A practical guide

Statistics and the Role of the Statistician in Clinical Trial Research

Ramesh Amatya, PhD, Associate Director, Biostatistics, Kendle

Dr Ramesh Amatya

Statistics seems to be the single most difficult area for most pharmaceutical physicians to tackle. In response to a high level of demand, I am delighted to be able to run a series on medical statistics which has been specially written for us by the Biometrics team at Kendle, whom I would like to thank for their enthusiasm and commitment to this series. – Ed

The Merriam-Webster Online Dictionary defines statistics as [1] "A branch of mathematics dealing with the collection, analysis, interpretation, and presentation of masses of numerical data." Similarly, a statistician is someone who is an expert in statistics and who deals with methods used in the collection, analysis, interpretation, and presentation of data.

Since statistics has widespread applications in many areas such as health, business administration, manufacturing, education and sociology, universities around the world offer different kinds of statistics courses applicable to these specific areas. Various statistics books tailored to these fields are available in the marketplace. For example, biostatistics focuses the study of statistics in areas applicable to biology, whereas business statistics covers the study of statistics relating to business management and decision making.

Basically, all applications involving the use of statistical methods may be grouped as descriptive statistics (inductive statistics) or inferential statistics (deductive statistics). Descriptive statistics involves description or the summarization of information to make it more understandable and usable. Inferential statistics refers to making generalizations about some population on the basis of a sample drawn from that population.

Descriptive Statistics

Descriptive statistics is concerned with summarizing and describing a given set of data. For example, if a practicing physician is interested in summarizing basic data regarding the number of patients visiting his clinic, patients' age distribution, gender, race, height, weight, and illness, descriptive statistics will include tabulation of these data (continuous and categorical) in summary tables and/or histograms. Summary statistics for continuous variables (i.e., age, weight, and height) will include number of patients, mean, standard deviation, median, minimum,

and maximum whereas summarization of categorical variables (i.e., gender, race, and type of illness) will include number and percent of these categories. In summarizing data by substituting a very few measures for many numbers, it is critical to understand the limitations of each summarizing measure so that the results can be interpreted correctly.

Inferential Statistics

Inferential statistics is based on various assumptions that may include continuity of variables, linearity of relationships, uncertainty of outcomes (probability), and sample distribution. Since the function of inductive statistics is that of induction or inferring properties of a population on the basis of known sample results (i.e., partial information obtained from a subset of the data of interest), proper knowledge and application of the probability theory is essential. The inferential results obtained using statistical methods could be misleading if the assumptions made in the statistical methods are not met. Therefore, it is important to check the validity of assumptions before drawing any conclusions from the results.

Role of the Statistician in Clinical Trial Research

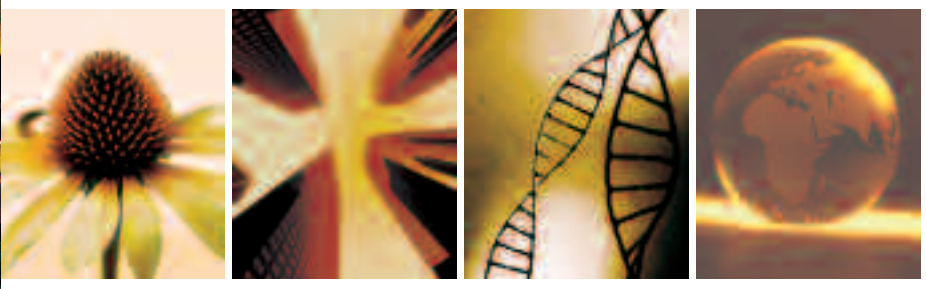
The goal of clinical trial research is to demonstrate the efficacy and safety of medicinal products as evidenced by clinical trials. The role of statistics in clinical trial design and analysis is acknowledged as essential in the guideline document "International Conference on Harmonisation of Technical Requirements: Registration of Pharmaceuticals for Human Use E6 (ICH E6, 5 February 1998) [2]." This guideline was developed by the appropriate International Conference on Harmonisation (ICH) Expert Working Group in consultation with the regulatory agencies, primarily to attempt to harmonise the principles of statistical methodology as applied to clinical trials for marketing applications submitted in Europe, Japan, and the United States.

