DIAGNOSTICS TESTS. SENSITIVITY, SPECIFICITY AND ROC CURVES

Santi Perez Hoyos
santi.perezhoyos@vhir.org

Miriam Mota Foix
miriam.mota@vhir.org

Vall d'Hebron
Institut de Recerca

OUTLINE

1. Diagnosis. Diagnostics tests
2. Sensitivity and specificity
3. Predictive values. Prevalence
4. Likelihood ratio
5. Receiver operator characteristic curves

OUTLINE

1. Diagnosis. Diagnostics tests
2. Sensitivity and specificity
3. Predictive values. Prevalence
4. Likelihood ratio
5. Receiver operator characteristic curves

Diagnosis

FOX

1. Diagnosis. Diagnostics tests

Diagnosis

Foto: FOX

1. Diagnosis. Diagnostics tests

Diagnosis

- Most important result of medical practice
- Key that leads to treatment and prognosis

Other valid definitions

- process of determining the nature of the morbid condition through examination.
- careful examination of the facts to determine the nature of something
- decision or opinion resulting from such examination or investigation..

Classification of individuals in healthy or sick

1. Diagnosis. Diagnostics tests

Process that leads to the diagnosis

1. Establish a presumption suspicion or hypothesis of existence of disease
2. Follow up on the clinical assumption and verify if this corresponds to the truth.

Uncertainty

- Symptoms of a patient may be compatible with more than one disease
- Biological variations from one individual to another
- Instruments are imprecise
- Patients are inaccurate to remember past events

1. Diagnosis. Diagnostics tests

Process that leads to the diagnosis

Different sources of information are used:

- anamnesis of the patient
- Physical exam
- results of diagnostics tests

DIAGNOSTIC TEST

Procedure by which a diagnosis is confirmed or discarded

1. Diagnosis. Diagnostics tests

Diagnostics tests. What we desire?

Positive results in patients and negative results in healthy

Conditions to require a diagnostic test (I):

- **Validity**: the degree to which a test measures what it is supposed to measure (sensitivity and specificity)
- **Reproducibility**: ability of the test to offer the same results when its application is repeated in similar circumstances (biological, instrumental and technical variability)

1. Diagnosis. Diagnostics tests

Diagnostic test

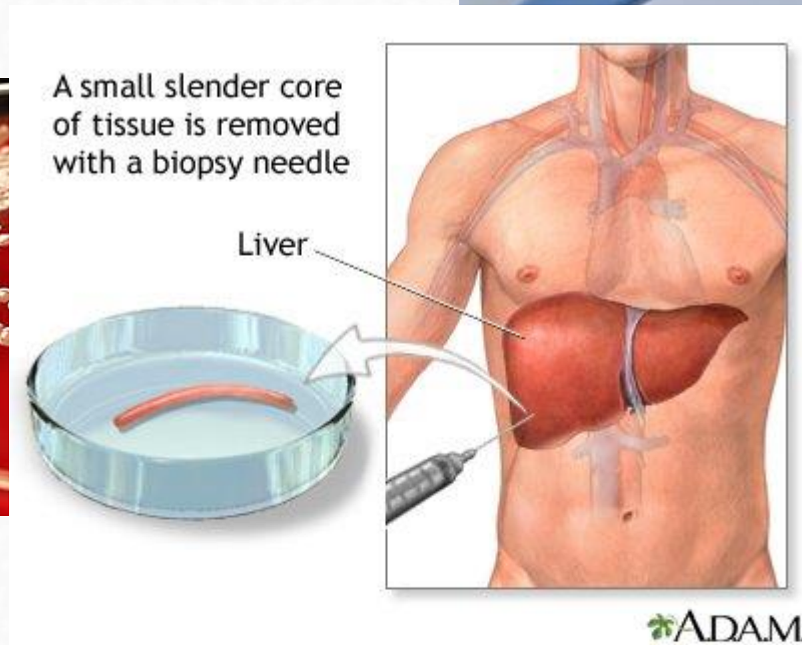
Conditions to require a diagnostic test (II):

- **Safety**: in the case of a positive result, what is the probability that this result indicates presence of the disease? (predictive values)
- **Easy** to apply
- **Accepted** by patients or the population in general
- Minimal **adverse effects**
- **Economically** bearable

1. Diagnosis. Diagnostics tests

Relationship between the results of the test and the authentic diagnosis

Authentic diagnosis → **Reference method**



1. Diagnosis. Diagnostics tests

Vall d'Hebron
Institut de Recerca

Relationship between the results of the test and the authentic diagnosis

Reference method

- expensive?
- traumatic for the patient?
- slow in getting results?

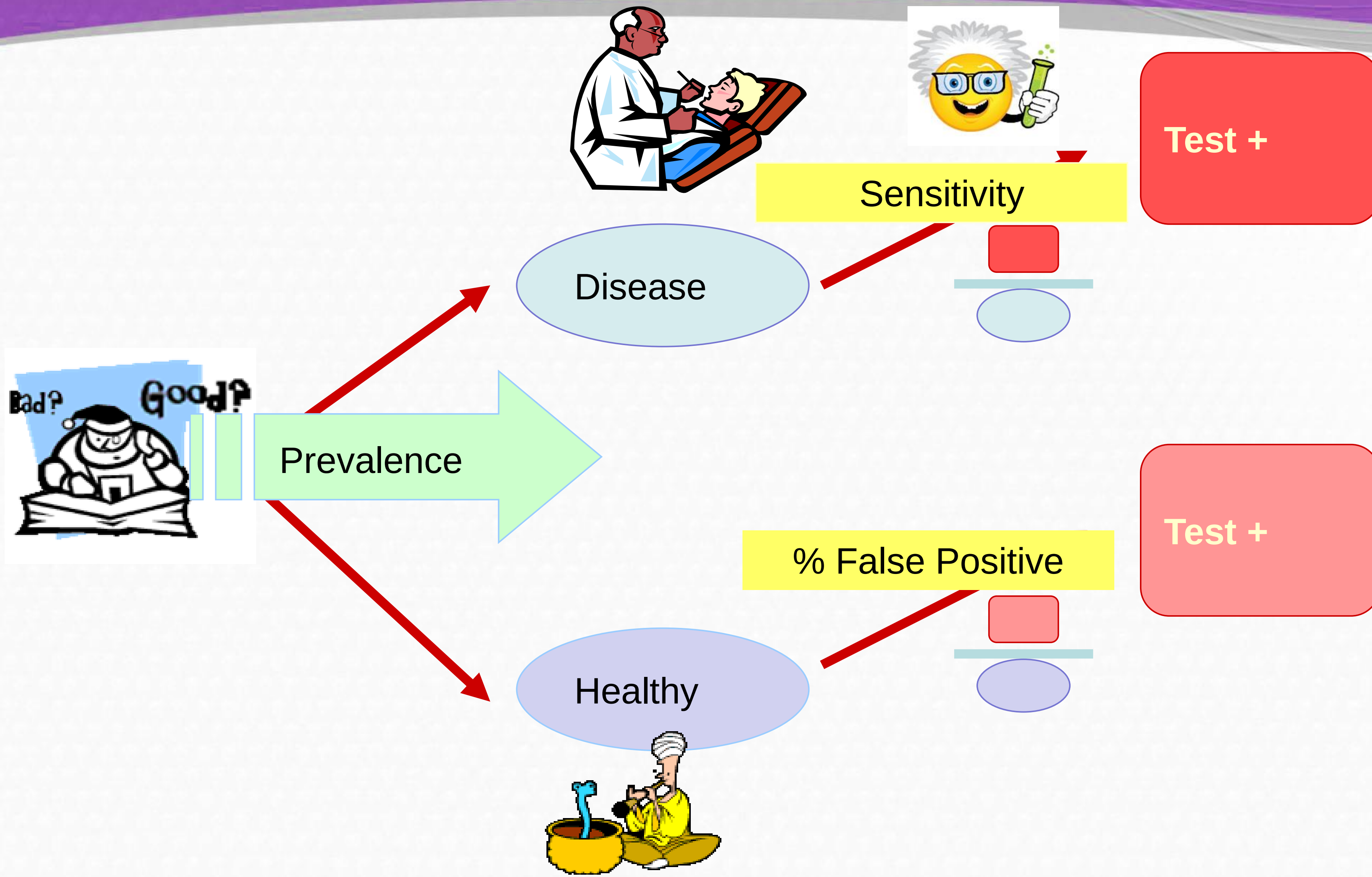
A more accessible method (diagnostic test) is needed.

Classify patients the same as the reference method?

OUTLINE

1. Diagnosis. Diagnostics tests
2. Sensitivity and specificity
3. Predictive values. Prevalence
4. Likelihood ratio
5. Receiver operator characteristic curves

