



ROC Curves Analysis

Introduction

Receiver operating characteristic (ROC) curves are used in medicine to determine a cutoff value for a clinical test. For example, the cutoff value of 4.0 ng/ml was determined for the prostate specific antigen (PSA) test for prostate cancer. A test value below 4.0 is considered to be normal and above 4.0 to be abnormal. Clearly there will be patients with PSA values below 4.0 that are abnormal (false negative) and those above 4.0 that are normal (false positive). The goal of an ROC curve analysis is to determine the cutoff value.

Assume that there are two groups of men and by using a 'gold standard' technique one group is known to be normal (negative), not have prostate cancer, and the other is known to have prostate cancer (positive). A blood measurement of prostate-specific antigen is made in all men and used to test for the disease. The test will find some, but not all, abnormal to have the disease. The ratio of the abnormal found by the test to the total number of abnormal known to have the disease is the true positive rate (also known as sensitivity). The test will find some, but not all, normals to not have the disease. The ratio of the normals found by the test to the total number of normals (known from the 'gold standard' technique) is the true negative rate (also known as specificity). The hope is that the ROC curve analysis of the PSA test will find a cutoff value that will, in some way, minimize the number of false positives and false negatives. Minimizing the false positives and false negatives is the same as maximizing the sensitivity and specificity.

For the PSA test abnormal values are large (> 4) and normal values are small (< 4). This is not always the case, however, so the present program allows for both conditions of abnormal being larger and abnormal being smaller.

The ROC curve is a graph of sensitivity (y-axis) vs. $1 - \text{specificity}$ (x-axis). An example is shown in Figure 1. Maximizing sensitivity corresponds to some large y value on the ROC curve. Maximizing specificity corresponds to a small x value on the ROC curve. Thus a good first choice for a test cutoff value is that value which corresponds to a point on the ROC curve nearest to the upper left corner of the ROC graph. This is not always true however. For example, in some screening applications it is important not to miss detecting an abnormal therefore it is more important to maximize sensitivity (minimize false negatives) than to maximize specificity. In this case the optimal cutoff point on the ROC curve will move from the vicinity of the upper left corner over toward the upper right corner. In prostate cancer screening, however, because benign enlargement of the prostate can lead to abnormal (high) PSA values, false positives are common and undesirable (expensive biopsy, emotional impact). In this case maximizing specificity is important (moving toward the lower left corner of the ROC curve).

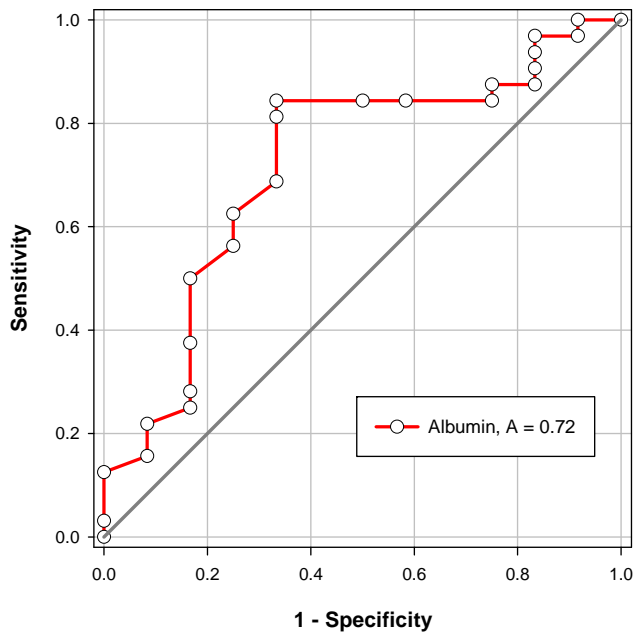


Figure 1: An example ROC curve.

An important measure of the accuracy of the clinical test is the area under the ROC curve. If this area is equal to 1.0 then the ROC curve consists of two straight lines, one vertical from 0,0 to 0,1 and the next horizontal from 0,1 to 1,1. This test is 100% accurate because both the sensitivity and specificity are 1.0 so there are no false positives and no false negatives. On the other hand a test that cannot discriminate between normal and abnormal corresponds to an ROC curve that is the diagonal line from 0,0 to 1,1. The ROC area for this line is 0.5. ROC curve areas are typically between 0.5 and 1.0 like shown in Figure 1.

Two or more tests can be compared by statistically comparing the ROC areas for each test. The tests may be correlated because they occurred from multiple measurements on the same individual. Or they may be uncorrelated because they resulted from measurements on different individuals. The ROC Curves Analysis Module refers to this as “Paired” and “Unpaired”, respectively, and can analyze either situation.

The test measurements may contain missing values and two methods are provided to handle missing values when comparing ROC areas – pairwise deletion and casewise deletion. This is described in detail later.

Given a value for the probability that the patient has the disease (pre-test probability) the probability that the patient has the disease, given the value of the test measurement, can be computed. Also, given a value for the false-positive/false-negative cost ratio (for the screening example above, the false-negative cost would be greater than the false-positive cost), an optimal test value cutoff can be computed. The present program allows entry of the pre-test probability and the false-positive/false-negative cost ratio.

Data Entry

Data can be entered in two formats in SigmaPlot – Indexed and Grouped.

