless the experimental design dictates otherwise), and the type I error level specified for the study. From this description, it should be obvious to the reader how many statistical tests were conducted and whether any attempts were made to control for multiple comparisons. The software used to conduct the analyses is also important to understanding what was conducted and should be reported. Statistical packages have different default settings and widely ranging capabilities, making this knowledge important to future studies.

Optimal statistical reporting includes all of these basic elements but also provides more information to aide in the interpretation of the study results. For example, knowledgeable readers will be able to differentiate between the names of parametric and nonparametric tests and will infer information about the underlying data structures; however, actually reporting the nature of the observed distributions, and whether they have met the assumptions necessary for parametric evaluation, can serve to facilitate the understanding of the analytical plan as well as guide future studies on the topic. This information can be provided by summary statistics on the distributions (*e.g.*, skewness, kurtosis), formally testing the form of the distributions where sample size permits, or simply commenting on the visual inspection or quality of the data.

When possible, a statistical power analysis should be reported. Although this is not a formal requirement of the Journal, conducting a power analysis before conducting a study is simply good practice.[3] Information regarding the available statistical power allows a reader to place any negative findings in context. Specifically, it allows a reader to differentiate between the nonexistence of an effect *versus* the lack of available power to detect it. The reporting of a statistical power analysis also allows a reader access to the authors' assumptions before conducting the study, including the effect sizes that were anticipated through the experimental manipulation(s). Simply stated, a power analysis gives a reader more tools with which to interpret the study's results.

Accepted for publication May 16, 2007. The author is not supported by, nor maintains any financial interest in, any commercial activity that may be associated with the topic of this article.

* International Committee of Medical Journal Editors. Uniform requirements for manuscripts submitted to biomedical journals: Writing and editing for biomedical publication, 2006. Available at: www.ICMJE.org. Accessed May 15, 2007

† CONSORT guidelines. Available at: http://www.consort-statement.org. Accessed May 15, 2007

‡ STARD guidelines. Available at: http://www.consort-statement.org/stard-statement.htm. Accessed May 15, 2007

Reporting a Solitary *P* Value

A common practice that frequently leads to suboptimal reporting of results is providing only a solitary *P* value. This occurs when a statistical result is described using a simple statement, and evidence for the statement is provided using only a *P* value. This practice demands a lot of the reader, for to properly interpret the meaning of the result, a reader must also be able to identify the conducted statistical test, the null hypothesis under evaluation, the sample size used for the comparison, and be assured that the involved procedures were conducted satisfactorily. Further, unless the statement is provided in context with other information, a reader will have no information regarding the size of the described effect, because *P* values are not valid measures of effect size (see Cohen[3] and Houle[4]). There are certainly instances where this practice is acceptable because sufficient information has been provided to the reader, but it usually should be avoided.

Optimal statistical reporting includes providing all of the information required for the proper interpretation of a result in as succinct a manner as possible. This could be done in several ways. In instances where the analytical plan is relatively straightforward, where there is a fixed sample size for each analysis (including missing data), and when the analytical plan has been well introduced, the results can be presented in the context of descriptive statements accompanied by descriptive statistics (*e.g.*, mean and SDs for parametric analyses; medians and range or interquartile range for nonparametric analyses) and/or effect size estimates (*e.g.*, mean or proportional differences with confidence intervals) *with* resulting *P* values. However, when this is not the case, which may be more often than is, a different strategy is optimal. In these cases, the presentation of the full statistical result along with descriptive information is suggested. For this reporting, the general form of

Test statistic (degrees of freedom) = value, P = value

provides the information required for proper interpretation in an elegant and succinct style. Examples of this form of reporting, which is actually a standard requirement for publishing in many social science journals,[5] is provided for several statistical tests in table 1. The strength of this style is that information regarding the conducted statistical test, the actual sample size, and the significance level can all be ascertained. Further, the degrees of freedom can be evaluated for issues related to the application of the test (*i.e.*, the extent of missing data) and for adjustments made as modifications to standard tests (*i.e.*, Greenhouse-Geisser correction). Finally, if desired, a standardized measure of effect size† can be appended to this style of reporting, adding a wealth of information for future meta-analyses.

Table 1. Examples of Full Statistical Reporting

Statistic	Example
Chi-square (2 × 2 table)	χ^2 (1, n = 80) = 0.10, P = 0.75
t Test	t (49) = 2.75, P = 0.01
Correlation	r (97) = 0.26, P = 0.01
Analysis of variance	$F_{3,42}$ = 1.0, P = 0.50

If available, exact *P* values are preferred.

Conclusion: Writing for Current and Future Readers

Although there is not one correct way to present a data analysis, the guiding principle of statistical reporting is to provide sufficient information for replication. This concept applies to the description of the statistical methods as well as to the reporting of results. However, to truly impact the field, we will need to craft our statistical reporting to both current and future readers, who will use this information to inform both their clinical practice and their research. This can be accomplished through thoughtful presentation of results and particularly by the consistent reporting of observed effect sizes (with confidence intervals). This type of information, which contains a wealth of information about the magnitude of the observed effects, can be assimilated far more easily than *P* values, which provide little information to future readers apart from statistical significance judgments. Successful implementation of these practices will likely yield articles that are easier to read and more likely to impact the field of anesthesiology.

Timothy T. Houle, Ph.D., Department of Anesthesiology, Wake Forest University School of Medicine, Statistical Editor, ANESTHESIOLOGY, Winston-Salem, North Carolina. thoule@wfubmc.edu

References

1. Knight J: Scientific literacy: Clear as mud. Nature 2003; 423:376–8
2. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, Moher D, Becker BJ, Sipe TA, Thacker SB: Meta-analysis of observational studies in epidemiology: A proposal for reporting. JAMA 2000; 283:2008–12
3. Cohen J: Statistical Power Analysis for the Behavioral Sciences, 2nd edition. Mahwah, New Jersey, Lawrence Erlbaum, 1988
4. Houle TT: Importance of effect sizes for the accumulation of knowledge. ANESTHESIOLOGY 2007; 103:415–7
5. American Psychological Association: Publication Manual of the American Psychological Association, 5th edition. Washington, DC, American Psychological Association, 2001