2. Sensitivity and specificity



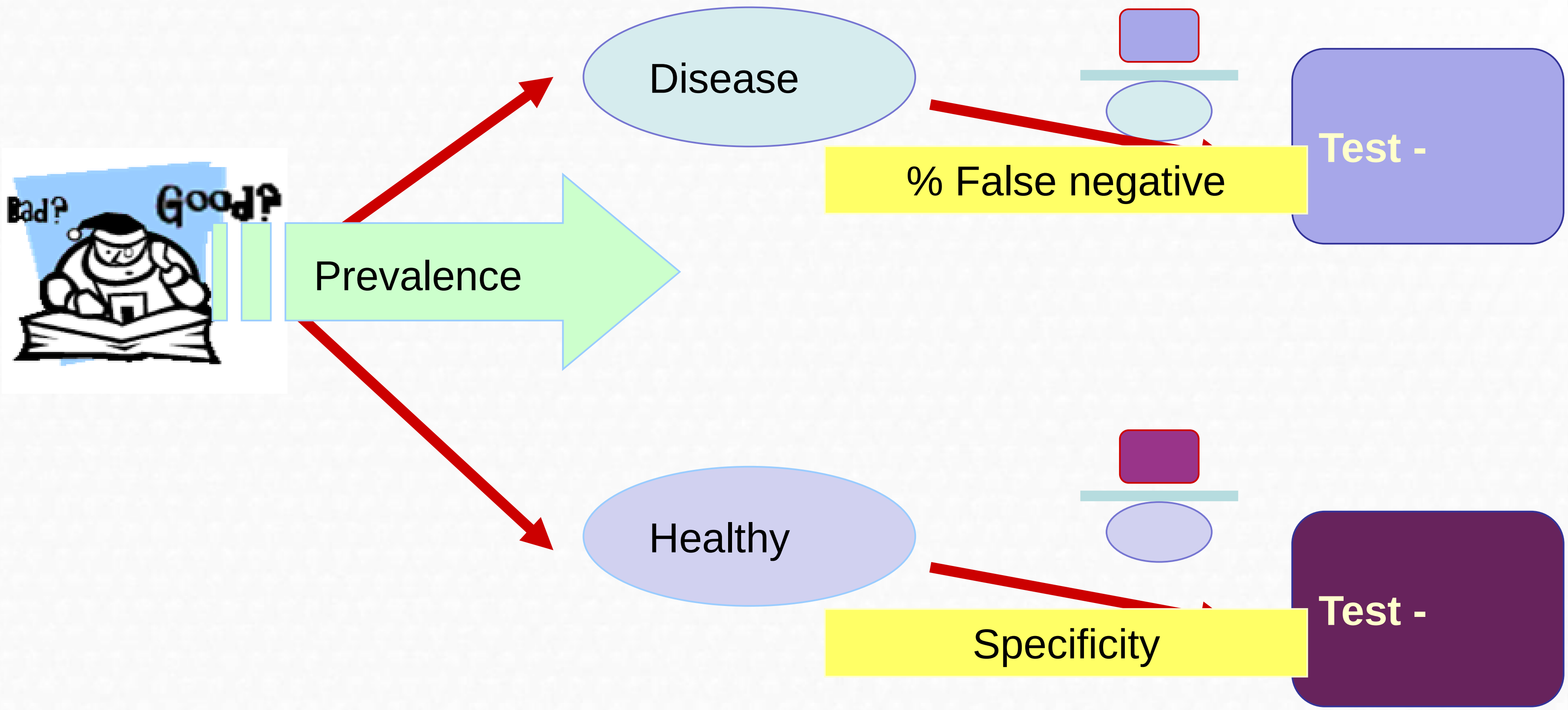
2. Sensitivity and specificity

Sensitivity (S): refers to the ability of the test to correctly identify those patients with the disease.

S = P(Test+/Disease)

- A test with 100% sensitivity correctly identifies all patients with the disease.
- A test with 80% sensitivity detects 80% of patients with the disease (true positives) but 20% with the disease go undetected (false negatives).

2. Sensitivity and specificity



2. Sensitivity and specificity

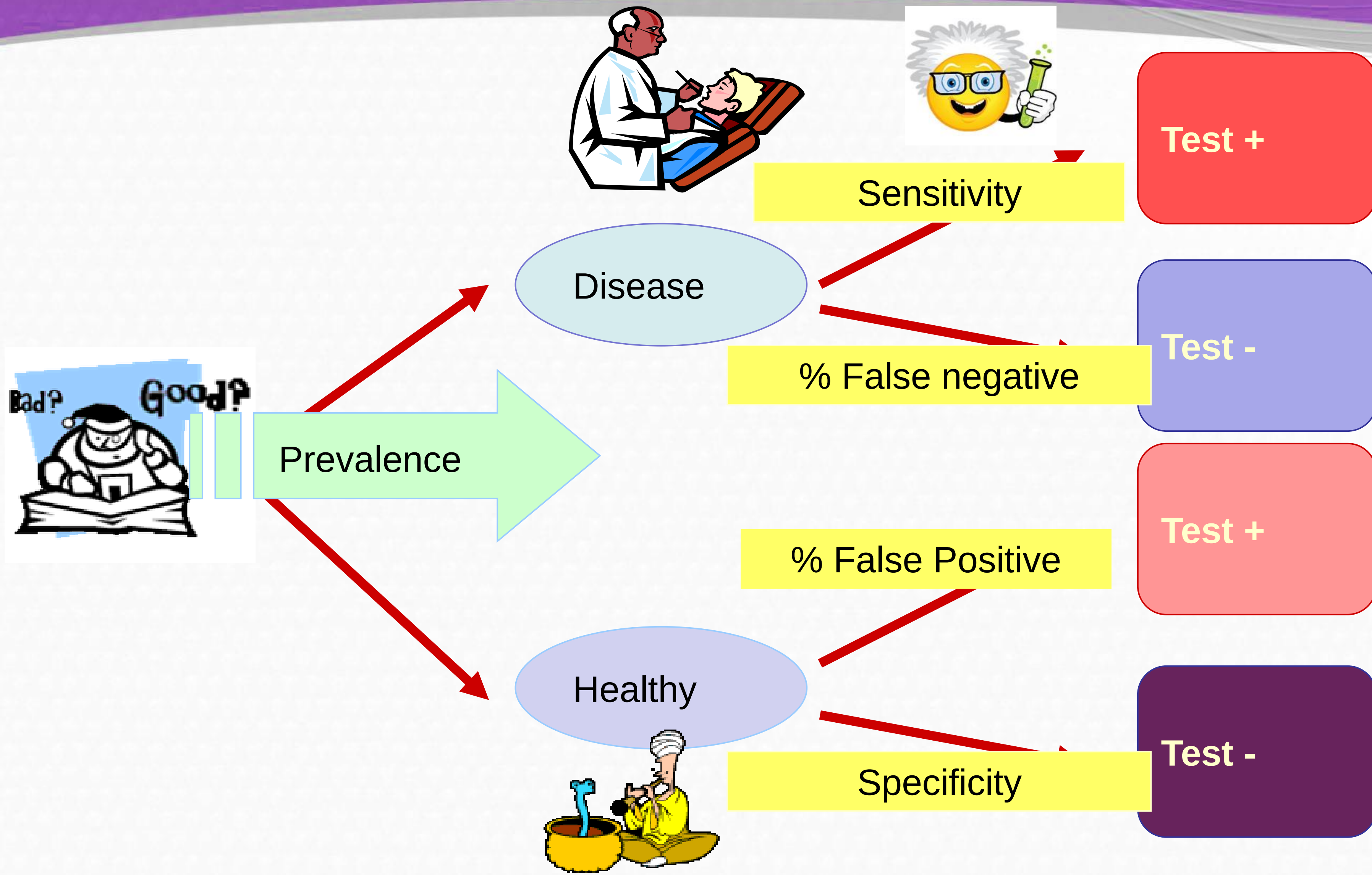
Vall d'Hebron Institut de Recerca

Specificity (E): refers to the ability of the test to correctly identify those patients without the disease.

$$E = P(\text{Test} -/\text{no disease})$$

- A test with 100% specificity correctly identifies all patients without the disease.
- A test with 80% specificity correctly reports 80% of patients without the disease as test negative (true negatives) but 20% patients without the disease are incorrectly identified as test positive (false positives).

2. Sensitivity and specificity



Relationship between the results of the test and the authentic diagnosis

		Reference method		
		Sick	Healthy	TOTAL
Diagnostic Test	Positive	a	b	a+b
	Negative	c	d	c+d
	TOTAL	a+c	b+d	a+b+c+d

a = True positives (TP)
b = False positives (FP)
c = False negative (FN)
d = True negative (TN)

Vall d'Hebron Institut de Recerca

Sentivity and specifity calculation,

		Reference method		
		Sick	Healthy	TOTAL
Diagnostic Test	Positive	TP	FP	TP+FP
	Negative	FN	TN	FN+TN
	TOTAL	TP+FN	FP+TN	TP+FP+FN +TN

		Reference method		
		Sick	Healthy	TOTAL
Diagnostic Test	Positive	a	b	a+b
	Negative	c	d	c+d
	TOTAL	a+c	b+d	a+b+c+d

Sensitivity

$$\frac{TP}{TP + FN} \qquad \frac{a}{a + c}$$

Specificity

$$\frac{TN}{TN + FP} \qquad \frac{d}{d + b}$$

2. Sensitivity and specificity

Vall d'Hebron Institut de Recerca

Example.

Sample: n= 2.641 patients with suspected prostate cancer
1st test: rectal examination
Reference method: prostate biopsy

		Biopsy result		
		Disease	Healthy	TOTAL
Rectal examination	Disease	634	269	903
	Healthy	487	1251	1738
	TOTAL	1121	1520	2641

Sensitivity = 634 / (634+487) = 0.5656 = 56.6% → 43.4% with cancer had a normal rectal examination (false negatives)

Specificity = 1251 / (269+1251) = 0.8230 = 82.3%

17.7% of the patients without disease were incorrectly diagnosed(false positives)

Other tests are needed to refine the diagnosis (Ex. PSA)

2. Sensitivity and specificity

Example with R

```
library(epiR)
table1 <- as.table(matrix(c(634,269,487,1251), nrow = 2, byrow = TRUE))
epi.tests(table1)
```

```
                Outcome +    Outcome -      Total
Test +              634          269          903
Test -              487         1251         1738
Total              1121         1520         2641

Point estimates and 95 % CIs:
---------------------------------------------------------
Apparent prevalence                    0.34 (0.32, 0.36)
True prevalence                        0.42 (0.41, 0.44)
Sensitivity                            0.57 (0.54, 0.59)
Specificity                            0.82 (0.80, 0.84)
Positive predictive value              0.70 (0.67, 0.73)
Negative predictive value              0.72 (0.70, 0.74)
Positive likelihood ratio              3.20 (2.83, 3.60)
Negative likelihood ratio              0.53 (0.49, 0.57)
---------------------------------------------------------
```

2. Sensitivity and specificity

Ideal values of sensitivity and specificity

- $0 \geq S$ and $E \leq 1$

- **High sensitivity when…**

  ✓the test is used to identify a serious but treatable disease (e.g. cervical cancer).
  ✓Screening of population.
  ✓First test in a battery of test

- **High specificity when…**

  ✓Diagnostic confirmation suggested by other data (no FP)
  ✓Avoid FP in expensive treatments or with high side effects

Vall d'Hebron
Institut de Recerca

OUTLINE

1. Diagnosis. Diagnostics tests
2. Sensitivity and specificity
3. Predictive values. Prevalence
4. Likelihood ratio
5. Receiver operator characteristic curves