Indexed Data Format

This is the format found in statistics programs such as SYSTAT and SigmaStat. “Indexed” is the terminology used in SigmaStat. It has one column that indexes another column (or other columns). It is also the format of the output of logistic regression where ROC curves are used to determine the ability of different logistic models to discriminate negative from positive test results (normals from abnormal). Each data set consists of a pair of columns – a classification variable and a test variable. The classification variable has a binary state that is either negative (normal) or positive (abnormal). Many programs use a value of 1 for positive and 0 for negative. The classification variable is required to be located in column 1 of the worksheet. The test variable is a continuous numeric variable and contains the test results. A single test variable will be located in column 2. Multiple test variables will be located in multiple columns starting in column 2. There is no built-in limit for the number of test variables. There is only one classification variable for multiple test variables and it is located in column 1. The test variable columns must be left justified and contiguous. Therefore no empty columns to the left of or within the data are allowed.

The following example shows a few rows of data for two data sets. The first column is the classification variable. It contains a column title “Thyroid Function” which is the classification variable name. It also contains the two classification states “Hypothyroid” and “Euthyroid” (normal thyroid function). Hypothyroid and Euthyroid are the abnormal and normal classification states, respectively. T4 and T5 are the names of different blood tests that will be used in the ROC analysis to discriminate between normal and abnormal and then compared to determine which is the better test. The classification variable must be in column 1 and the two test variables in the two columns adjacent to it

The classification variable name will be obtained from the column 1 column title if it exists. The test names will be obtained from the column titles of the test variable columns if they exist. The classification state names will be obtained from the entries in the cells of column 1. If no column titles have been entered for the test variables then default names for the tests, “Test 1”, “Test 2”, etc., will be used and displayed in the graphs and reports. The test variable names should be unique but the program will subscript any identical names that are not.

T4 & T5 Analysis - Indexed			
	1-Thyroid Function	2-T4	3-T5
1	Hypothyroid	0.90000	1.20822
2	Hypothyroid	1.00000	1.49901
3	Hypothyroid	1.90000	1.87303
4	Hypothyroid	2.00000	2.25433
5	Hypothyroid	2.10000	2.41396
6	Euthyroid	5.00000	5.84037
7	Euthyroid	5.60000	6.85851
8	Euthyroid	5.70000	7.03701
9	Euthyroid	5.80000	7.17566
10	Euthyroid	6.10000	7.19916
11	Euthyroid	6.20000	7.23872
12	Euthyroid	6.30000	7.57394

Figure 2: Indexed data format for two tests. The test names are T4 and T5, the classification states are Euthyroid and Hypothyroid and the Classification variable name is Thyroid Function. The index column is always column 1 and data columns must be left adjusted.

There must be two or more non-missing data points for each test for each classification state. Missing values are handled automatically by the analysis. For data columns, missing values are everything but numeric values (blank cells, the SigmaPlot double-dash missing value symbol, "+inf", "-inf", "NaN", etc.). Missing values are ignored for all computations except the Paired area comparison (see the Missing Value Method section) where they are handled using one of two possible algorithms.

Grouped Data Format

The grouped data format consists of pairs of data columns – one pair for each test. One column in a data pair consists of the negative (normal) data values and the other column for positive (abnormal) values. So, for example, if two tests are to be compared, the worksheet will contain four columns of data – the first two columns for the first test and the third and fourth column for the second test.

A specific column title format is used to identify the test associated with the data column pair and the classification states within each pair. The user is encouraged to use this format since it clearly identifies the data in the data worksheet and will annotate all the graphs and reports generated. It is not necessary to use column titles as the program will identify column pairs starting in column 1 with the generated test names "Test 1", "Test 2", etc., and will arbitrarily assign "1" and "0" classification state names to the first and second columns, respectively, but this is clearly not the best way to organize the data. Since the test names and classification states are numerical it is also more difficult to interpret the results.

Column Title Convention for Grouped Data

This column title convention is a simple way to identify worksheet data for the Grouped data format. The following example shows a few rows for two data sets. The first two columns contain the data for the T4

test. The first column “T4 - Euthyroid” is the column with the normal data for test T4. The column title consists of the test name followed by a minus sign followed by the classification state. Spaces on either side of the minus sign are ignored. The second column “T4 - Hypothyroid” is the column with the abnormal data for test T4. The third and fourth column titles are the same as the first two except the second test name T5 is used.

T4 & T5 Analysis				
	1-T4 - Euthyroid	2-T4 - Hypothyroid	3-T5 - Euthyroid	4-T5-Hypothyroid
1	5.00000	0.90000	5.84037	1.20822
2	5.60000	1.00000	6.85851	1.49901
3	5.70000	1.90000	7.03701	1.87303
4	5.80000	2.00000	7.17566	2.25433
5	6.10000	2.10000	7.19916	2.41396
6	6.20000	3.00000	7.23872	3.62332
7	6.30000	3.80000	7.57394	3.63255
8	6.60000	3.90000	7.70231	3.64476
9	6.65000	3.95000	7.71128	3.86069
10	6.70000	4.00000	7.98808	4.04146

