Paper 166-2009

Talking Past Each Other? How to Communicate with Medical Writers When Preparing Clinical Research Manuscripts for Journal Submission

Scott C. Thompson, Medtronic, Inc., Memphis, TN

Stephanie R. Thompson, Rochester Institute of Technology, Rochester, NY

ABSTRACT

Clinical research manuscripts are often a blend of qualitative prose and quantitative data. And, the people who help the author(s) prepare each section — introduction, methods and materials, results, and discussion — are typically from very different parts of the research organization and have different backgrounds. Speaking a common language makes the process easier. This paper will present some tips on how to improve communication between medical writers and SAS® programmers, such as biostatisticians. Presented from both points of view, what does and doesn't work will be discussed in a point, counter-point fashion.

INTRODUCTION

Is statistics too specialized in vocabulary for someone with only a cursory knowledge of the topic to comprehend? Clinically relevant statistical results are rendered meaningless when communicated ineffectively. In fact, this axiom holds true for decision making in all professions, not just health care.

The statistical results of clinical research must be communicated in a manner that allows clinicians to assess critically the quality and reliability of both the study design and any conclusions that might affect clinical practice. According to Lang and Secic (1997), however, since the 1980s, studies of the statistical quality of journal articles have consistently found high error rates in the "application, reporting, and interpretation of statistical information, in even the most respected medical journals." Despite this, as described by Cockburn (2006), statistics remain "the vehicle that describes and quantifies research outcomes," but statistical methodology has become increasingly complex, so many clinical researchers now rely on biostatisticians to design their studies and on medical writers to describe their results.

After all, it is exceedingly difficult to explain many statistical concepts in terms that are both technically accurate and easily understood by those with only a cursory knowledge of the topic. As a consequence, clinical research manuscripts are often a blend of qualitative prose and quantitative data. And, the people who help the author(s) prepare each section — introduction, methods and materials, results, and discussion — are typically from very different parts of the research organization and have different backgrounds. So the potential for miscommunication remains high. For example, what a medical writer might call a "study question," a biostatistician might call a "statistical hypothesis." Speaking a common language, therefore, makes the whole process easier.

In this paper, we present some tips on how to improve communication between medical writers and SAS® programmers, such as biostatisticians. Although study design, including the selection of statistical methods, is beyond the scope of this paper, how to effectively describe the various attributes of a study design is not (e.g., timeframe, selection criteria, group assignment, duration of follow-up, masking).

In the *Uniform Requirements for Manuscripts Submitted to Biomedical Journals: Writing and Editing for Biomedical Publication*, with regard to statistics, the International Committee of Medical Journal Editors (ICMJE, 2007) states the following:

> 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 the use of P values, which fails 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.

In this paper, we have used these requirements to provide the overall structure for how we present our tips on how to effectively communicate clinically relevant statistics:

- Original data
- Methods
- Findings
- Measurement error and uncertainty indicators
- Hypothesis testing
- References cited
- Terminology, abbreviations, acronyms, and symbols
- Computer software used

SAS programmers and medical writers alike should refer to these communication tips frequently when assisting authors with preparing clinical research manuscripts for journal submission. Table 1 summarizes these tips in checklist form.

TABLE 1. CHECKLIST OF COMMUNICATION TIPS FOR ASSISTING AUTHORS WITH THE PREPARATION OF CLINICAL RESEARCH MANUSCRIPTS FOR JOURNAL SUBMISSION

Does the manuscript...

ORIGINAL DATA

☐ Specify how the original data were collected and measured?
☐ Specify the formats in which the data were stored?
☐ Specify any quality control methods that were used to assure completeness and accuracy of data collection?

METHODS

☐ Describe the basic design of the study?
☐ Indicate whether the study was observational or experimental, retrospective or prospective, randomized or controlled?
☐ State the timeframe, in years, of the study and the duration of follow-up?
☐ Describe the patient population studied, the inclusion and exclusion criteria used to select the population, and how this population was assigned to study groups (treatment or control)?
☐ Describe the treatment under study and exactly how it was administered (treatment protocol)?
☐ Describe the study setting and the source of study participants?
☐ Specify the sample size, sampling technique, and, if the size was determined in combination with a power calculation, specify the details of the calculation?
☐ Indicate, if relevant, how the study participants, and in some instances the data collectors, were masked (blinded) from the study groupings (treatment or control)?