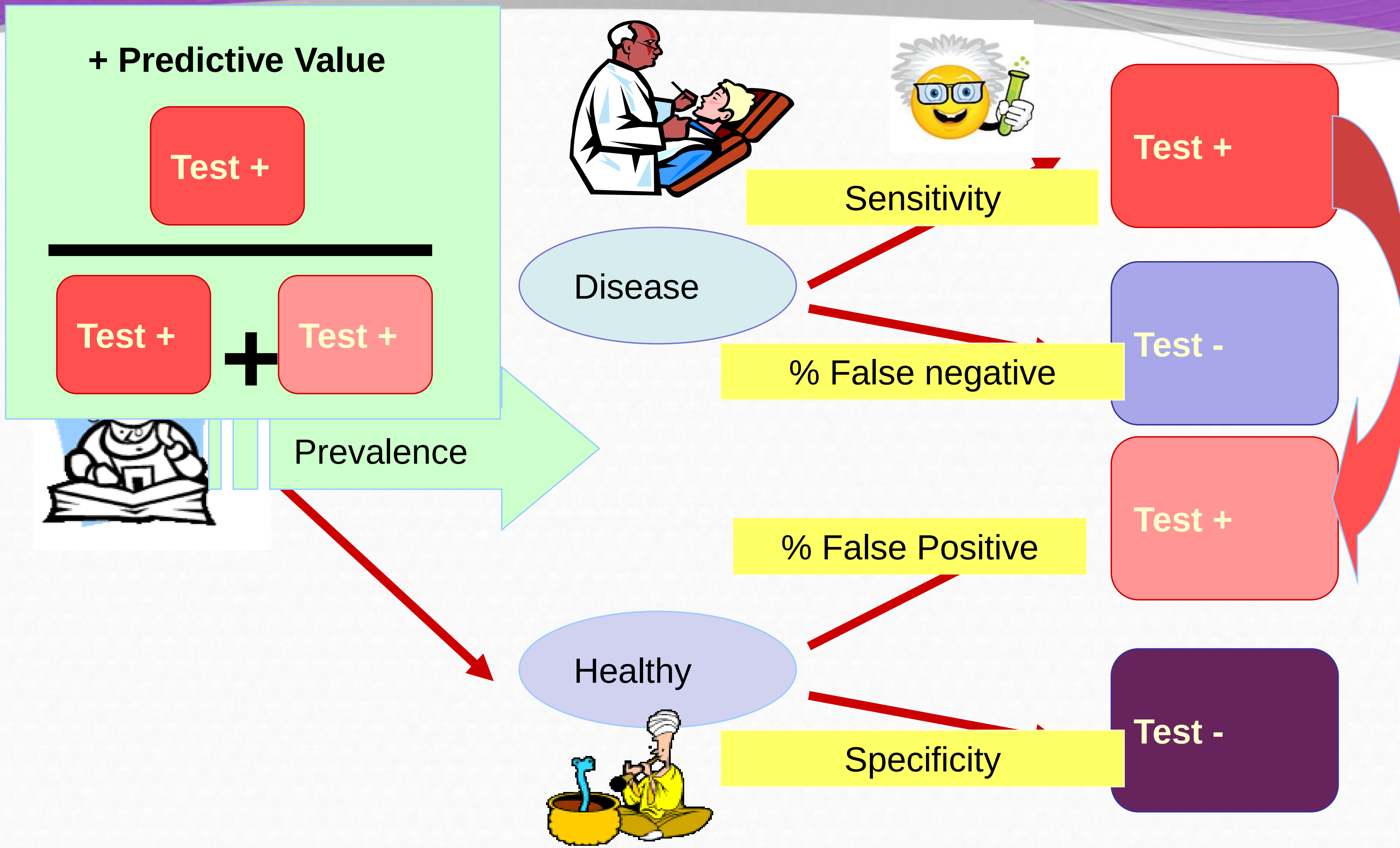
3. Predictive values. Prevalence

But…the patient is sick or not?

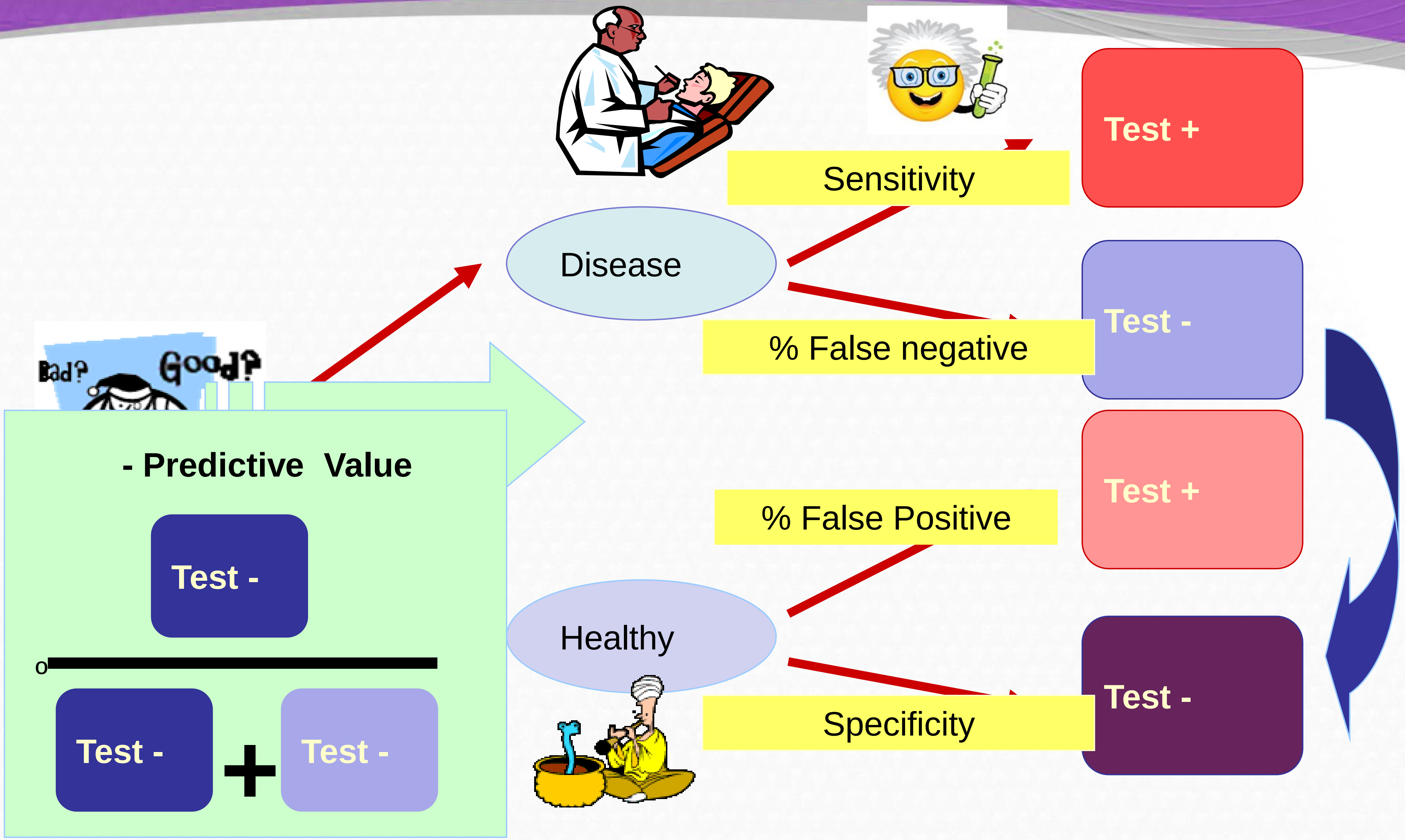
S and E help to decide if a test should be used or not, but once the result is known they do not provide any information

they are based on the prior knowledge (sick/healthy) of the individual

3. Predictive values. Prevalence



3. Predictive values. Prevalence



3. Predictive values. Prevalence

But…the patient is sick or not?

If my patient has had a positive / negative test result, what is the probability that patient suffer / not suffer from the disease

Positive predictive value

How likely is it that this patient has the disease given that the test result is positive?

$$PPV = P(\text{Disease} / \text{Test} +)$$

Negative predictive value

How likely is it that this patient does not have the disease given that the test result is negative?

$$NPV = P(\text{no Disease} / \text{Test} -)$$

3. Predictive values. Prevalence

	Disease	Healthy
Test +	**True + Sety**	**False + % FP**
Test -	**False - % FN**	**True – Especificity**

PPV

$$\frac{\text{Test +}}{\text{Test +} + \text{Test +}}$$

PPN

$$\frac{\text{Test -}}{\text{Test -} + \text{Test -}}$$

3. Predictive values. Prevalence

Example.

Sample: n= 2.641 patients with suspected prostate cancer
1st test: rectal examination
Reference method: prostate biopsy

		Biopsy result		
		Disease	Healthy	TOTAL
Rectal examination	Disease	634	269	903
	Healthy	487	1251	1738
	TOTAL	1121	1520	2641

Positive predictive value = 634 / (269+634) = 70.21% → 70.21% of the patients with abnormal digital rectal examination were correctly diagnosed

Negative predictive value = 1251 / (487+1251) = 71.98%

↓

71.98% with normal digital rectal examination were healthy

3. Predictive values. Prevalence

Example with R

```
Disease positive Disease negative Total
Test positive                634                269   903
Test negative                487               1251  1738
Total                       1121               1520  2641

 Point estimates and 95 % CIs:
---------------------------------------------------------
                                    Estimation Lower CI Upper CI
Apparent prevalence                      0.342    0.324    0.360
True prevalence                          0.424    0.406    0.444
Sensitivity                              0.566    0.536    0.595
Specificity                              0.823    0.803    0.842
Positive predictive value                0.702    0.671    0.732
Negative predictive value                0.720    0.698    0.741
Diagnostic accuracy                      0.714    0.696    0.731
Likelihood ratio of a positive test      3.196    2.835    3.603
Likelihood ratio of a negative test      0.528    0.492    0.567
---------------------------------------------------------
```

3. Predictive values. Prevalence

Prevalence

Proportion of people in a defined population who suffer from the disease at a specific time

