Sun Y, Gobikrishna A, Hadid A, Lucariello RJ, Chiaramida SA. Software for receiver operating characteristic (ROC) analysis and evaluation of sensitivity and specificity. JACC 29(2): 335A, presented at ACC 46th Ann. Meeting, Anaheim, CA, 1997.

Receiver Operating Characteristic (ROC)

ROC

Linear regression and analysis of variance are the two common statistical methods used in cardiovascular research as well as many other areas. Medical decisions, however, often involve a nonlinear process of thresholding or grading and thus require a different type of analysis approach. The purpose of this study is to develop a computer program for analyzing outcomes of clinical decisions in terms of sensitivity and specificity. The program is also capable of estimating the receiver operating characteristic (ROC) curves which can be used to assess performance in a multiple-grading study or to determine the optimal threshold for making a medical decision. In the case where a decision is based on thresholding, the outcome of the decision is a true positive, a true negative, a false positive, or a false negative. If sufficient data are available, sensitivity (or true positive rate) can be estimated by 1 - FN/AP, where FN is the number of false negatives and AP is the number of actual positives; specificity (or true negative rate) can be estimated by 1 - FP/AN, where FP is the number of false positives and AN is the number of actual negatives. A ROC curve is obtained by plotting sensitivity vs. (1 - specificity) as the threshold is systematically varied. The theoretically optimal operating point on the ROC plane is the upper-left corner that corresponds to 100% sensitivity and 100% specificity. However, real-world data are usually insufficient to produce a complete ROC curve. By assuming the underlying probability distribution, e.g. the normal distribution, a ROC curve can be estimated from a relatively small set of data. The program also provides the goodness-of-fit test on the assumption of probability distribution. The software has been developed for the Macintosh computers. The software receives a text file as input which can be generated by a standard spread-sheet program. A tutorial has also been developed to demonstrate the use of this software based on examples of diagnosis and management for cardiac patients. The software should be useful to perform sensitivity-specificity and ROC analyses in clinical studies involving threshold-based decisions.

Introduction

In the late 1940's, signal detection theory was developed by statisticians and electrical engineers to describe probabilistically the detection of signals degraded by noise.[1] Signal detection theory is applicable to many different disciplines such as clinical chemistry, psychology, information retrieval, weather forecasting, polygraph lie detection, medical imaging, and diagnostics.[2-4]

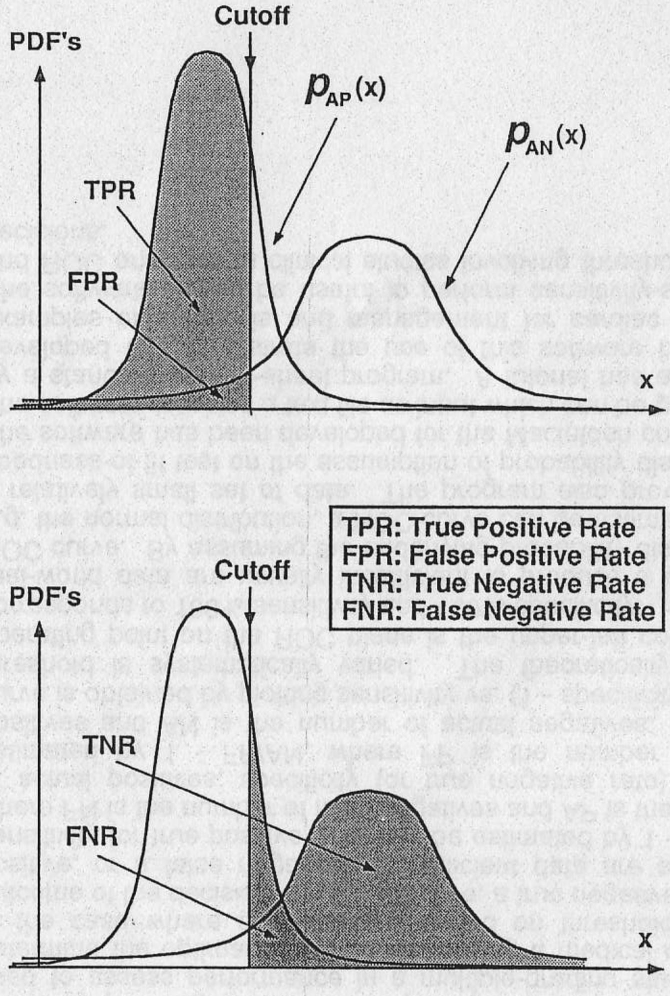
In the case of a binary decision, the null hypothesis H_0 is the hypothesis that the signal is absent (negative) and the alternative hypothesis H_1 is the hypothesis that the signal is present (positive). There are four possible outcomes for a binary decision as shown below:

		Signal	
		Present	Absent
Decision	Positive	TP	FP
	Negative	FN	TN

where
TP: true positive,
FP: false positive,
TN: true negative,
FN: false negative

[1]Egan JP. *Signal Detection Theory and ROC Analysis.* New York: Academic Press, 1975.
[2]Zweig JA, Campbell G. Receiver-operating characteristic ROC plots: a fundamental evaluation tool in clinical medicine. *Clinical Chemistry* 39: 561-577, 1993.
[3]Swets JA. Form of empirical ROCs in discrimination and diagnostic tasks: implications for theory and measurement of performance. *Psychological Bulletin* 99: 181-198, 1986.
[4]Turner DA. An intuitive approach to receiver operating characteristic curve analysis. *J Nuclear Med* 19: 213-220, 1978.

Probability Density Functions for Positive and Negative Samples



Terminology

Interdisciplinary equivalent terminology for outcomes

True positive	True negative	False positive	False negative
Correct rejection of H_0	Correct acceptance of H_0	Incorrect rejection of H_0 (Type I error)	Incorrect acceptance of H_0 (Type II error)
Hit	-	False alarm	Miss

ROC: receiver operating characteristic or relative operating characteristic

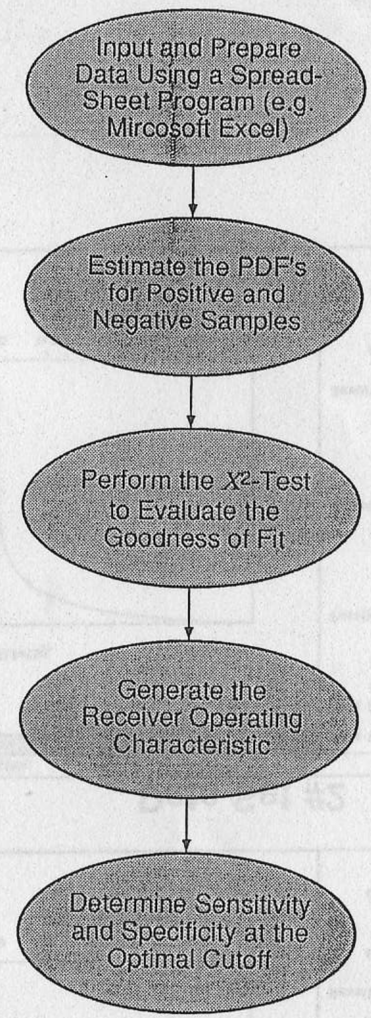
Interdisciplinary equivalent terminology for ROC plot

Ordinate	Abscissa
True positive rate	False positive rate
Sensitivity	1 - Specificity
Positivity	1 - Negativity
Hit rate	False alarm rate
True detection rate	False detection rate

Definitions

Actual Positive (AP)	Number of positive samples defined by some sort of gold standard
Actual Negative (AN)	Number of negative samples defined by the gold standard
True Positive (TP)	Number of positive samples that are correctly assigned as positive
True Negative (TN)	Number of negative samples that are correctly assigned as negative
False Positive (FP)	Number of negative samples that are incorrectly assigned as positive
False Negative (FN)	Number of positive samples that are incorrectly assigned as negative
False Positive Rate (FPR)	$FPR = \frac{FP}{AN} = \frac{FP}{TN + FP}$
False Negative Rate (FNR)	$FNR = \frac{FN}{AP} = \frac{FN}{TP + FN}$
Sensitivity	True positive rate (TPR) $TPR = 1 - FNR$
Specificity	True negative rate (TNR) $TNR = 1 - FPR$
Diagnostic Accuracy	1 - Total error rate = $1 - \sqrt{FNR^2 + FPR^2}$

Procedure for ROC Analysis



Example

We conducted a study to determine the sensitivity and specificity of detecting abnormal LV function based on ejection fraction determined by echocardiographic automated border detection (EF-ABD).