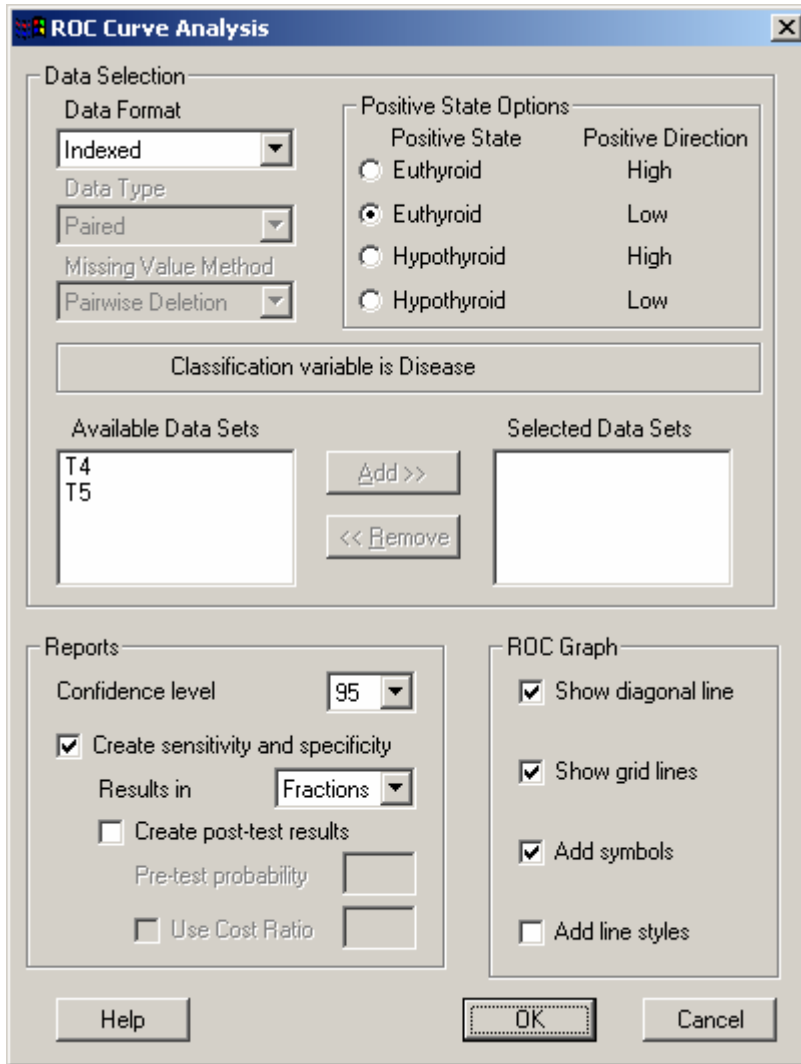
Figure 3. Grouped data format for two tests. This is the same data as in Figure 1. There are two tests T4 and T5. Each test consists of a pair of data columns. In this case T4 is in columns 1 and 2 and T5 in columns 3 and 4. The “Test-State” column title format is used to identify the two tests and the normal (Euthyroid) and abnormal (Hypothyroid) states.

The test names in both columns of a column pair must be the same. Also there must be exactly two classification states in the column titles.

Like the Indexed format, missing values in the worksheet cells are ignored except for special handling when comparing ROC areas (see the Missing Value Method section).

Program Options

Selecting ROC Curves from the SigmaPlot Toolbox menu opens the dialog



Test and classification state names from the indexed data shown in Figure 2 of the Data Entry section are displayed in this dialog.

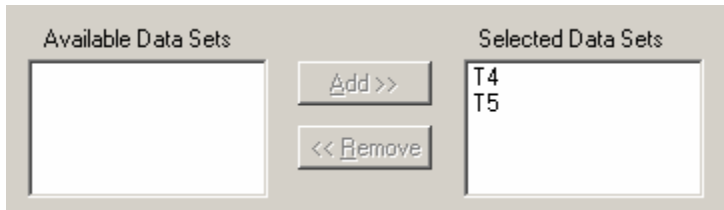
Data Selection Options

Data Format (Automatic Determination)

In most cases the program will identify the data format from the information in the data worksheet. In the dialog above the format was identified as Indexed. You may select from the two formats - Indexed and Grouped.

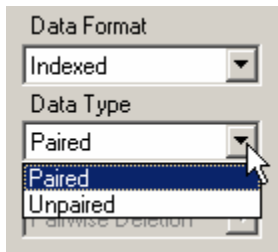
Available Data Sets – Selected Data Sets

Select one or more of the available data sets by clicking on them in the Available Data Sets window and then clicking on the Add button. If desired, you may then select a test name in the Selected Data Sets window and click Remove to deselect the test.



Data Type

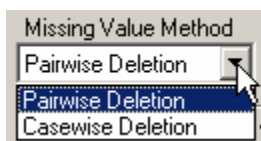
If two or more data sets are selected then the Data Type option for correlated tests is made available



You may select either Paired, for correlated tests, or Unpaired. If Paired is selected the ROC areas and area comparisons are determined using the DeLong, Delong and Clarke-Pearson method⁽²⁾. If Unpaired is selected the areas are computed using the Hanley and McNeil method⁽³⁾ and the areas are compared using a Z test.

Missing Value Method

If missing values exist then two options are available for the pairwise comparison of ROC areas – Pairwise Deletion and Casewise Deletion. This option is not available if no missing values exist.



Pairwise deletion only deletes rows containing missing values for the particular pair being analyzed – not for an entire row of data. Fewer data values are deleted using this method. There are situations when pairwise deletion will fail but this is the option to use when it is possible. Casewise deletion deletes all cells in any row of data containing a missing value. Much more data may be deleted using this option. To better understand the difference, consider a simple example of two data columns of equal length one of which has no missing values and the other has one missing value. When ROC areas are being compared, certain computations on these two columns will be done pairwise – the first column with itself, the first column with the second column and the second column with itself. When the column without a missing value is being compared with itself no row deletions occur for pairwise deletion. For casewise deletion, however, the row that contains the missing value will be deleted from *both* data sets. So, for casewise deletion, the computation involving the column without a missing value with itself will be done with one row deleted (the row corresponding to the missing value in the other data set). The program determines when pairwise deletion is not valid and informs the user when this is the case.

Positive State Options - Classification State and Direction

The two classification states are referred to as “Negative” (normal) or “Positive” (abnormal). The ROC analysis software must be informed which state is “Positive” and whether the test measurement values for the positive state are “High”, meaning higher than those of the negative state, or “Low”, meaning lower than those of the negative state.

