Vall d'Hebron
Institut de Recerca

DIAGNOSTICS TESTS. SENSITIVITY, SPECIFICITY AND ROC CURVES

Curs d'Estadística Bàsica per a la Recerca Biomèdica

UEB – VHIR
Santiago Pérez-Hoyos. Ricardo Gonzalo

Vall d'Hebron Hospital · Generalitat de Catalunya Departament de Salut · Institut Català de la Salut · UAB Universitat Autònoma de Barcelona · Generalitat de Catalunya Departament d'Economia i Coneixement · Institució CERCA Centres de Recerca de Catalunya · Instituto de Salud Carlos III

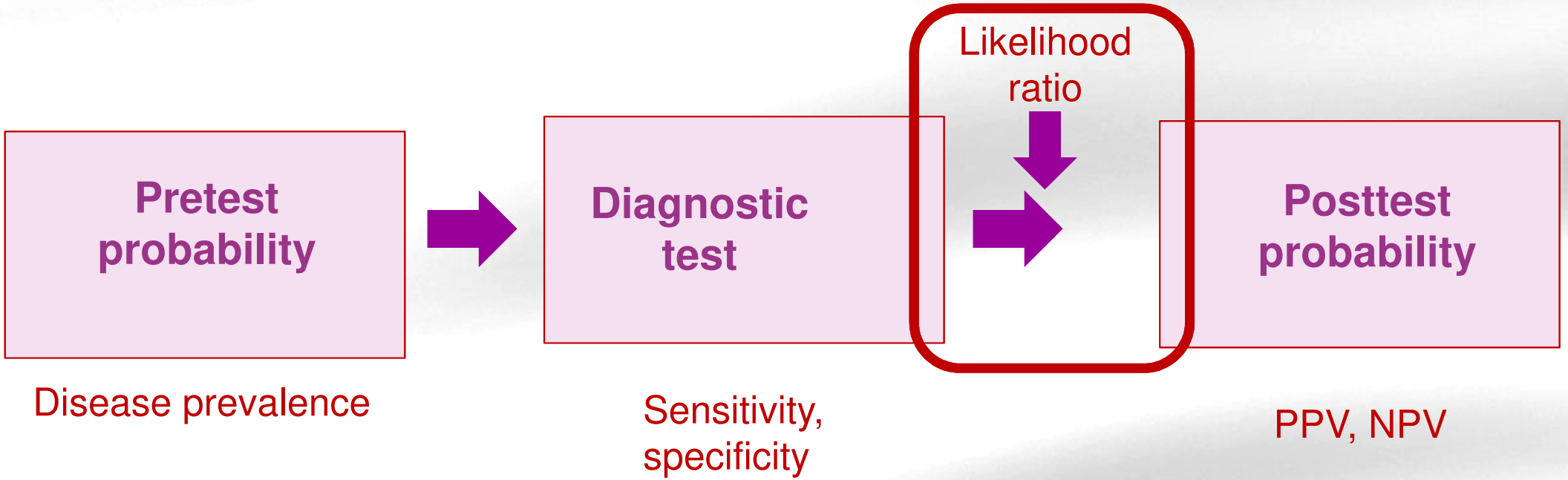
OUTLINE

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

Diagnostic


“If we did this 333 times, we’d be giving false hope to 332 people for approximately four hours and we’d save one life. Isn’t that worth it?”

Diagnostic Process



Diagnostic Measures

Disease

- Prevalence

Healthy

Diagnostic Measures

Test +

Disease

- Prevalence

Test +

Healthy

Diagnostic Measures

Test +

Sensitivity

Disease

- Prevalence

% False Positive

Test +

Healthy

Diagnostic Measures

Disease

Healthy

\-
Prevalence

Test -

Test -

Diagnostic Measures

Prevalence

Disease

% False negative

Test -

Healthy

Specificity

Test -

Diagnostic Measures

Test +

Sensitivity

Disease

Test -

% False negative

- Prevalence

Test +

% False Positive

Healthy

Test -

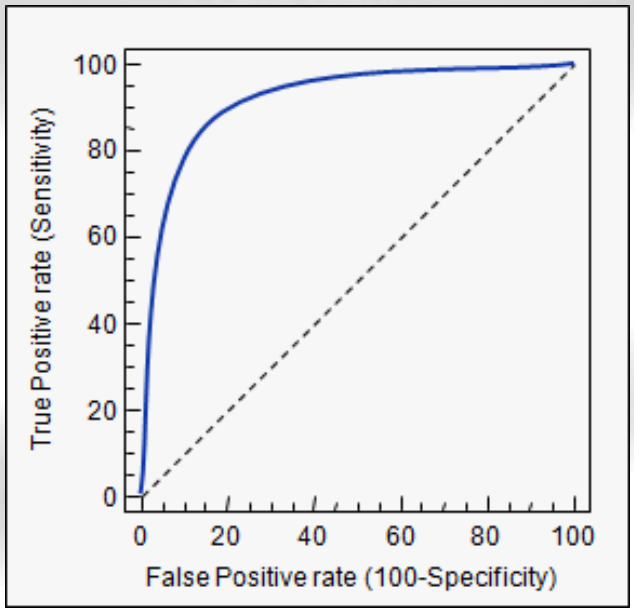
Specificity

Relationship between the results of the test and the authentic diagnosis

In a diagnostic test with possible results:

- Positive or Negative: Diagnostic table (contingency table)
- Numerical values: ROC curve

		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**



1. Diagnosis. Diagnostics tests

		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)

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

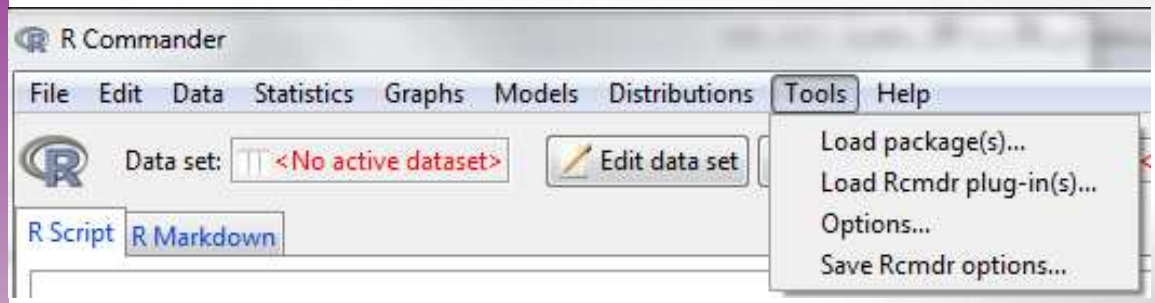
Sensitivity = 634 / (634+487) = 0.5656 = 56.6% → 43.4% with cancer had a normal rectal examination