Probability of having the disease before knowing the result of the test,

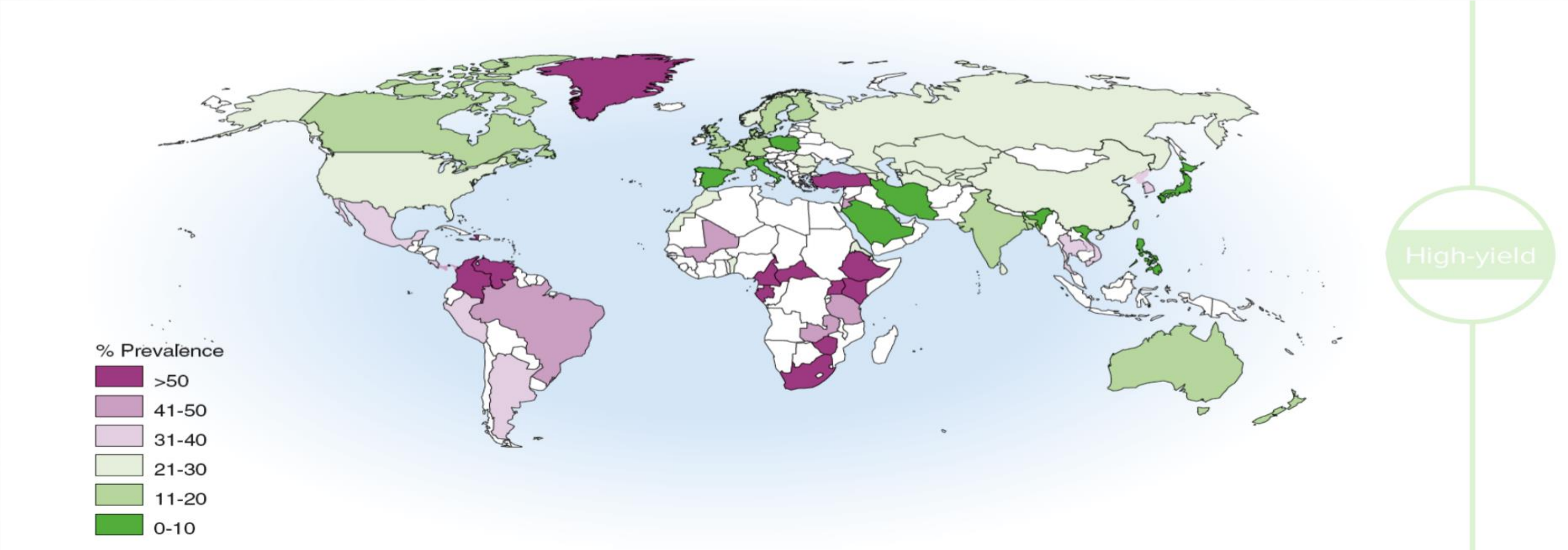


FIGURE 138-1 Global seroprevalence of herpes simplex virus type 2 infection.

3. Predictive values. Prevalence

Prevalence calculation.

		Biopsy result		
		Disease	Healthy	TOTAL
Rectal examination	Disease	634	269	903
	Healthy	487	1251	1738
	TOTAL	1121	1520	2641

Prevalence = All patient sick / All population = 1121/2641 = 42.45%

3. Predictive values. Prevalence

Example with R

```
     Disease positive Disease negative Total
Test positive              634              269   903
Test negative              487             1251  1738
Total                     1121             1520  2641

 Point estimates and 95 % CIs:
---------------------------------------------------------
                                   Estimation Lower CI Upper CI
Apparent prevalence                     0.342    0.324    0.360
True prevalence                         0.424    0.406    0.444
Sensitivity                             0.566    0.536    0.595
Specificity                             0.823    0.803    0.842
Positive predictive value               0.702    0.671    0.732
Negative predictive value               0.720    0.698    0.741
Diagnostic accuracy                     0.714    0.696    0.731
Likelihood ratio of a positive test     3.196    2.835    3.603
Likelihood ratio of a negative test     0.528    0.492    0.567
---------------------------------------------------------
```

3. Predictive values. Prevalence

Dependence of PPV and NPV on disease prevalence

Example: VIH diagnosis. Population 1.

		Correct diagnosis		
		VIH+	VIH-	TOTAL
Test result	Positive	5.970	13.970	19.940
	Negative	30	2.780.030	2.780.060
	TOTAL	6.000	2.794.000	2.800.000

Example: VIH diagnosis. Population 2.

		Correct diagnosis		
		VIH+	VIH-	TOTAL
Test result	Positive	796.000	10.000	806.000
	Negative	4.000	1.990.000	1.994.000
	TOTAL	800.000	2.000.000	2.800.000

Dependence of PPV and NPV on disease prevalence

Example: VIH diagnosis. Population 1.

		Correct diagnosis		
		VIH+	VIH-	TOTAL
Test result	Positive	5.970	13.970	19.940
	Negative	30	2.780.030	2.780.060
	TOTAL	6.000	2.794.000	2.800.000

Prevalence = 6.000 / 2.800.000 **= 0.21%**

Sensitivity = 5970 / 6000 = 99.5%

Specificity = 2.780.030 /2.794.000 = 99.5%

PPV = 5970 / 19940 = **29.9%** → Only 29.9% of the individuals who tested positive, would be ill

NPV = 2.780.030 / 2.780.60 = 99.9%

Dependence of PPV and NPV on disease prevalence

Example: VIH diagnosis. Population 2.

		Correct diagnosis		
		VIH+	VIH-	TOTAL
Test result	Positive	796.000	10.000	806.000
	Negative	4.000	1.990.000	1.994.000
	TOTAL	800.000	2.000.000	2.800.000

Prevalence = 800.000 / 2.800.000 **= 28.6%**

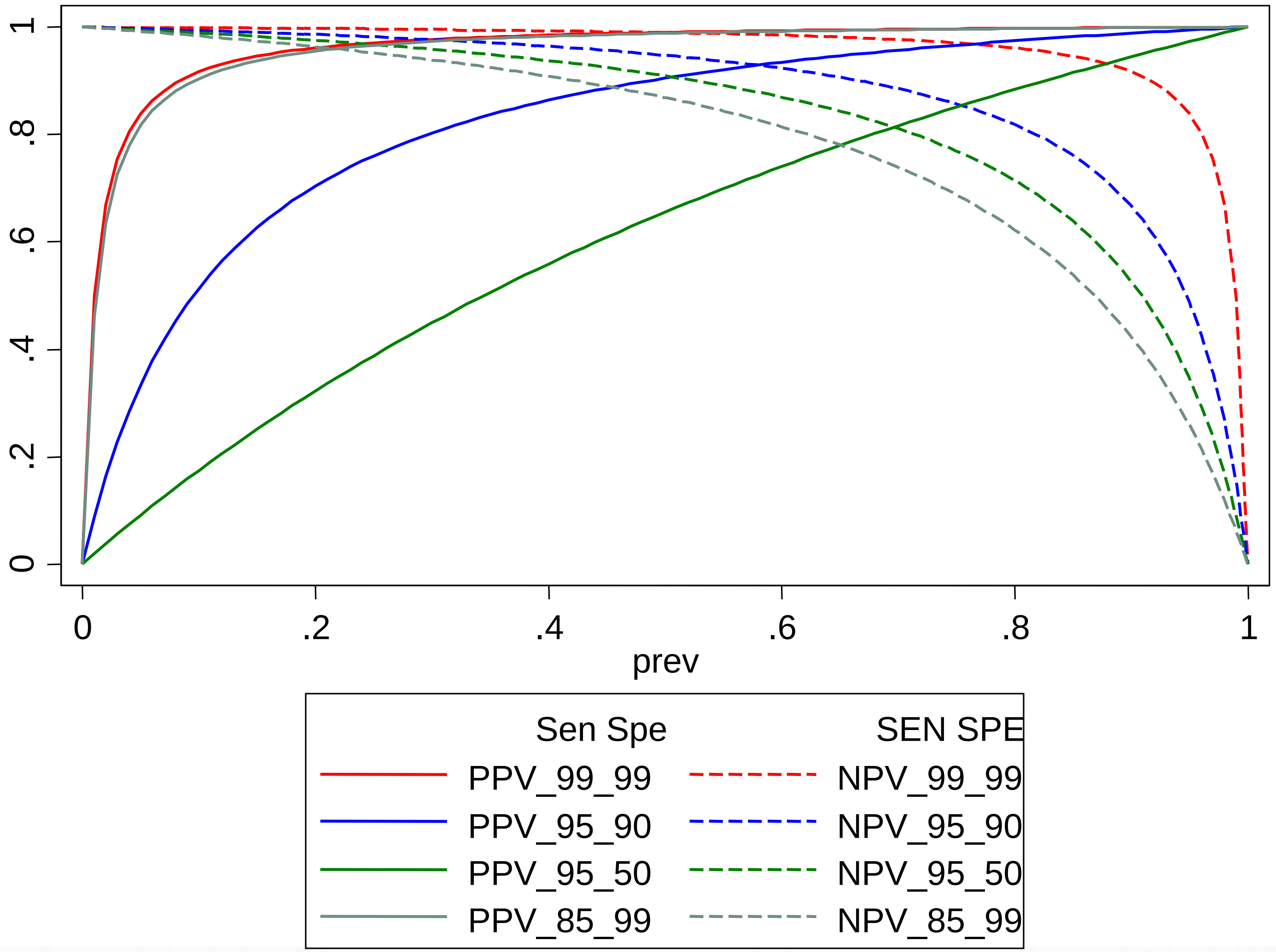
Sensitivity = 796.000 / 800.000 = 99.5%

Specificity = 1.990.000 /2.000.000 = 99.5%

PPV = 796.000 / 806.000 = **98.7%** → If the prevalence is high, a positive result tends to confirm the presence of the disease