Accepted normal values for the PSA (prostate specific antigen) test are less than 4 ng/ml and abnormal values are higher than this. Thus if the two classification states names are “positive” and “negative” then the Positive state is “positive” and the Positive Direction is “High”. In this case you would select the radio button next to “positive” and “High”.

Positive State Options	
Positive State	Positive Direction
<input checked="" type="radio"/> positive	High
<input type="radio"/> positive	Low
<input type="radio"/> negative	High
<input type="radio"/> negative	Low

On the other hand, for the T4 (thyroxine) test for hypothyroidism the T4 values are lower in the abnormal state than for the normal state. In this case the abnormal Positive State is “Hypothyroid” and the Positive Direction is “Low”. So you would select the radio button next to “Hypothyroid” and “Low”.

Positive State Options	
Positive State	Positive Direction
<input type="radio"/> Euthyroid	High
<input type="radio"/> Euthyroid	Low
<input type="radio"/> Hypothyroid	High
<input checked="" type="radio"/> Hypothyroid	Low

What happens if you select the incorrect option? Sensitivity (specificity) is defined in terms of the positive (negative) state. So if the positive state is incorrectly selected then sensitivity and specificity will be incorrectly defined (switched) and the ROC curve will have the X and Y axes switched. This will result in an ROC curve that appears below the diagonal unity line. It will have an area less than 0.5. The program will detect this and give you the options

ROC area < 0.5

An ROC area is less than 0.5. It is very likely that a classification state with an incorrect positive direction has been selected. Do you want to

Abort the computation?
Retry with different selections?
Ignore the warning and continue?

It is possible that there is something wrong with the data so you can Abort the analysis and correct the problem. More likely you have selected the incorrect positive state or direction so you can Retry the analysis with correct selections. In rare occasions for multiple tests some tests will have areas greater than 0.5 and one or more will have areas less than 0.5. In this case you can Ignore this warning and continue with the analysis.

Report Options

Confidence Intervals

Confidence intervals are computed for statistics in both the Sensitivity & Specificity and Area Comparison reports. You can generate 90, 95 and 99% confidence intervals.

Create Sensitivity and Specificity Report

Cutoff values are created between each test data value in the (sorted) data set. If there are a large number of data points and several tests then there will be a large number of cutoff values and the Sensitivity & Specificity Report can be very long. The checkbox

Create sensitivity and specificity

allows you to turn off this report. If you turn off this report then all report options in the dialog below this are not required and are disabled.

Fractions/Percents

You may display sensitivities, specificities and probabilities in either fraction or percent format. Selecting Percents also requires the pre-test probability to be entered as a percent.

Create Post-Test Results

Selecting this option allows entry of the pre-test probability. It also enables the possible entry of the false-positive/false-negative cost ratio. Given a pre-test probability the program will create post-test probabilities, both the positive predictive value (PV + = probability of disease given a positive test result) and the negative predictive value (PV - = probability of no disease given a negative test result), for each cutoff value. If the cost ratio option is selected then the optimal cutoff value will be computed. All of these results are displayed for each test in the Sensitivity & Specificity report.

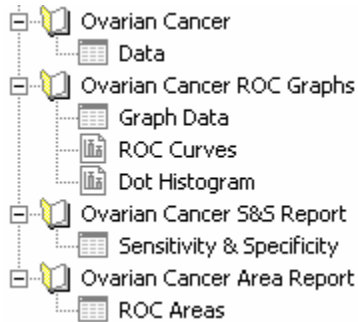
ROC Graph Options

All of the graph options in the dialog apply to the ROC graph. They allow you to add a diagonal line to the graph, add grid lines, add symbols for sensitivity and specificity at each cutoff point and change the ROC plot lines from solid to different line styles.

Analysis Results

Introduction

Typical results of the ROC analysis are shown in the following example from the Notebook Manager.



The first section entitled 'Ovarian Cancer' contains the worksheet containing the raw data. The program created the next three sections that contain two graphs and two reports. The contents of the two graphs

ROC Curves
Dot Histogram

and the two reports

Sensitivity & Specificity
ROC Areas

are described in the next sections.

ROC Curves Graph

The ROC curves graph for three data sets is shown in Figure 4. These graphs are derived from numerical results in the worksheet entitled Graph Data. The graph title is obtained from the section name containing the raw data. The legend shows the test names and the ROC areas for each curve. The diagonal line and grids options were selected for this graph.

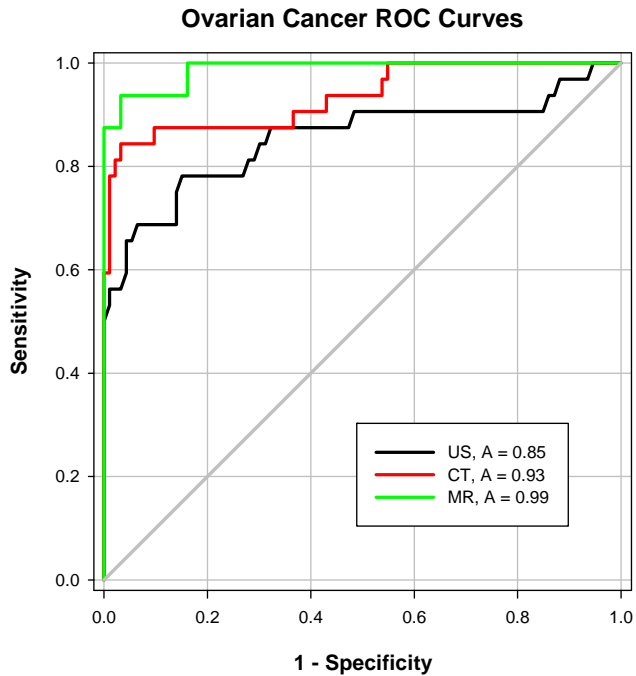
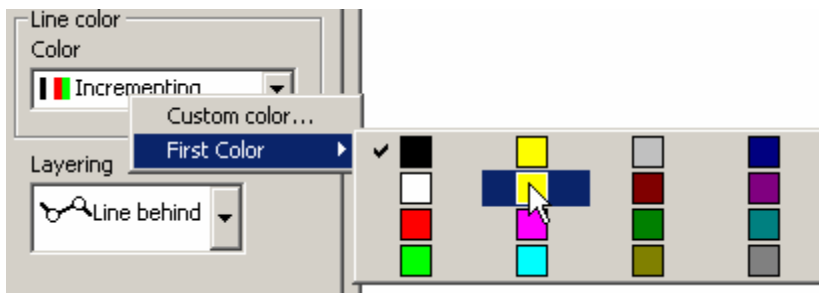