FINDINGS

☐ Report all numerical findings with the appropriate degree of precision?
☐ Provide the numerical findings not only as derivatives (e.g., percentages) but also as the absolute numbers from which the derivatives were calculated (numerator and denominator)?
☐ Specify the statistical methods used to analyze the numerical findings?
☐ Include variability where applicable (e.g., mean [standard deviation] or median [interquartile range])?
☐ Report all P values as exact numbers to 2 digits past the decimal point, regardless of significance, unless they are lower than 0.01, in which case they should be presented to 3 digits?
☐ Provide findings in logical sequence in the text, tables, and figures, giving the main or most important findings first?
☐ Summarize the numerical findings in tables?
☐ Identify the statistical measures of variations (e.g., standard deviation, standard error of the mean)?
☐ Use figures as an alternative to tables with many entries?
☐ Restrict the tables and figures to those needed to explain the hypothesis and assess its support?

MEASUREMENT ERROR AND UNCERTAINTY INDICATORS

☐ Provide measurement error and uncertainty indicators?

HYPOTHESIS TESTING

- ☐ Identify what was studied and the reasons for studying it?
- ☐ State the purpose of the study in the form of a testable question (hypothesis)?
- ☐ State the choice of statistic used to test the hypothesis, as well as the null hypothesis?
- ☐ State how significance was determined?
- ☐ Provide information that was part of the analyses, including P values and alpha values?

REFERENCES CITED

- ☐ Reference standard works when possible (with pages stated) for the design of the study and statistical methods?

TERMINOLOGY, ABBREVIATIONS, ACRONYMS, AND SYMBOLS

- ☐ Communicate statistical terminology, abbreviations, acronyms, and symbols in terms that are both technically accurate and easily understood by those with only a cursory knowledge of the topic?
- ☐ Spell out abbreviations and acronyms at first mention in the text, followed by the abbreviation or acronym in parentheses?
- ☐ Define statistical symbols at first mention in the text, followed by the symbol in parentheses?

COMPUTER SOFTWARE USED

- ☐ Specify the computer software used to perform the statistical analyses?
- ☐ When referencing SAS software, specify the version of the software in addition to the name?

ORIGINAL DATA

Reproducibility is the hallmark of good science Nancekivell (2004). And clinical research manuscripts are more than just reports of research; they are also a source of data for other researchers Lang and Secic (1997). In its *Instructions for Authors*, the Journal of the American Medical Association (JAMA, 2008) goes one step further than the ICMJE uniform requirements by requiring authors to agree to the following as a condition of publication: "If requested, authors should be prepared to provide the data or cooperate fully in obtaining and providing the data on which the manuscript is based for examination by the editors or their assignees." Accordingly, ensure that a clinical research manuscript specifies how original data were collected, measured, and the formats in which the data were stored. In addition, ensure that the manuscript specifies any quality control methods that were used to assure completeness and accuracy of data collection.

METHODS

With regard to statistical methods, ensure that a clinical research manuscript:

- Describes the basic design of the study. At a minimum, the manuscript should indicate whether the study was:
  - Observational or experimental
  - Retrospective or prospective
  - Randomized or controlled Lang and Secic (1997)
- States the timeframe, in years, of the study and the duration of follow-up
- Describes the patient population studied — the group of patients for which the findings are to be generalized, the inclusion and exclusion criteria used to select the population, and how this population was assigned to study groups (treatment or control)
- Describes the treatment under study and exactly how it was administered (treatment protocol)
- Describes the study setting and the source of study participants to assist clinicians with determining the applicability of the study to other circumstances
- Specifies the sample size, sampling technique, and, if the size was determined in combination with a power calculation, specifies the details of the calculation

