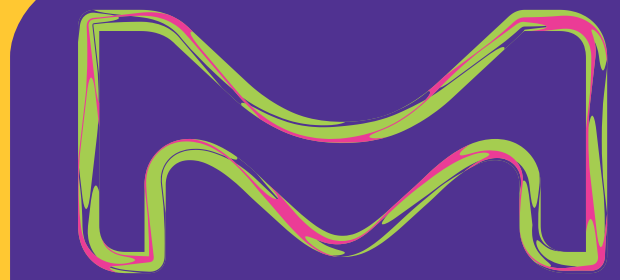


ROC CURVE ANALYSIS USING SAS

Zheng Yao

Sr. Statistical Programmer

MERCK

Outline

- **Background**
- **Examples:**
  - **Accuracy assessment**
  - **Compare ROC curves**
  - **Cut-off point selection**
- **Summary**

Outline

- **Background**
- Examples:
  - Accuracy assessment
  - Compare ROC curves
  - Cut-off point selection
- Summary

Background

- Biomarkers (e.g. PD-1/L1) draw lots of attention nowadays.
- It is often of interest to use biomarker for disease screening, diagnosis and prediction.
- The fundamental for use of biomarkers in clinical practice is the **accuracy** and the **optimal cut-off point selection**
- The receiver operating characteristic (ROC) curve is a procedure that can aid in the **accuracy assessment, ROC curve comparison and cut-off point selection**.

Background

	Golden Standard Outcome	
Cut Point Outcome	Positive	Negative
Positive	**TP (a)**	FP (b)
Negative	FN (c)	**TN (d)**

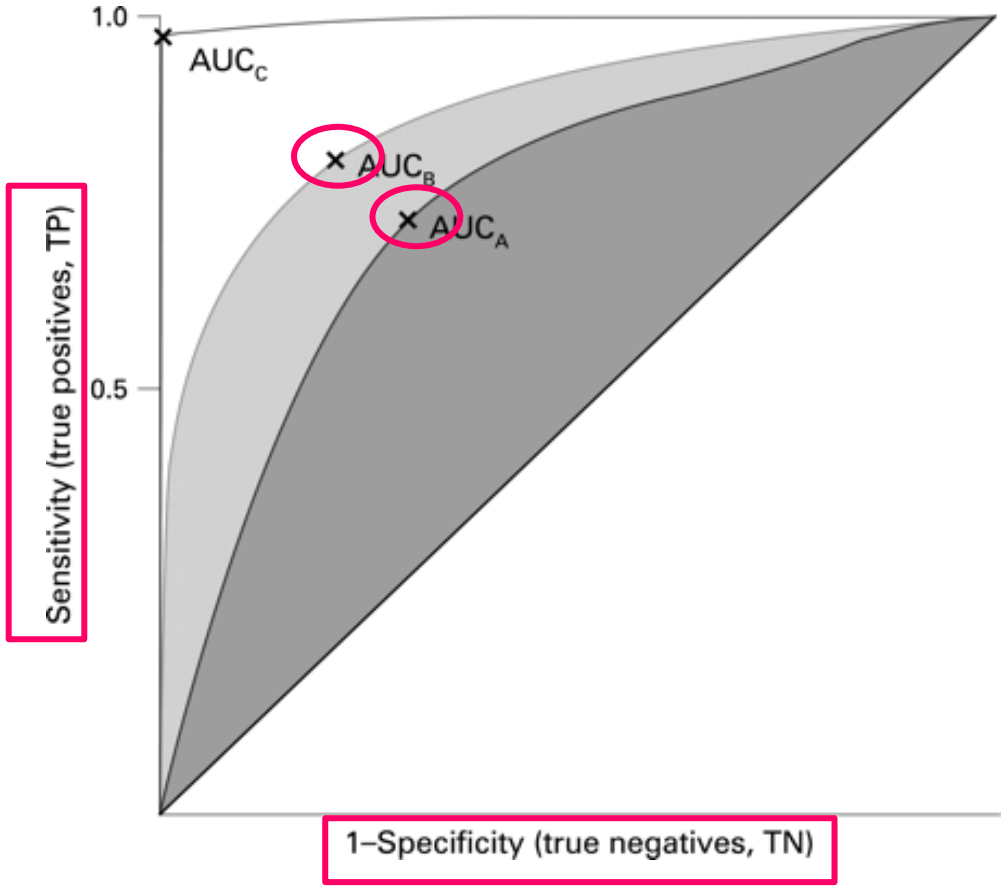
*Note: **TP**=True Positive, **FP**=False Positive, **FN**=False Negative, and **TN**=True Negative.*

