Handout # 11

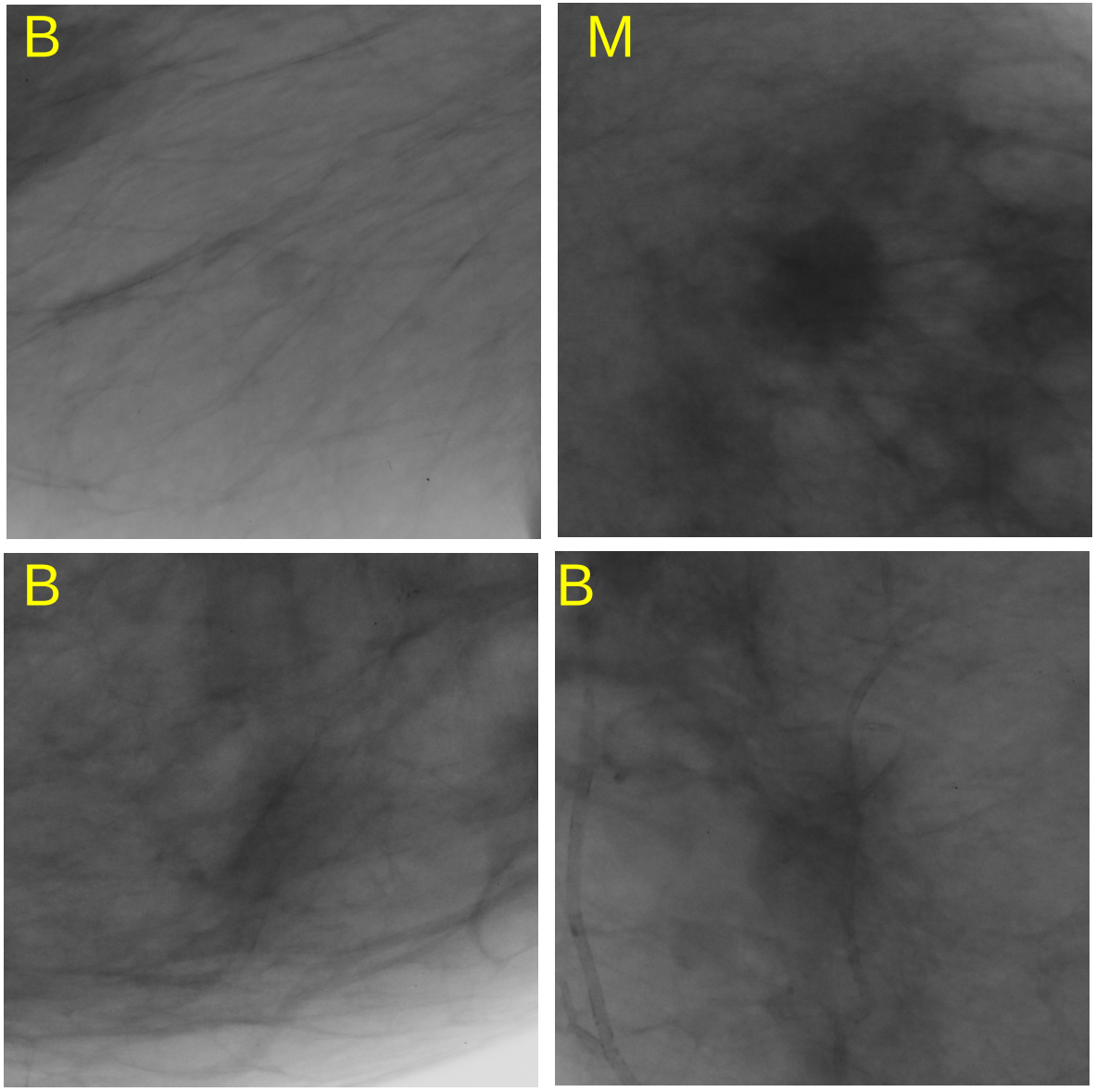
Receiver Operating Characteristic Analysis (I)

11. 1 Background

The **bottom line of medical imaging** is the extents to which images help the physician arrive at correct diagnoses?