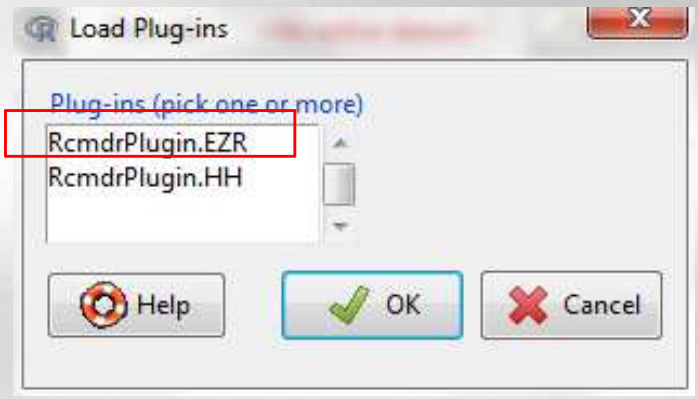
Specificity = 1251 / (269+1251) = 0.8230 = 82.3%

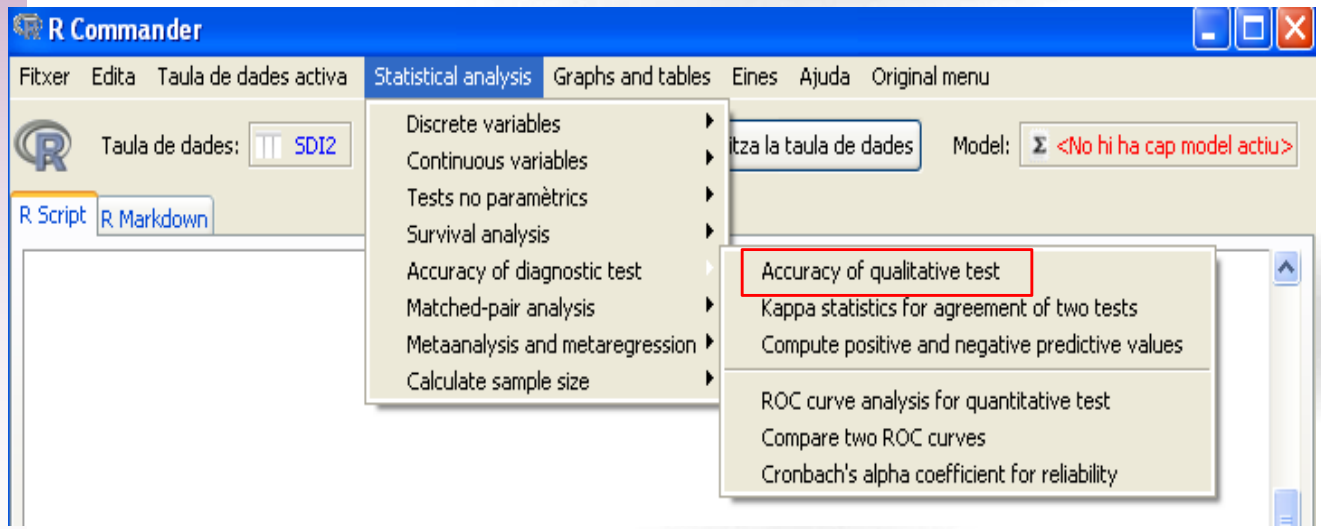
↓ 17.7% of the patients without disease were incorrectly diagnosed

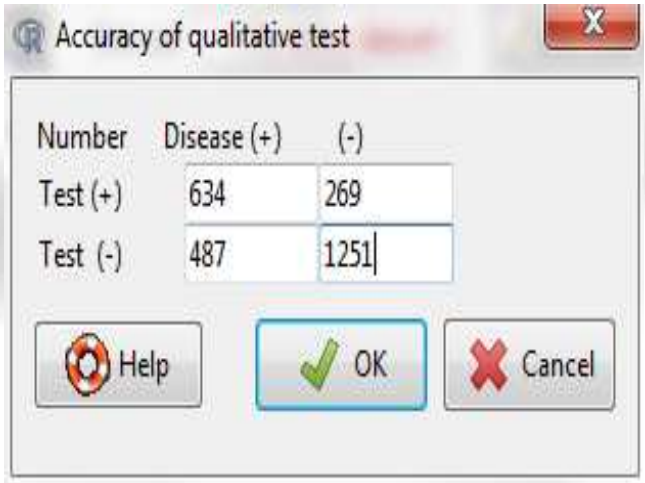
Other tests are needed to refine the diagnosis (Ex. PSA)

Example with R Commander. Plug-in "EZR"









Example with R Commander. Plug-in "EZR"

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

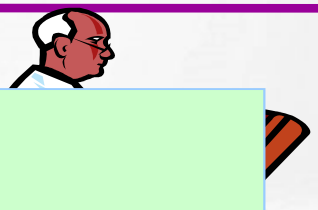
Point estimates and 95 % CIs:

	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

Positive/Negative predictive value

If the test is positive which is the probability the patient is really infected

Diagnostic Measures



+ Predictive Value

Test +

Test + ± Test +

Sensitivity

% False negative

% False Positive

Specificity

Test +

Test -

Test +

Test -

Healthy

Diagnostic Measures

Sensitivity

Test +

Disease

% False negative

Test -

Test +

% False Positive

Test -

Specificity

Test -

- Predictive Value

Test -

Test -

+

Test -

PPV, NPV Example with R Commander. Plug-in "EZR"