	Golden Standard Outcome	
Cut Point Outcome	Positive	Negative
Positive	**60**	20
Negative	40	**80**
Total	100	**100**

- **Sensitivity:** the proportion of positive observations that are measured as positive, i.e. true positive rate (TPR), **Sensitivity = a / (a + c)**
- **Specificity:** the proportion of negative observations that are measured as negative, i.e. true negative (TNR), **Specificity=d / (b + d);**
- **Youden's Index: (sensitivity + specificity) - 1**
  - Sensitivity = 60 / (60+ 40) = 0.6
  - Specificity = 80 / (20+ 80) = 0.8
  - Youden's Index = (sensitivity + specificity) – 1 = 0.6 + 0.8 - 1 = 0.4

Background

- ROC(**R**eceiver-**O**perating **C**haracteristic) Curve: constructed with **sensitivity (FP)** on the vertical axis and **1-specificity (TP)** on the horizontal
- Area under curve (**AUC**): a measure of overall **accuracy**
- ROC curves comparison
- Cut-off point selection
  - Youden's Index: (sensitivity + specificity) - 1



FP

Outline

- Background
- **Examples:**
  - **Accuracy assessment**
  - **Compare ROC curves**
  - Cut-off point selection
- Summary

Compare ROC curves

- 156 cases of patients with positive PD-L1 expression were selected and treated with anti-PD-L1 agents (drug A).
- To explores the potential association between PD-L1 expression in tumor tissue and anti-PD-L1 response (i.e. BOR, PFS or OS)
- The **expression of PD-L1** was measured with with two method (**TC** and **IC**) for each case
- Best Overall Response (**BOR**) is used as the **predicitve measure** according to (RECIST 1.1, 1). The BOR variable is coded as '1' and '0' (i.e. effective and noneffective) .

	BOR	
PD-L1 expression	Effective	Non-effective
Positive	**TP (a)**	FP (b)
Negative	FN (c)	**TN (d)**

Obs	USUBJID	TC	IC	BOR
1	001	15.814	2.823	0
2	002	8.991	15.975	1
3	003	4.477	5.46	0
4	004	17.55	12.447	1
5	005	0.604	3.217	0

NOTE: *Partial list output of the BOR dataset (obs=5)*

Compare ROC curves

- **Is IC better than TC?**

#:for example: Fertility: Hand in hand in Europe, GM: Ferrara IME program

* Forum include scientific workshops and HCP trainings

Compare ROC curves

The PROC LOGISTIC procedure for ROC curve comparison

- **TC** and **IC** are both independent variables in the model statement.
- the **ROC** statement produces a ROC
- the **ROCCONTRAST** statement produces a significance test for the ROC curve.
- the **REFERENCE**(‘**TC**’) statement means that TC is set as a reference when comparing with IC in the significance test.

```
proc logistic data = BOR;
  model BOR(event='1') = TC IC;
  roc"TC" TC;
  roc"IC" IC;
  roccontrast reference("TC")/estimate;
run;
```

Note. *Reference code for ROC curve comparison*

Compare ROC curves

Results from Logistic model

- The significance test demonstrates that TC (p = 0.0014) and IC (p< 0.0001) are statistically significant for use in ROC curve.
- Logistic regression equation: $\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x_1 + \cdots \beta_n x_n$