Receiver Operating Characteristic (ROC) analysis is by far the most accepted empirical method to compare different systems, different techniques (dosage, pulse sequences, etc), different image processing algorithms (before and after image processing procedures)…

❖ Example of medical images



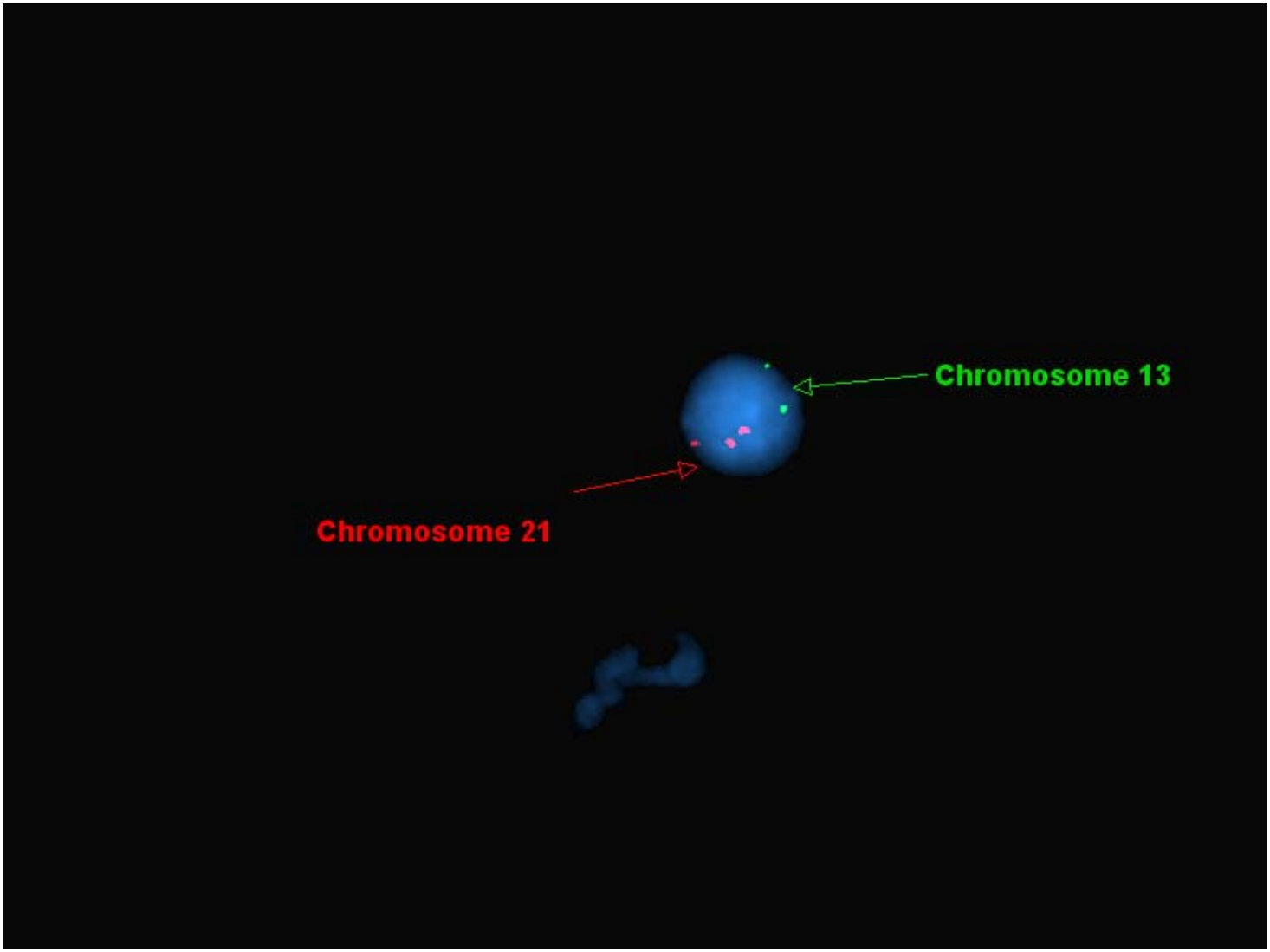
Medical images to assist breast cancer diagnosis