NPV = 1.990.000 / 2.000.000 = 99.8%

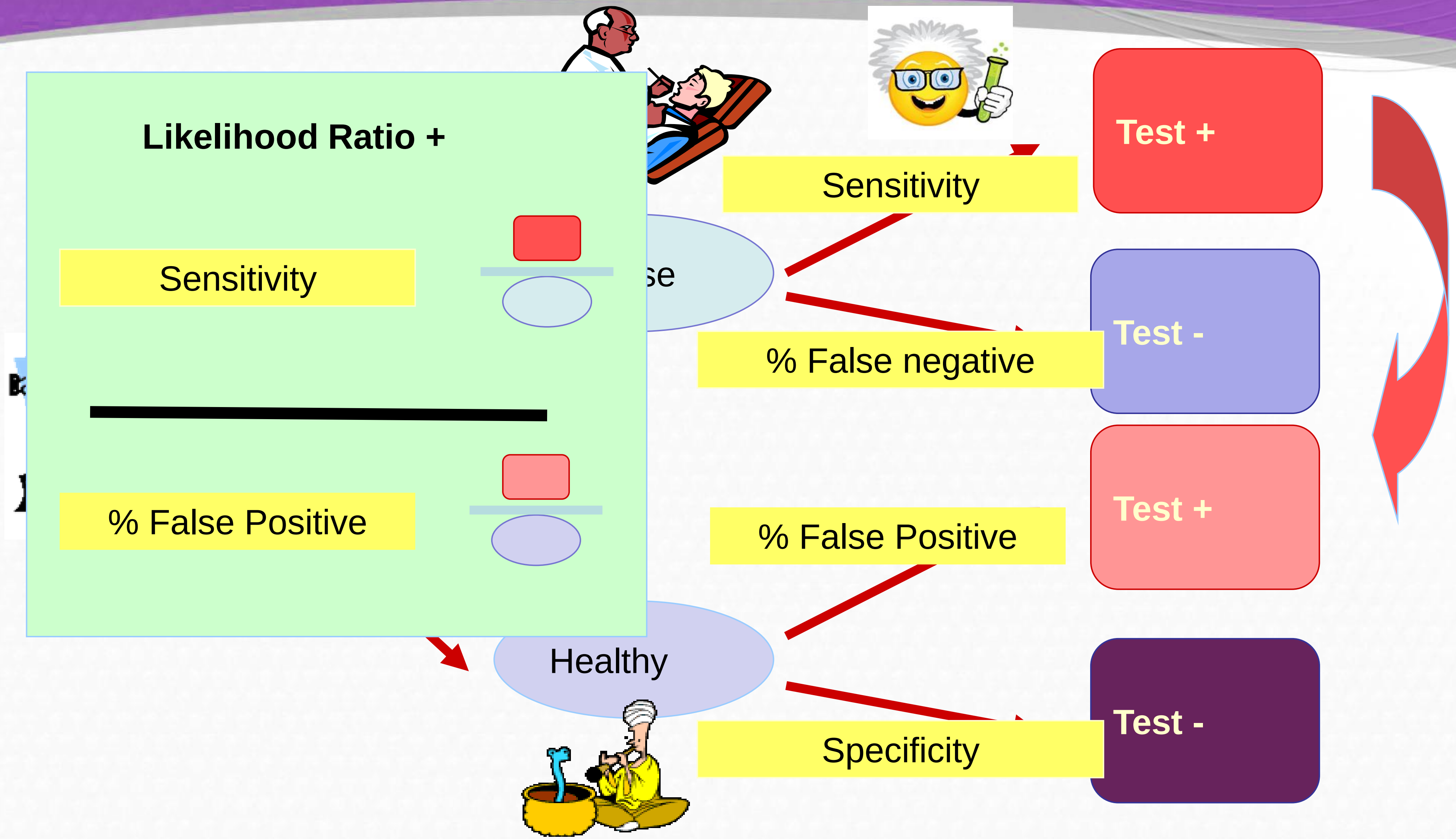
PPV & NPV depend on prevalence



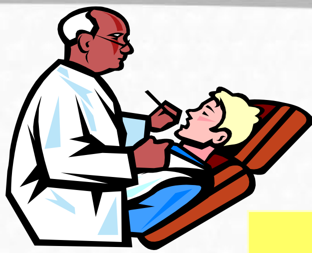
OUTLINE

1. Diagnosis. Diagnostics tests
2. Sensitivity and specificity
3. Predictive values. Prevalence
4. Likelihood ratio
5. Receiver operator characteristic curves

4. Likelihood ratio



4. Likelihood ratio



Disease

Sensitivity

Test +

% False negative

Test -

% False Positive

Test +

Specificity

Test -

Likelihood Ratio -

$$\frac{\text{\% False negative}}{\text{Specificity}}$$

Likelihood ratio

- How much more likely is it that a patient who tests positive has the disease compared with one who test negative,
- Index to compare two methods (diagnostic tests) that do not depend on the prevalence in the population.

  - **LR+** = P (positive test in ill) / P (positive test in healthy) =

    = TP / FP =

    = **sensitivity / (1-specificity)**

  - **LR-** = P (negative test in ill) / P (negative test in healthy) =

    = FN / TN =

    = **(1-sensitivity) / specificity**

Likelihood ratio

- Interpretation: Positive Likelihood Ratio (LR+)

  - LR+ over 5 - 10: Significantly increases likelihood of the disease
  - LR+ between 0.2 to 5 (especially if close to 1): Does not modify the likelihood of the disease
  - LR+ below 0.1 - 0.2: Significantly decreases the likelihood of the disease.

4. Likelihood ratio

Example. **Sample:** n= 2.641 patients with suspected prostate cancer
1st test: rectal examination
Reference method: prostate biopsy

		Biopsy result		
		Disease	Healthy	TOTAL
Rectal examination	Disease	634	269	903
	Healthy	487	1251	1738
	TOTAL	1121	1520	2641

Sensitivity = 634 / (634+487) = 0.5656 = 56.6%

Specificity = 1251 / (269+1251) = 0.8230 = 82.3%

Prevalence = 1121/2641 = 42.45%

PPV = 634 / (269+634) = 70.21%

NPV = 1251 / (487+1251) = 71.98%

LR+ = 0.566 / (1-0.823) = **3.19** → Hay x3 veces más probabilidades de observar un tacto anormal en enfermos que en sanos

LR- = (1-0.566) / 0.823 = **0.53**

4. Likelihood ratio

Vall d'Hebron Institut de Recerca

Example. **Sample:** n= 2.641 patients with suspected prostate cancer
1st test: rectal examination
Reference method: prostate biopsy

	Disease positive	Disease negative	Total
Test positive	634	269	903
Test negative	487	1251	1738
Total	1121	1520	2641

Point estimates and 95 % CIs:

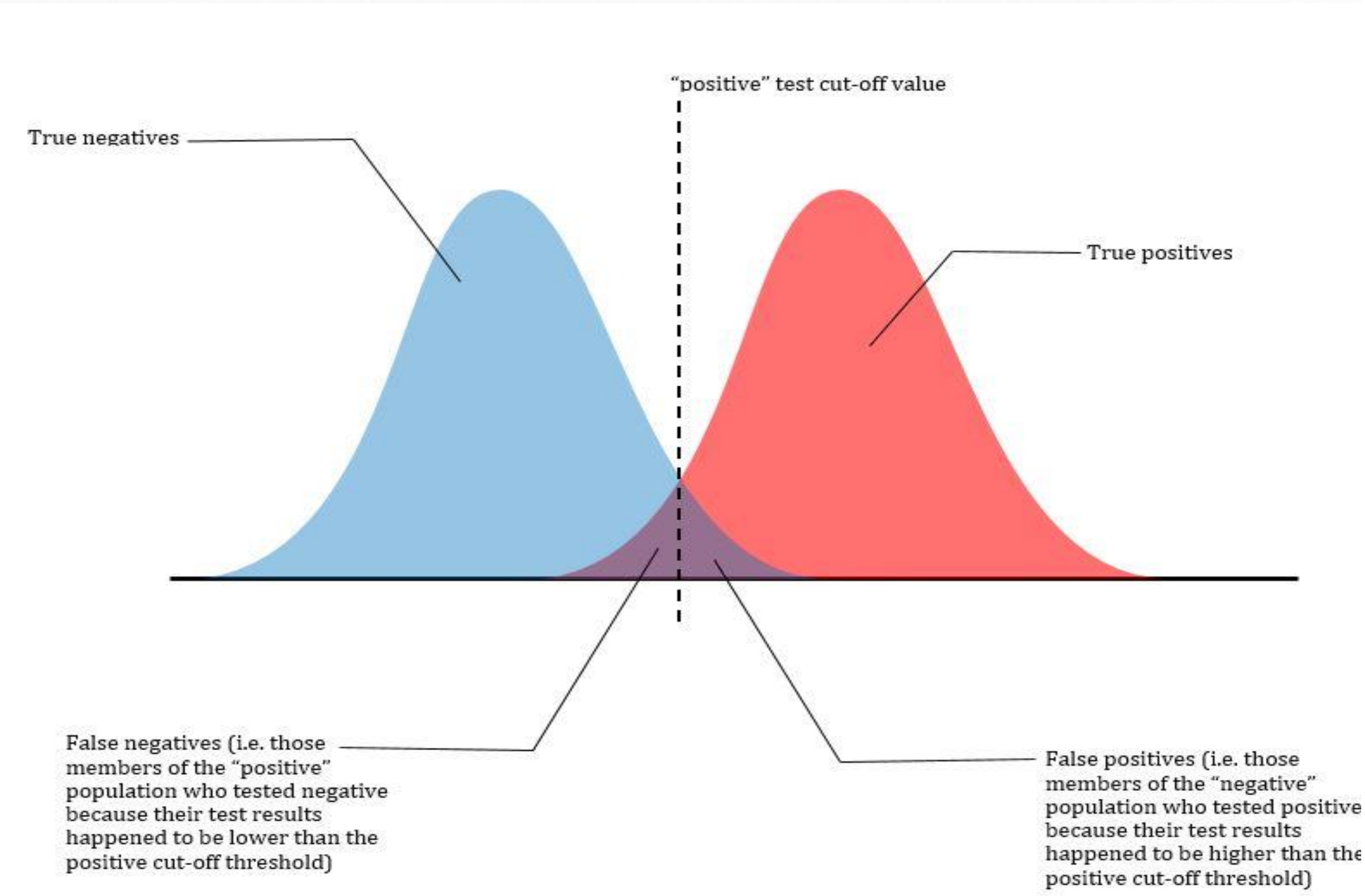
	Estimation	Lower CI	Upper CI
Apparent prevalence	0.342	0.324	0.360
True prevalence	0.424	0.406	0.444
Sensitivity	0.566	0.536	0.595
Specificity	0.823	0.803	0.842
Positive predictive value	0.702	0.671	0.732
Negative predictive value	0.720	0.698	0.741
Diagnostic accuracy	0.714	0.696	0.731
Likelihood ratio of a positive test	**3.196**	**2.835**	**3.603**
Likelihood ratio of a negative test	0.528	0.492	0.567