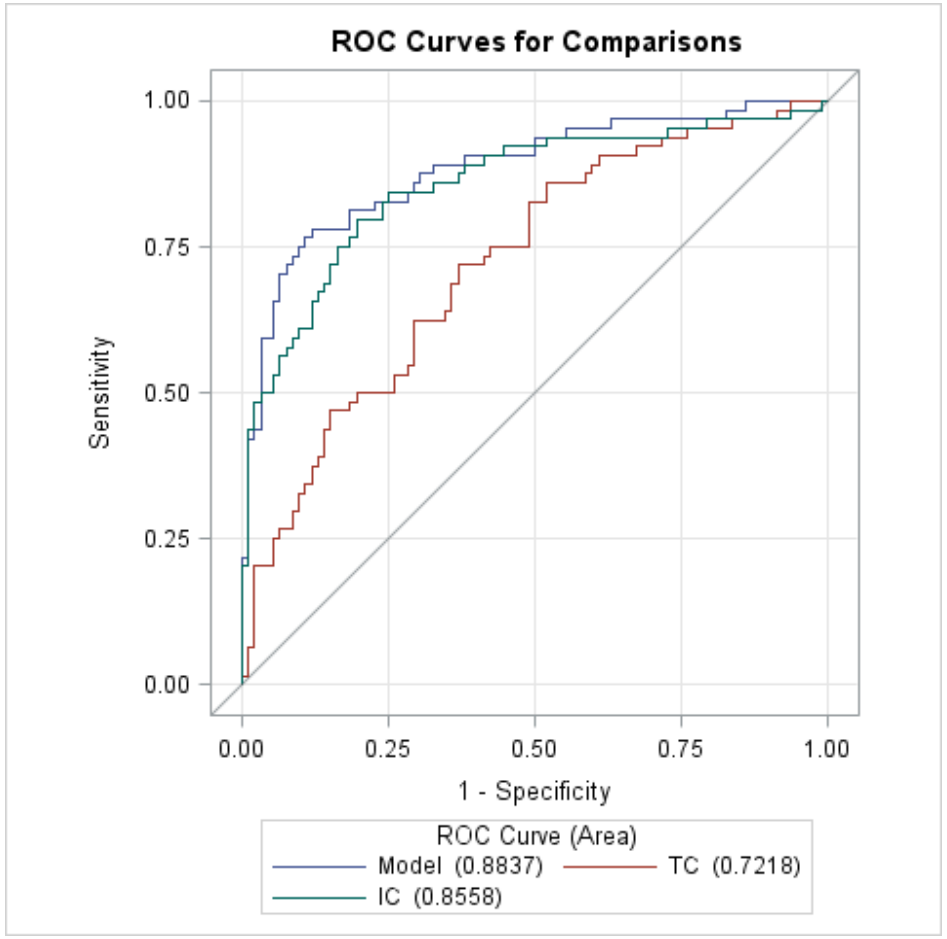
Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-3.9048	0.6434	36.8311	<.0001
TC	1	0.1104	0.0344	10.2640	0.0014
IC	1	0.1905	0.0364	27.3634	<.0001

Compare ROC curves

AUC statistical test and ROC curve

- **ROC Curve:** all of three ROC curves are above the diagonal line.
- **AUC:** all of three 95% CIs do not contain 0.5. Therefore, we can conclude that all these three AUC are significantly better than chance.

ROC Association Statistics							
	Mann-Whitney						
ROC Model	Area	Standard Error	95% Wald Confidence Limits		Somers' D (Gini)	Gamma	Tau-a
Model	0.8837	0.0285	0.8277	0.9396	0.7673	0.7673	0.3737
TC	0.7218	0.0410	0.6414	0.8022	0.4436	0.4436	0.2160
IC	0.8558	0.0324	0.7923	0.9193	0.7116	0.7116	0.3466



