


Proc LOGISTIC ROCs! Let's see how...

Colleen E McGahan

Lead Biostatistician, Surveillance & Outcomes Unit,
BC Cancer Agency, Vancouver

Receiver Operating Characteristic Curve



SUMMARY

- Brief overview of ROC curves
- ROC curve statements/options available in `proc LOGISTIC`
- Assumes use of SAS 9.2
- Assumes basic knowledge of logistic regression
- Does not cover model selection techniques

Introduction

- Logistic regression provides the estimated probability that the event of interest will happen.
- It can be used as a decision making tool whereby, given the probability of the event happening you decide to take action or not
- In order to do this, a probability cut-off is required – a probability higher than the cut-off you predict the event will happen, below the cut-off you predict the event won't happen.

- The accuracy of a predictive model can be assessed by comparing the predicted outcome with the actual known outcome.
- And the predicted outcome depends on the probability cut-off specified

For example,

Low grade breast cancer patients require radiation therapy (RT) after surgical removal of the tumour.

- Standard treatment: Whole breast
 - 4 to 6 weeks treatment
- New treatment: Partial breast
 - 1 week treatment
 - reduced side effects
 - takes 2 hours longer per patient to plan
 - some women are not suitable to receive a partial plan

The question is:

- Can these women be identified ahead of time so 2 hours is not spent on planning their treatment and they can make arrangements to be receiving treatment for 4-6 weeks.

Logistic Regression

- Model Outcome:
 - probability of not suitable for a Partial Plan
- Patient characteristics:
 - age
 - breast laterality
 - radiation dose volume
 - breast volume
 - surgical cavity volume
 - ratio of the surgical cavity volume to breast volume
 - ratio of radiation dose volume to breast volume
 - breast quadrant the tumour was located in
 - number of weeks between final breast surgery and the CT scan used for radiation planning purposes

Predictive Model Criteria

Want the predictive model to save planning a partial plan for at least 25% of patients that would not be suitable to have one.

Would accept a cost of not giving a partial plan to 5% of patients that could potentially receive one.

Probability cut-off for not suitable for Partial Plan = 0.2

| PREDICTED OUTCOME | KNOWN OUTCOME | |
|----------------------------------|---|---|
| | Not suitable | Suitable |
| Not suitable ($p \geq 0.2$) | TRUE POSITIVE rate=0.762 (n=48) | FALSE POSITIVE rate=0.272 (n=72) |
| Suitable ($p < 0.2$) | FALSE NEGATIVE rate=0.238 (n=15) | TRUE NEGATIVE rate=0.728 (n=193) |
| | 1 | 1 |

SENSITIVITY

Probability cut-off for not suitable for Partial Plan = 0.45

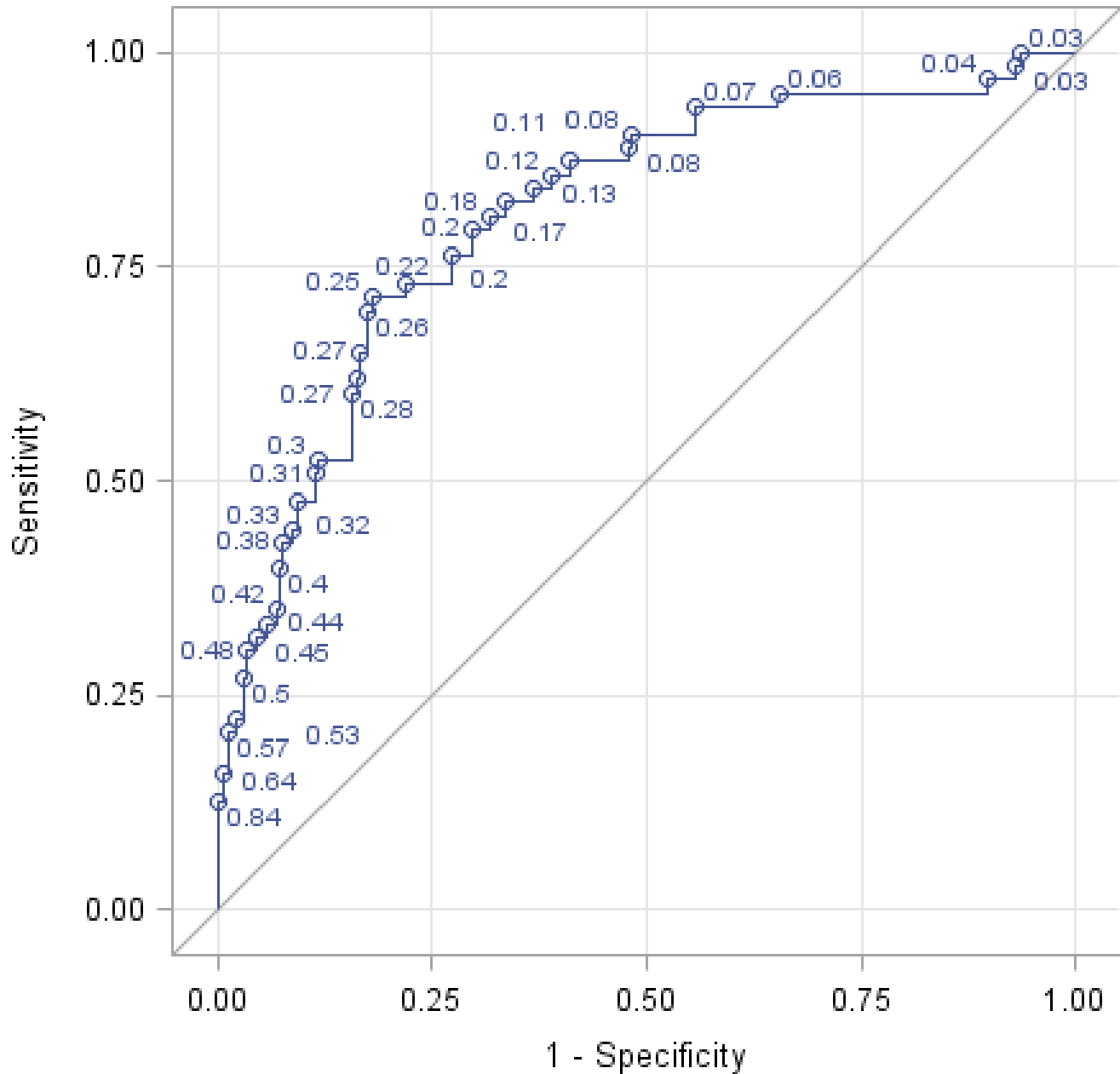
1-SPECIFICITY

| PREDICTED OUTCOME | KNOWN OUTCOME | |
|-----------------------------------|---|---|
| | Not suitable | Suitable |
| Not suitable ($p \geq 0.45$) | TRUE POSITIVE rate=0.318 (n=20) | FALSE POSITIVE rate=0.049 (n=13) |
| Suitable ($p < 0.45$) | FALSE NEGATIVE rate=0.682 (n=43) | TRUE NEGATIVE rate=0.951 (n=252) |
| | 1 | 1 |

SPECIFICITY

ROC Curve for Model

Area Under the Curve = 0.8090



Proc LOGISTIC ROCs!
Let's see how...



Proc LOGISTIC ROCs!

- In order to produce a ROC curve in proc LOGISTIC, ODS graphics needs to be turned on.

```
ods graphics on;
```

The ROC curve can then be requested in the `proc LOGISTIC` statement using the `PLOTS` option.

```
ods graphics on;  
proc logistic DATA=dset  
    PLOTS(ONLY)=(ROC(ID=prob) EFFECT);  
    CLASS quadrant / PARAM=glm;  
    MODEL partplan = quadrant cavnobr;  
run;
```

The `ONLY` option suppresses the default plots and only the requested plots are displayed.

