

The Internet Medical Journal

ISSN 1093-7935

[Community Forum](http://www.medjournal.com/forum/index.php) (<http://www.medjournal.com/forum/index.php>)

- [Articles](http://www.medjournal.com/forum/forumdisplay.php?f=8) (<http://www.medjournal.com/forum/forumdisplay.php?f=8>)

- - [Statistical analysis in clinical studies: An introduction to fundamentals for physicians](http://www.medjournal.com/forum/showthread.php?t=954) (<http://www.medjournal.com/forum/showthread.php?t=954>)

rjtheodorou

September 13th, 2004 08:04

Statistical analysis in clinical studies: An introduction to fundamentals for physicians

Several studies have shown that most physicians have scant familiarity with statistical analysis of the results of their research. The aim of this review is to acquaint the reader with selected aspects of statistical analysis important for the conduct of clinical trials and to dispel the fears concerning the use of statistics.

Statistics together with study design form the foundation of clinical epidemiology. Although there are many biostatistics textbooks they most commonly fail to convey an understanding of statistical analysis to the average physician. In this paper basic concepts and principles of statistics are highlighted in an effort to enable physicians to collaborate more effectively with biostatisticians and to critically evaluate the merits of published studies.

BIAS IN CLINICAL STUDIES

Observer, subject, or instrument bias may affect accuracy. Observer bias ensues when the observer has a preconceived opinion of the study results or knows the diagnosis a priori. A blinded to the results of surgery or pathology examination observer may reduce this type of bias. When the study subject distorts measurements, subject bias occurs. Blinding of the subject as well as unobtrusive measurements reduce bias. Instrument bias, which is decreased by instrument calibration, may occur if the machine drifts over time (Lijmer et al, 1999, Mower, 1999, Brealey and Scally, 2001, Brealey et al, 2002).

SENSITIVITY AND SPECIFICITY

Sensitivity and specificity represent the measures of accuracy of a certain diagnostic test. In fact, the measurements have to be sensitive in order to detect differences that are important to the research question, and specific enough to show only the feature of interest. Sensitivity describes how well a diagnostic test identifies those people with the disease, and it is defined as the proportion of subjects with the disease who have a positive test. By definition, sensitivity represents true positive/ true positive + false negative = $a/a + c$ results (see Table). Similarly, specificity describes how well a diagnostic test identifies those people without the disease, and it is defined as the proportion of subjects without the disease who have a negative test. Specificity represents the true negative/ true negative + false positive = $d/d + b$ results (see Table). In general, screening tests need to have high sensitivity, whereas diagnostic tests should be characterized by high specificity.

In addition, two more distinctive principles characterizing measurements and thus, study

methods are recognized. Reliability defined as a measure of the extent to which a measurement or observation is reproducible and validity, defined as the vicinity of a measurement or observation to the reality play an important role in the execution of an efficient research study.

PREDICTIVE VALUE OF A DIAGNOSTIC TEST

Despite the fact that a test may have high accuracy (sensitivity and specificity) it may yet perform poorly in the clinical setting (low positive predictive value) where the disease prevalence is low. Thus, the predictive value of a test is an important index of actual test performance. The positive predictive value of a test indicates the probability that a disease is actually present when the test is positive, and can be calculated as follows: Positive predictive value = true positive/ true positive + false positive = $a / a + b$ (see Table). The negative predictive value of a test indicates the probability that a disease is actually absent if the test is negative, and also can be calculated in the following formula that is, Negative predictive value = true negative/ true negative + false negative = $d / d + c$ (see Table).

		DISEASE	
		PRESENT	ABSENT
TEST RESULT	POSITIVE	a	b
	NEGATIVE	c	d

THE ROC CURVES

An efficient way to display the relationship between sensitivity and specificity and the cut-off point for positive and negative tests is with receiver operating characteristic (ROC) curves (Obuchowski, 2000, Wagner et al, 2002, Gur et al, 2003). The ROC curve is a plot of the sensitivity and the 1-specificity. Each point on the curve represents a different cut-off value for the test indicated. Each cut-off value results in a percentage of true positive (y-axis) and false positive (x-axis) ratios. The test that yields the greatest number of true positives with the smallest number of false positives, representing the ROC curve, which tends upwards and to the left, is preferred. A poor diagnostic test has a low ROC curve approaching the diagonal. Under the diagonal, true positives and false positives are equal at every cut-off points leading to an indifferent test.

PROBABILITY AND CONFIDENCE INTERVALS

It is commonplace that the probability principle is of utmost importance in statistics. A normal or gaussian distribution of values is a bell-shaped curve with its x-axis representing the measurement of frequency of measurements and the y-axis representing the relative number of repetitions with the individual x values. The area under a portion of the curve is the probability that the true value is at or greater than the value of x at the line. In the normal distribution the measurements, which occur with the greatest frequency occur at the center of the distribution and are known as the central tendency.