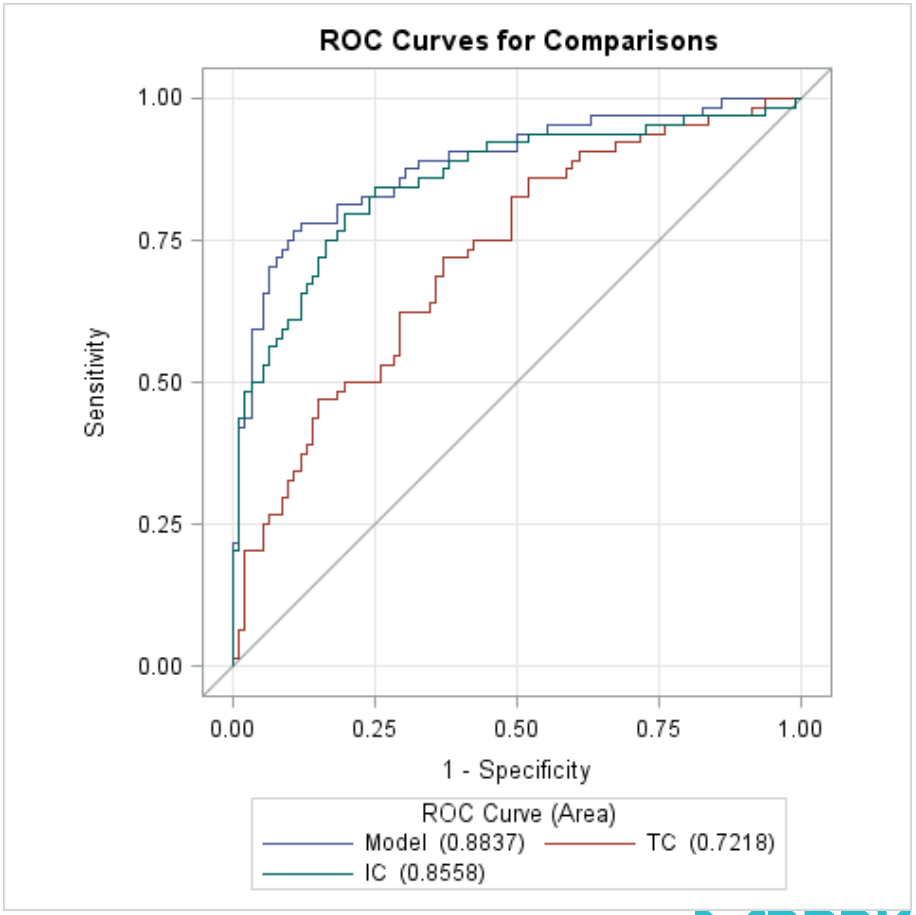
Compare ROC curves

ROC comparison test

- IC vs. TC (p=0.0116). Therefore, the AUC of IC (0.8558) is statistically larger than that of TC (0.7218).
- the PD-L1 scoring methodology of IC is better than TC.

ROC Contrast Estimation and Testing Results by Row						
Contrast	Estimate	Standard Error	95% Wald Confidence Limits		Chi-Square	Pr > ChiSq
Model - TC	0.1619	0.0405	0.0825	0.2412	15.9746	<.0001
IC - TC	0.1340	0.0531	0.0300	0.2380	6.3740	0.0116

Note.



Outline

- Background
- **Examples:**
  - Accuracy assessment
  - Compare ROC curves
  - **Cut-off point selection**
- Summary