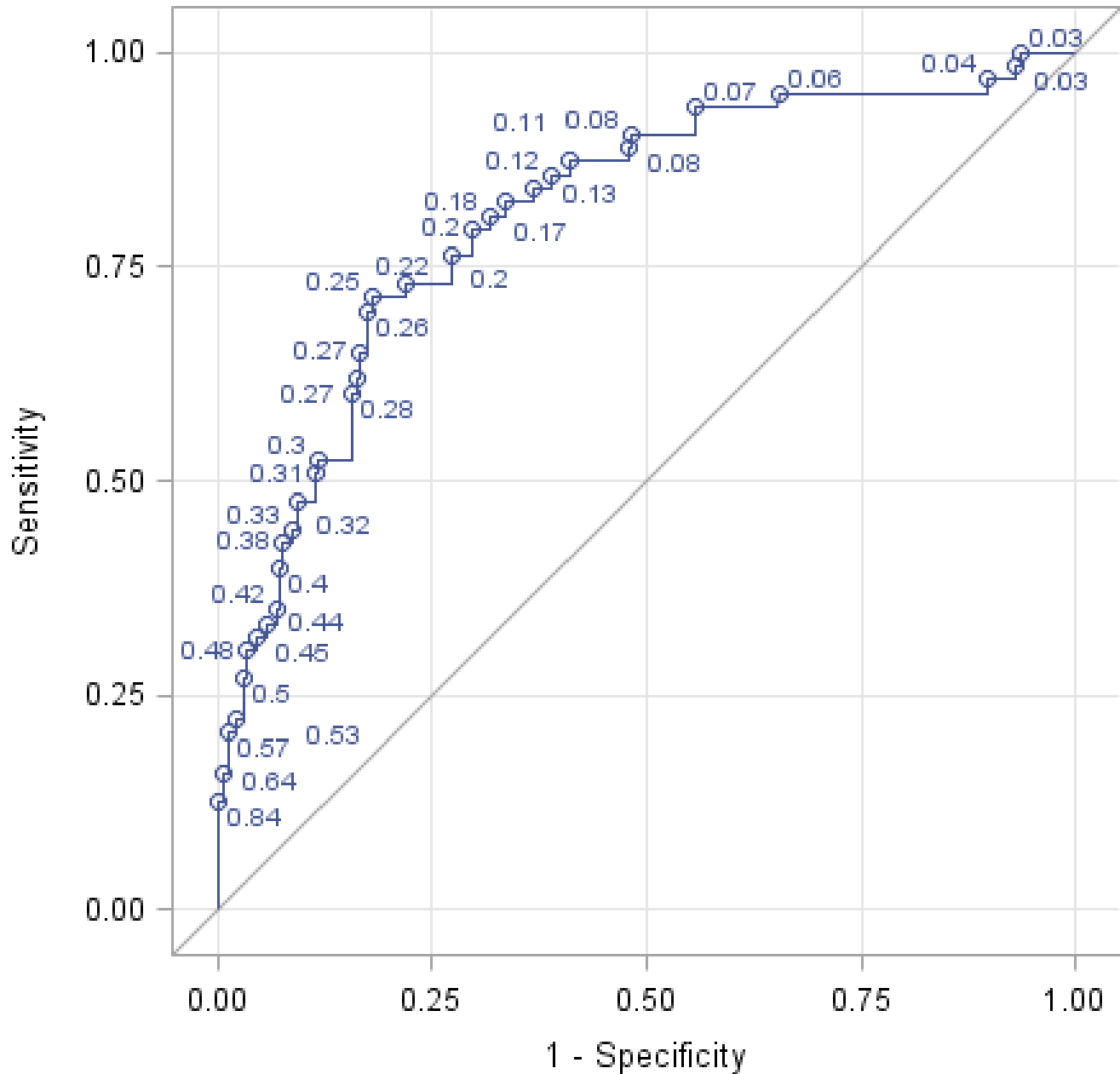
In this case only the ROC curve and the `EFFECT` plot will be displayed.

```
ods graphics on;  
proc logistic DATA=dset  
    PLOTS (ONLY) = (ROC (ID=prob) EFFECT) ;  
    CLASS quadrant / PARAM=glm;  
    MODEL partplan = quadrant cavtobr;  
run;
```

ID=prob option requests that the predicted probabilities are displayed on the ROC plot

ROC Curve for Model

Area Under the Curve = 0.8090

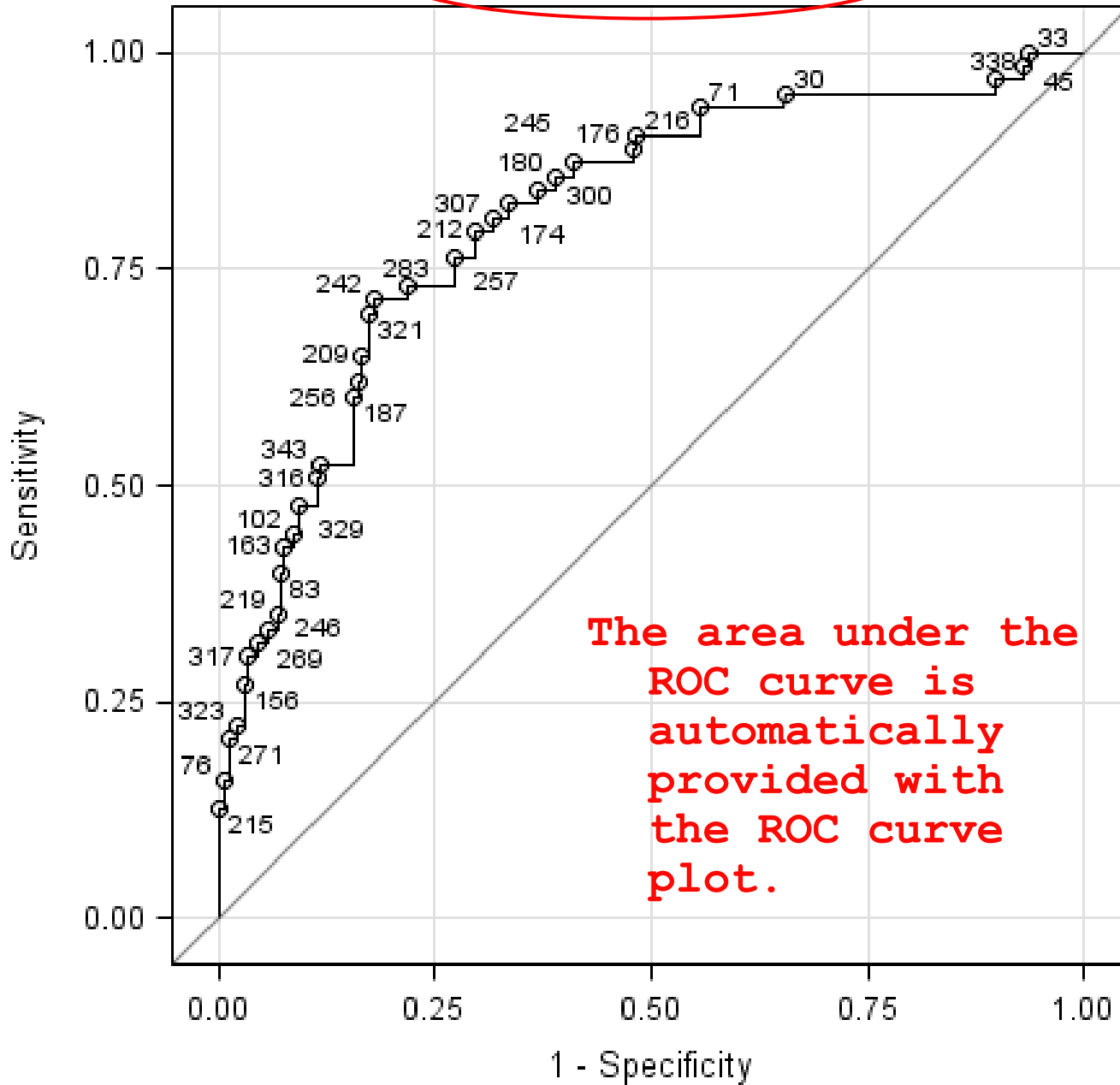


```
ods graphics on;  
proc logistic DATA=dset  
    PLOTS (ONLY) = (ROC (ID=casenum) EFFECT);  
    CLASS quadrant / PARAM=gim;  
    MODEL partplan = quadrant cavtoabr;  
run;
```

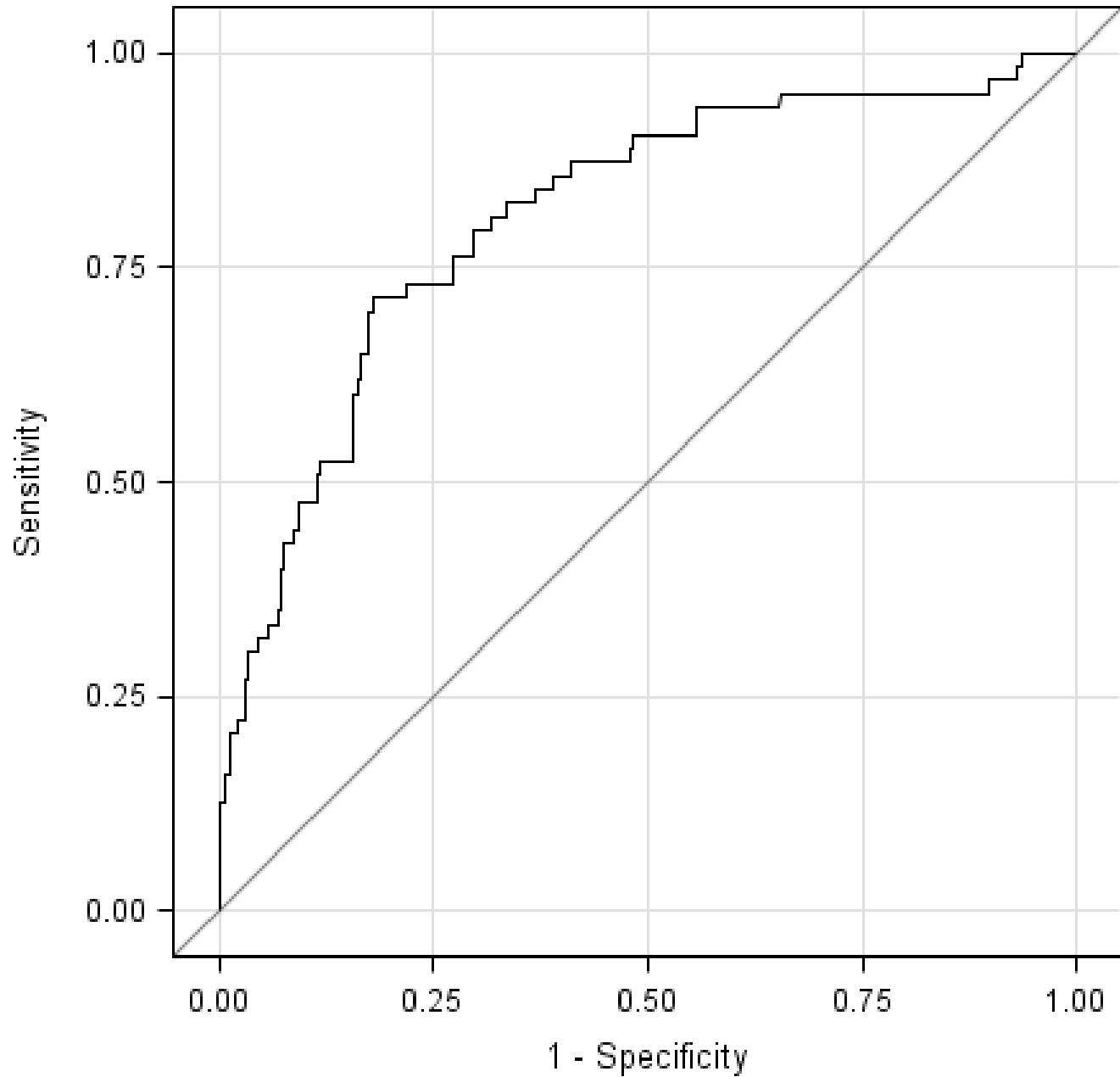
- ID=prob option requests that the predicted probabilities are displayed on the ROC plot
- If instead you wanted the observation number displayed, ID=casenum could be used.