	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

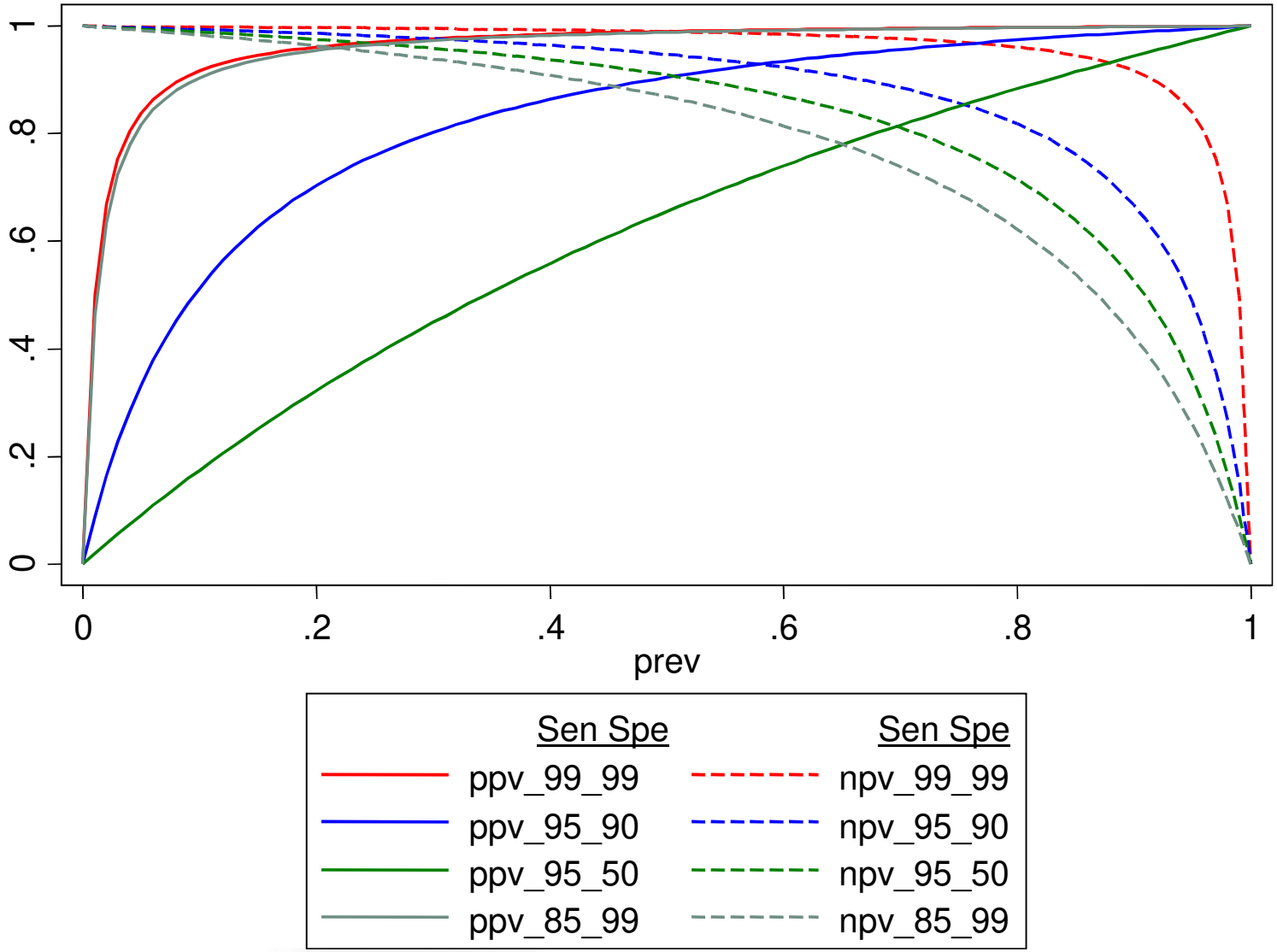
PPV, NPV with more prevalence

	Disease positive	Disease negative	Total
Test positive	6340	259	6599
Test negative	4870	1251	6121
Total	11210	1510	12720

Point estimates and 95 % CIs:

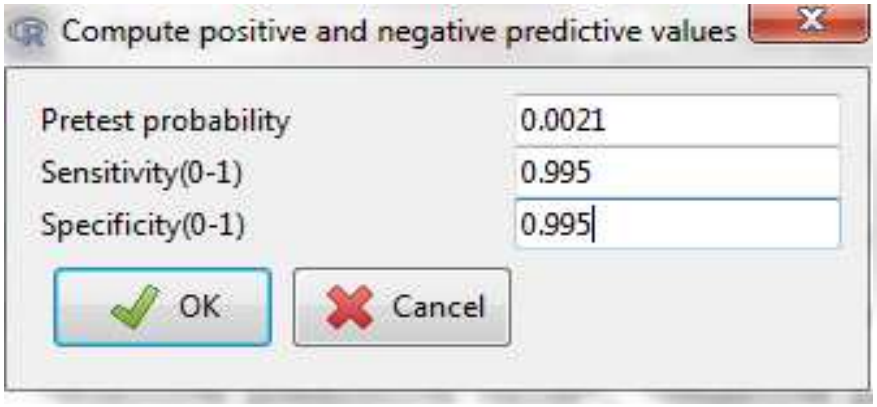
	Estimation	Lower CI	Upper CI
Apparent prevalence	0.519	0.510	0.528
True prevalence	0.881	0.876	0.887
Sensitivity	0.566	0.556	0.575
Specificity	0.828	0.808	0.847
Positive predictive value	**0.961**	**0.956**	**0.965**
Negative predictive value	**0.204**	**0.194**	**0.215**
Diagnostic accuracy	0.597	0.588	0.605
Likelihood ratio of a positive test	3.297	2.948	3.688
Likelihood ratio of a negative test	0.524	0.508	0.541