Location Uncertainty

Truth Uncertainty

Reader Variance

❖ Examples of medical images



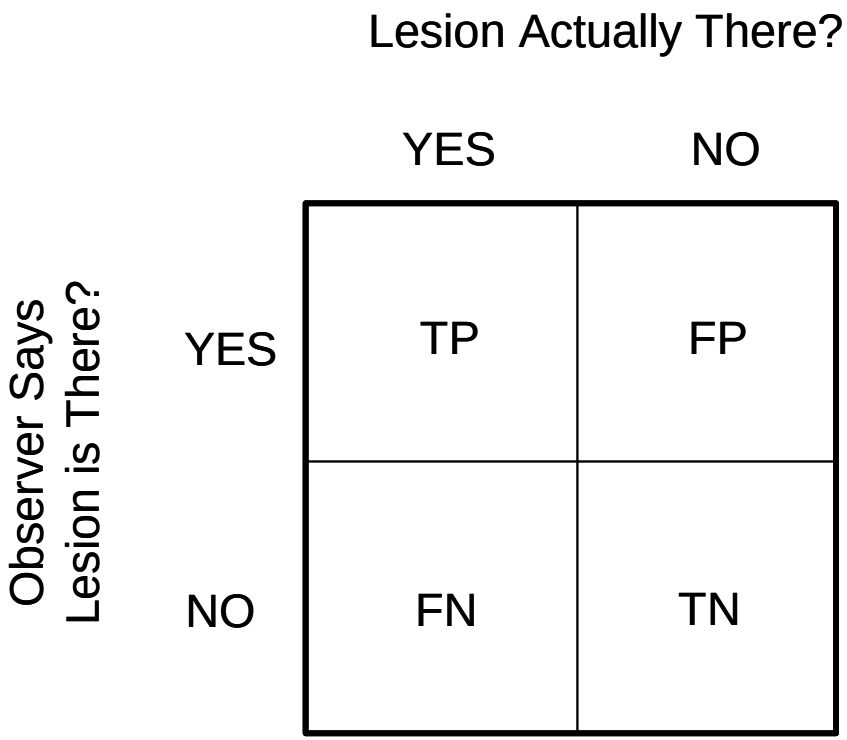
11.2 Several Concepts:

True positive (TP): You say the lesion is there when it is really there (correct positive assessment);

False positive (FP): You say the lesion is there when it is really NOT there (wrong positive assessment);

True negative (TN): You say the lesion is NOT there when it is really NOT there (correct negative assessment);

False negative (FN): You say the lesion is NOT there when it is really there (wrong negative assessment).



❖ More about TP, FP, FN, TN

	Object is present	Object is not present
Object is observed	**True Positive**	**False Positive**
Object is not observed	**False Negative**	**True Negative**

Sensitivity is the ability to detect a lesion when it is really there.

Sensitivity = number of correct positive assessments / number of truly positive cases

= TP / (TP+FN)

Sensitivity can be expressed as **True Positive Fraction (TPF)**

Specificity: is the ability to say that the lesion is absent when it is really not there.

Specificity = number of correct negative assessments / number of truly negative cases

= TN / (TN+FP)

Specificity can be expressed as **True negative Fraction (TNF)**

Accuracy = number of correct assessments / total number of cases

= TP+TN / (TP+TN+FP+FN)

Numbers game !!!

High sensitivity --- how many positive cases in your pool?

High specificity --- how many negative cases in your pool?

More about Sensitivity (TPF), and Specificity (TNF)

$$TPF = \frac{\textbf{\textit{True Positive Cases}}}{\textbf{\textit{Total Actual Positive Cases}}} = \frac{TP}{TP + FN}$$

$$TNF = \frac{\textbf{\textit{True Negative Cases}}}{\textbf{\textit{Total Actual Negative Cases}}} = \frac{TN}{TN + FP}$$

$$FNF = \frac{\textbf{\textit{False Negative Cases}}}{\textbf{\textit{Total Actual Positive Cases}}} = \frac{FN}{TP + FN}$$

$$FPF = \frac{\textbf{\textit{False Positive Cases}}}{\textbf{\textit{Total Actual Negative Cases}}} = \frac{FP}{FN + FP}$$