ROC Curve for Model

Area Under the Curve = 0.8090



ROC Curve for Model
Area Under the Curve = 0.8090



```
ods graphics on;  
proc logistic DATA=dset;  
  CLASS quadrant / PARAM=glm;  
  MODEL partplan = quadrant cavto br /  
    OUTROC=dset_name;  
run;
```

The option OUTROC= produces an output dataset that holds all the data that is used to produce the ROC plot.

VIEWTABLE: Receiver Operating Characteristics

| | Probability Level | No. of Correctly Predicted Events | No. of Correctly Predicted Nonevents | No. of Nonevents Predicted as Events | No. of Events Predicted as Nonevents | Sensitivity | 1 - Specificity |
|----|-------------------|-----------------------------------|--------------------------------------|--------------------------------------|--------------------------------------|--------------|-----------------|
| 1 | 0.9996098411 | 1 | 265 | 0 | 62 | 0.0158730159 | 0 |
| 2 | 0.9890844658 | 2 | 265 | 0 | 61 | 0.0317460317 | 0 |
| 3 | 0.9756835305 | 3 | 265 | 0 | 60 | 0.0476190476 | 0 |
| 4 | 0.9434518496 | 4 | 265 | 0 | 59 | 0.0634920635 | 0 |
| 5 | 0.9353801923 | 5 | 265 | 0 | 58 | 0.0793650794 | 0 |
| 6 | 0.9236486803 | 6 | 265 | 0 | 57 | 0.0952380952 | 0 |
| 7 | 0.8919756696 | 7 | 265 | 0 | 56 | 0.1111111111 | 0 |
| 8 | 0.8353664426 | 8 | 265 | 0 | 55 | 0.126984127 | 0 |
| 9 | 0.8003994659 | 8 | 264 | 1 | 55 | 0.126984127 | 0.0037735849 |
| 10 | 0.783013909 | 8 | 263 | 2 | 55 | 0.126984127 | 0.0075471698 |
| 11 | 0.7264493317 | 9 | 263 | 2 | 54 | 0.1428571429 | 0.0075471698 |
| 12 | 0.6431072195 | 10 | 263 | 2 | 53 | 0.1587301587 | 0.0075471698 |
| 13 | 0.6234507017 | 10 | 262 | 3 | 53 | 0.1587301587 | 0.0113207547 |
| 14 | 0.6044447286 | 11 | 262 | 3 | 52 | 0.1746031746 | 0.0113207547 |
| 15 | 0.5855376375 | 12 | 262 | 3 | 51 | 0.1904761905 | 0.0113207547 |
| 16 | 0.5700183463 | 13 | 262 | 3 | 50 | 0.2063492063 | 0.0113207547 |
| 17 | 0.568987181 | 13 | 261 | 4 | 50 | 0.2063492063 | 0.0150943396 |
| 18 | 0.5599176113 | 13 | 260 | 5 | 50 | 0.2063492063 | 0.0188679245 |
| 19 | 0.5499463684 | 13 | 259 | 6 | 50 | 0.2063492063 | 0.0226415094 |
| 20 | 0.5310464599 | 14 | 259 | 6 | 49 | 0.2222222222 | 0.0226415094 |
| 21 | 0.5069313824 | 14 | 258 | 7 | 49 | 0.2222222222 | 0.0264150943 |
| 22 | 0.5020175756 | 14 | 257 | 8 | 49 | 0.2222222222 | 0.0301886792 |
| 23 | 0.5000354978 | 15 | 257 | 8 | 48 | 0.2380952381 | 0.0301886792 |
| 24 | 0.4994565271 | 16 | 257 | 8 | 47 | 0.253968254 | 0.0301886792 |
| 25 | 0.4951223106 | 17 | 257 | 8 | 46 | 0.2698412698 | 0.0301886792 |
| 26 | 0.4930484167 | 17 | 256 | 9 | 46 | 0.2698412698 | 0.0339622642 |
| 27 | 0.4924959025 | 18 | 256 | 9 | 45 | 0.2857142857 | 0.0339622642 |
| 28 | 0.4760355107 | 19 | 256 | 9 | 44 | 0.3015873016 | 0.0339622642 |
| 29 | 0.4699692178 | 19 | 255 | 10 | 44 | 0.3015873016 | 0.0377358491 |
| 30 | 0.4679401888 | 19 | 254 | 11 | 44 | 0.3015873016 | 0.041509434 |

```
ods graphics on;  
proc logistic DATA=dset;  
  CLASS quadrant / PARAM=glm;  
  MODEL partplan = quadrant cavgobr /  
    OUTROC=dset_name;  
run;
```

If the PLOTS option is not specified the OUTROC= option will still produce a ROC curve.

But, the predicted probabilities or observation number will not be shown on the plot.

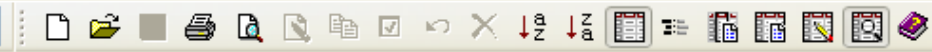
The OUTROC dataset is not provided if you only request the ROC curve using the PLOTS option.

```
ods graphics on;  
ods output classification=classification_table;  
proc logistic DATA=dset;  
  CLASS quadrant / PARAM=glm;  
  MODEL partplan = quadrant cavnobr /  
    OUTROC=dset_name  
    CTABLE PPROB=(.3 .5 to .9 by .1);  
run;
```

CTABLE with PPROB= option can be used to obtain the Classification Table

Can also request the classification table dataset with the ods output statement

If PPROB= is not specified then the default is to show classifications for probability cut-offs from 0 to 1 at intervals of .02



Classification Table

| Prob Level | Correct | | Incorrect | | Correct | Percentages | | | | |
|------------|---------|-----------|-----------|-----------|---------|--------------|--------------|-----------|-----------|--|
| | Event | Non-Event | Event | Non-Event | | Sensi-tivity | Speci-ficity | False POS | False NEG | |
| 0.300 | 28 | 227 | 38 | 35 | 77.7 | 44.4 | 85.7 | 57.6 | 13.4 | |
| 0.500 | 13 | 256 | 9 | 50 | 82.0 | 20.6 | 96.6 | 40.9 | 16.3 | |
| 0.600 | 10 | 262 | 3 | 53 | 82.9 | 15.9 | 98.9 | 23.1 | 16.8 | |
| 0.700 | 9 | 263 | 2 | 54 | 82.9 | 14.3 | 99.2 | 18.2 | 17.0 | |
| 0.800 | 8 | 263 | 2 | 55 | 82.6 | 12.7 | 99.2 | 20.0 | 17.3 | |
| 0.900 | 6 | 265 | 0 | 57 | 82.6 | 9.5 | 100.0 | 0.0 | 17.7 | |

VIEWTABLE: Classification Table

| | Probability Level | Number of Correct Events | Number of Correct Nonevents | Number of Incorrect Events | Number of Incorrect Nonevents | Percentage of Correct Classification | Sensitivity in Percent | Specificity in Percent | Percentage of False Positive | Percentage of False Negative |
|---|-------------------|--------------------------|-----------------------------|----------------------------|-------------------------------|--------------------------------------|------------------------|------------------------|------------------------------|------------------------------|
| 1 | 0.300 | 28 | 227 | 38 | 35 | 77.7 | 44.4 | 85.7 | 57.6 | 13.4 |
| 2 | 0.500 | 13 | 256 | 9 | 50 | 82.0 | 20.6 | 96.6 | 40.9 | 16.3 |
| 3 | 0.600 | 10 | 262 | 3 | 53 | 82.9 | 15.9 | 98.9 | 23.1 | 16.8 |
| 4 | 0.700 | 9 | 263 | 2 | 54 | 82.9 | 14.3 | 99.2 | 18.2 | 17.0 |
| 5 | 0.800 | 8 | 263 | 2 | 55 | 82.6 | 12.7 | 99.2 | 20.0 | 17.3 |
| 6 | 0.900 | 6 | 265 | 0 | 57 | 82.6 | 9.5 | 100.0 | 0.0 | 17.7 |

```
ods graphics on;  
proc logistic DATA=dset  
  PLOTS (ONLY) = (ROC (ID=prob) ) ;  
  CLASS quadrant / PARAM=glm;  
  MODEL partplan = quadrant cavtobr /  
                                OUTROC=dset_name  
                                ROCEPS=0.1;  
run;
```

The ROCEPS= option requires a number between 0 and 1 which specifies the criterion for how you want the event probabilities grouped on the ROC curve or output into the OUTROC dataset.