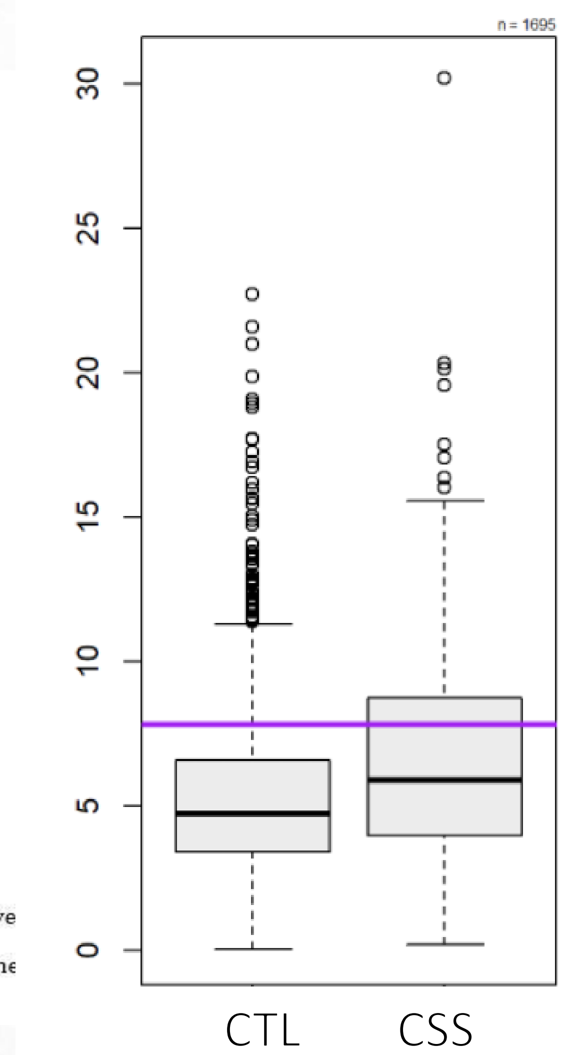
OUTLINE

1. Diagnosis. Diagnostics tests
2. Sensitivity and specificity
3. Predictive values. Prevalence
4. Likelihood ratio
5. Receiver operator characteristic curves

5. Receiver operator characteristic curves (ROC)

Building of ROC curves





http://www.derangedphysiology.com/main/cicm-primary-exam/required-reading/research-methods-and-statistics/Chapter%203.0.5/receiver-operating-characteristic-roc-curve

5. Receiver operator characteristic curves (ROC)

Building of ROC curves

Until now we had tests with dichotomous results $\begin{cases} \text{Positive} \\ \text{Negative} \end{cases}$

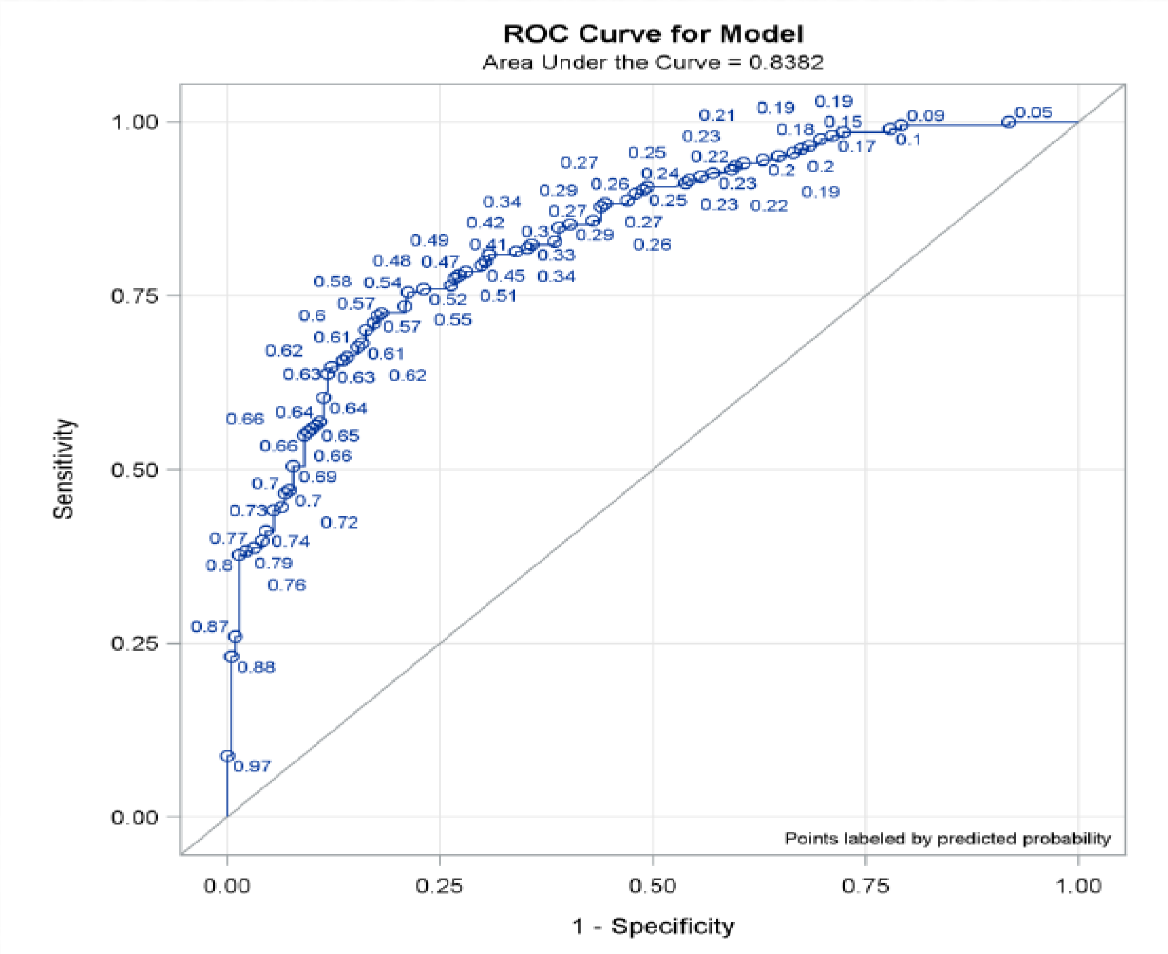
generates a sensitivity and specificity value

Test with continuous results → It has to be chosen different cutting points that allow a dichotomous classification

Many values of sensitivity and specificity that vary according to the cut point chosen → Graphic representation

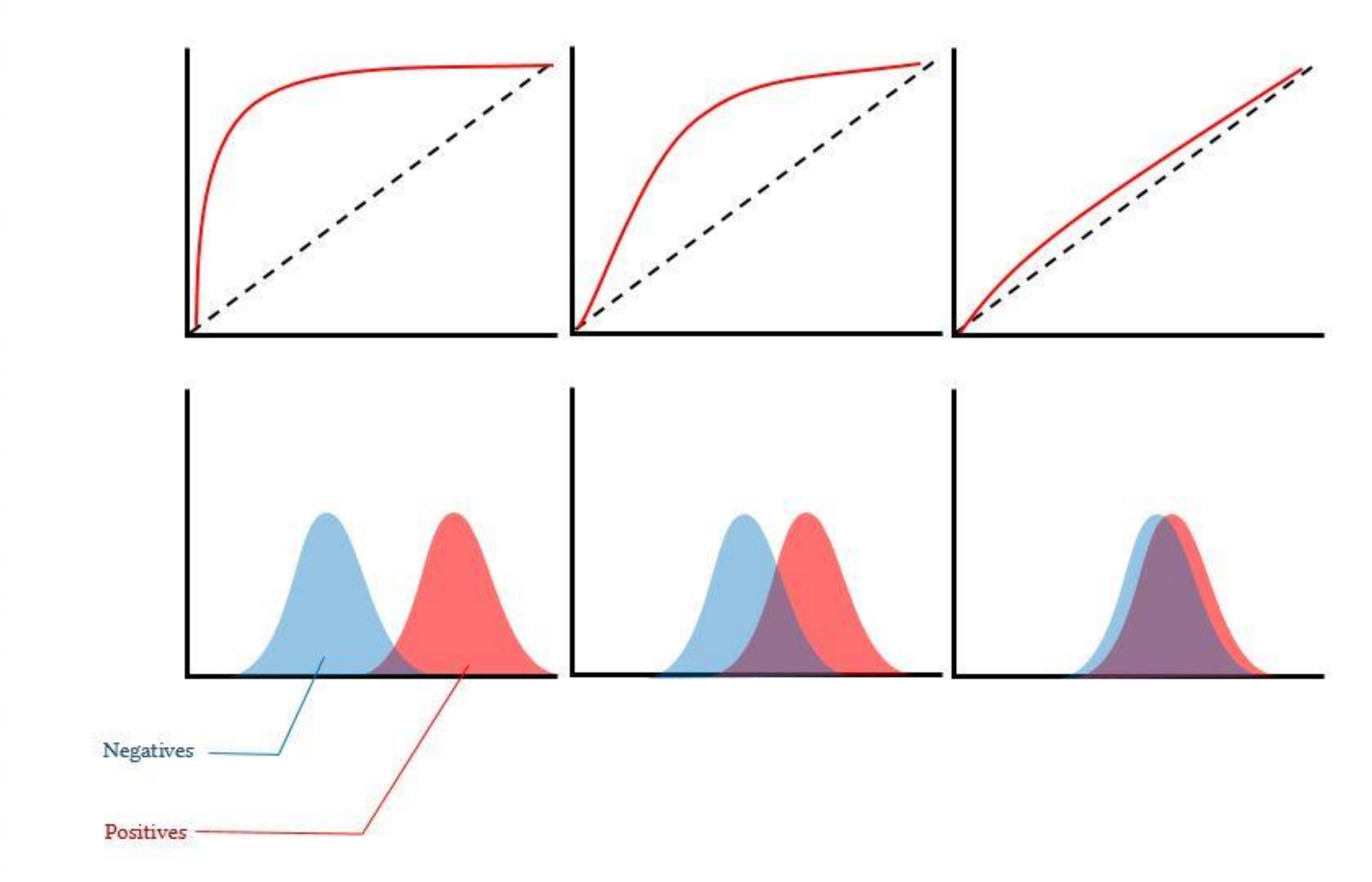
5. Receiver operator characteristic curves (ROC)