Figure 4. The ROC curves graph for three tests.

Of course this graph can be edited in any way you wish. You might want to change the starting color of the color scheme used for the line colors. You can do this by double clicking on one of the ROC plot lines and then right clicking on the Line Color listbox as shown next.



Dot Histogram Graph

Dot histograms for the data associated with the ROC curves in Figure 4 are shown in Figure 5 below.

Ovarian Cancer Data

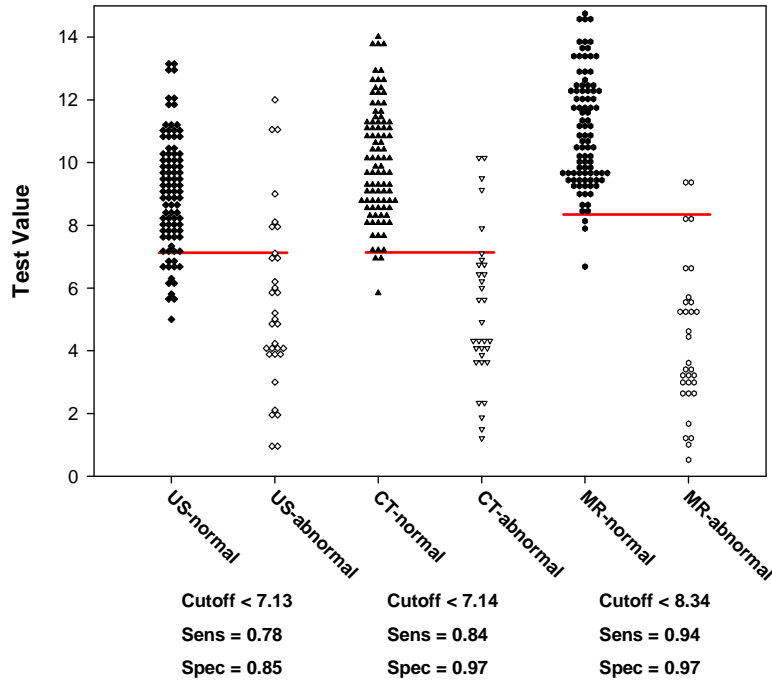


Figure 5: Dot histogram pairs for each test. The horizontal lines and the tables below the graph show the optimal cutoff values determined from the pre-test probability and cost ratio.

The graph title is obtained from the title of the section containing the raw data. The x-axis tick labels are obtained from the test names and the classification state names. The tick labels will rotate if they are too long to fit horizontally. The symbol layout design allows for symbols to touch horizontally and nest vertically.

If values for pre-test probability and false-positive/false-negative cost ratio are entered then the optimal cutoff values for each test are computed and represented as a horizontal line across the two dot histograms for each test. The numeric values for the optimal cutoff parameters are shown as tables below the x-axis.

Sensitivity & Specificity Report

The sensitivity & Specificity report contains results for all tests with additional tests results placed in report rows below those of prior tests. The results for each test can be separated into three parts: 1) optimal cutoff value, 2) sensitivity and specificity versus cutoff values and 3) likelihood ratios and post-test probabilities.

If values for both pre-test probability and cost ratio have been entered then the optimal cutoff is calculated. A slope of the tangent to the ROC curve m is defined in terms of the two entered values ($P =$ pre-test probability)⁽¹⁾

$$m = \left(\frac{\text{false - positive cost}}{\text{false - negative cost}} \right) \left(\frac{1 - P}{P} \right) \quad (1)$$

The optimal cutoff value is computed from sensitivity and specificity using the slope m by finding the cutoff that maximizes the function ⁽¹⁾

$$\text{Sensitivity} - m(1 - \text{Specificity}) \quad (2)$$

The results of this computation in the Sensitivity & Specificity report are shown in Table 1.

Optimal Operating Point	
Pre-test Probability	0.50000
Cost Ratio	1.00000
Cutoff <	7.12500
Sensitivity	0.78130
Specificity	0.84950

Table 1: Optimal cutoff results in the Sensitivity & Specificity report.

For this data set, the optimal cutoff is 7.125 for a pre-test probability of 0.5 and cost ratio of 1.0.

Sensitivities, specificities and their confidence intervals are listed as a function of cutoff value in the second part of the report. A portion of these results is shown in Table 2. These results can be expressed as fractions or percents by using the Fractions/Percents option.