(The default is 1E-8)

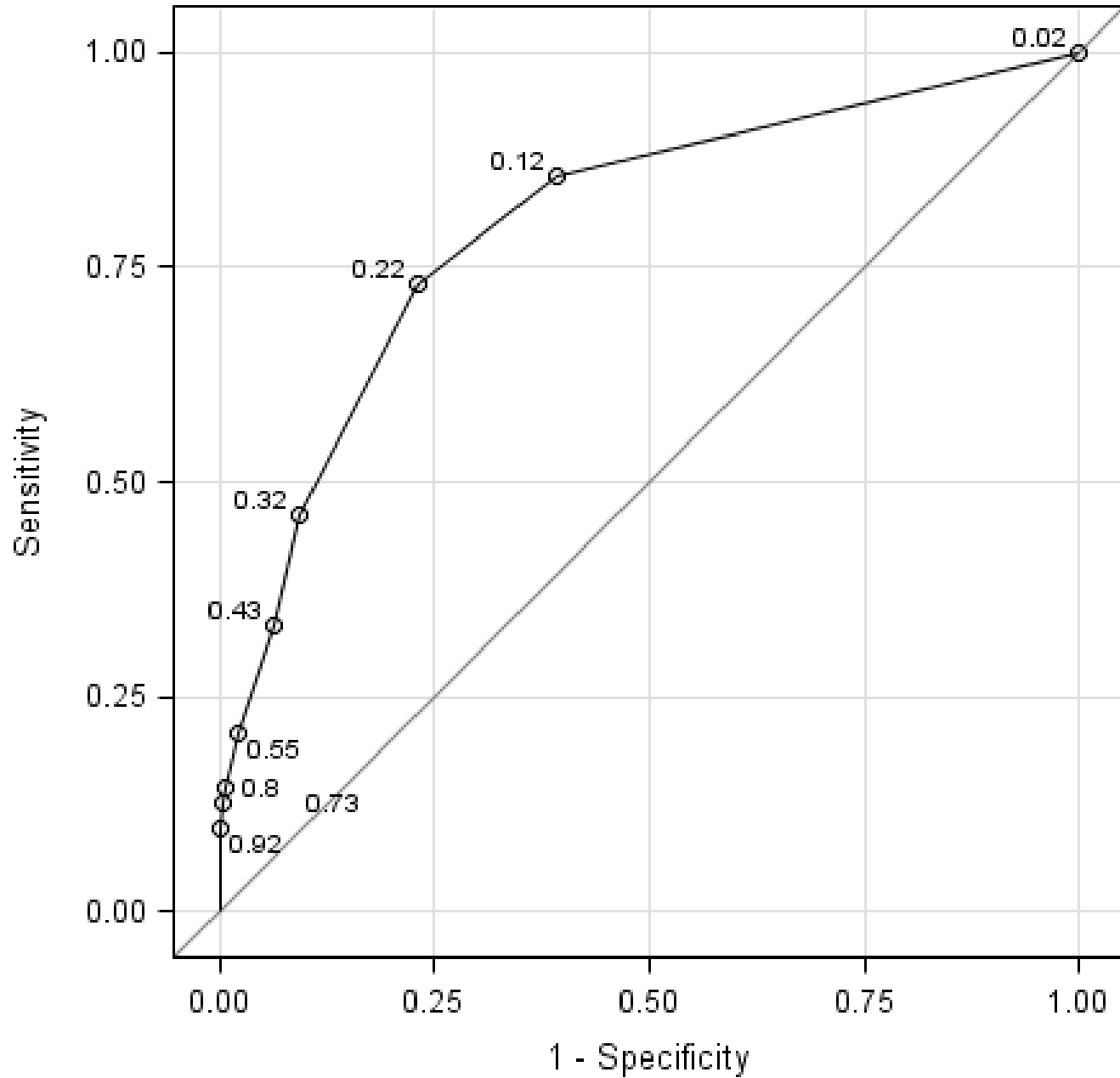
The AUC is adjusted accordingly.

- In each group, the difference between the largest and the smallest estimated event probabilities does not exceed the given value.
- The smallest estimated probability in each group serves as a cut point for predicting an event response.
- The OUTROC= option needs to be specified for ROCEPS= to have an effect

VIEWTABLE: Predicted Values and Diagnostic Statistics

| Cavity to Breast Ratio | cat | quadnew | Incavtoibr | centre | dset | radonc | bcs | ctdt | bcstoc | laterality | dosevol | dosetobr | Response Value | Estimated Probability |
|------------------------|-----|-------------|--------------|--------|-------|--------|-----------|-----------|--------|------------|---------|----------|----------------|-----------------------|
| 17.5 | 2+ | Lower Outer | 2.8635054038 | CSI | valid | | . | . | . | . | . | . | N | 0.9996098411 |
| 11.7 | 2+ | Upper Outer | 2.4589670708 | CSI | valid | | . | . | . | . | . | . | N | 0.9890844658 |
| 10.5 | 2+ | Upper Outer | 2.3511148265 | VIC | plan | C | 28MAY2008 | 23JUL2008 | 8.00 | L | 511.15 | 29.06 | N | 0.9756835305 |
| 10.1 | 2+ | Lower Outer | 2.3160081133 | VIC | plan | A | 31JUL2007 | 18SEP2007 | 7.00 | R | 435.4 | 32.69 | N | 0.9434518496 |
| 9.93 | 2+ | Lower Outer | 2.2952296603 | VIC | plan | B | 24JUL2007 | 20SEP2007 | 8.29 | L | 286.88 | 27.30 | N | 0.9353801923 |
| 5.84 | 2+ | Lower Inner | 1.7640324777 | VIC | plan | C | 16OCT2006 | 07DEC2006 | 7.43 | L | 369.18 | 18.21 | N | 0.9236486803 |
| 5.66 | 2+ | Upper Inner | 1.7331232827 | VC | valid | | . | . | . | . | . | . | N | 0.8919756696 |
| 4.94 | 2+ | Upper Inner | 1.5981158276 | VIC | plan | E | 22NOV2007 | 14JAN2008 | 7.57 | R | 274.54 | 19.44 | N | 0.8353664426 |
| 6.92 | 2+ | Central | 1.9348089832 | VC | valid | | . | . | . | . | . | . | N | 0.8003994659 |
| 4.06 | 2+ | Lower Inner | 1.401354915 | VC | valid | | . | . | . | . | . | . | N | 0.783013909 |
| 6.32 | 2+ | Central | 1.8433926629 | VIC | plan | A | 08MAY2007 | 20JUN2007 | 6.14 | L | 136.24 | 17.93 | N | 0.7264493317 |
| 5.75 | 2+ | Central | 1.7489946455 | CSI | valid | | . | . | . | . | . | . | N | 0.6431072195 |
| 2.92 | 2+ | Lower Inner | 1.0706760855 | VC | valid | | . | . | . | . | . | . | N | 0.6234507017 |
| 6.63 | 2+ | Lower Outer | 1.8911679748 | CSI | valid | | . | . | . | . | . | . | N | 0.6044447286 |
| 3.07 | 2+ | Upper Inner | 1.1208061072 | VIC | plan | B | 30AUG2007 | 04OCT2007 | 5.00 | R | 102.96 | 12.12 | N | 0.5855376375 |
| 2.97 | 2+ | Upper Inner | 1.0898938009 | CSI | valid | | . | . | . | . | . | . | N | 0.5700183463 |
| 2.97 | 2+ | Upper Inner | 1.0878161851 | VIC | plan | D | 27JUL2006 | 06SEP2006 | 5.86 | R | 298.22 | 11.29 | N | 0.568987181 |
| 5.24 | 2+ | Central | 1.6559490995 | VC | valid | | . | . | . | . | . | . | N | 0.5599176113 |
| 5.37 | 2+ | Upper Outer | 1.68145018 | VIC | plan | A | 25APR2006 | 06JUL2006 | 10.29 | R | 368.2 | 25.69 | N | 0.5499463684 |
| 5.26 | 2+ | Upper Outer | 1.6604449761 | CSI | valid | | . | . | . | . | . | . | N | 0.5310464599 |
| 2.60 | 2+ | Upper Inner | 0.955862992 | VIC | plan | C | 14MAR2007 | 20APR2007 | 5.29 | L | 219.19 | 14.11 | N | 0.5069313824 |
| 2.57 | 2+ | Upper Inner | 0.944709863 | VC | valid | | . | . | . | . | . | . | N | 0.5020175756 |
| 5.08 | 2+ | Upper Outer | 1.6251889879 | VIC | plan | D | 04SEP2007 | 24OCT2007 | 7.14 | L | 456.37 | 19.85 | N | 0.5000354978 |
| 2.56 | 2+ | Upper Inner | 0.9388477184 | VIC | plan | D | 07JUN2007 | 31JUL2007 | 7.71 | R | 100.5 | 9.76 | N | 0.4994565271 |
| 2.53 | 2+ | Upper Inner | 0.9288475763 | VIC | plan | C | 27APR2007 | 05JUL2007 | 9.86 | L | 178.8 | 15.69 | N | 0.4951223106 |
| 2.52 | 2+ | Upper Inner | 0.924026458 | VIC | plan | C | 05SEP2006 | 16OCT2006 | 5.86 | R | 295.27 | 17.50 | N | 0.4930484167 |
| 2.13 | 2+ | Lower Inner | 0.757688494 | CSI | valid | | . | . | . | . | . | . | N | 0.4924959025 |
| 2.04 | 2+ | Lower Inner | 0.7112868089 | VC | valid | | . | . | . | . | . | . | N | 0.4760355107 |
| 4.71 | 2+ | Central | 1.5492909389 | VIC | plan | D | 23OCT2006 | 08DEC2006 | 6.57 | R | 174.7 | 15.18 | N | 0.4699692178 |
| 2.37 | 2+ | Upper Inner | 0.8636365753 | VIC | plan | D | 29JUN2006 | 16AUG2006 | 6.86 | L | 70.4 | 13.58 | N | 0.4679401888 |