Sensitivity= **TPF** = 1 – **FNF**

Specificity = **TNF** =1 – **FPF**

$$\textbf{Accuracy} = \frac{\text{True Positive and Negative Cases}}{\text{Total Cases}}$$

11.3 Rating Your Observations: Confidence Levels

(1) A single decision threshold: True or False; Present or Absent

(2) A continuously-distributed scoring method,

For examples, observers can rate any scores from 0 to 100. in the reading (e.g., 45, 68, and 95). The larger score means the higher probability of presence of the targets

(3) Five confidence levels

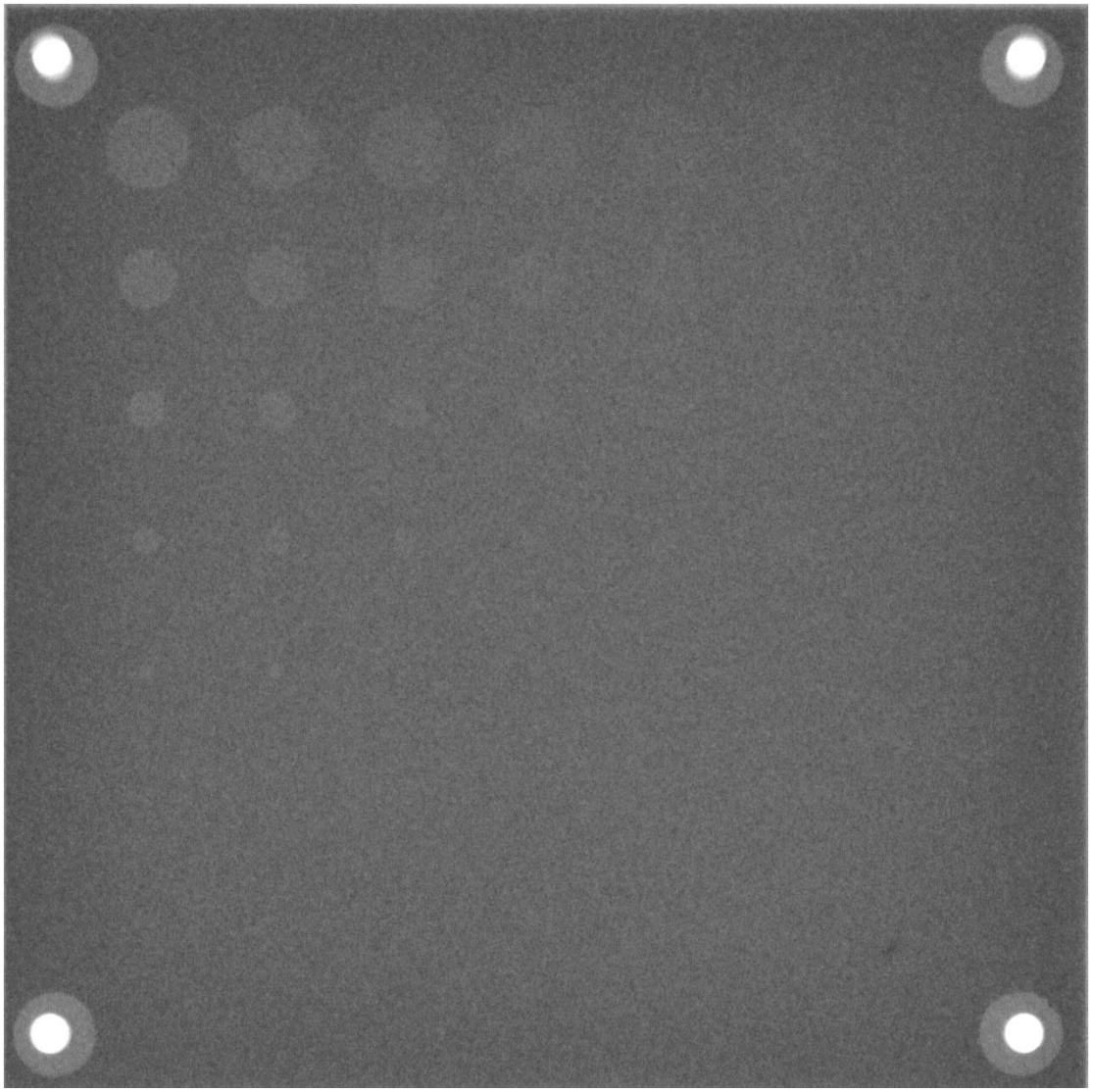
Five confidence levels:

- Definitely not present 1
- Probably not present 2
- Uncertain 3
- Probably present 4
- Definitely present 5

The choice of **decision threshold** determines the trade-off between **sensitivity** and **specificity.**

High sensitivity may be achieved at the cost of **low specificity.**

❖ **More information about five confidence levels:**



11.4 ROC Curves

An ROC curve is a plot of the sensitivity versus one minus specificity (i.e. TPF vs FPF) .

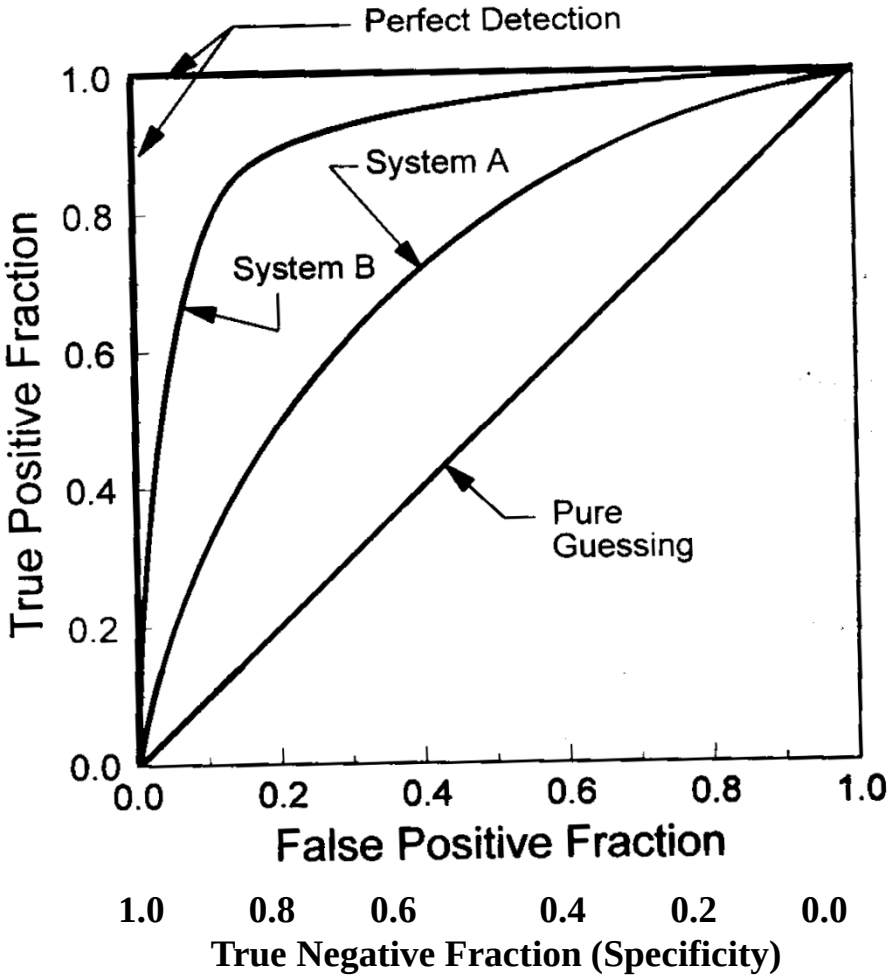
- **A single decision threshold** allows the assessment of **TP**, **TN**, **FP**, and **FN** as defined above.
- From these values the true-positive fraction is calculated as TP/(TP + FN), the true-positive fraction is the **sensitivity.**
- Then the false-positive fraction is calculated as FP/(FP + TN), and the false-positive fraction is equal to **one minus specificity.**
- **A single decision threshold would produce only a single point along the ROC curve**.