Cutoff <	Sensitivity	95% CI	Specificity	95% CI
0.95000	0.03125	0.000791 to 0.1622	1.00000	0.9611 to 1.000
1.45000	0.06250	0.007661 to 0.2081	1.00000	0.9611 to 1.000
1.95000	0.09375	0.01977 to 0.2502	1.00000	0.9611 to 1.000
2.05000	0.12500	0.03513 to 0.2899	1.00000	0.9611 to 1.000
2.55000	0.15630	0.05275 to 0.3279	1.00000	0.9611 to 1.000
3.40000	0.18750	0.07208 to 0.3644	1.00000	0.9611 to 1.000
3.85000	0.21880	0.09277 to 0.3997	1.00000	0.9611 to 1.000
3.92500	0.25000	0.1146 to 0.4340	1.00000	0.9611 to 1.000
3.97500	0.28130	0.1375 to 0.4675	1.00000	0.9611 to 1.000
4.02500	0.31250	0.1612 to 0.5001	1.00000	0.9611 to 1.000
4.07500	0.34380	0.1857 to 0.5319	1.00000	0.9611 to 1.000
4.12500	0.37500	0.2110 to 0.5631	1.00000	0.9611 to 1.000
4.17500	0.40630	0.2370 to 0.5936	1.00000	0.9611 to 1.000
4.50000	0.43750	0.2636 to 0.6234	1.00000	0.9611 to 1.000
4.85000	0.46880	0.2909 to 0.6526	1.00000	0.9611 to 1.000
4.95000	0.50000	0.3189 to 0.6811	1.00000	0.9611 to 1.000
5.10000	0.53130	0.3474 to 0.7091	0.98920	0.9415 to 0.9997
5.40000	0.56250	0.3766 to 0.7364	0.98920	0.9415 to 0.9997
5.65000	0.56250	0.3766 to 0.7364	0.97850	0.9245 to 0.9974

Table 2: Sensitivity and specificity results in the Sensitivity & Specificity report.

The third part of the Sensitivity & Specificity report contains the likelihood ratios and post-test probabilities.

The positive and negative likelihood ratios are defined respectively as

$$LR_{+} = \frac{\text{Probability of a positive test given the presence of disease}}{\text{Probability of a positive test given the absence of disease}} = \frac{\text{Sensitivity}}{1 - \text{Specificity}} \quad (3)$$

$$LR_{-} = \frac{\text{Probability of a negative test given the presence of disease}}{\text{Probability of a negative test given the absence of disease}} = \frac{1 - \text{Sensitivity}}{\text{Specificity}} \quad (4)$$

The post-test probabilities are the probability of disease given a positive test (PV+) and the probability of no disease given a negative test (PV-). These will be computed when a pre-test probability has been entered. Using P = pre-test probability, the equations used for these probabilities are

$$PV_{+} = \frac{\text{Sensitivity} \times P}{\text{Sensitivity} \times P + (1 - \text{Specificity}) \times (1 - P)} \quad (5)$$

$$PV_{-} = \frac{\text{Specificity} \times (1 - P)}{\text{Specificity} \times (1 - P) + (1 - \text{Sensitivity}) \times P} \quad (6)$$

A portion of the report showing the likelihood and post-test probabilities results is shown in Table 3.

Cutoff <	LR +	LR -	PV +	PV -
0.9500		0.9690	1.0000	0.5080
1.4500		0.9380	1.0000	0.5160
1.9500		0.9060	1.0000	0.5250
2.0500		0.8750	1.0000	0.5330
2.5500		0.8440	1.0000	0.5420
3.4000		0.8130	1.0000	0.5520
3.8500		0.7810	1.0000	0.5610
3.9250		0.7500	1.0000	0.5710
3.9750		0.7190	1.0000	0.5820
4.0250		0.6880	1.0000	0.5930
4.0750		0.6560	1.0000	0.6040
4.1250		0.6250	1.0000	0.6150
4.1750		0.5940	1.0000	0.6270
4.5000		0.5630	1.0000	0.6400
4.8500		0.5310	1.0000	0.6530
4.9500		0.5000	1.0000	0.6670
5.1000	49.1940	0.4740	0.9800	0.6790
5.4000	52.0830	0.4420	0.9810	0.6930
5.6500	26.1630	0.4470	0.9630	0.6910
5.7500	17.4150	0.4520	0.9460	0.6890

Table 3: Positive and negative likelihood ratios, LR+ and LR-, and post-test probabilities, PV+ and PV-, in the Sensitivity & Specificity report.

The positive likelihood ratio is not defined for some cutoff values since specificity = 1.

ROC Areas Report

The ROC Area report consists of two parts: 1) ROC areas and their associated statistics and 2) pairwise comparison of ROC areas. An example of a report is shown in Table 4.

ROC Areas				
	1	2	3	4
1	ROC Curve Areas			
2	Paired Analysis, Pairwise Deletion			
3				
4	Test Variables	Albumin	Total Protein	Total Score
5	ROC Curve Area	0.7188	0.6478	0.6869
6	Standard Error	0.09283	0.1000	0.1051
7	95% Confidence Interval	0.5368 To 0.9007	0.4518 To 0.8439	0.4809 To 0.8929
8	P Value	0.02686	0.1364	0.05364
9	Sample Size - success	12	12	12
10	Sample Size - failure	32	31	37
11	Missing - success	0	0	0
12	Missing - failure	5	6	0
13				
14	ROC Curve Area Comparison			
15				
16	Pair	Albumin, Total Protein	Albumin, Total Score	Total Protein, Total Score
17	Area Difference	0.07090	0.03181	-0.03909
18	Standard Error	0.05602	0.09803	0.1069
19	95% Confidence Interval	-0.03890 to 0.1807	-0.1603 to 0.2240	-0.2485 to 0.1704
20	ChiSquare, DF = 1	1.602	0.1053	0.1338
21	P Value	0.2057	0.7456	0.7145

Table 4: An example ROC Areas report. From top to bottom it shows the type of analysis used together with the missing value method, the ROC areas and associated statistics and a pairwise comparison of ROC areas.