Building of ROC curves



5. Receiver operator characteristic curves (ROC)

Building of ROC curves



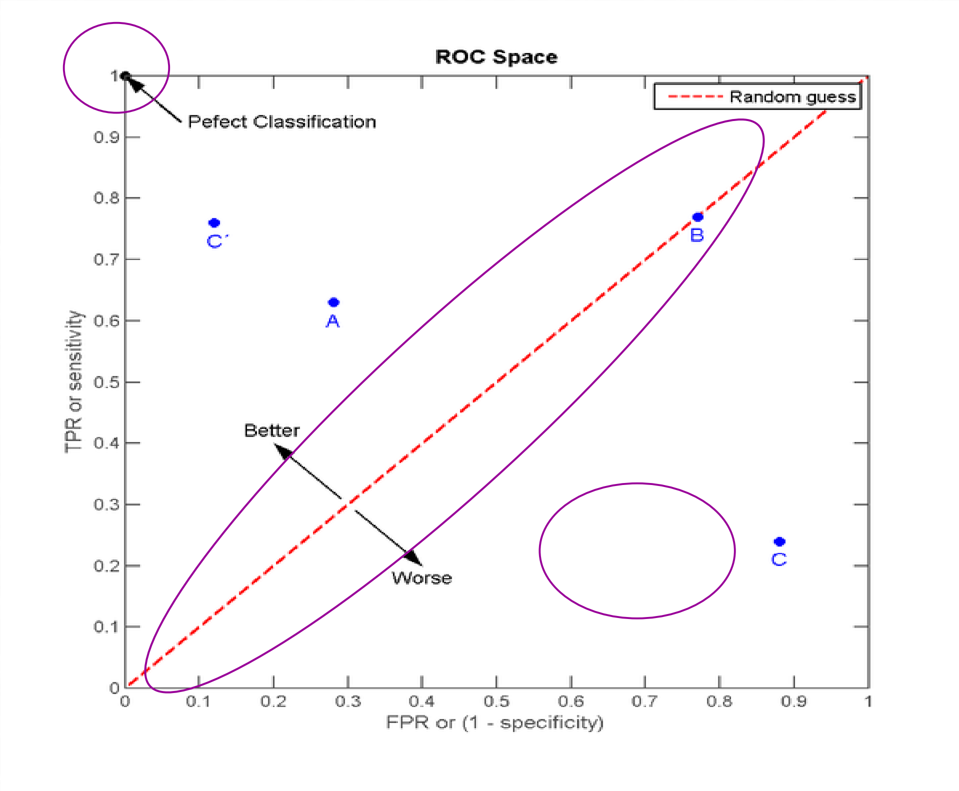
http://www.derangedphysiology.com/main/cicm-primary-exam/required-reading/research-methods-and-statistics/Chapter%203.0.5/receiver-operating-characteristic-roc-curve

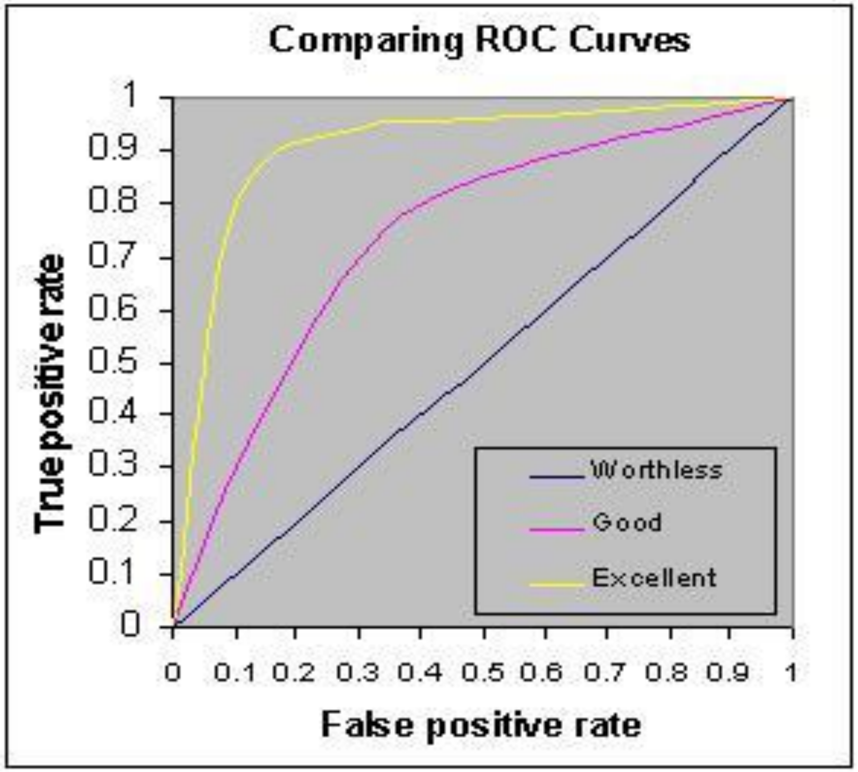
5. Receiver operator characteristic curves (ROC)

Vall d'Hebron Institut de Recerca

Building of ROC curves

Receiver Operator Characteristic





5. Receiver operator characteristic curves (ROC)

Example

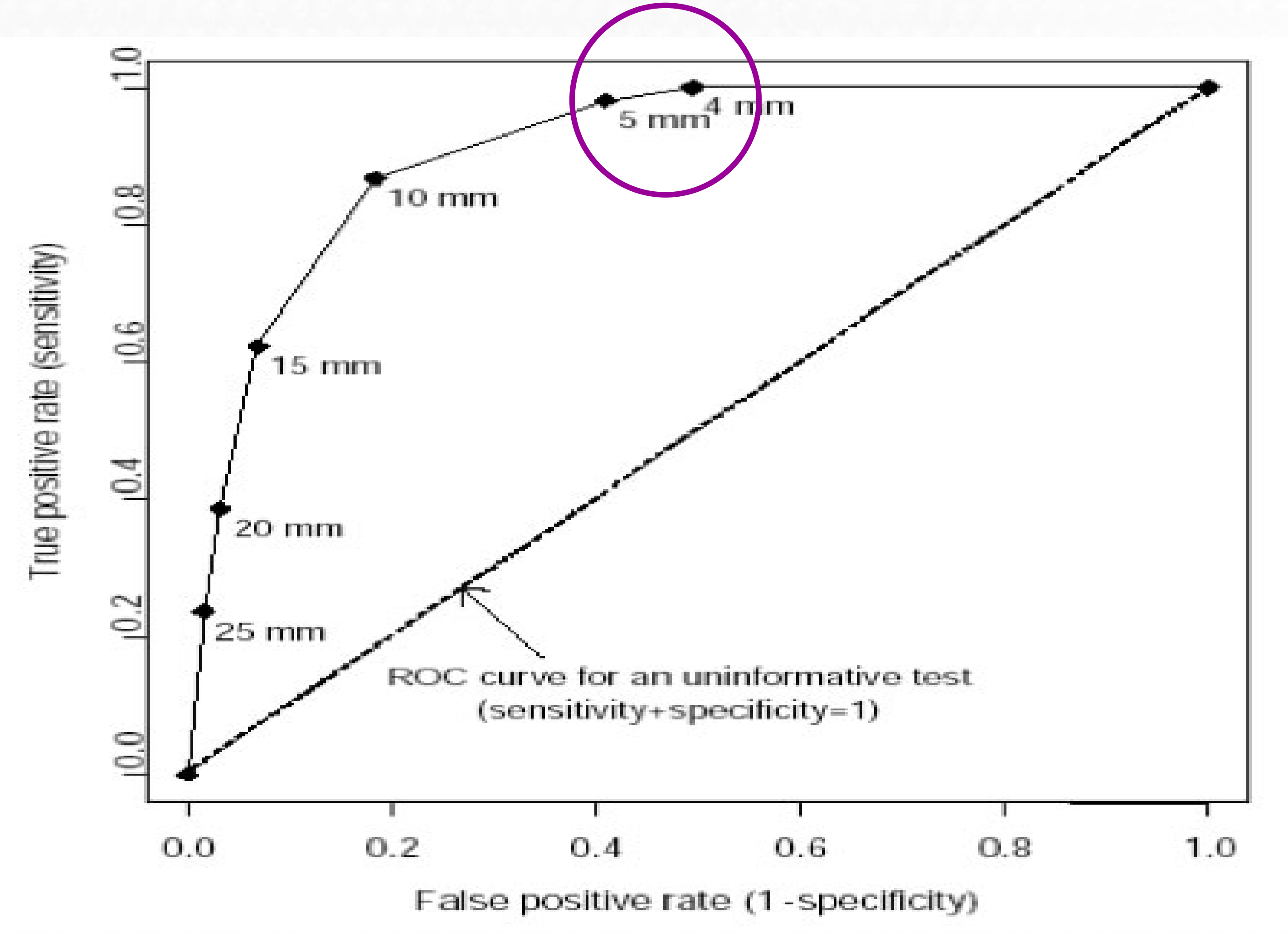
Ultrasounds can be used to detect thinning of the uterus Wall (indicative of posible tumor). If the result is positive a biopsy is required

Cutoff for abnormal wall thickness	Sentivity (%)	Specificity (%)	1-Specificity(%)
>4 mm	99	50	50
>5 mm	97	61	39
>10 mm	83	80	20
>15 mm	60	90	10
>20 mm	40	95	5
>25 mm	20	98	2

Objetive: To maximize the number of VP (correct diagnosis of cancer) with an aceptable number of FP (biopsies made when there was no cancer)

5. Receiver operator characteristic curves (ROC)

Example.



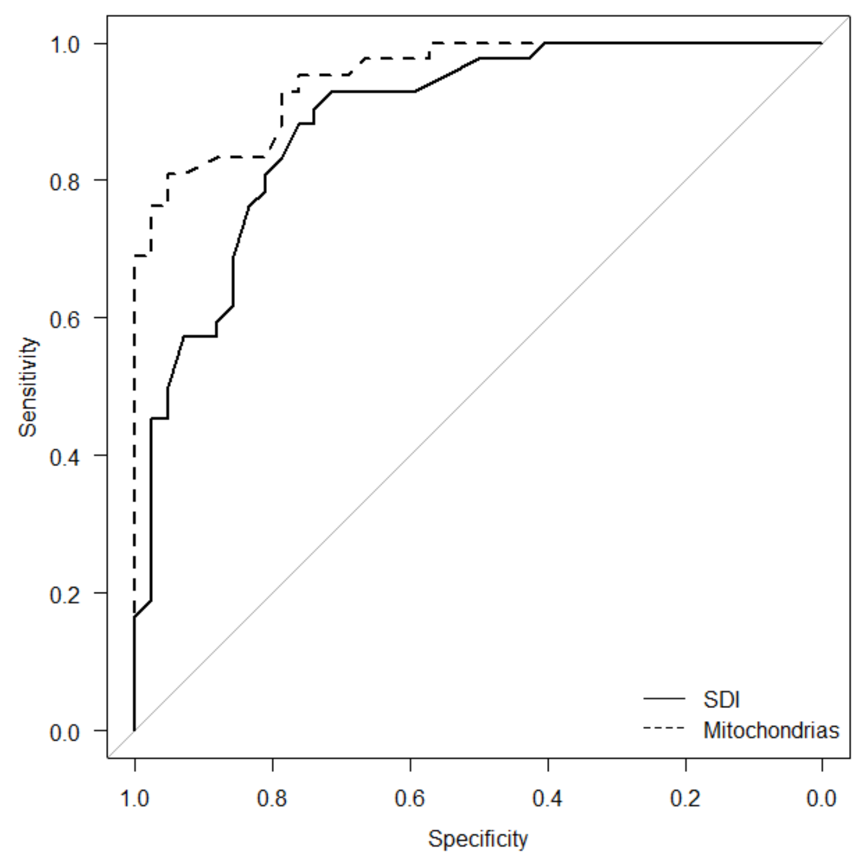
5. Receiver operator characteristic curves (ROC)

Comparison of ROC curves

Which of the two biomarkers is better?