ROC Curve for Model
Area Under the Curve = 0.8005



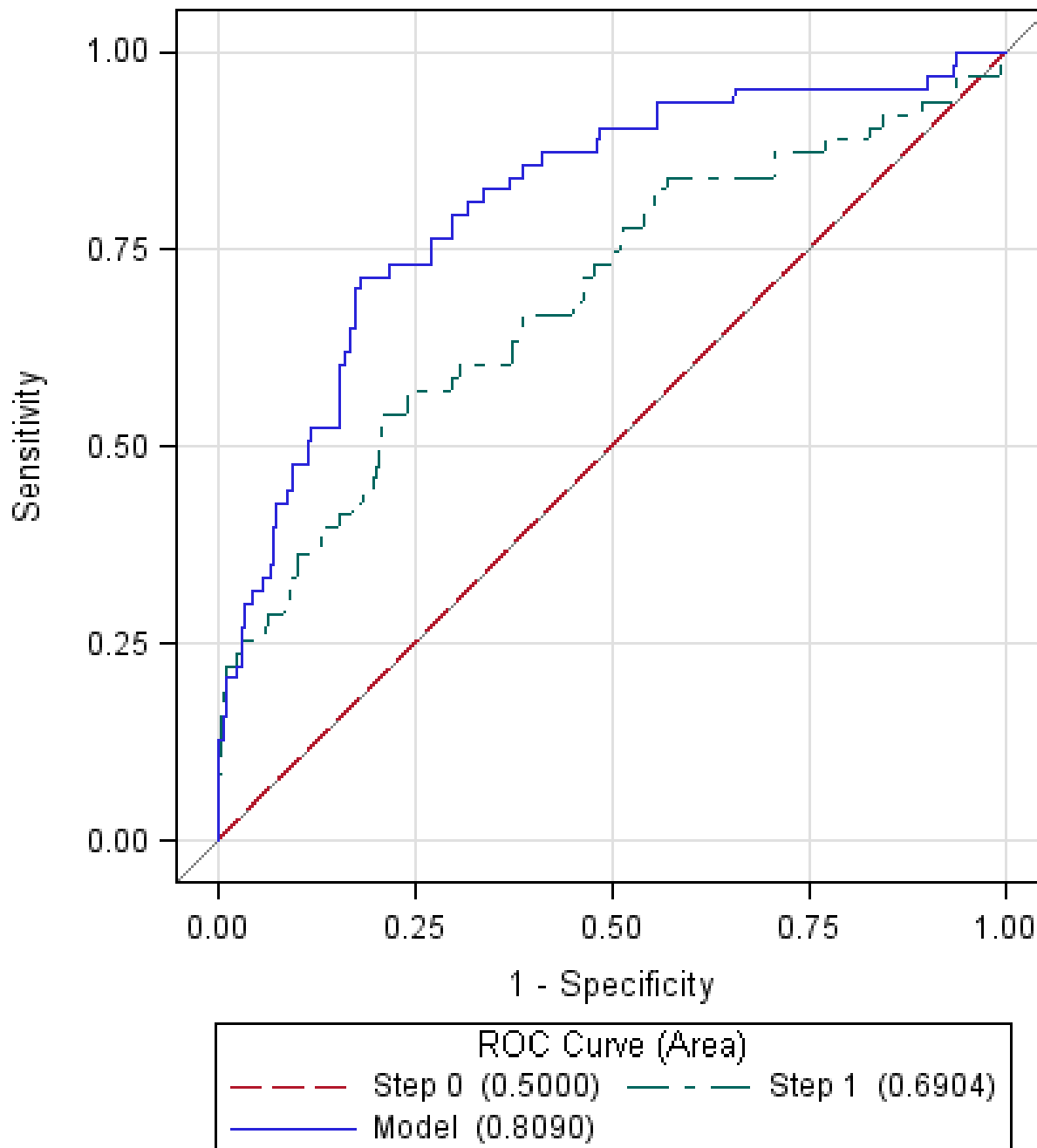
```
ods graphics on;  
  proc logistic DATA=dset PLOTS (ONLY) = (ROC (ID=prob) ) ;  
    CLASS quadrant / PARAM=glm;  
    MODEL partplan = quadrant cavtobr /  
        SELECTION=forward;  
run;
```

If an automated model selection method is specified in the
MODEL statement,

and a ROC plot has been requested in the PLOTS option,
or the OUTROC= option has been specified in the MODEL
statement

then a ROC plot for the final selected model is provided as well as
an overlaid plot of each of the ROC curves for each stage of the
model selection process.

ROC Curves for All Model Building Steps



Although the PLOTS option requests predicted probabilities to be shown, it does not include them on this overlaid plot. One can imagine how messy it could get.

```
ods graphics on;  
proc logistic DATA=dset PLOTS (ONLY) = (ROC (ID=prob) ) ;  
  CLASS quadrant / PARAM=glm;  
  MODEL partplan = quadrant cavtoibr / NOFIT;  
  ROC 'Quadrant' quadrant;  
  ROC 'Cavity to Breast Ratio' cavtoibr;  
run;
```

- If an automated model selection process is not used then the ROC statement can be used to obtain a ROC curve for each predictor of interest
- A ROC curve in which all of them are overlaid is also provided
- Effects used in any of the ROC statements must be specified in the MODEL statement

```
ods graphics on;
  proc logistic DATA=dset PLOTS (ONLY) = (ROC (ID=prob) ) ;
    CLASS quadrant / PARAM=glm;
    MODEL partplan = quadrant cavtoabr / NOFIT;
    ROC 'Quadrant' quadrant;
    ROC 'Cavity to Breast Ratio' cavtoabr;
run;
```

- The NOFIT option can be specified to instruct SAS to ignore fitting the model specified in the MODEL statement.
- A model for each of the specified ROC statements is fitted.
- An observation with a missing value for any of the variables specified in the MODEL statement is excluded, even if it is missing for a variable not specified in the ROC statement. This gives a common sample size for each ROC model.



Results

- Results
 - Logistic: The SAS System
 - Model Information
 - Observations Summary
 - Response Profile
 - Class Level Information
 - Global Score
 - ROC Model: Quadrant
 - Convergence Status
 - Fit Statistics
 - Global Tests
 - Type 3 Tests
 - Parameter Estimates
 - Odds Ratios
 - ROC Curve
 - ROC Model: Cavity to Breast Volume Re
 - Convergence Status
 - Fit Statistics
 - Global Tests
 - Parameter Estimates
 - Odds Ratios
 - ROC Curve
 - ROC Comparisons
 - ROC Curves
 - ROC Association Statistics

Output - (Untitled)

The SAS System 15:42 Friday, November

The LOGISTIC Procedure

Model Information

| | | |
|---------------------------|------------------|----------------------|
| Data Set | WORK.TEMP | |
| Response Variable | partplan | Partial Plan Suitabl |
| Number of Response Levels | 2 | |
| Model | binary logit | |
| Optimization Technique | Fisher's scoring | |

| | |
|-----------------------------|-----|
| Number of Observations Read | 347 |
| Number of Observations Used | 328 |

Response Profile

| Ordered Value | partplan | Total Frequency |
|---------------|----------|-----------------|
| 1 | N | 63 |
| 2 | Y | 265 |

Probability modeled is partplan='N'.

NOTE: 19 observations were deleted due to missing values for the response or explanatory variables.

