Determination of Cutoff Score for a Diagnostic Test

G Singh

Citation

G Singh. *Determination of Cutoff Score for a Diagnostic Test*. The Internet Journal of Laboratory Medicine. 2006 Volume 2 Number 1.

Abstract

The cutoff value for a new diagnostic test for classifying cases as positive or negative may be determined utilizing some statistical techniques also in addition to clinical and other criteria. Mean ± 2SD, ROC curve, discriminant score methods may prove to be useful statistical tools for such situation.

When a new diagnostic test is developed or when a diagnostic test is to be used in a clinical condition different from the one for which the test was developed, test's cutoff score may require re-determination. This determination or re-determination may usually be based on biological, clinical or demographic situations. Some statistical methods may also be used or may be used in addition to the clinical experiences, analytical and empirical evidences for finding more reliable and valid cutoff point for classifying cases as positive or negative. 95% confidence interval (CI) of mean i.e. Mean ± 2SD method, ROC curve, discriminant function analysis may prove to be helpful statistical tools for such situation.

MEAN ± 2SD

An easy, crude and commonly used method is the application of 95%CI of mean for choosing a cutoff. A sample of adequate size of diagnosed cases (known positive cases) suffering from particular disease may be chosen. Then the diagnostic test is administered and values observed for cases are recorded. Mean and standard deviation (SD) of test values are calculated. Now an interval obtained by subtracting 2 x SD from mean and by adding 2 x SD to mean (that is, $\mu \pm 2\sigma$) shows that the chance of a test value coming outside this interval will be less than 5%. The lower limit of this interval (i.e. mean - 2SD) may be considered as cutoff point. If a subject's test value comes less than this cutoff then may be considered negative (normal) and if value comes greater than or equal to cutoff value then considered positive (diseased). This method may carry a chance of declaring some false negative cases, which can lower its sensitivity. Alternatively, this method may be carried out on a sample of known negative cases. In this case the upper limit of its

95%CI (i.e. mean + 2SD) may be taken as cutoff value. If a subject's test value comes greater than this cutoff value then may be considered positive (diseased). This approach may carry a chance of declaring some false positive cases that can lower its specificity. Depending upon the seriousness of the loss that may incur on lower sensitivity or specificity, a suitable approach may be chosen.

For illustration, take an example of CA-125 a glycoprotein, which is commonly used tumor marker in ovarian carcinoma. In a small study[1] mean ± sd of CA-125 for healthy volunteers was reported 8.08 ± 3.26 U/ml. Assuming if it is a large sample and drawn randomly from a specific population then, 95%CI comes 1.56 to 14.60 (i.e. 8.08 – 2x3.26 to 8.08 + 2x3.26). This shows that the chance of the value of CA-125 in such situation coming outside the interval 1.56 to 14.60 is less than 5%. In simple words, there is a high chance that test value for such healthy subjects may come in between 1.56 to 14.60. Now the upper limit of this interval i.e. 14.60 may be taken as a cutoff score. To determine cutoff score of this marker for specific diseased group, mean and sd of randomly drawn large sample from such diseased population can be used to obtain 95% confidence interval whose lower limit may be chosen as cutoff value. A subject having higher value than this cutoff may be considered as positive (diseased).

This method may work as a preliminary exercise in determining cutoff score of a test under varying conditions. The drawback here is that upper limit found for negative (healthy) subjects may not coincide with the lower limit found for positive (diseased) cases. Sometimes there may be a gap between the two and sometimes they overlap.

ROC CURVE

Receiver operating characteristics (ROC) curve between sensitivity and 1-specificity is a useful method to evaluate the performance of a diagnostic test in classification of subjects into two categories (say) positive and negative.[2] ROC curve may be used to judge how well the test performs. If area under the curve is near 1 it has higher chance of correct classification and when it is near 0, higher chance of incorrectly classifying in opposite group. The value 0.5 shows the test is no better than just tossing a coin for classification into positive or negative.[3] Some statistical software (like SPSS ver.10 or onwards) may be used for ROC curve. For determination of cutoff value this method may also be used. A sample of adequate size may be taken with known positive and negative cases. Then administer the diagnostic test and note the values observed. For every observed value this method displays sensitivity and 1-specificity of the test. Now one may choose a particular observed value of the test as cutoff value, which corresponds to the desired sensitivity and specificity (or 1-specificity).