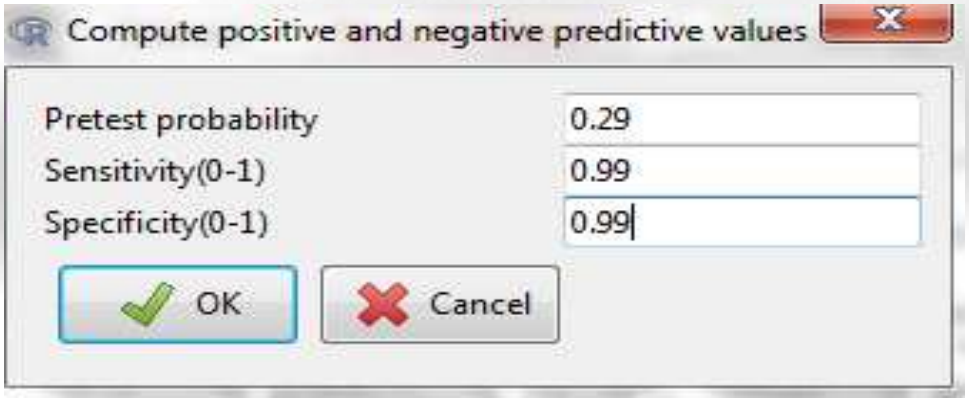
PPV & NPV depend on prevalence



Dependence of PPV and NPV on disease prevalence

Example: VIH diagnosis of two populations with RCommander






```
> predictive.value
                              Assumptions
Pretest probability                0.0021
Sensitivity                        0.995
Specificity                        0.995

                                 Estimated
Positive predictive value            0.295
Negative predictive value                1
```

```
> predictive.value
                              Assumptions
Pretest probability                  0.29
Sensitivity                          0.99
Specificity                          0.99

                                 Estimated
Positive predictive value            0.976
Negative predictive value            0.996
```

Diagnostic Measures

Likelihood Ratio +

Sensitivity

% False Positive

Sensitivity

% False negative

Test +

Test -

% False Positive

Test +

Healthy

Specificity

Test -

Diagnostic Measures

Test +

Sensitivity

Disease

Test -

% False negative

Likelihood Ratio -

% False negative

Test +

% False Positive

Specificity

Test -

Specificity

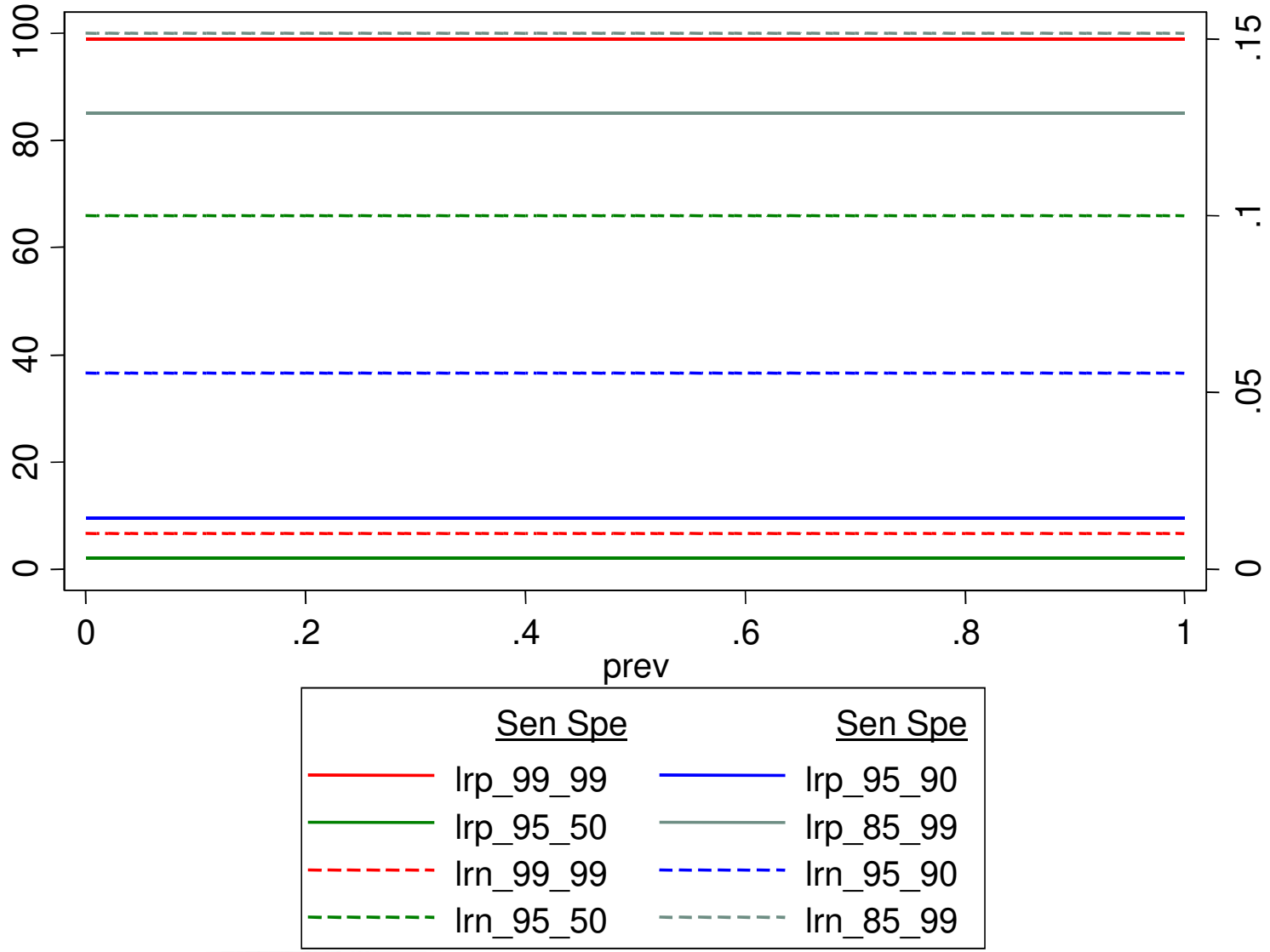
LR+,LR-Example with R Commander. Plug-in "EZR"

	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**

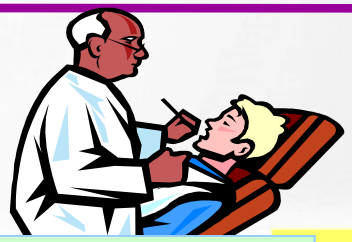
LR+ & LR- do NOT depend on prevalence



Interpretation:

- Positive Likelihood Ratio (LR+)
- LR+ over 5 - 10: Significantly increases likelihood of the disease
- LR+ between 0.2 to 5 (esp 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