In this case there are three correlated tests. Row two of the report shows that a Paired Analysis was performed and, since there were missing values in the data, Pairwise Deletion of missing values was selected to compare the areas.

The first section of the report shows the ROC curve areas for the three tests. This is followed by the standard error of the area estimate, the 95% confidence interval (90% and 99% are also available) and the P value that determines if the area value is significantly different from 0.5. The sample size and the number of missing values for each classification state are given. The number of missing values reflects only what is seen in the data and does not give the number used for each computation-pair in the pairwise-deleted comparison of areas.

The second section shows the results of the pairwise comparison of areas. The method of DeLong, DeLong and Clarke-Pearson⁽²⁾ is used to compare areas when the Paired data type option is selected. When the Unpaired data type is selected, areas are compared using a Z test. The report shows results for all pairs of data sets. The difference of each area pair and its standard error and 95% confidence interval are computed. This is followed by the chi-square statistic for the area comparison (or Z statistic if Unpaired is selected) and its associated P value.

Formatted Full Precision Display

This report presents the numeric results in a four significant digit format with full precision available. Double click on any cell (except the confidence intervals) to display the number at full precision.

Additional Graphs

Results data in both reports can be used to create additional graphs. Some examples seen in the literature are shown here.

Sensitivity and Specificity vs. Cutoff

The data for the graph in Figure 6 is from the Sensitivity & Specificity report in columns 1, 2 and 4. Use the Data Sampling option in Graph Properties, Plots, Data to specify the row range for the graph (you can also drag select the rows in the worksheet to do this).

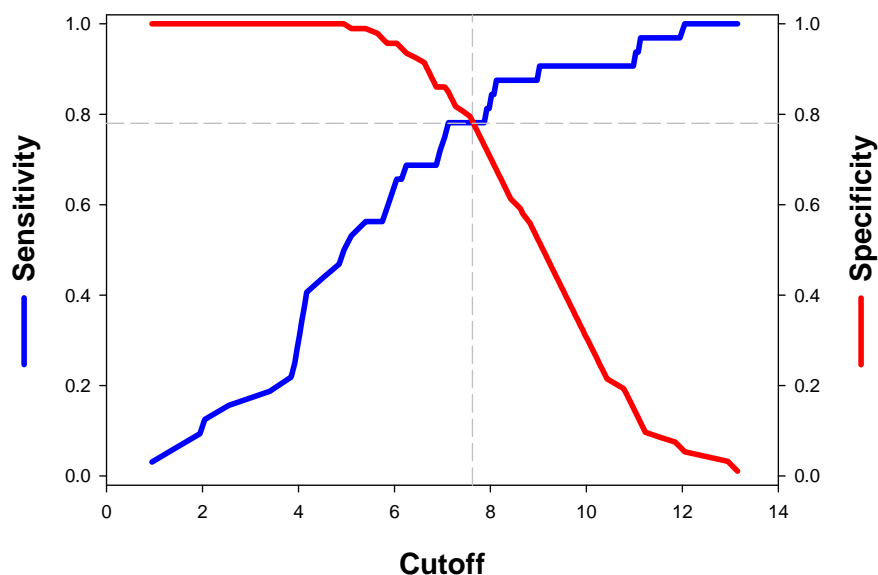


Figure 6: Graph of sensitivity and specificity vs. cutoff for one test using data from columns 1,2 and 4 of the Sensitivity & Specificity report.

Likelihood Ratios

The positive and negative likelihood ratios for three different imaging modalities are shown in Figure 7 (the data is artificial). The data is in columns 1, 6 and 7 of the Sensitivity & Specificity report. The values associated with the optimal cutoff are shown as solid symbols. The largest positive likelihood and smallest negative likelihood at the optimal cutoff is associated with magnetic resonance imaging (MR).

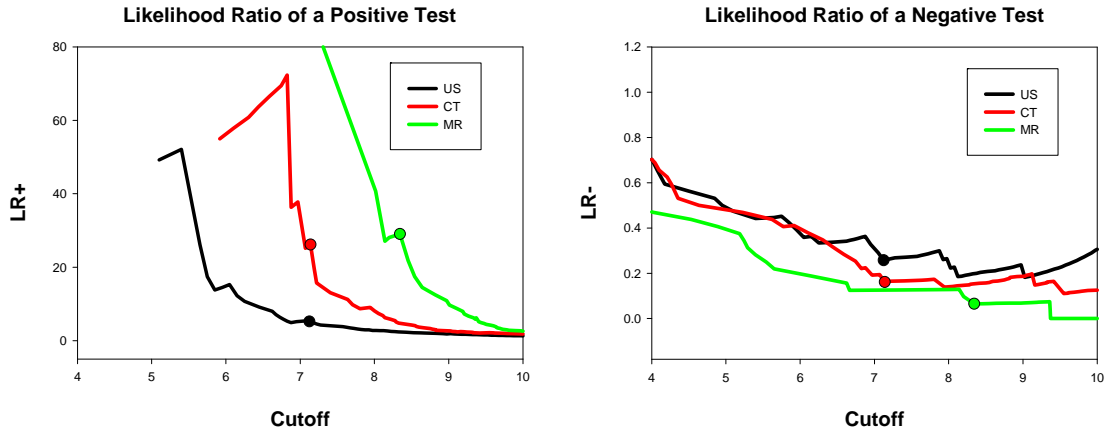


Figure 7: Positive and negative likelihood ratios graphed from data in the Sensitivity & Specificity report from columns 1, 6 and 7. The results for three tests are shown together with values associated with the optimal cutoff (solid symbols).