Class Level Information

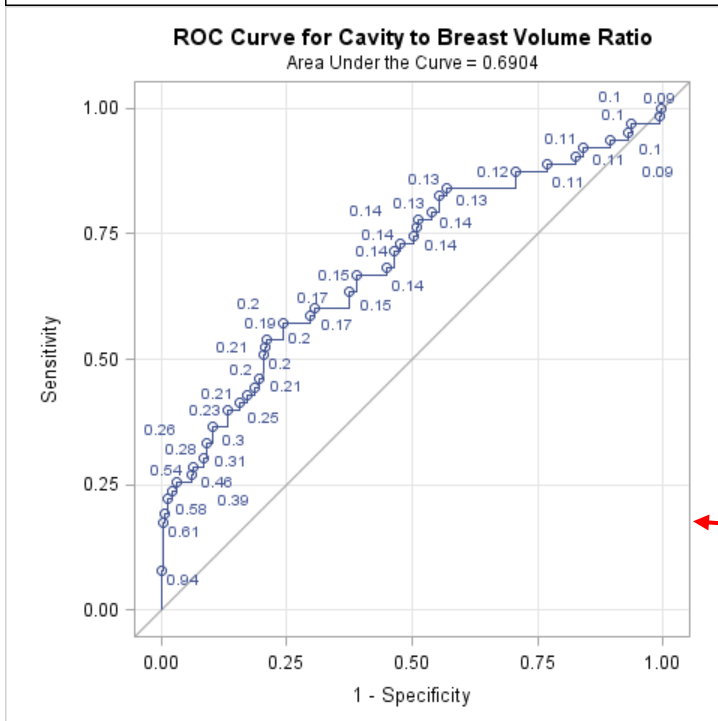
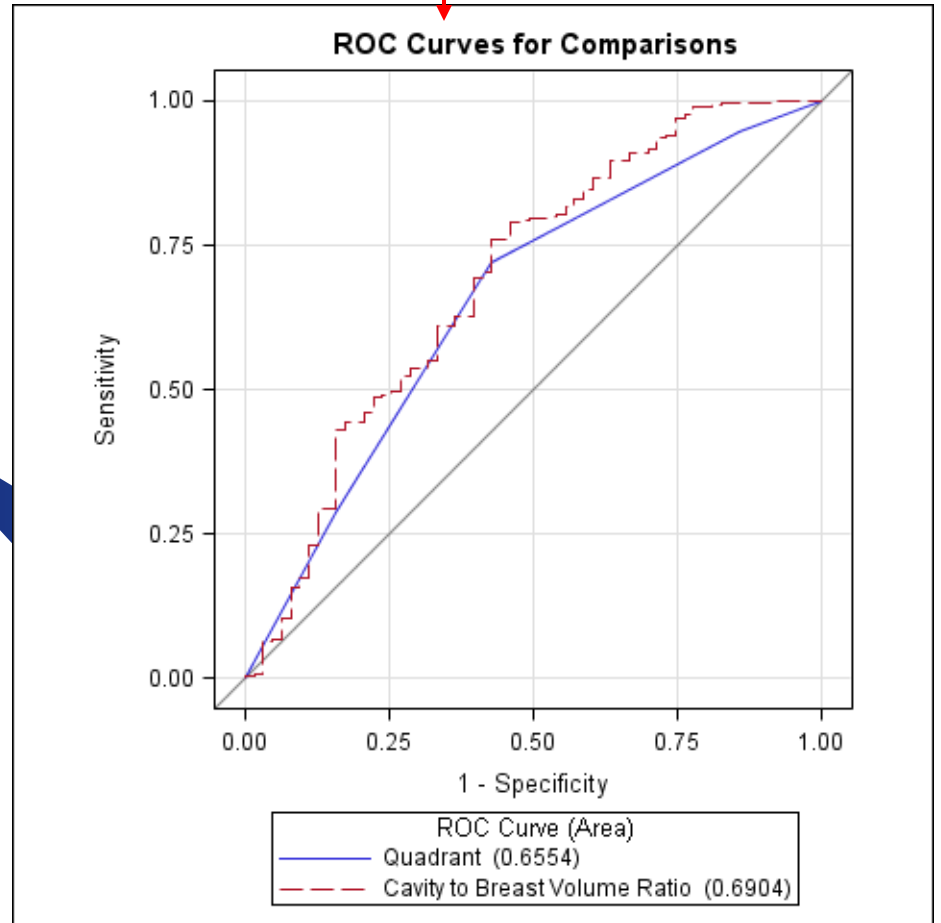
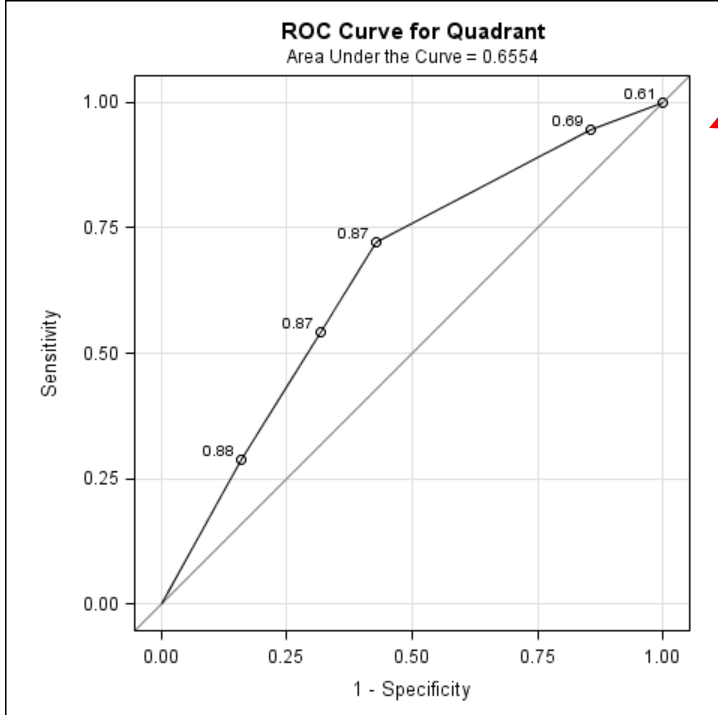
| Class | Value | Design Variables | | | | |
|----------|-------------|------------------|---|---|---|---|
| quadrant | Central | 1 | 0 | 0 | 0 | 0 |
| | Lower Inner | 0 | 1 | 0 | 0 | 0 |
| | Lower Outer | 0 | 0 | 1 | 0 | 0 |
| | Upper Inner | 0 | 0 | 0 | 1 | 0 |
| | Upper Outer | 0 | 0 | 0 | 0 | 1 |

Score Test for Global Null Hypothesis

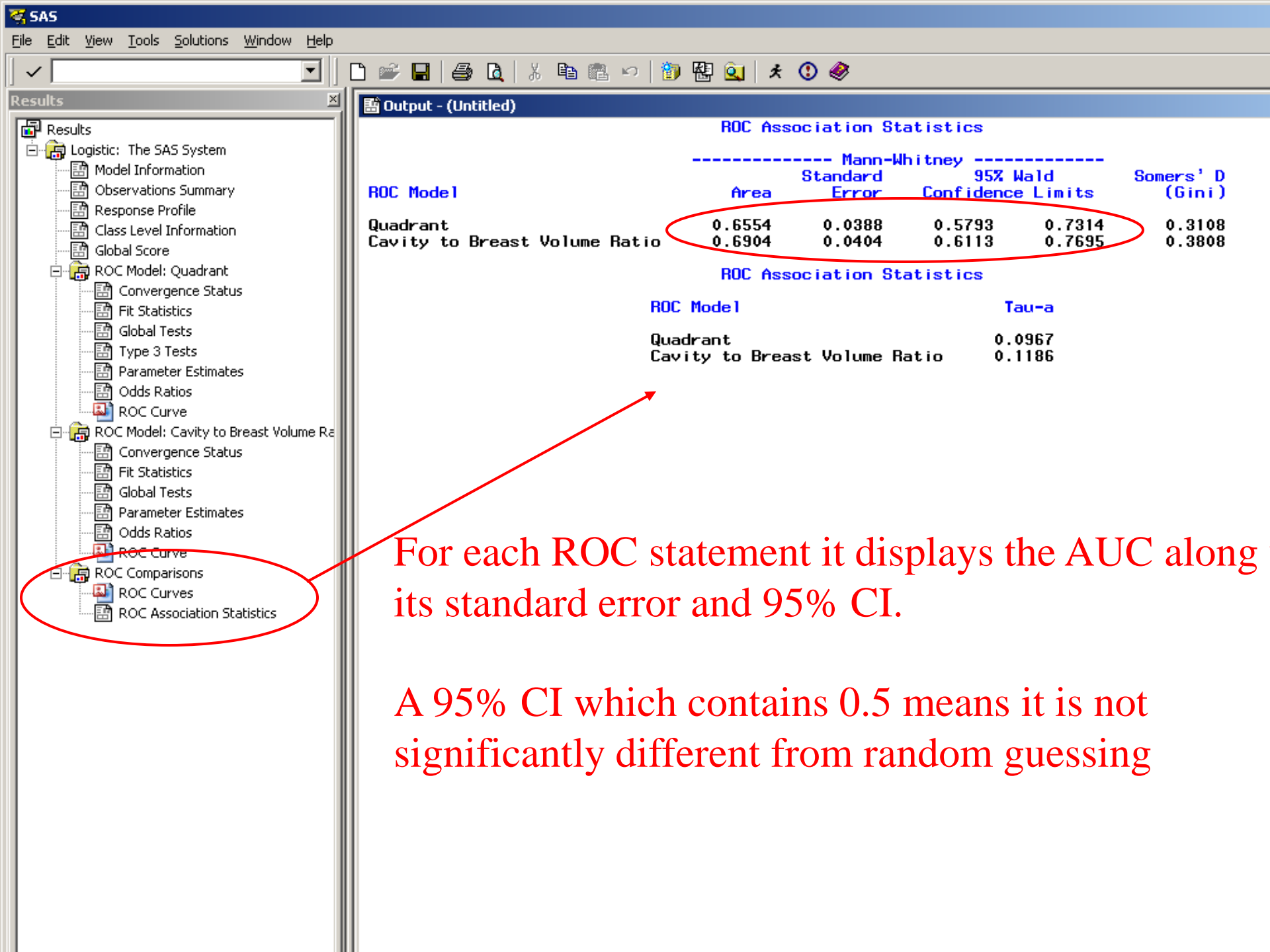
| Chi-Square | DF | Pr > ChiSq |
|------------|----|------------|
| 73.1051 | 5 | <.0001 |

From 1st ROC statement

Overlaid plot of each ROC statement plot



From 2nd ROC statement



- Results
 - Logistic: The SAS System
 - Model Information
 - Observations Summary
 - Response Profile
 - Class Level Information
 - Global Score
 - ROC Model: Quadrant
 - Convergence Status
 - Fit Statistics
 - Global Tests
 - Type 3 Tests
 - Parameter Estimates
 - Odds Ratios
 - ROC Curve
 - ROC Model: Cavity to Breast Volume Ratio
 - Convergence Status
 - Fit Statistics
 - Global Tests
 - Parameter Estimates
 - Odds Ratios
 - ROC Curve
 - ROC Comparisons
 - ROC Curves
 - ROC Association Statistics

Output - (Untitled)

ROC Association Statistics

----- Mann-Whitney -----

| ROC Model | Area | Standard Error | 95% Wald Confidence Limits | Somers' D (Gini) |
|-------------------------------|--------|----------------|----------------------------|------------------|
| Quadrant | 0.6554 | 0.0388 | 0.5793 0.7314 | 0.3108 |
| Cavity to Breast Volume Ratio | 0.6904 | 0.0404 | 0.6113 0.7695 | 0.3808 |

ROC Association Statistics

| ROC Model | Tau-a |
|-------------------------------|--------|
| Quadrant | 0.0967 |
| Cavity to Breast Volume Ratio | 0.1186 |

For each ROC statement it displays the AUC along with its standard error and 95% CI.