Diagnostic Measures



Accuracy=% classified right

Test + ± Test -

Disease ± Healthy

Sensitivity

% False negative

% False Positive

Specificity

Test +

Test -

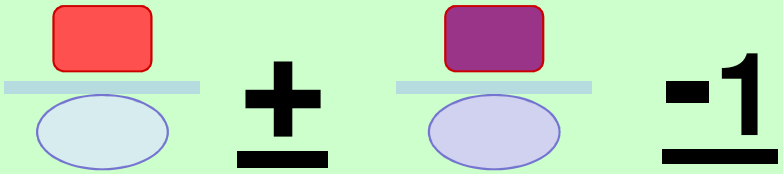
Test +

Test -

Diagnostic Measures

Youden's Index =

sensitivity + specificity -1



Sensitivity

Test +

% False negative

Test -

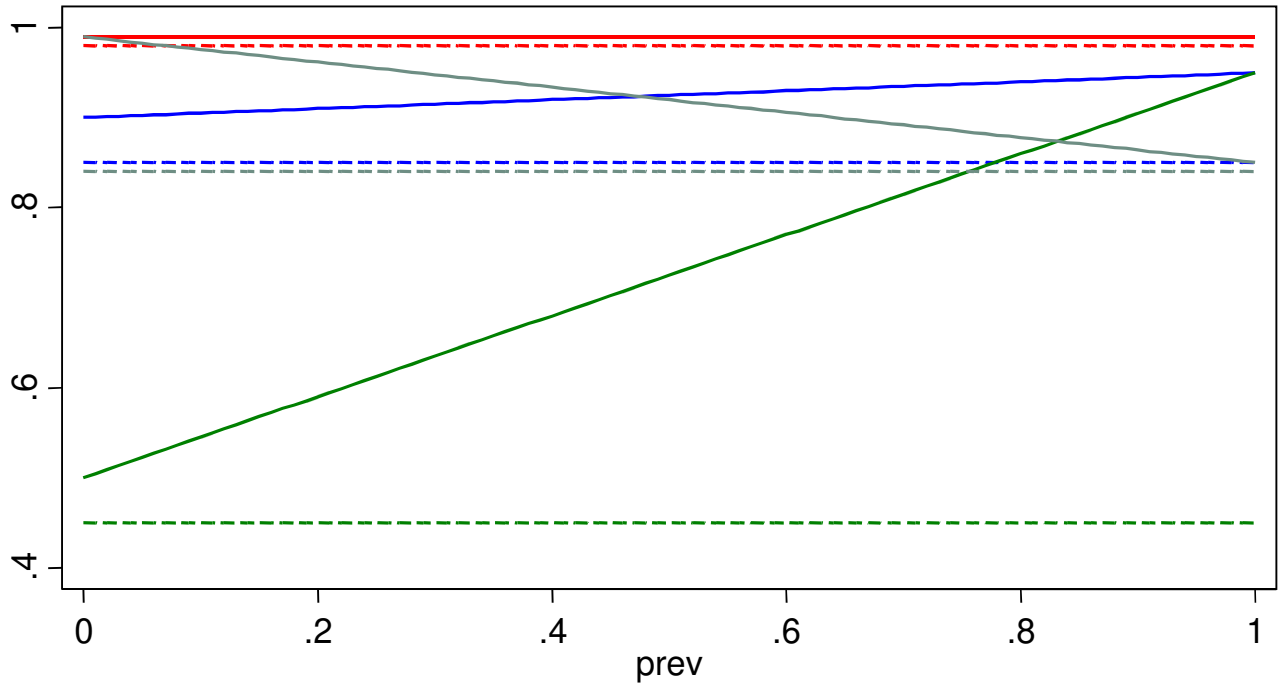
% False Positive

Test +

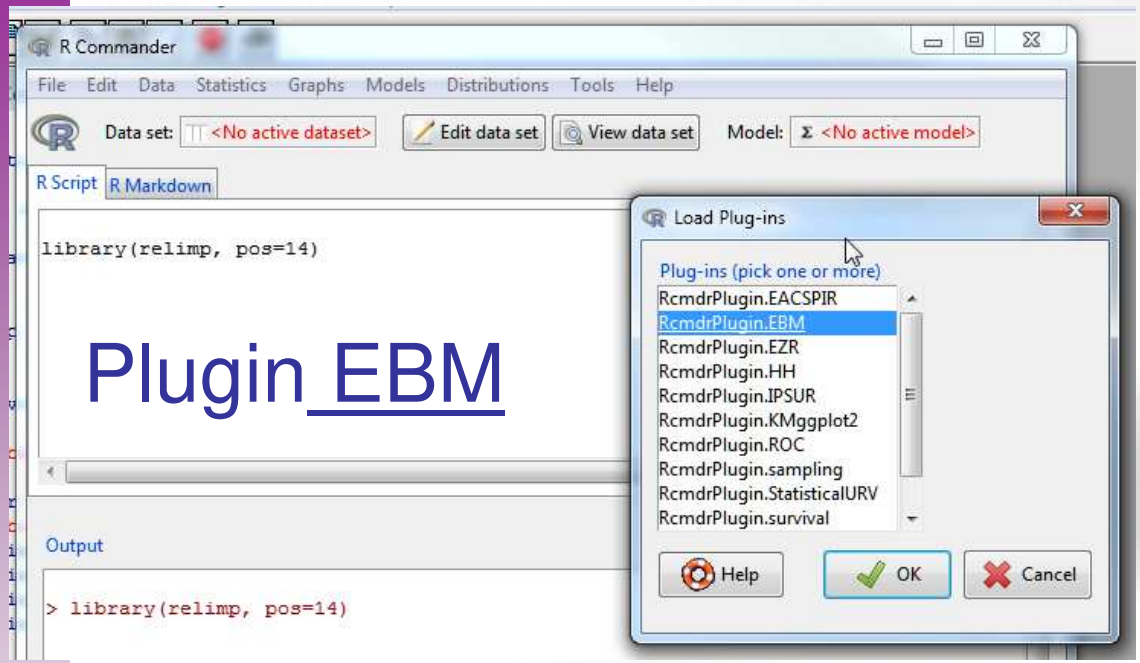
Specifity

Test -

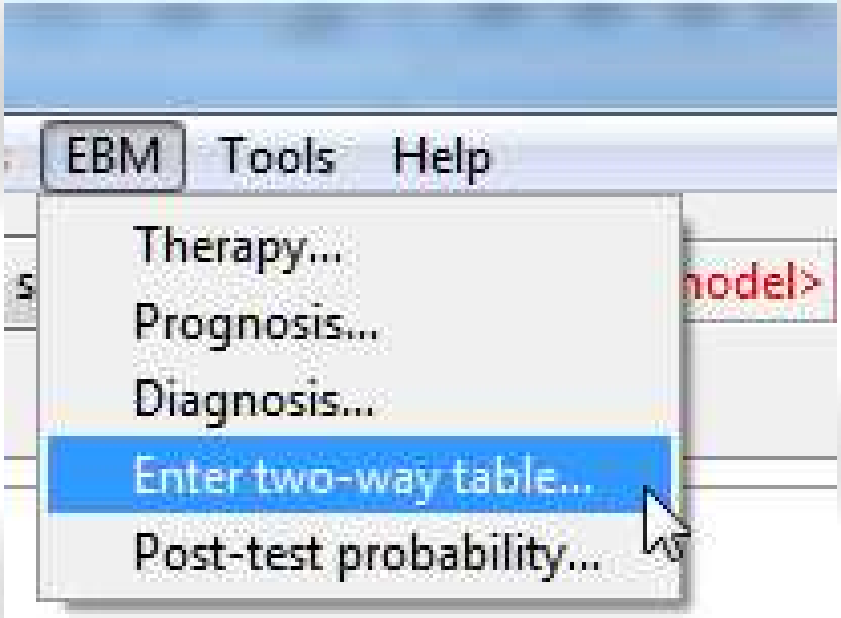
Accuracy depend on prevalence