In a study[4] to determine the cutoff value of serum I-CaD determined by ELISA for differentiating between patients with or without a glioma, a receiver-operating characteristics analysis was done. Taking each observed value as cutoff, ROC gives sensitivity and specificity of the test. This study reported the following result;

Figure 1

Serum I-CaD	Sensitivity	Specificity
Test values observed		
-	-	-
-	-	-
40	91 %	82 %
-	-	-
45	91 %	84 %
-	-	-
50	88 %	85 %
-	-	-
-	-	-

It is to be seen here that when 50 is taken as cutoff sensitivity decreased though specificity increased slightly. So a serum I-CaD value 45 may be thought of optimally better cutoff for differentiating patients. However, in a situation where high sensitivity is more desirable than high specificity, one may look for cutoff value even smaller than 40 similarly, when high specificity is more important, cutoff value greater than 50 may be chosen for classification.

A high sensitivity results in low number of false negative cases while high specificity leads to low number of false positive cases. Therefore, depending on the situation, requirement and seriousness of loss due to misclassification optimal value of sensitivity and specificity is decided and the test value corresponding to this may be taken as cutoff score for classification.

DISCRIMINANT SCORE

Discriminant function analysis is a popular tool in solving classification problem. A function is generated from a sample of known positive and negative cases then, the function is used for new cases with observed diagnostic test values to classify them as positive or negative. This method gives for each case discriminant score and predicted group membership corresponding to the observed test value. More elaborately, let a sample of known negative (normal) and positive (suffering with the particular disease) cases be chosen. The diagnostic test (for which cutoff is to be determined) may now be administered and the test value observed may be recorded. Perform discriminant function analysis using some statistical software. Observed test values will be entered in one column and known group membership (positive or negative) in another column. Result of the analysis will display discriminant score and predicted group membership for each case. If the test is such that low value indicates normal (negative) and high value as disease (positive) then, sort the cases according to discriminant score in ascending order. When it is arranged in increasing order of discriminant score, locate the case (row) from where the predicted group membership changes. The test value corresponding to this discriminant score from where predicted group membership changes (from normal to disease) might be taken as cutoff score. A value coming greater than or equal to this cutoff will be treated as positive while less than cutoff value as negative. If a test is such that low value indicates disease (positive) and high value as normal (negative) then also this method can be used to find a cutoff. Discriminant score may prove to be a suitable method in such type of situation requiring determination of cutoff value of a diagnostic test. It is also to be mentioned here that this method may be used for determining cutoff scores of more than one diagnostic test administered at a time. That is, when classification is to be made based on scores of more than one diagnostic tests together.

Lastly, it is worth mentioning here that determination of cutoff value may not be considered as prerogative of a statistician. A value derived by analyzing data may

sometime be clinically unacceptable on certain grounds.

CORRESPONDENCE TO

Dr.Girish Singh 6, Ganga Bagh Colony, Lanka Varanasi – 221005 India Phone : 91-542-2367304 e-mail : drgirishsingh@yahoo.com

References

1. Chaube A., Tewari M., Singh U. and Shukla HS.CA 125: A potential tumor marker for gallbladder cancer. Journal of Surgical Oncology 2006; 93(8): 665-669.
2. Biao Z A semiparametric hypothesis testing procedure for the ROC curve area under a density ratio model. Computational statistics and data analysis 2006; 50 (7) 1855-1876.
3. Altman DG and Bland JM. Statistics Notes: Diagnostic tests 3: receiver operating characteristic plots BMJ 1994; 309: 188.
4. Ping-Pin Zheng, Wim C. Hop, Peter A.E. Sillevis Smitt, Martin J. van den Bent, Cees J.J. Avezaat, Theodorus M. Luider and Johan M. Kros. Low-Molecular Weight Caldesmon as a Potential Serum Marker for Glioma. Clinical Cancer Research 2005; Vol.11(June 15): 4388-4392.

Author Information

Girish Singh, Ph.D.

Biostatistics Unit, Department of Basic Principles, Institute of Medical Sciences, Banaras Hindu University