Editor's note:
Biostatistics is a growing topic with a continuous development of new techniques. With a computer and the aid of many websites, even the most sophisticated statistical analyses can be done. These technical revolutions mean that the boundary between the essential statistics and the more advanced statistical methods has been blurred.

The understanding of biostatistics is important to all thoracic surgeons, as most of them received some statistics lessons in their training. Nevertheless, I think that few surgeons sit down to read statistics books. What thoracic surgeons need is to take very small doses of biostatistics, absorbed in a few minutes.

Therefore, the Statistic Corner in the *Journal of Thoracic Disease* (*JTD*) should keep the emphasis on enabling the reader to confront which method applies and when. Thus, in the corner, we could write about the analyses of different types of outcomes variable, the linking analyses of study design, the measures of association and impact, and the general strategies for the statistical analysis.

Statistic Corner

The chicken-and-egg debate about statistics and research

Luca Bertolaccini[1], Andrea Viti[2], Alberto Terzi[1]

[1]Thoracic Surgery Unit, Sacro Cuore Research Hospital, Negrar, Verona, Italy; [2]Thoracic Surgery Unit, S. Croce e Carle Hospital, Cuneo, Italy
Correspondence to: Luca Bertolaccini, MD, PhD, FCCP. Thoracic Surgery Unit, Sacro Cuore Research Hospital, Via Don Angelo Sempreboni 5, 37024 Negrar, Verona, Italy. Email: luca.bertolaccini@gmail.com.

Submitted Jul 20, 2014. Accepted for publication Jul 26, 2014.
doi: 10.3978/j.issn.2072-1439.2014.08.26
View this article at: http://dx.doi.org/10.3978/j.issn.2072-1439.2014.08.26

Statistics as a science began in the second half of the XVII century with the aim to collect data in order to lay down laws as a rational foundation of decision-making. The word statistics derives from the Latin word, *Status*. In Hamlet, William Shakespeare first used the word *statist* with a political meaning (*Devised a new commission, wrote it fair: I once did hold it, as our statists do, a baseness to write fair and labour'd much/how to forget that learning, but, sir, now, it did me yeoman's service: wilt thou know. The effect of what I wrote?*). Nevertheless, it is only in the last century that a few statisticians were active in developing new methods of analysis, theories, and applications of statistics. Nowadays, many branches of surgeries are completely penetrated by statistics and decision-making is often based on statistical analyses and accompanies the life of thoracic surgeons.

The goal of statistical analysis is to gain a better understanding of measurements; however, the inappropriate use of statistics can be confusing. In the 1860, Benjamin Disraeli, British Prime Minister, said that *there are three types of lies: lies, damned lies, and statistics.* Personal and subjective "good" judgment are not fact, and do not constitute substantive evidence (1). Statistical analyses make possible the elaboration of complex data and provide a mathematical basis with which to draw conclusions.

Despite the wide use of statistics, thoracic surgeons should carefully guard against pitfalls that can produce misleading conclusion. As a matter of facts, Sir Douglas G. Altman affirmed that general standard of statistics in medical journals is poor (2). Truthfully, properly used statistical methods can reject a hypothesis, but the statistics alone can never establish that a hypothesis is certainly true. Among the statistical methods, tests of significance have a prominent position. A test of significance is a statistical procedure by which one determines whether collected data are consistent with a specific hypothesis under investigation. The correct interpretation of P values, ubiquitous in surgical literature, is of paramount importance. An understanding of the meanings of the null and alternative hypotheses is fundamental. The null hypothesis of a study states that no difference exists between the study groups; in a two-armed

randomized controlled trial, the null hypothesis is that there is no difference between arms for the endpoint under investigation. On the contrary, the alternative hypothesis is that a difference exists between arms. The P value represents the probability that the difference observed between studies arms could occurs only by chance. The magnitude of the P value depends, among other factors, on sample size. If the sample size is sufficiently large, even tiny differences between study groups will become statistically significant. The question is whether small differences are of clinical relevance or not. A significant P value not necessarily reflects a clinical relevant difference and a not significant P value might mask clinically important results (for instance a serum level of potassium of 4.2 mEq/L can be significantly lower that a 4.4 mEq/L level if a large sample size is used but its relevance in clinical practice is of no meaning). Therefore, the distinction between statistical significance and clinical relevance will become even more important (3). Thus, a procedure may be found to be not statistically significant because of inadequate sample size (3,4).

According to Doug Altman, the unperceived misuse of statistics could interest the patients, the resources, and the consequences of publishing misleading results (5).

The development in computing technologies and the great availability of statistical software packages joined to the lack of a control system to validate the competence of people who perform statistical analysis can explain this prevalent misuse of statistics (6). Basic knowledge about medical statistics is invaluable for critical assessment of scientific findings. The learning curve for appropriate interpretation of biostatistics is sharp and the process highly interactive (7). Although the errors in research methods are mainly authors' responsibility, a clear attitude taken by the editorial boards of medical journals is also required to minimize this problem in forthcoming years (4).

Unappropriated or wrong statistical analysis, words of great concern when we read them in reviewers' comments. Hence, the *Statistic Corner in the Journal of Thoracic Disease (JTD)* intends to launch a series of invited reviews about statistics in thoracic surgery research. Obviously, these articles will only scratch the surface of medical statistics. Nonetheless, we hope that will provide a stimulus to enhance the skills to interpret statistical analyses. We welcome ideas and suggestions, from readers as well as potential authors, regarding other topics within the field of medical statistics. I will coordinate these reviews and, therefore, please feel free to contact me (preferably by e-mail).

Cite this article as: Bertolaccini L, Viti A, Terzi A. The chicken-and-egg debate about statistics and research. J Thorac Dis 2014;6(9):1349-1350. doi: 10.3978/j.issn.2072-1439.2014.08.26

Acknowledgements

Disclosure: The authors declare no conflict of interest.

References

1. Hickey RJ, Allen IE. Surgeons General's reports on smoking and cancer: uses and misuses of statistics and of science. Public Health Rep 1983;98:410-1.
2. Altman DG. Statistics in medical journals: developments in the 1980s. Stat Med 1991;10:1897-913.
3. Guller U. Caveats in the interpretation of the surgical literature. Br J Surg 2008;95:541-6.
4. Lucena C, Lopez JM, Pulgar R, et al. Potential errors and misuse of statistics in studies on leakage in endodontics. Int Endod J 2013;46:323-31.
5. Altman DG. Statistics and ethics in medical research. Misuse of statistics is unethical. Br Med J 1980;281:1182-4.
6. Ludbrook J. Statistics in biomedical laboratory and clinical science: applications, issues and pitfalls. Med Princ Pract 2008;17:1-13.
7. Guller U, DeLong ER. Interpreting statistics in medical literature: a vade mecum for surgeons. J Am Coll Surg 2004;198:441-58.