The statistician plays a key role in the clinical trial research process from beginning to end. Some of the key functions of the statistician throughout the life of the trial are:

1. Protocol Development – The statistician is responsible for the design of the study, determination of randomization procedures, sample size calculation, and writing statistical analysis methods for the protocol. Close collaboration between the clinician and statistician at this stage is necessary, so that the hypothesis to be tested is clearly stated, the endpoints chosen to test the hypothesis have clinical relevance, and the number of subjects to be examined is adequate. Failure to meet the objectives of a study may not be attributable to a drug or device without an efficacious effect, but to an improperly planned study. Basically, the statistician assumes "co-authorship" over the protocol and signs off on it.

2. Review of Case Report Forms (CRFs) – Although CRFs are designed by the CRF specialist, the statistician is responsible for reviewing the CRFs to ensure that data captured in the CRFs is analyzable and adequate for addressing the objectives of the study.

3. Review of Data Management Plan (DMP) and Data Validation Plan (DVP) – The DMP and DVP are documents prepared by clinical data managers that outline data management and data validation procedures. The statistician is responsible for reviewing these documents to ensure that raw data is collected and processed efficiently, errors are checked, queries are generated, and data is corrected and prepared.



Descriptive statistics is concerned with summarizing and describing a given set of data.

4. Statistical Analysis Plan (SAP) Development – The statistician develops the Statistical Analysis Plan in collaboration with the clinical team (medical writer, clinician, data manager, and lead programmer). The SAP is the statistical document that includes text regarding the study design, objectives, randomization procedure, sample size justification, visit schedules and assessments, and statistical methods for the analyses as well as mock-up tables, figures, and listings for data presentation. This document follows or uses the study protocol, CRFs, and ICH Guidelines E9.

5. Development of Derived Data Set Specifications – The statistician develops derived data set specifications in collaboration with the project programmer. The purpose of these specifications is to outline the general considerations of datasets including variable name, label, type, code, and comments.

6. Review of Programming Specifications – The statistician reviews specifications for programming the tables, figures, and listings, which are prepared by the programmer. The purpose of this document is to describe what variables are to be used in the programming of individual tables, figures, and listings.

7. Preparation of the Quality Control (QC) Plan – It is important that the tables, figures, and listings generated are accurate. The QC plan outlines the detailed process and methods to be used by the independent statistician and/or independent programmer in validating the outputs generated by the programmer. The statistician prepares this QC document in collaboration with the programmer.

8. Programming and Quality Control of Analyses Output – The statistician works with the programmer to create and conduct a QC analyses of the tables, figures, and listings. This is an important step to ensure that all analyses outputs are properly programmed, validated, and documented.

9. Review of all Tables, Figures, and Listings – The statistician is responsible for reviewing all tables, figures, and listings for accuracy. It is the responsibility of the statistician to ensure that statistical methods are applied correctly as outlined in the SAP to generate these tables, figures, and listings. The statistician works closely with the programmer to ensure that all outputs generated are correct.

10. Writing Clinical Study Report – The statistician is responsible for writing statistical methods, writing the inferential results, and reviewing the clinical trial report. The report follows ICH Guideline E3 [3]. The statistician works with the medical writer to finalize the clinical trial report.

Conclusion

If used correctly, statistics can serve as a powerful and useful tool in the field of clinical research as well as other areas. Newly formulated and beneficial drugs or devices can be identified more efficiently and approved sooner, ultimately benefiting ailing patients.

In contrast, statistical methods applied incorrectly can lead to erroneous conclusions about drug safety and efficacy. The appropriate method for the given problem should be selected carefully and the validity of assumptions underlying these methods should be verified to ensure legitimate conclusions. Thus, consultation with the statistician throughout the duration of the clinical study can produce more accurate and useful results.

Ramesh Amatya, PhD is the Associate Director of Biostatistics at Kendle. He has more than 26 years of experience in biostatistics.

Editor's Note: This is the first in a series of nine articles on basic statistics and clinical research. Up next: Increasing Statistical Comprehension with Two Useful Methods: Confidence Intervals and the t-Test Procedure

REFERENCES

1. Statistics. "Merriam-Webster Online Dictionary, 2005. http://www.merriam-webster.com (8 Sept. 2005)
2. International Conference on Harmonization ICH E9: Statistical principles for clinical trials (February 1998 (Europe), September 1998 (FDA))
3. International Conference on Harmonization ICH E3: Guideline on Structure and Content of Clinical Study Reports (November 1995 (Europe), July 1996 (FDA))