- **The entire curve is generated by sliding the decision threshold along the decision parameters.**

- At each position of the decision threshold, a separate sensitivity and specificity can be calculated.

- **An ROC curve is a plot of all those points linked together.**

- Because an ideal imaging system has 100% sensitivity and 100% specificity, it is represented by a line running along the left and top borders of the ROC plot.
- Pure guessing would result in the diagonal line shown.
- **Better imaging systems have curves that come closer to the upper left corner in the ROC curve.**
- System B is therefore better than system A, as shown here.



J.T. Bushberg, et al, *The Essential Physics of Medical Imaging*, Edition 2, p288-291, 2002

11.5 Methods to Plot an ROC Curve

- The empirical ROC Curve.
- The conventional ROC Curve.

Plot ROC curve with the empirical method, An Example

Assuming there are a total of 200 cases, among them, 100 benign cases (no lesion) and 100 malignant cases (lesion present), an observation study resulted the following

	Observer's Interpretation/ (BI-RADs Score)				
	Definitely no lesion	Probably no lesion	Not sure	Probably with lesion	Definitely with lesion
For 100 Benign cases	55	15	10	12	8
For 100 malignant cases	15	10	5	50	20
Total	70	25	15	62	28

Benign Diagnosis **Malignant diagnosis**

1

Using the threshold 1:

		P Lesion is present	N Lesion is not present
P	Lesion is observed	**10+5+50+20=85**	15+10+12+8=45
N	Lesion is not observed	15	**55**

Sensitivity (TPF) = **85**/100 FNF = 1- TPF = 15/100

Specificity (TNF) = **55**/100 FPF = 1-TNF = 45/100

Accuracy = (**85**+**55**)/200 = 70%

An Example: Plot ROC curve with the empirical method

Assuming there are a total of 200 cases, among them, 100 benign cases (no lesion) and 100 malignant cases (lesion present), an observation study resulted the following

	Observer's Interpretation/ (BI-RADs Score)				
	Definitely no lesion	Probably no lesion	Not sure	Probably with lesion	Definitely with lesion
For 100 Benign cases	**55**	**15**	10	12	8
For 100 malignant cases	15	10	**5**	**50**	**20**
Total	70	25	15	62	28

Benign Diagnosis | **Malignant diagnosis**

2

Using the threshold 2:

		P Object is present	N Object is not present
P	Object is observed	**5+50+20=75**	**10+12+8=30**
N	Object is not observed	15+10 = 25	**55+15 = 70**

Sensitivity (TPF) = **75**/100 FNF = 1- TPF = 25/100

Specificity (TNF) = **70**/100 FPF = 1-TNF = 30/100

Accuracy = (**75**+**70**)/200 = 72.5%

An Example: Plot ROC curve with the empirical method

Assuming there are a total of 200 cases, among them, 100 benign cases (no lesion) and 100 malignant cases (lesion present), an observation study resulted the following

	Observer's Interpretation/ (BI-RADs Score)				
	Definitely no lesion	Probably no lesion	Not sure	Probably with lesion	Definitely with lesion
For 100 Benign cases	**55**	**15**	**10**	12	8
For 100 malignant cases	15	10	5	**50**	**20**
Total	70	25	15	62	28

Benign Diagnosis | **Malignant diagnosis**

3

Using the threshold 3:

		P Object is present	N Object is not present
P	Object is observed	**50+20=70**	**12+8=20**
N	Object is not observed	5+15+10 = 30	**55+15 +10= 80**

Sensitivity (TPF) = **70**/100 FNF = 1- TPF = 30/100

Specificity (TNF) = **80**/100 FPF = 1-TNF = 20/100

Accuracy = (**70**+**80**)/200 = 75%

An Example: Plot ROC curve with the empirical method

Assuming there are a total of 200 cases, among them, 100 benign cases (no lesion) and 100 malignant cases (lesion present), an observation study resulted the following

	Observer's Interpretation/ (BI-RADs Score)				
	Definitely no lesion	Probably no lesion	Not sure	Probably with lesion	Definitely with lesion
For 100 Benign cases	**55**	**15**	**10**	**12**	8
For 100 malignant cases	15	10	5	50	**20**
Total	70	25	15	62	28