- Indicates, if relevant, how the study participants, and in some instances the data collectors, sometimes called study observers, were masked (blinded) from the study groupings (treatment or control)

FINDINGS

As described by Vickers (2007), the results of testing hypotheses are all too often put in terms of "P values, hazard ratios, regression coefficients, concordance indices, and a whole slew of other numbers" that have little or no meaning to either the patient or the practicing physician. To be clinically relevant, statistical findings must be expressed in terms that mean something in clinical practice — treatments, recurrences, days of survival — and that can be used in decisions (e.g., absolute differences rather than ratios, averages [means] rather than medians). Stated differently, in a practice setting, physicians must speak in terms of health care, not statistics.

When data are summarized in a clinical research manuscript, report the numbers in a manner that has meaning to the manuscript reader (i.e., Can a distinction be made at the hundredths level or is tenths commonly used?). Ensure the manuscript communicates all numerical findings with the appropriate degree of precision. Numerical data should be rounded when communicated, not when analyzed. Information is lost when numbers are rounded, and this loss can affect the quality of the results. For example, if pain were measured on a Likert scale ranging from 1 to 10, reporting a mean of 7.345674 provides no additional value to the reader than 7.3. Using the single decimal provides some indication if the mean tends toward 7 or 8, but 6 decimals are not helpful since no distinction can be made at the 0.000001 scale.

Furthermore, provide the numerical findings not only as derivatives (e.g., percentages) but also as the absolute numbers from which the derivatives were calculated (numerator and denominator), and specify the statistical methods used to analyze them ICMJE (2007). On a pain scale ranging from 1 to 10, for example, patients who start out with scores of 10/10 who then report a 1/10 during follow-up demonstrate a greater degree of clinical benefit than patients who start out with scores of 3/10 who then report a 1/10 during follow-up. Include variability where applicable (e.g., mean [standard deviation] or median [interquartile range]). All P values should be reported as exact numbers to 2 digits past the decimal point, regardless of significance, unless they are lower than 0.01, in which case they should be presented to 3 digits. Express any P values lower than 0.001 as P<0.001. P values can never equal 0 or 1 JAMA (*Instructions for Table Creation*, 2008).

Findings should be provided in logical sequence in the text, tables, and figures, giving the main or most important findings first. Do not repeat in the clinical research manuscript text all the data in the tables ICMJE (2007). A "well-structured table" is the most efficient way to convey a large amount of data in a clinical research manuscript JAMA & Archives Journals (2007). Tables capture information concisely, and display it efficiently; they also provide information at any desired level of detail and precision. Including data in tables rather than text frequently makes it possible to reduce the length of the text. Identify statistical measures of variations, such as standard deviation and standard error of the mean. Use figures as an alternative to tables with many entries; do not duplicate data in figures and tables. Restrict the tables and figures to those needed to explain the hypothesis and assess its support ICMJE (2007).

The Output Delivery System (ODS) introduced in SAS 8 makes producing publication-ready tables virtually effortless for SAS programmers. The output of an analyses performed with SAS no longer has to be re-typed to meet the format requirements for the particular journal to which an associated clinical research manuscript is to be submitted. Besides requiring extra effort, re-typing data is always a potential source for errors. Numbers can be transposed or mistyped altogether. The ODS enables SAS programmers to define a style that meets the table format requirements for a particular journal. Once the style is defined, it can be applied to the ODS output of one table or many — ensuring that all meet the table format requirements on which the style is based, thereby eliminating the need for a medical writer to verify that these tables meet the journal's requirements. Creating a style from scratch or overriding a default style are both excellent ways to create these "well-structured tables."

MEASUREMENT ERROR AND UNCERTAINTY INDICATORS

Besides sample size, P values, and alpha values, when communicating statistical findings in a clinical research manuscript, provide other relevant statistical information, such as confidence intervals. Measurements of the characteristics of a sample are used to estimate the same characteristics of the population for which the findings are to be generalized. Because these measurements contain error, the uncertainty associated with an estimate can be communicated with a confidence interval. If a clinical research manuscript were to state, for example, that an 80% confidence interval on the mean showed positive results, the reader might view this finding differently than if it were to state a 99% confidence interval. The confidence intervals, stated as such, would be important to the manuscript reader in making a determination as to the clinical relevance of the study findings.