Youden's Index do NOT depend on prevalence

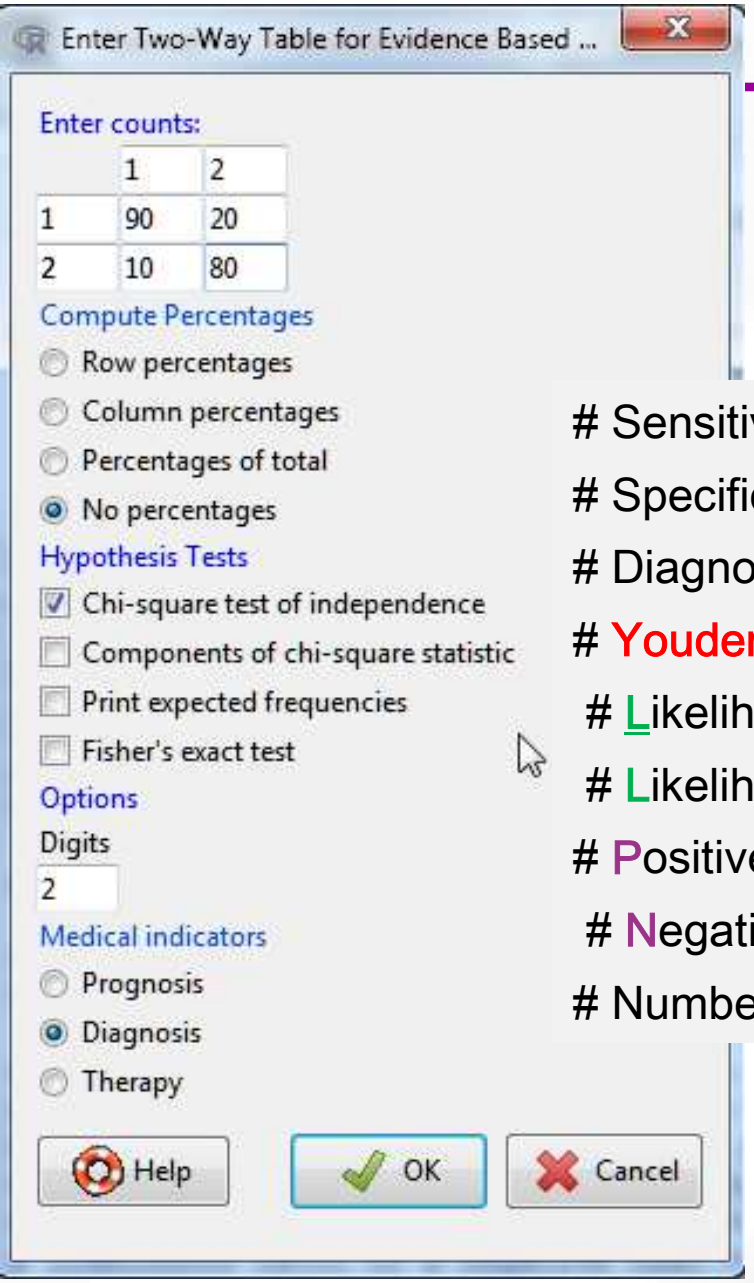


Diagnostic Measures in R Commander



Plugin EBM





	Disease	Non Disease	Total
Biomarker +	90	20	110
Biomarker -	10	80	90
Total	100	100	200

Sensitivity (Se) = 90 (95% CI 82.38 - 95.1) %

Specificity (Sp) = 80 (95% CI 70.82 – 87.33)%

Diagnostic acuracy (% of all correct results) = 85 (95% CI 79.28- 89.65) %.

Youden's index = 0.7 (95% CI 0.53 - 0.82).

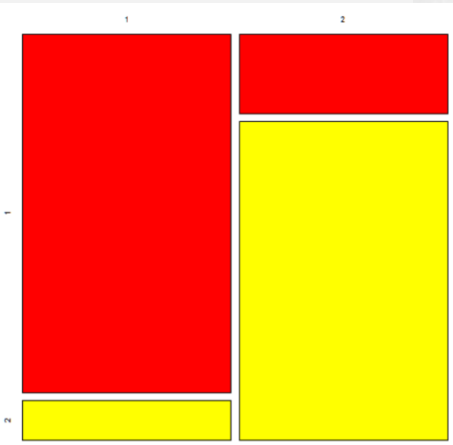
Likelihood ratio of a positive test = 4.5 (95% CI 3.02 – 6.7).