Select a rational cut-off point in ROC curve analysis

The PROC LOGISTIC procedure for ROC curve analysis

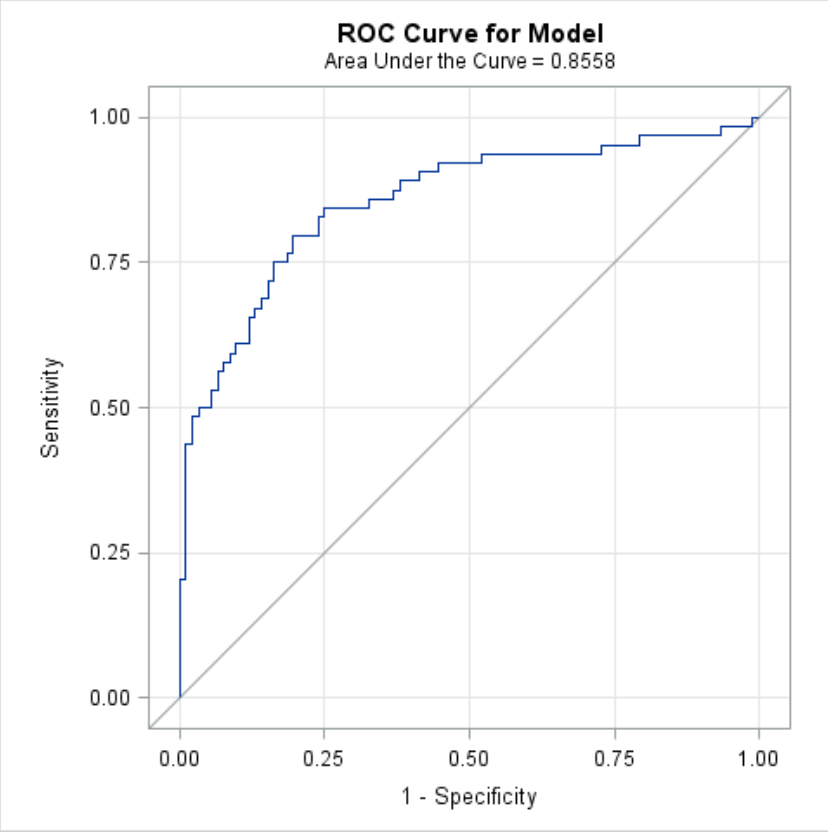
- The **OUTROC=** option creates a dataset containing sensitivity and specificity data which here is called ROCDATA.
- The **ROC** statement produces a ROC
- the **ROCCONTRAST** statement produces a significance test for the ROC curve.
- The **PREDICTED=** option creates a dataset containing *estimated event probabilities* (i.e. pred) for each subject.

```
proc logistic data = BOR;
  model BOR(event='1') = IC/outroc=rocdata;
  output out=pred predicted=pred;
  roc "BOR";
  roccontrast;
run;
```

Select a rational cut-off point in ROC curve analysis

The ROC curve

- The diagonal line, from (0,0) to (1,1), is indicative of an independent variable that discriminates no different from guessing (50/50 chance).
- The **AUC** is 0.8558 as compared to that of the diagonal line which is always 0.5.



Select a rational cut-off point in ROC curve analysis

AUC statistical test

- The **95% confidence interval** (0.6414, 0.8022) does not contain 0.5, therefore our AUC is significantly better than chance.
- A **Chi-square test** provides a p-value ($p < .0001$) associated with the null hypothesis (AUC =0.5).

ROC Association Statistics

ROC Model	Mann-Whitney Area	Mann-Whitney Standard Error	Mann-Whitney 95% Wald Confidence Limits		Somers' D (Gini)	Gamma	Tau-a
Model	0.8558	0.0324	0.7923	0.9193	0.7116	0.7116	0.3466
BOR	0.5000	0	0.5000	0.5000	0	.	0

ROC Contrast Test Results

Contrast	DF	Chi-Square	Pr > ChiSq
Reference = Model	1	120.4840	<.0001

***Note**. If the 95% confidence interval does not include 0.5, then we can conclude that ROC curve is statistically significant.*

Select a rational cut-off point in ROC curve analysis

The model's intercept and regression coefficients

- Partial output from the PROC LOGISTIC procedure
- Logistic regression equation: $\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x_1 + \cdots \beta_n x_n$

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-2.6197	0.4261	37.8058	<.0001
IC	1	0.1880	0.0342	30.1969	<.0001

NOTE: *The model's **intercept** (-1.6866) and **regression coefficients** (0.1122) are needed for computing the cutoff point.*