HYPOTHESIS TESTING

Ensure a clinical research manuscript identifies what was studied and the reasons for studying it. It needs to state the purpose of the study in the form of a testable question (hypothesis), preferably at the end of its introduction. Based on one of the authors' experience, for example, in clinical studies involving medical devices, the hypothesis typically takes the following structure: "This study examined the clinical relevance of using X for treating Y," where X equals the particular medical device and Y equals the particular medical condition under study.

State the choice of statistic used to test the hypothesis, as well as the null hypothesis. It is not sufficient to state that the results were significant without stating how significance was determined. Provide information that was part of the analyses, including P values and alpha values. Different studies require different tests, yet the purpose of this paper is not to provide guidance on proper statistical methodology. That said, ensure, at a minimum, that the tests performed are specified in the study design.

REFERENCES CITED

Include citations in a clinical research manuscript that support the statistical approach used in the study to provide the reader another way to evaluate the study and to look for commonality with previously reviewed and published literature. We have found the following works invaluable references with regard to communicating study design and statistical methods:

1. Lang Thomas A., Secic, Michelle. 1997. *How to Report Statistics in Medicine: Annotated Guidelines for Authors, Editors, and Reviewers*. Philadelphia, PA: American College of Physicians.
2. Journal of the American Medical Association. January 2, 2008. *Glossary of Methodologic Terms*. Chicago, IL: American Medical Association. <http://jama.ama-assn.org/content/vol295/issue1/images/data/103/DC4/JAMA_auinst_term.dtl> (June 22, 2008).
3. Journal of the American Medical Association. January 2, 2008. *Instructions for Table Creation*. Chicago, IL: American Medical Association. http://jama.ama-assn.org/misc/ tablecreationinst.pdf (June 22, 2008).

Although SAS offers several publications on the use of its software in clinical research, we have come to rely on the software's help section as well as the online help that SAS provides. These references have provided the answers to all the questions (e.g., syntax and method algorithms) that we have had when collaborating on clinical research manuscripts together. We did not see a need to reference the SAS help in this paper, however, just simply to point its use out.

TERMINOLOGY, ABBREVIATIONS, ACRONYMS, AND SYMBOLS

Statistical terminology, abbreviations, acronyms, and symbols are part of the standard vocabulary of the biostatistician; however, these expressions are typically not commonly used by many of the other professions involved in clinical research. Since a clinical research manuscript is targeted for a more general audience, the use of such jargon and symbols should be limited to the greatest extent possible. The goal is to communicate statistical concepts in terms that are both technically accurate and easily understood by those with only a cursory knowledge of the topic.

The following sections describe some commonly used terminology, abbreviations, acronyms, and symbols. These sections are by no means an exhaustive list; they are simply intended to highlight common usages and serve as a means for others to evaluate how they express the statistical concepts in their own manuscripts.

Terminology — Statistics has its own vocabulary, as does SAS. Use of the term "proc," for example, is second nature to a SAS programmer, but it does not specifically mean anything to a medical writer. Combine it with the term "univariate" and the communication divide just widens. That said, simply state the statistical method that was used rather than the SAS procedure. Conveying to the medical writer that an analysis of variance was performed and that a 95% confidence interval was calculated around the mean allows the writer to compare this methodology to the study design and to complete one quality check. Conversely, for clarity, avoid non-technical uses of statistical terms, such as "random," "normal," "significant," "correlations," and "sample" ICMJE (2007).

Abbreviations and Acronyms — Do not use abbreviations and acronyms in the title or abstract and limit their use in the text JAMA (2008). Spell out all abbreviations and acronyms at first mention in the text, followed by the abbreviation or acronym in parentheses. If you define an abbreviation or acronym, ensure that it is used later in the document. It serves no purpose to establish the abbreviated version if it will not be used later. Remember to remind the medical writer that SAS is no longer an acronym. Occasionally, it is still parenthetically referenced as such, which is not correct.