Benign Diagnosis | **Malignant diagnosis**

4

Using the threshold 4:

		P Object is present	N Object is not present
P	Object is observed	**20**	8
N	Object is not observed	50+5+15+10 =80	**55+15+10+12=92**

Sensitivity (TPF) = **20**/100 FNF = 1- TPF = 80/100

Specificity (TNF) = **92**/100 FPF = 1-TNF = 8/100

Accuracy = (**20**+**92**)/200 = 56%

An Example: Plot ROC curve with the empirical method

Assuming there are a total of 200 cases, among them, 100 benign cases (no lesion) and 100 malignant cases (lesion present), an observation study resulted the following

	Observer's Interpretation/ (BI-RADs Score)				
	Definitely no lesion	Probably no lesion	Not sure	Probably with lesion	Definitely with lesion
For 100 Benign cases	**55**	**15**	**10**	**12**	**8**
For 100 malignant cases	15	10	5	50	20
Total	70	25	15	62	28

Benign Diagnosis

Malignant diagnosis

Beginning Point (0,0)

Beginning Point (0,0):

		P Object is present	N Object is not present
P	Object is observed	0	0
N	Object is not observed	50+5+15+10+20 =100	55+15+10+12+8=100

Sensitivity (TPF) = 0/100 FNF = 1- TPF = 100/100

Specificity (TNF) = 100/100 FPF = 1-TNF = 0/100

Accuracy = (0+100)/200 = 50%

An Example: Plot ROC curve with the empirical method

Assuming there are a total of 200 cases, among them, 100 benign cases (no lesion) and 100 malignant cases (lesion present), an observation study resulted the following

	Observer's Interpretation/ (BI-RADs Score)				
	Definitely no lesion	Probably no lesion	Not sure	Probably with lesion	Definitely with lesion
For 100 Benign cases	55	15	10	12	8
For 100 malignant cases	**15**	**10**	**5**	**50**	**20**
Total	70	25	15	62	28

Benign Diagnosis | **Malignant diagnosis**

End point (1,1)

End point (1,1):

		P Object is present	N Object is not present
P	Object is observed	**100**	100
N	Object is not observed	0	**0**

Sensitivity (TPF) = **100**/100 FNF = 1- TPF = 0/100

Specificity (TNF) = **0**/100 FPF = 1-TNF = 100/100

Accuracy = (**100**+**0**)/200 = 50%

Table: Summary

	Sensitivity (TPF)	Specificity (TNF)	FPF=1-TNF
Point (1,1)	100/100	0/100	100/100
Threshold 1	(10+5+50+20)/100=85/100	55/100	1-55/100=45/100
Threshold 2	(5+50+20)/100=75/100	(55+15)/100=70/100	1-70/100=30/100
Threshold 3	(50+20)/100 = 70/100	(55+15+10)/100=80/100	1-80/100=20/100
Threshold 4	20/100	(55+15+10+12)=92/100	1-92/100=8/100
Point (0,0)	0/100	100/100	0/100