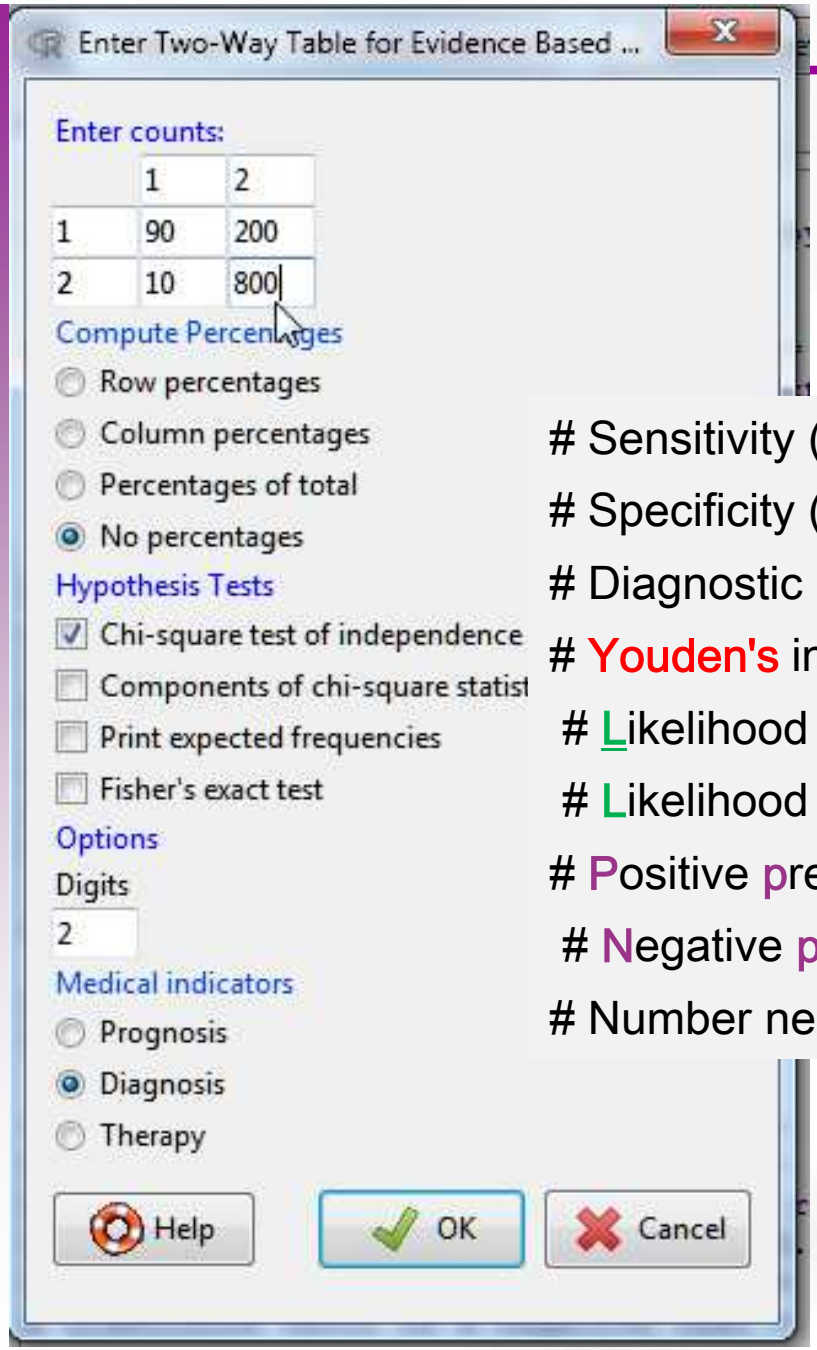
Likelihood ratio of a negative test = 0.12 (95% CI 0.07 - 0.23).

Positive predictive value = 81.82 (95% CI 73.33 – 88.53) %.

Negative predictive value = 88.89 (95% CI 80.51- 94.54) %.

Number needed to diagnose = 1.43 (95% CI 1.21 – 1.88)





	Disease	Non Disease	Total
Biomarker +	90	200	290
Biomarker -	10	800	910
Total	100	1000	1100

Sensitivity (Se) = 90 (95% CI 82.38 - 95.1) %

Specificity (Sp) = 80 (95% CI 77.38 - 82.44)%

Diagnostic acuracy (% of all correct results) = 80.91 (95% CI 78.46 - 83.19) %.

Youden's index = 0.7 (95% CI 0.6 - 0.78).

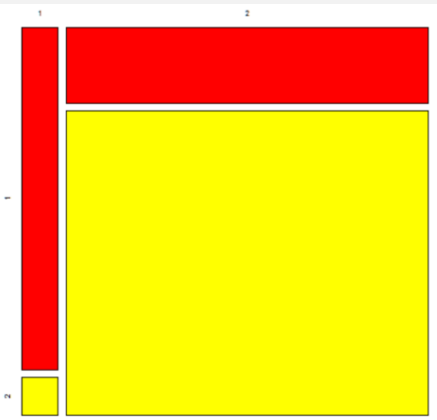
Likelihood ratio of a positive test = 4.5 (95% CI 3.91 - 5.18).

Likelihood ratio of a negative test = 0.12 (95% CI 0.07 - 0.23).