A 95% CI which contains 0.5 means it is not significantly different from random guessing

```

ods graphics on;
proc logistic DATA=dset
    PLOTS (ONLY) = (ROC (ID=prob) )
    ROCOPTIONS (NODETAILS OUT=roc_dset);
    CLASS quadrant / PARAM=glm;
    MODEL partplan = quadrant cavtobr / NOFIT;
    ROC 'Quadrant' quadrant;
    ROC 'Cavity to Breast Ratio' cavtobr;
run;

```

- ROCOPTIONS can be used to specify global options that apply to every ROC statement
- NODETAILS suppresses the model fitting information for models specified in the ROC statements
- OUT=<*dset name*> appends the ROC data for each ROC statement into one dataset (the alias for ROCOUT= in the MODEL statement)

```
ods graphics on;
proc logistic DATA=dset
    PLOTS (ONLY) = (ROC (ID=prob))
    ROCOPTIONS (NODETAILS) ;
    CLASS quadrant / PARAM=glm;
    MODEL partplan = quadrant cavtoibr cavvol brvol
                / NOFIT;
    ROC 'Ratio' quadrant cavtoibr;
    ROC 'Quadrant' quadrant;
    ROC 'Separate' quadrant cavvol brvol;
    ROCCONTRAST REFERENCE('Ratio') E ESTIMATE;
run;
```

- ID=prob asks for probability cut-offs to be shown on the ROC curves
- Use NODETAILS as not interested in the model fitting information for each of the ROC statements
- Requesting 3 models each with a separate ROC statement

```

ods graphics on;
proc logistic DATA=dset
    PLOTS (ONLY) = (ROC (ID=prob) )
    ROCOPTIONS (NODETAILS) ;
    CLASS quadrant / PARAM=glm;
    MODEL partplan = quadrant cavtoibr cavvol brvol
        / NOFIT;
    ROC 'Ratio' quadrant cavtoibr;
    ROC 'Quadrant' quadrant;
    ROC 'Separate' quadrant cavvol brvol;
    ROCCONTRAST REFERENCE('Ratio') E ESTIMATE;
run;

```

- Using the ROCCONTRAST statement to compare the ROC curves specified by the ROC statements
- The ROC curve labelled 'Ratio' will be the REFERENCE comparison curve
- E requests the contrast coefficients to be displayed

Output - (Untitled)

ROC Association Statistics

| ROC | ----- Mann-Whitney ----- | | | | Somers' D (Gini) | Gamma | Tau-a |
|----------|--------------------------|-------------------|-------------------------------|--------|---------------------|--------|--------|
| | Area | Standard Error | 95% Wald Confidence Limits | | | | |
| Quadrant | 0.6554 | 0.0388 | 0.5793 | 0.7314 | 0.3108 | 0.3936 | 0.0967 |
| Ratio | 0.8090 | 0.0315 | 0.7472 | 0.8707 | 0.6179 | 0.6179 | 0.1924 |
| Addition | 0.8025 | 0.0319 | 0.7400 | 0.8650 | 0.6050 | 0.6050 | 0.1883 |

ROC Contrast Coefficients

| ROC Model | Row1 | Row2 |
|-----------|------|------|
| Quadrant | 1 | 0 |
| Ratio | -1 | -1 |
| Addition | 0 | 1 |

ROC Contrast Test Results

| Contrast | DF | Chi-Square | Pr > ChiSq |
|-------------------|----|------------|------------|
| Reference = Ratio | 2 | 18.9817 | <.0001 |

ROC Contrast Rows Estimation and Testing Results

| Contrast | Estimate | Standard Error | 95% Wald Confidence Limits | Chi-Square | Pr > ChiSq | |
|------------------|----------|-------------------|-------------------------------|------------|---------------|--------|
| Quadrant - Ratio | -0.1536 | 0.0360 | -0.2242 | -0.0830 | 18.1930 | <.0001 |
| Addition - Ratio | -0.00647 | 0.0109 | -0.0279 | 0.0150 | 0.3497 | 0.5543 |

ROC Association
Statistics produced
when ROC statements
are specified

ROC Association Statistics

| ROC | Area | Mann-Whitney Standard Error | 95% Wald Confidence Limits | | Somers' D (Gini) | Gamma | Tau-a |
|----------|--------|--------------------------------|-------------------------------|--------|---------------------|--------|--------|
| Quadrant | 0.6554 | 0.0388 | 0.5793 | 0.7314 | 0.3108 | 0.3936 | 0.0967 |
| Ratio | 0.8090 | 0.0315 | 0.7472 | 0.8707 | 0.6179 | 0.6179 | 0.1924 |
| Addition | 0.8025 | 0.0319 | 0.7400 | 0.8650 | 0.6050 | 0.6050 | 0.1883 |

ROC Contrast Coefficients

| ROC Model | Row1 | Row2 |
|-----------|------|------|
| Quadrant | 1 | 0 |
| Ratio | -1 | -1 |
| Addition | 0 | 1 |

Requested by specifying E in the ROCCONTRAST statement

ROC Contrast Test Results

| Contrast | DF | Chi-Square | Pr > ChiSq |
|-------------------|----|------------|------------|
| Reference = Ratio | 2 | 18.9817 | <.0001 |

ROC Contrast Rows Estimation and Testing Results

| Contrast | Estimate | Standard Error | 95% Wald Confidence Limits | | Chi-Square | Pr > ChiSq |
|------------------|----------|----------------|----------------------------|---------|------------|------------|
| Quadrant - Ratio | -0.1536 | 0.0360 | -0.2242 | -0.0830 | 18.1930 | <.0001 |
| Addition - Ratio | -0.00647 | 0.0109 | -0.0279 | 0.0150 | 0.3497 | 0.5543 |

ROC Association Statistics

| ROC | Mann-Whitney | | | | Somers' D (Gini) | Gamma | Tau-a |
|----------|--------------|----------------|----------------------------|--------|------------------|--------|--------|
| | Area | Standard Error | 95% Wald Confidence Limits | | | | |
| Quadrant | 0.6554 | 0.0388 | 0.5793 | 0.7314 | 0.3108 | 0.3936 | 0.0967 |
| Ratio | 0.8090 | 0.0315 | 0.7472 | 0.8707 | 0.6179 | 0.6179 | 0.1924 |
| Addition | 0.8025 | 0.0319 | 0.7400 | 0.8650 | 0.6050 | 0.6050 | 0.1883 |

This tells us whether the reference group is different from at least one of the others at the 5% sig. level

ROC Contrast Coefficients

| ROC Model | Row1 | Row2 |
|-----------|------|------|
| Quadrant | 1 | 0 |
| Ratio | -1 | -1 |
| Addition | 0 | 1 |

If no options are specified in the ROCCONTRAST statement the Wald test for ROC contrast is provided by default

ROC Contrast Test Results

| Contrast | DF | Chi-Square | Pr > ChiSq |
|-------------------|----|------------|------------|
| Reference = Ratio | 2 | 18.9817 | <.0001 |

ROC Contrast Rows Estimation and Testing Results

| Contrast | Estimate | Standard Error | 95% Wald Confidence Limits | | Chi-Square | Pr > ChiSq |
|------------------|----------|----------------|----------------------------|---------|------------|------------|
| Quadrant - Ratio | -0.1536 | 0.0360 | -0.2242 | -0.0830 | 18.1930 | <.0001 |
| Addition - Ratio | -0.00647 | 0.0109 | -0.0279 | 0.0150 | 0.3497 | 0.5543 |

```
ods graphics on;
proc logistic DATA=dset
    PLOTS (ONLY) = (ROC (ID=prob) )
    ROCOPTIONS (NODETAILS) ;
    CLASS quadrant / PARAM=glm;
    MODEL partplan = quadrant cavtoibr cavvol brvol
        / NOFIT;
    ROC 'Ratio' quadrant cavtoibr;
    ROC 'Quadrant' quadrant;
    ROC 'Separate' quadrant cavvol brvol;
    ROCCONTRAST REFERENCE('Ratio') E ESTIMATE;
run;
```

- ESTIMATE requests that the comparison be made between the REFERENCE ROC and each ROC statement
- ESTIMATE=ALLPAIRS could be used to do all pair comparisons

ROC Association Statistics

| ROC | Area | Mann-Whitney Standard Error | 95% Wald Confidence Limits | | Somers' D (Gini) | Gamma | Tau-a |
|----------|--------|--------------------------------|-------------------------------|--------|---------------------|--------|--------|
| Quadrant | 0.6554 | 0.0388 | 0.5793 | 0.7314 | 0.3108 | 0.3936 | 0.0967 |
| Ratio | 0.8090 | 0.0315 | 0.7472 | 0.8707 | 0.6179 | 0.6179 | 0.1924 |
| Addition | 0.8025 | 0.0319 | 0.7400 | 0.8650 | 0.6050 | 0.6050 | 0.1883 |