Accuracy1 = (85+55)/200 = 70%

Accuracy2 = (75+70)/200 = 72.5%

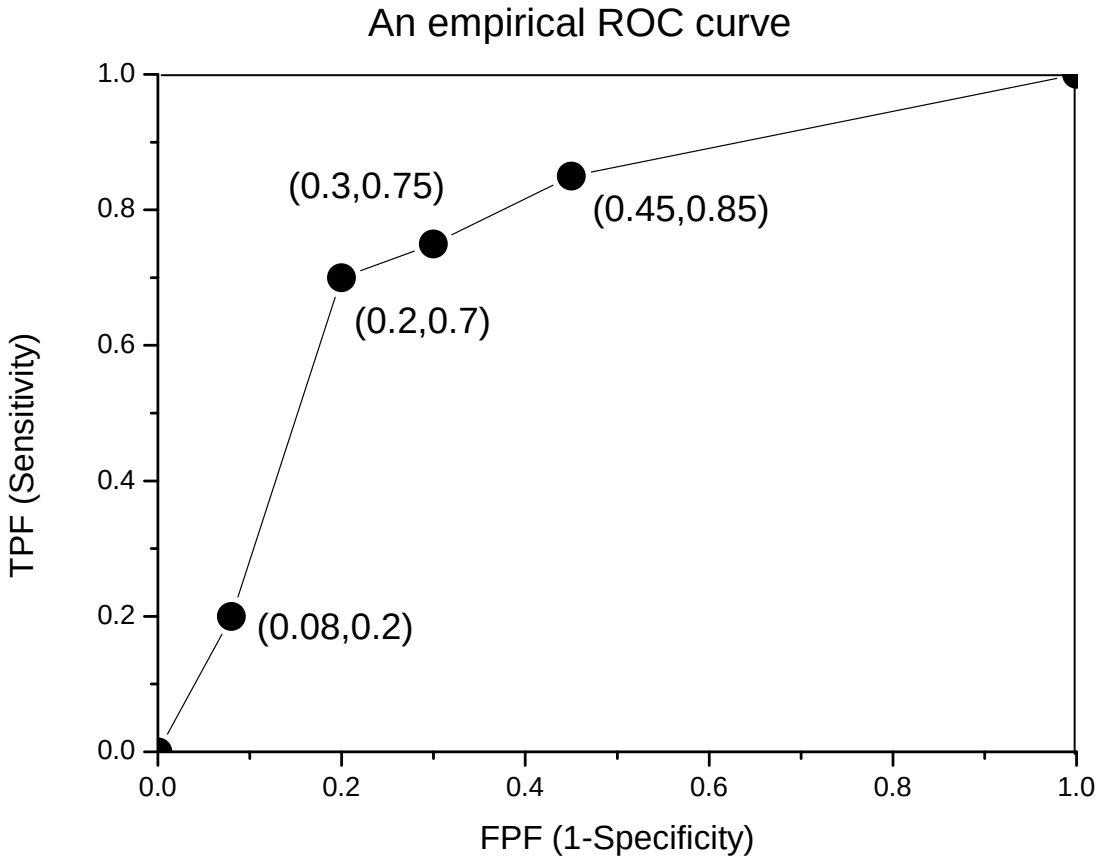
Accuracy3 = (70+80)/200 = 75%

Accuracy4 = (20+92)/200 = 56%

The empirical ROC Curve

The ROC curve begins at the (0, 0) coordinate, then connects all operating points, finally ends at (1, 1) coordinates.

	Sensitivity (TPF)	FPF =1-TNF
Threshold 1	0.85	0.45
Threshold 2	0.75	0.3
Threshold 3	0.7	0.2
Threshold 4	0.2	0.08



Summary

- Definition of sensitivity, specificity, and accuracy
- **ROC** analysis is a useful tool to evaluate the performance of diagnostic tests or medical imaging systems.
- More discussions on the statistical principles and the method of plotting conventional “smooth” ROC curve follows.

Homework #10-A

1.Please define sensitivity, specificity, and accuracy in medical image perception.

2. A physician read 50 mammograms and sent 6 patients to surgical biopsy procedures (6 positive cases according to his/her judgment). The pathology report following the biopsy indicated that among these 6 cases, only 4 patients have breast cancer. An expert panel reviewed the mammograms of the rest 44 patients (those who were found negative by the physician and were not sent for biopsy) and found another 1 positive cases and the follow-up biopsy/pathology verified that these one patients have breast cancer. A 5-year follow-up indicated that no more new breast cancer cases were found among all patients.

(1) What is the sensitivity of the physician's diagnosis?

(2) What is the specificity of the physician's diagnosis?

3. An observer study for comparing the performance of two medical imaging systems has resulted the following data:

For system 1

	Sensitivity (TPF)	Specificity (TNF)
Threshold 1	85/100	55/100
Threshold 2	75/100	70/100
Threshold 3	70/100	80/100
Threshold 4	20/100	92/100

For system 2

	Sensitivity (TPF)	Specificity (TNF)
Threshold 1	88/100	53/100
Threshold 2	78/100	74/100
Threshold 3	73/100	85/100
Threshold 4	25/100	93/100

(1) Please plot their ROC curves respectively for two system.
(2) Which system is better?