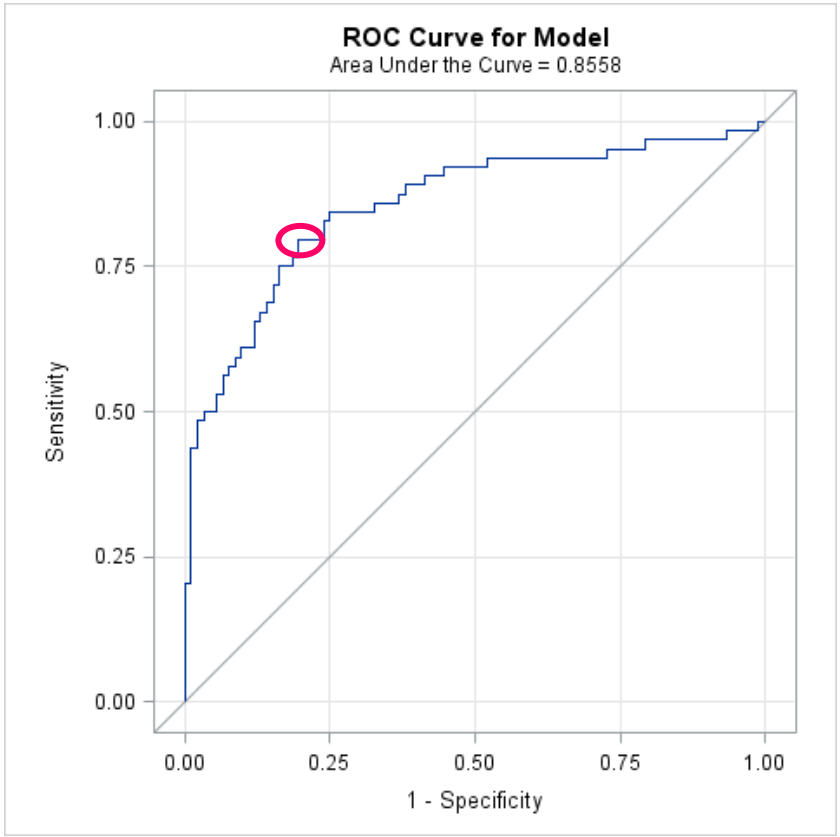
Select a rational cut-off point in ROC curve analysis

Determine an optimal PD-L1 cutoff point for BOR

- Logistic regression equation: $\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x_1 + \cdots \beta_n x_n$
- $\ln\left(\frac{p}{1-p}\right)$ = **intercept** + **slope**(X) = **-2.6197** + **0.1880***cutoff

Obs	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	0.93930	1	92	0	63	0.015625	0.000000
2	0.93369	1	91	1	63	0.015625	0.010870
3	0.87087	2	91	1	62	0.031250	0.010870
4	0.85808	3	91	1	61	0.046875	0.010870
5	0.84359	4	91	1	60	0.062500	0.010870

Note: *Partial list output of the* ***rocdata*** *dataset*



Select a rational cut-off point in ROC curve analysis

Determine an optimal cutoff point for BOR

- Logistic regression equation: $\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x_1 + \cdots \beta_n x_n$
- logit = **intercept** + **slope**(X) = **-2.6197** + **0.1880***cutoff

```
data rocdata2(keep=cutoff prob Sensitivity Specificity
Youden);
set rocdata;
  logit=log(_prob_/(1-_prob_));
  cutoff=(logit+2.6197)/0.1880;

  prob= _prob_;
  Sensitivity = _SENSIT_;
  Specificity = 1-_1MSPEC_;
  Youden= _SENSIT_ + (1-_1MSPEC_)-1;
run;
```

***Note**. The cutoff variable is formed by re-arranging logistic regression model to solve for X. The model is: logit = intercept + slope(x).*

Obs	cutoff	prob	Sensitivity	Specificity	Youden
1	10.38	0.37240	0.797	0.804	0.601
2	7.62	0.30321	0.844	0.750	0.594
3	9.18	0.34148	0.797	0.793	0.590
4	7.98	0.31185	0.828	0.761	0.589
5	12.54	0.43045	0.750	0.837	0.587

***Note**. Partial output of cutoff value and Youden Index by sorting the dataset in **descending** Youden. A large Youden may be one criteria for deciding an appropriate cutoff score.*

Outline

- Background
- Examples:
  - Accuracy assessment
  - Compare ROC curves
  - Cut-off point selection
- **Summary**
  - ***Accuracy assessment***
  - ***ROC curve comparison***
  - ***Cut-off point selection***

BACK UP SLIDES