Optimal Cutoff vs. Cost Ratio

Frequently it can be difficult to determine a value for the false-positive/false-negative cost ratio. So it is worth performing a sensitivity analysis (sensitivity here means how much one variable changes with changes in a second variable) to see whether the cutoff value changes significantly in the range of cost-ratio values of interest. The ROC Curves Module was run multiple times for different cost ratios and a graph of optimal cutoff vs. cost ratio for the three imaging modality tests is shown below.

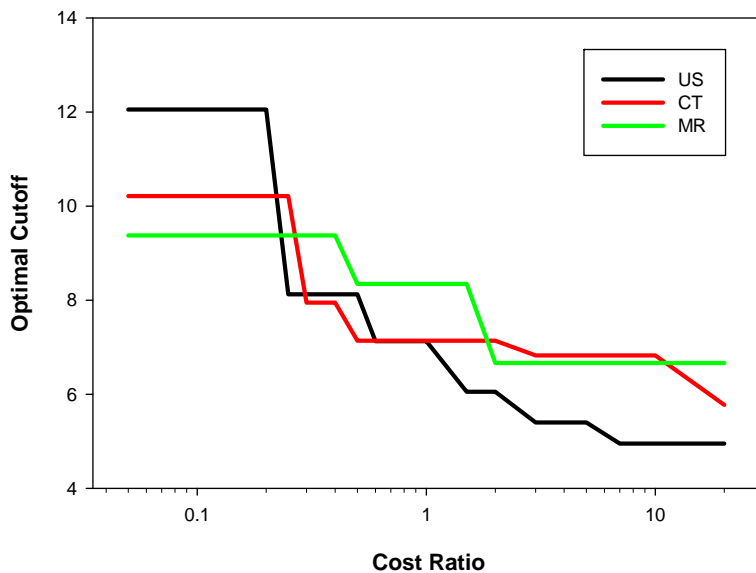


Figure 8: Optimal cutoff values obtained from multiple runs of the program. Regions of insensitivity, or strong sensitivity, to cost ratio can be identified.

If the relative cost of a false-positive is much greater than that of a false-negative then the cost ratio is greater than 1. But lets assume that we don't know exactly how much greater it is but have some idea that it should be in the range of 2 to 5, say. Looking at the optimal cutoff for the best imaging modality

(MR, green line) we find that it doesn't change for cost ratios from 2 to 20. So the optimal cutoff is insensitive to cost ratio and, in this case, it is not important to know a precise value for cost-ratio.

Post-Test Probability vs. Pre-Test Probability

Given values of sensitivity and specificity associated with the optimal cutoff a graph of post-test probabilities as a function of pre-test probability can be created using equations (5) and (6). The post-test probability of disease when the test is positive, blue lines in Figure 9, was obtained from equation (5) and the post-test probability of disease when the test was negative, red lines, was obtained from 1.0 minus equation (6). A transform was written in SigmaPlot implementing these two equations that generated the post-test probabilities for a range of pre-test probabilities. The results for the best test, MR, and worst test, US, are shown. The MR test is clearly better since the post-test probability range, from negative test to positive test, is larger. Thus given a positive test the patient is more likely to have the disease using the MR test rather than the US test. Similarly, given a negative test it is less likely that the patient has the disease using the MR test.

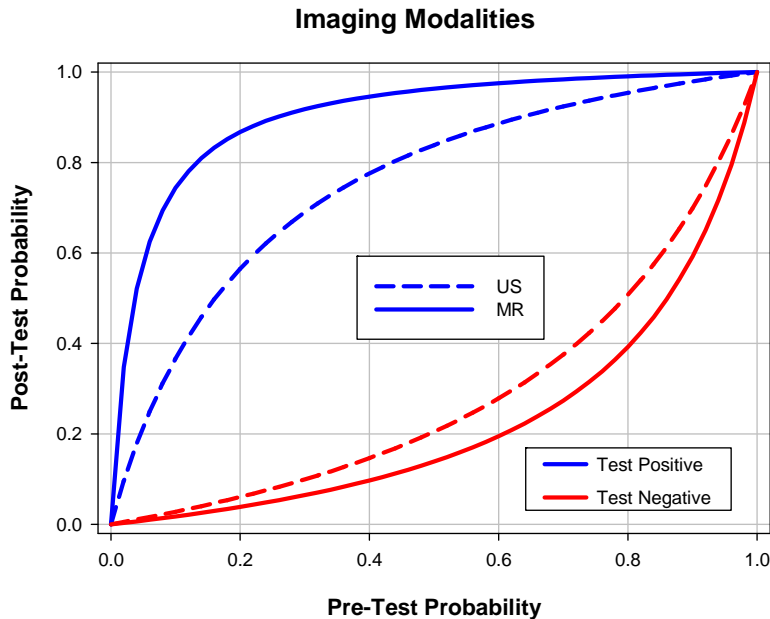


Figure 9: Post-test probabilities of disease given positive and negative test results. The MR test is based on sensitivity = 0.94 and specificity = 0.97 whereas the US test used sensitivity = 0.78 and specificity = 0.85.

References

1. Zweig, MH, Campbell, G. Receiver-operating characteristic (ROC) plots: A fundamental evaluation tool in clinical medicine. *Clin Chem* 1993;39/4, 56-577.
2. DeLong, ER, DeLong, DM, Clarke-Pearson, DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;44, 837-845.
3. Hanley, JA, McNeil, BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982, 143, 29-36.