Confidence intervals express the variation around the mean of a measurement, or a frequency. If a series of identical studies were performed on different samples from the same populations and a 95% confidence interval for the difference between the sample means existed, then 95% of these confidence intervals would include the population difference between means. The

researcher may select the degree of confidence, with 95% being the most common choice just as 5% level of statistical significance is widely used.

THE P-VALUE

One of the most commonly used statistic terms is the null hypothesis (H_0), which states that there is no difference between study groups except the one that is attributable to random phenomena. The alternate hypothesis (H_a) is the statement that there is a difference that cannot be explained by chance. The alternate hypothesis is proved by the exclusion of H_0 . The p-value is the probability on the assumption that H_0 is true of obtaining a measurement equal to or more extreme than that actually observed. In the graph of the normal distribution the p-value is represented by the area under the curve at and above the observed value marked by the line on the x-axis.

The level of statistical significance, also called type I error or false-positive result is the probability of rejecting H_0 when H_0 is actually true. It has been arbitrarily set at 0.05 as the threshold for statistical significance to distinguish whether an observed change in a set of measurements or frequencies may have arisen by chance or it represented something other than random variation. A type II error or false-negative result is the probability of accepting H_0 as true when H_0 is actually false, and as such missing a clinically significant difference. It is set at 0.1-0.2 as acceptable by most researchers. Practically, small p-values mean p-values of 0.05, which represent moderate evidence against to strong evidence; and those less than 0.001 represent strong to very strong evidence.

CONCLUSIONS

The statistical analysis, which depends on the research question, determines what type of variables and how these will be measured. Specifically, the statistical analysis includes the choice of study design, calculation of sample size and power calculations, as well as analysis of the outcomes (Lijmer et al, 1999, Mower, 1999). It is obvious that an understanding of the fundamentals of design of clinical studies by clinicians is important for interpreting the published results of diagnostic tests, or before clinical decision-making.

REFERENCES

Brealey S, Scally A (2001) Bias in plain film reading performance studies. *Br J Radiol*. Vol. 74, 307-316.

Brealey S, Scally A, Thomas N (2002) Review article: methodological standards in radiographer plain film reading performance studies. *Br J Radiol*. Vol. 75, 107-113.

Gur D, Rockette H, Armfield D, et al (2003) Prevalence effect in a laboratory environment. *Radiology*. Vol. 228, 10-14.

Lijmer J, Mol B, Heisterkamp S, et al (1999) Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA*. Vol. 282, 1061-1066.

Mower W (1999) Evaluating bias and variability in diagnostic test reports. *Ann Emerg Med*. Vol. 33, 85-91.

Obuchowski N (2000) Sample size tables for receiver operating characteristic studies. AJR Am J Roentgenol. Vol. 175, 603-608.

Wagner R, Beiden S, Campbell G, et al (2002) Assessment of medical imaging and computer-assist systems: lessons from recent experience. Acad Radiol. Vol. 9, 1264-1277.

KEY POINTS

§ Screening tests need to have high sensitivity, whereas diagnostic tests should have high specificity.

§ Data presentation and interpretation of results is aided by the use of statistics.

§ Definition and analysis of basic terms including the sensitivity and specificity of a diagnostic test, the predictive value, the ROC curves, the probability and confidence intervals, and the p-value is important for explaining statistical concepts to physicians

TITLE:

STATISTICAL ANALYSIS IN CLINICAL STUDIES: AN INTRODUCTION TO FUNDAMENTALS FOR PHYSICIANS

AUTHORS:

S.J. Theodorou, MD (1,2), D.J. Theodorou, MD (1,2), Y. Kakitsubata, MD (3)

AFFILIATIONS:

1. Department of Radiology, School of Medicine, University of California, San Diego Medical Center, San Diego, CA, USA
2. Department of Radiology, Veterans Administration Medical Center, San Diego, CA, USA
3. Department of Radiology, Miyazaki Medical College, Miyazaki, Japan

CORRESPONDING AUTHOR:

Stavroula J. Theodorou, M.D.

13 Papadopoulos Street

Ioannina 45444

Greece

Tel/FAX: +30-2651-0-38003

Email: rjtheodorou@hotmail.com daphne_theodorou@hotmail.com

CITATION

Theodorou SJ, Theodorou DJ, Kakitsubata Y (2004, September 13). Statistical analysis in clinical studies: An introduction to fundamentals for physicians. Internet Medical Journal. Retrieved September 13, 2004 from <http://www.medjournal.com/forum/showthread.php?t=954>

All times are GMT. The time now is 21:46.