ROC Contrast Coefficients

| ROC Model | Row1 | Row2 |
|-----------|------|------|
| Quadrant | 1 | 0 |
| Ratio | -1 | -1 |
| Addition | 0 | 1 |

ROC Contrast Test Results

| Contrast | DF | Chi-Square | Pr > ChiSq |
|-------------------|----|------------|------------|
| Reference = Ratio | 2 | 18.9817 | <.0001 |

ROC Contrast Rows Estimation and Testing Results

| Contrast | Estimate | Standard Error | 95% Wald Confidence Limits | | Chi-Square | Pr > ChiSq |
|------------------|----------|----------------|-------------------------------|---------|------------|------------|
| Quadrant - Ratio | -0.1536 | 0.0360 | -0.2242 | -0.0830 | 18.1930 | <.0001 |
| Addition - Ratio | -0.00647 | 0.0109 | -0.0279 | 0.0150 | 0.3497 | 0.5543 |

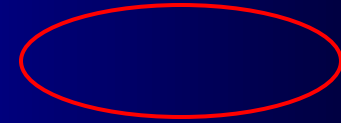
Output produced by specifying the ESTIMATE option in the ROCCONTRAST statement

Using a validation dataset

- The dataset that model is determined from is often called the ‘test’ dataset.
- The model can then be applied to a ‘validation’ dataset to see how well it performs on the validation dataset
- This can be done using the SCORE statement in proc logistic

```
ods graphics on;  
proc logistic DATA=dset OUTMODEL=model_name;  
    CLASS quadrant / PARAM=glm;  
    MODEL partplan = quadrant cavtobr;  
run;  
proc logistic INMODEL=model_name;  
    SCORE DATA=partplan OUT=score_dset CLM  
        OUTROC=score_roc;  
run;
```

- OUTMODEL



SUMMARY

- Using PLOTS= option to:
 - Produce ROC curve/AUC
 - Show predicted probs/observation # on plot
- Using OUTROC= option to:
 - Produce ROC curve/AUC
 - Obtaining ROC plot dataset
 - Utilise ROCEPS= option
- Using automated model selection process to obtain overlaid ROC plots for each model selection process

Type 3 Analysis of Effects

| Effect | DF | Wald Chi-Square | Pr > ChiSq |
|------------------|----|--------------------|------------|
| quadrant | 4 | 29.1317 | <.0001 |
| cav_breast_ratio | 1 | 32.5279 | <.0001 |

Analysis of Maximum Likelihood Estimates

| Parameter | | DF | Estimate | Standard Error | Wald Chi-Square | Pr > ChiSq |
|------------------|-------------|----|----------|-------------------|--------------------|------------|
| Intercept | | 1 | -3.4610 | 0.4697 | 54.2957 | <.0001 |
| quadrant | Central | 1 | 0.1326 | 0.5459 | 0.0590 | 0.8081 |
| quadrant | Lower Inner | 1 | 1.9773 | 0.6096 | 10.5215 | 0.0012 |
| quadrant | Lower Outer | 1 | -0.6308 | 0.6847 | 0.8486 | 0.3569 |
| quadrant | Upper Inner | 1 | 1.7165 | 0.4698 | 13.3508 | 0.0003 |
| quadrant | Upper Outer | 0 | 0 | . | . | . |
| cav_breast_ratio | | 1 | 0.6814 | 0.1195 | 32.5279 | <.0001 |

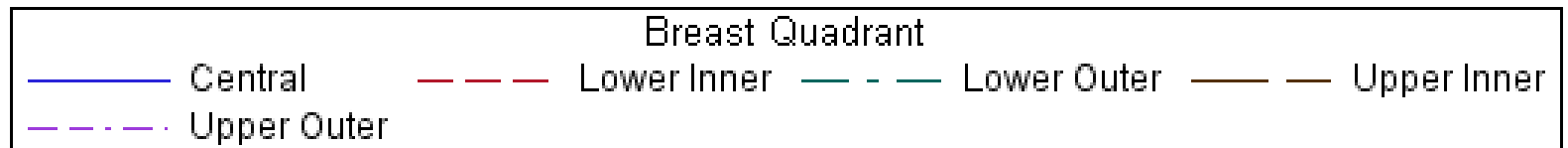
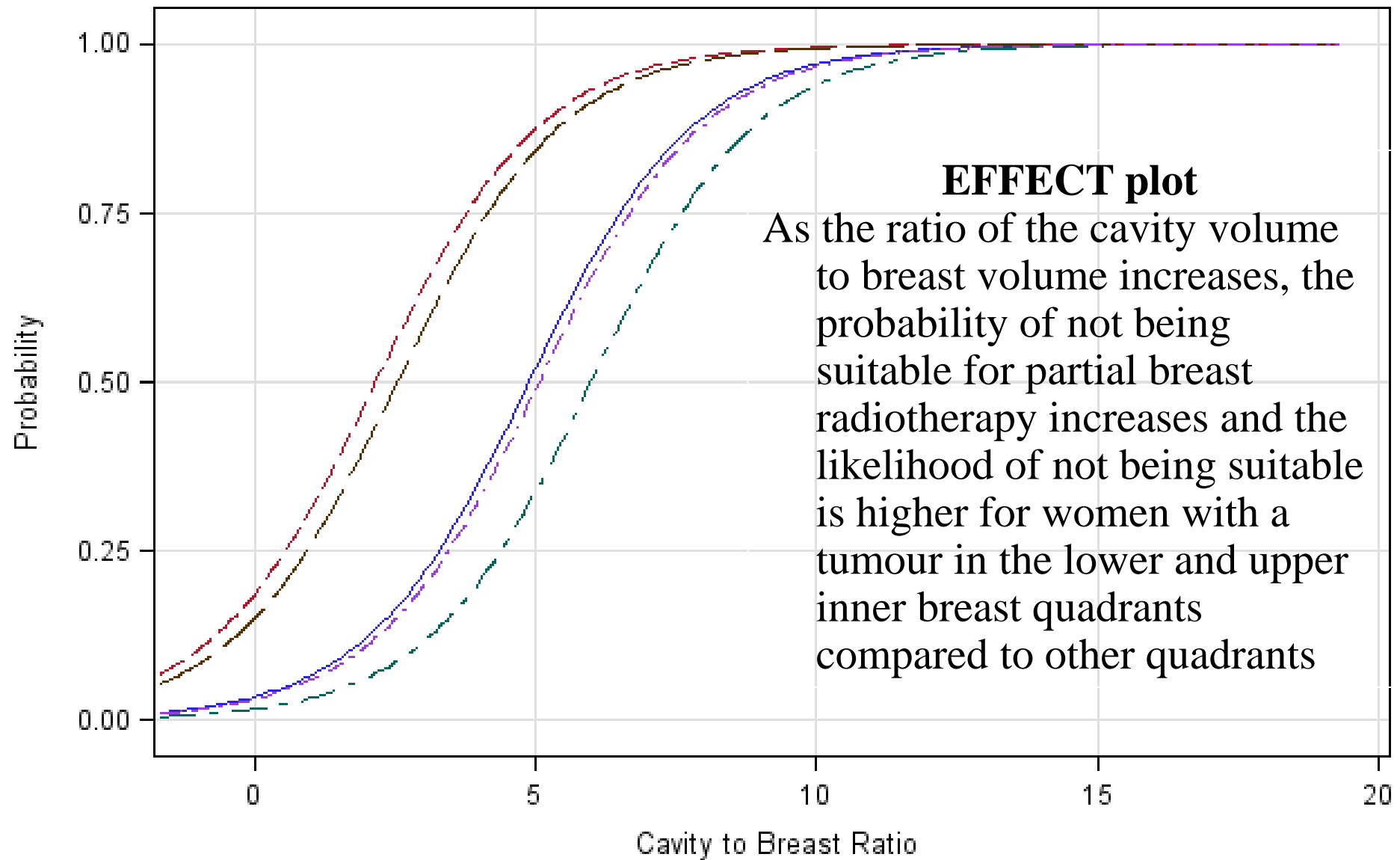
Odds Ratio Estimates

| Effect | | | Point Estimate | 95% Wald Confidence Limits | |
|------------------|-------------|----------------|-------------------|-------------------------------|--------|
| quadrant | Central | vs Upper Outer | 1.142 | 0.392 | 3.329 |
| quadrant | Lower Inner | vs Upper Outer | 7.223 | 2.187 | 23.857 |
| quadrant | Lower Outer | vs Upper Outer | 0.532 | 0.139 | 2.037 |
| quadrant | Upper Inner | vs Upper Outer | 5.565 | 2.216 | 13.974 |
| cav_breast_ratio | | | 1.977 | 1.564 | 2.498 |

Association of Predicted Probabilities and Observed Responses

| | | | |
|--------------------|-------|-----------|-------|
| Percent Concordant | 80.9 | Somers' D | 0.618 |
| Percent Discordant | 19.1 | Gamma | 0.618 |
| Percent Tied | 0.0 | Tau-a | 0.192 |
| Pairs | 16695 | c | 0.809 |

Predicted Probabilities for partplan=N



```
ods graphics on;  
proc logistic DATA=dset PLOTS (ONLY) = (ROC (ID=prob) ) ;  
  CLASS quadrant / PARAM=glm;  
  MODEL partplan = quadrant cav_vol /  
    OUTROC=dset_name;  
run;
```

If the ROC curve is requested in the PLOTS option but the OUTROC= option is not specified then the OUTROC dataset will not be provided.