AREA UNDER THE CURVE

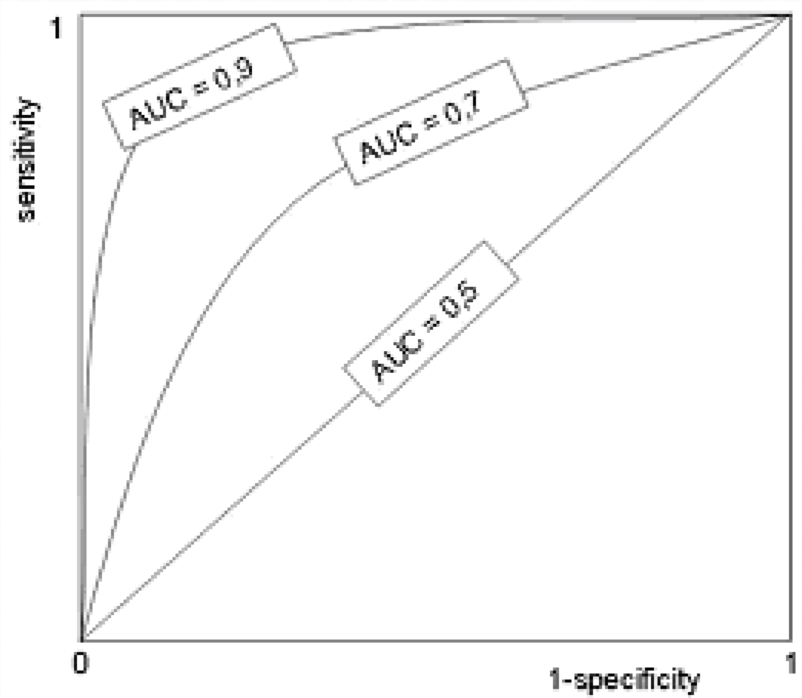
0<AUC<1



5. Receiver operator characteristic curves (ROC)

Comparison of ROC curves

AREA UNDER THE CURVE



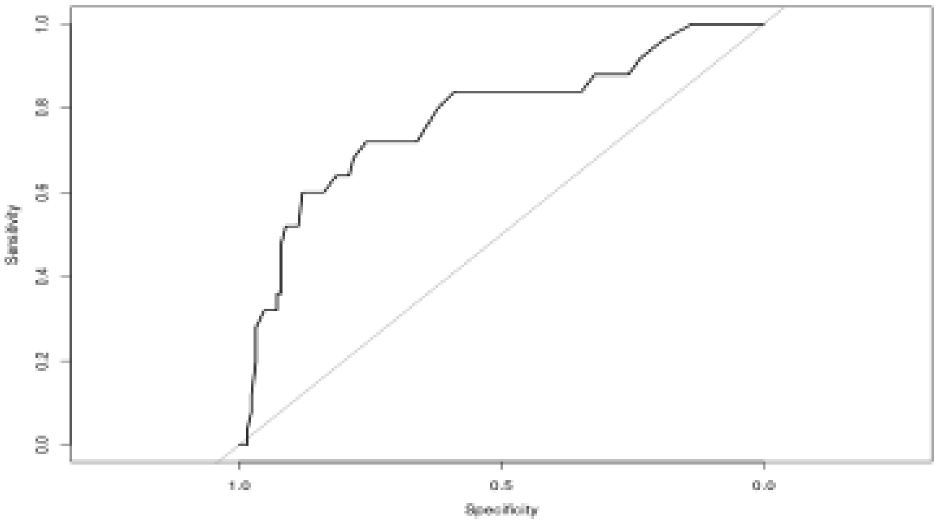
$$\text{AUC} = \begin{cases} 0.5 & \text{no discrimination} \\ 0.6\text{–}0.7 & \text{poor} \\ 0.7\text{–}0.8 & \text{accceptable (fair)} \\ 0.8\text{–}0.9 & \text{excellent (good)} \\ > 0.9 & \text{outstanding} \end{cases}$$

#Area under curve

```r
diabetes <- read_sav("datasets/diabetes.sav")
print(auc(factor(diabetes$MORT),diabetes$EDAT))
```

```
## Area under the curve: 0.775
```

```r
plot(roc(factor(diabetes$MORT),diabetes$EDAT,auc=TRUE))
```



5. Receiver operator characteristic curves (ROC)

Vall d'Hebron Institut de Recerca

Exercise. In R

Dataset: Osteoporosis.

To predict the clasificable variable, what diagnostic test do you think is best?

* body mass index
* bone density

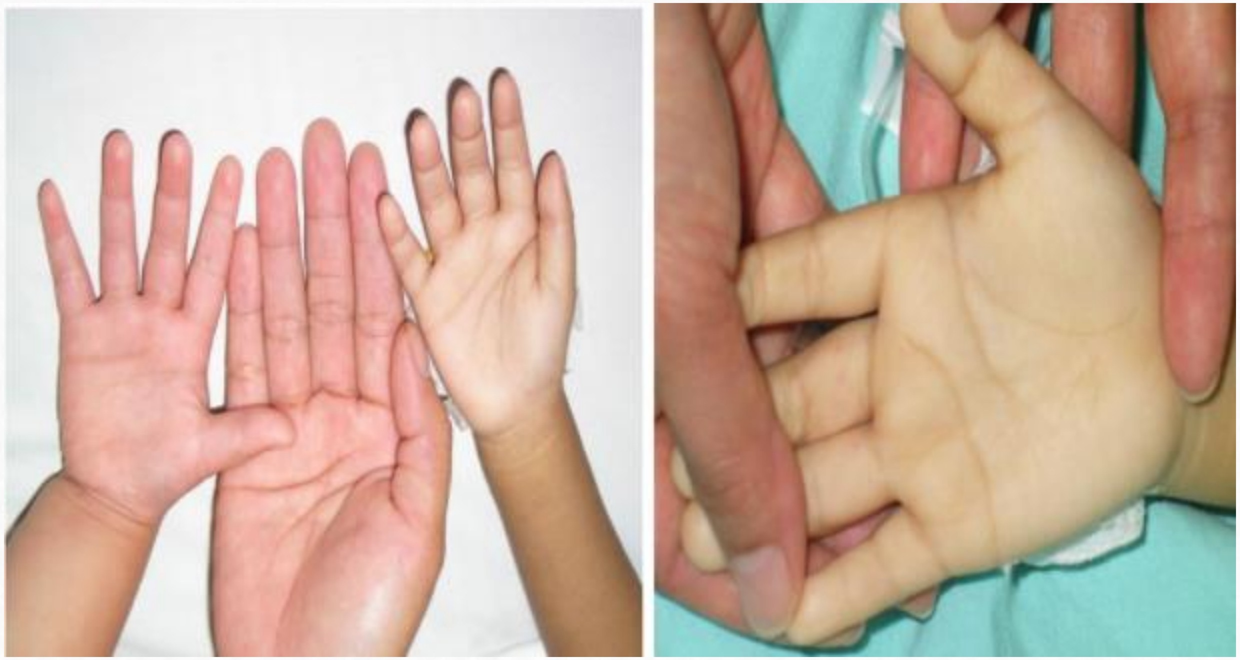
5. Receiver operator characteristic curves (ROC)

Example with R

```
library(pROC)
roc_bua <- roc(osteoporosis$menop, osteoporosis$bua)
roc_imc <- roc(osteoporosis$menop, osteoporosis$imc)
#review the roc object
roc_bua
roc_imc
#plot the ROC
plot(roc_bua)
plot(roc_imc)
#get the "best" "threshold"
coords(roc_bua, "best", "threshold", transpose = TRUE)
coords(roc_imc, "best", "threshold", transpose = TRUE)
#test to compare the two ROC
roc.test(roc_bua, roc_imc)
```

4. Likelihood ratio

Exercise: Calculate the same indices for the "palmar pallor sign" that was evaluated in a jungle region of Colombia to see if it could be of interest in the diagnosis of anemia. A blood count was taken in 167 children, so it was known whether they had anemia or not (48 yes, 119 no). The palmar pallor sign was positive in 16 anemics. In 95 non-anemic was negative. It is useful?



4. Likelihood ratio

Exercise: Calculate the same indices for the "palmar pallor sign" that was evaluated in a jungle region of Colombia to see if it could be of interest in the diagnosis of anemia. A blood count was taken in 167 children, so it was known whether they had anemia or not (48 yes, 119 no). The palmar pallor sign was positive in 16 anemics. In 95 non-anemic was negative. It is useful?

		Anemia		
		Disease	Healthy	Total
Sign	Positive	16	24	40
	Negative	32	95	127
	Total	48	119	167

4. Likelihood ratio

Vall d'Hebron Institut de Recerca

Exercise: Calculate the same indices for the "palmar pallor sign" that was evaluated in a jungle region of Colombia to see if it could be of interest in the diagnosis of anemia. A blood count was taken in 167 children, so it was known whether they had anemia or not (48 yes, 119 no). The palmar pallor sign was positive in 16 anemics. In 95 non-anemic was negative. It is useful?

	Disease positive	Disease negative	Total
Test positive	16	32	48
Test negative	24	95	119
Total	40	127	167

Point estimates and 95 % CIs:

	Estimation	Lower CI	Upper CI
Apparent prevalence	0.287	0.220	0.362
True prevalence	0.240	0.177	0.312
Sensitivity	0.400	0.249	0.567
Specificity	0.748	0.663	0.821
Positive predictive value	0.333	0.204	0.484
Negative predictive value	0.798	0.715	0.866
Diagnostic accuracy	0.665	0.588	0.736
Likelihood ratio of a positive test	1.588	0.979	2.575
Likelihood ratio of a negative test	0.802	0.611	1.053