Patient group: 50 consecutive patients

Gold standard: ejection fraction determined by radioventriculography (EF-RVG)

Positive: defined by EF-RVG < 50%

Negative: defined by EF-RVG ≥ 50%

Echo view plane: apical 4-chamber view and short-axis view

Linear regression:

short-axis view r = 0.75

EF-ABD = 0.83 • EF-RVG + 21%

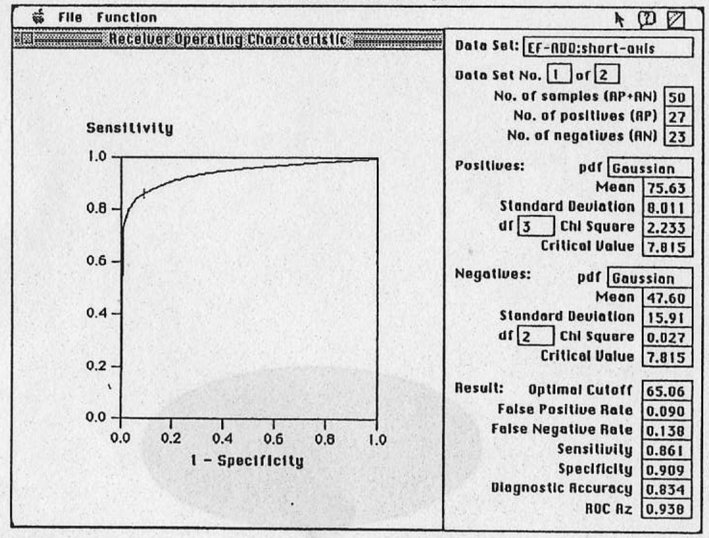
4-chamber view r = 0.79

EF-ABD = 0.84 • EF-RVG + 14%

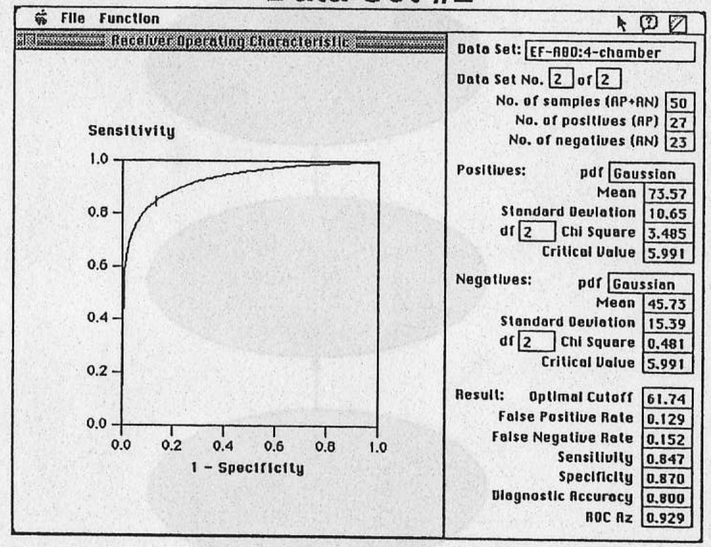
ROC Analysis:

	Short-axis	4-chamber
Optimal cutoff	65%	62%
Sensitivity	86%	85%
Specificity	91%	87%
Diag. accuracy	83%	80%