Symbols — Certain symbols make up much of the statistical vocabulary. The chi-square test, for example, is used in many applications to test a hypothesis. One rarely sees the name of the test but rather its symbol ($\chi 2$). Although this symbol is distinctive, it may not render in certain fonts or software, which could leave the manuscript reader wondering what test was performed. Another commonly used symbol is the letter "n," which represents the number of study participants. Simply stating "n" (or "N") can lead to confusion regarding whether the reference is to a sample, a population, or the N function in SAS. As with abbreviations and acronyms, when first referenced, use the name of the test or statistic followed by the symbol in parentheses.

COMPUTER SOFTWARE USED

A clinical research manuscript needs to specify the computer software used to perform the statistical analyses. The statistical methods chosen for a particular study will more than likely dictate the computer software used. According to Lang and Secic (1997), it is also important to specify the computer software used because not all statistical software uses the same algorithms or default options to compute the same statistics and, as a consequence, the results may vary from software package to package.

When referencing SAS software in a clinical research manuscript, specify the version of the software in addition to the name. SAS 9.1.3 Service Pack 4 leaves no doubt as to what software was used. This reference to SAS software could lead to questions about method. As the software has evolved, additional statistical methods, tests, and options have been added. Does the version used include the test you referenced? The manuscript reader can verify the test if you provide the version, since the software documentation is provided online.

CONCLUSIONS

Statistics should help clinicians interpret, assess, and apply the results of effectively designed clinical research, the outcomes of which influence health care. SAS programmers and medical writers alike can facilitate this process by establishing a common language for explaining statistical concepts in a manner that is both technically accurate and easily understood by those with only a cursory knowledge of the topic.

REFERENCES

Cockburn, David M. 2006. "How to Make Clinical Decisions from Statistics." *Clinical and Experimental Optometry* 89(3):176–183.

International Committee of Medical Journal Editors. 2007. *Uniform Requirements for Manuscripts Submitted to Biomedical Journals: Writing and Editing for Biomedical Publication*. <http://www.icmje.org/index.html> (June 21, 2008).

JAMA & Archives Journals. 2007. *AMA Manual of Style: A Guide for Authors and Editors*, 10th ed. New York, NY: Oxford University Press, USA.

Journal of the American Medical Association. January 2, 2008. *Instructions for Authors*. Chicago, IL: American Medical Association. <http://jama.ama-assn.org/misc/ifora.dtl> (June 21, 2008).

Journal of the American Medical Association. January 2, 2008. *Instructions for Table Creation*. Chicago, IL: American Medical Association. http://jama.ama-assn.org/misc/ tablecreationinst.pdf (June 22, 2008).

Lang Thomas A., Secic, Michelle. 1997. *How to Report Statistics in Medicine: Annotated Guidelines for Authors, Editors, and Reviewers*. Philadelphia, PA: American College of Physicians.

Nancekivell, Sharon. 2004. "Writing a Publishable Journal Article: A Perspective from the Other Side of the Desk." *Science Magazine* [online]. <http://sciencecareers.sciencemag.org/career_development/previous_issues/articles/3010/writing_a_publishable_journal_article_a_perspective_from_the_other_side_of_the_desk/(parent)/> (June 21, 2008).

Vickers, Andrew J. 2007. "Does Medical Research Help Us Make Better Clinical Decisions?" *Medscape Business of Medicine* [online]. http://www.medscape.com/viewarticle/ 564710_print (June 18, 2008).

ACKNOWLEDGMENTS

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Scott C. Thompson
Principal Technical Writer
Medtronic, Inc.
1800 Pyramid Place
Memphis, TN 38132
Work Phone: (901) 344-0891
E-mail: scott.thompson@medtronic.com

Stephanie R. Thompson
Rochester Institute of Technology
Nine Lomb Memorial Drive
Rochester, NY 14623-5603
Work Phone: (585) 475-7237
E-mail: stephanie.thompson@rit.edu