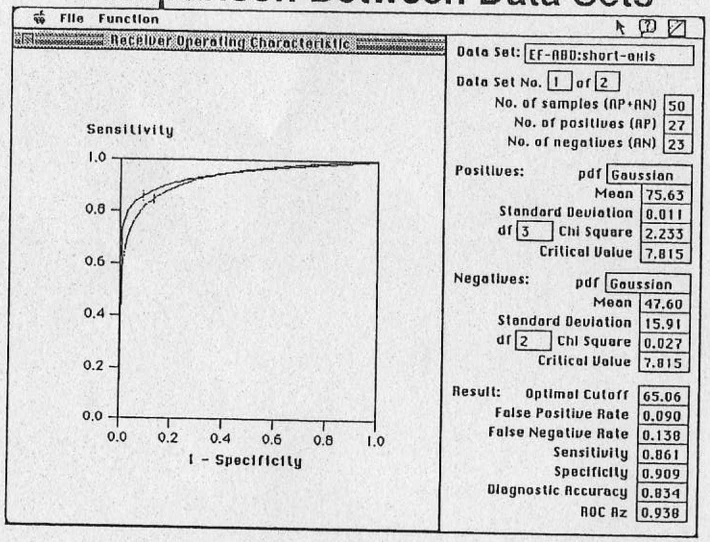
Data Set #1



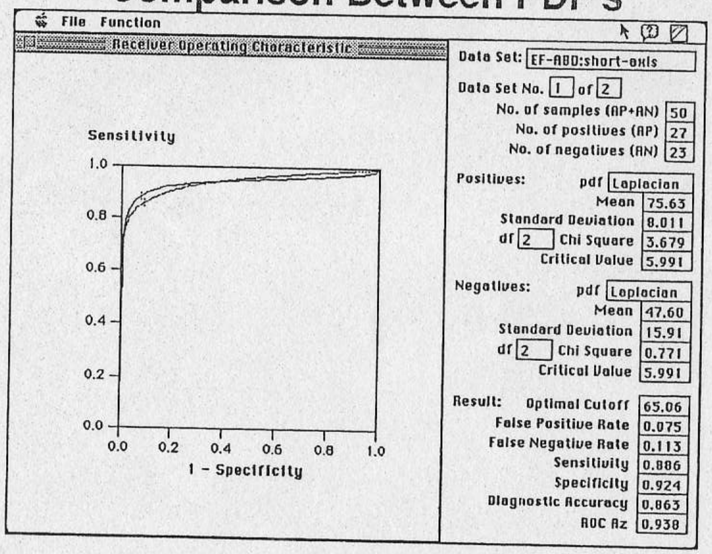
Data Set #2

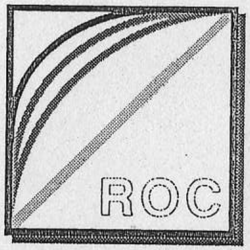


Comparison Between Data Sets



Comparison Between PDF's





Conclusions

A computer program for ROC analysis and determination of sensitivity and specificity was developed. The program should be useful for evaluating diagnostic methods and procedures in cardiology as well as other medical areas.

- **Assumption.** The ROC analysis implemented in this program is based on the assumption of the underlying probability density functions (pdf's) for the positive and negative samples. Currently available pdf's include the Guassian and the Laplacian distribution. The validity of the assumption is evaluated by performing a goodness-of-fit test based on the X^2 distribution.

- **Utility.** The software is useful for evaluating a diagnostic decision from a set of data of which the actual outcomes are known. The ROC analysis provides information about sensitivity and specificity, optimal cutoff point, and tradeoff between false positive and false negative.

- **Computer requirement.** The program is implemented in the C++ language for a Macintosh-type computer (Apple Computer, Inc.). The program should run on a PowerPC or a Macintosh (floating-point coprocessor is required for 68030 or lower processors).

- **Software availability.** The program is available as a shareware. For more information contact Ying Sun, Department of Electrical & Computer Engineering, University of Rhode Island (401-874-2515 or sun@ele.uri.edu).

- **Future directions.** We will continue the development of this program, especially in the direction of incorporating more pdf's. We also plan to make the software available via the World Wide Web.