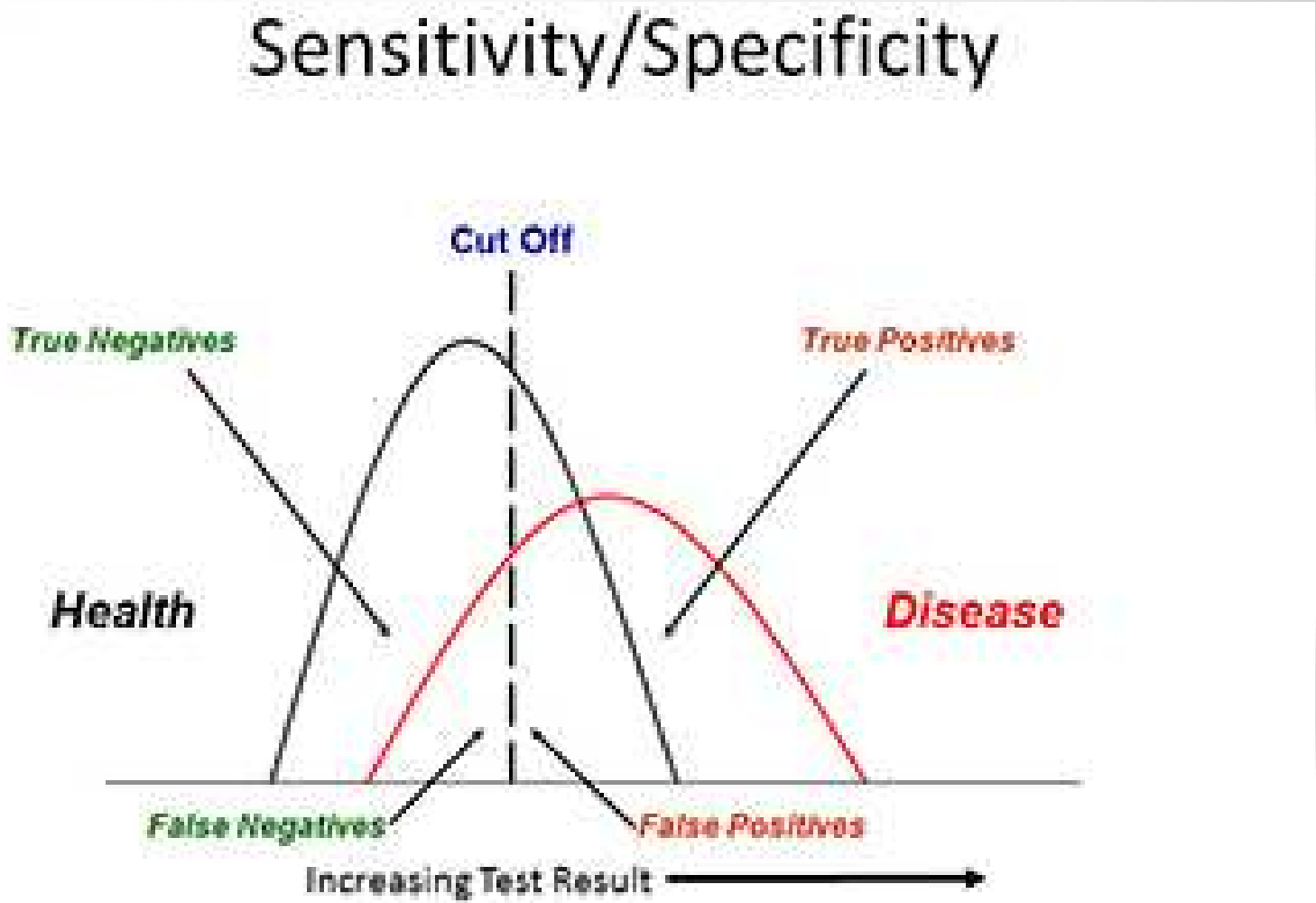
Positive predictive value = 31.03 (95% CI 25.76 - 36.71) %.

Negative predictive value = 98.77 (95% CI 97.74 - 99.41) %.

Number needed to diagnose = 1.43 (95% CI 1.29 - 1.67).

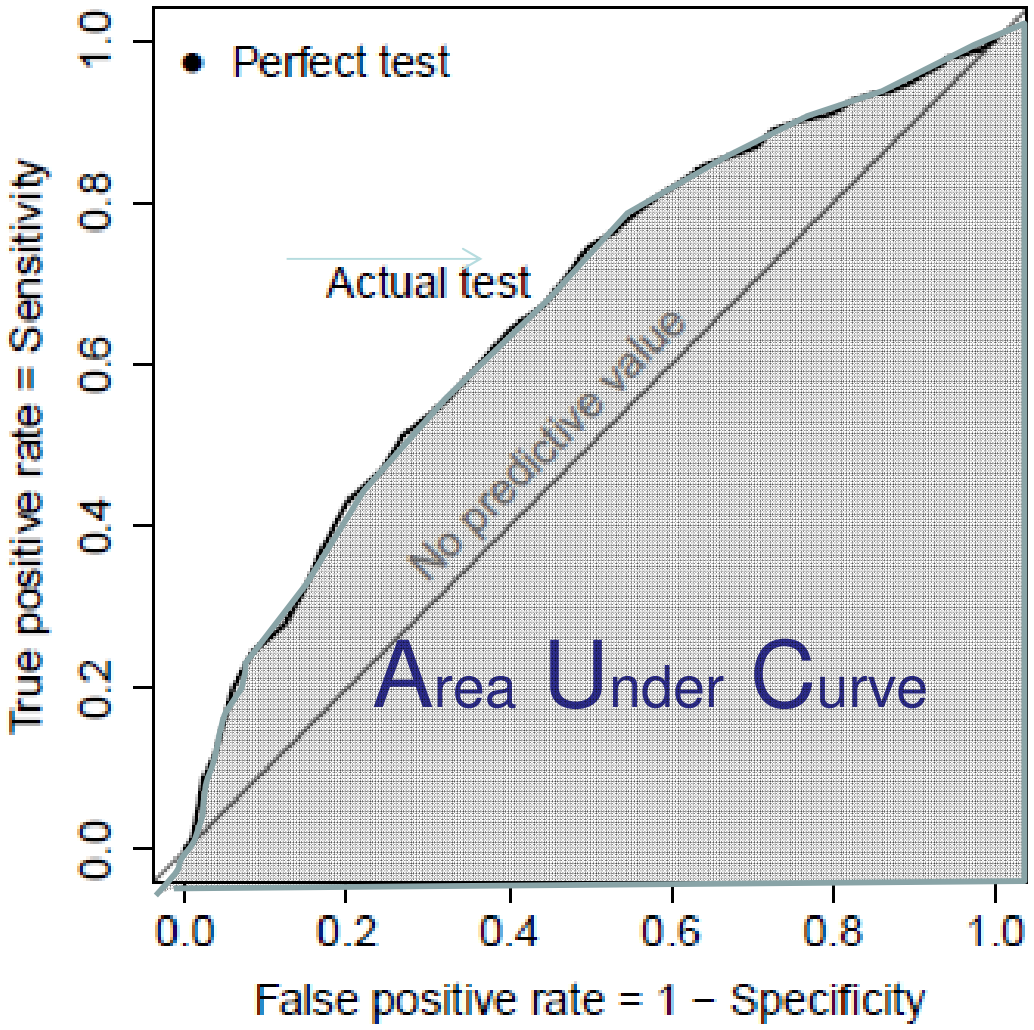


Discrimination for quantitative test o biomarker



- If it is not a single biomarker or test you can build one by fitting a statistical model such as logistic regression
- _Select the best cutpoint to categorize biomarker or test in two groups (ROC Curve, Area Under Curve)
- Proceed as before (Diagnostic Measures)

Receiver Operating Characteristic curves

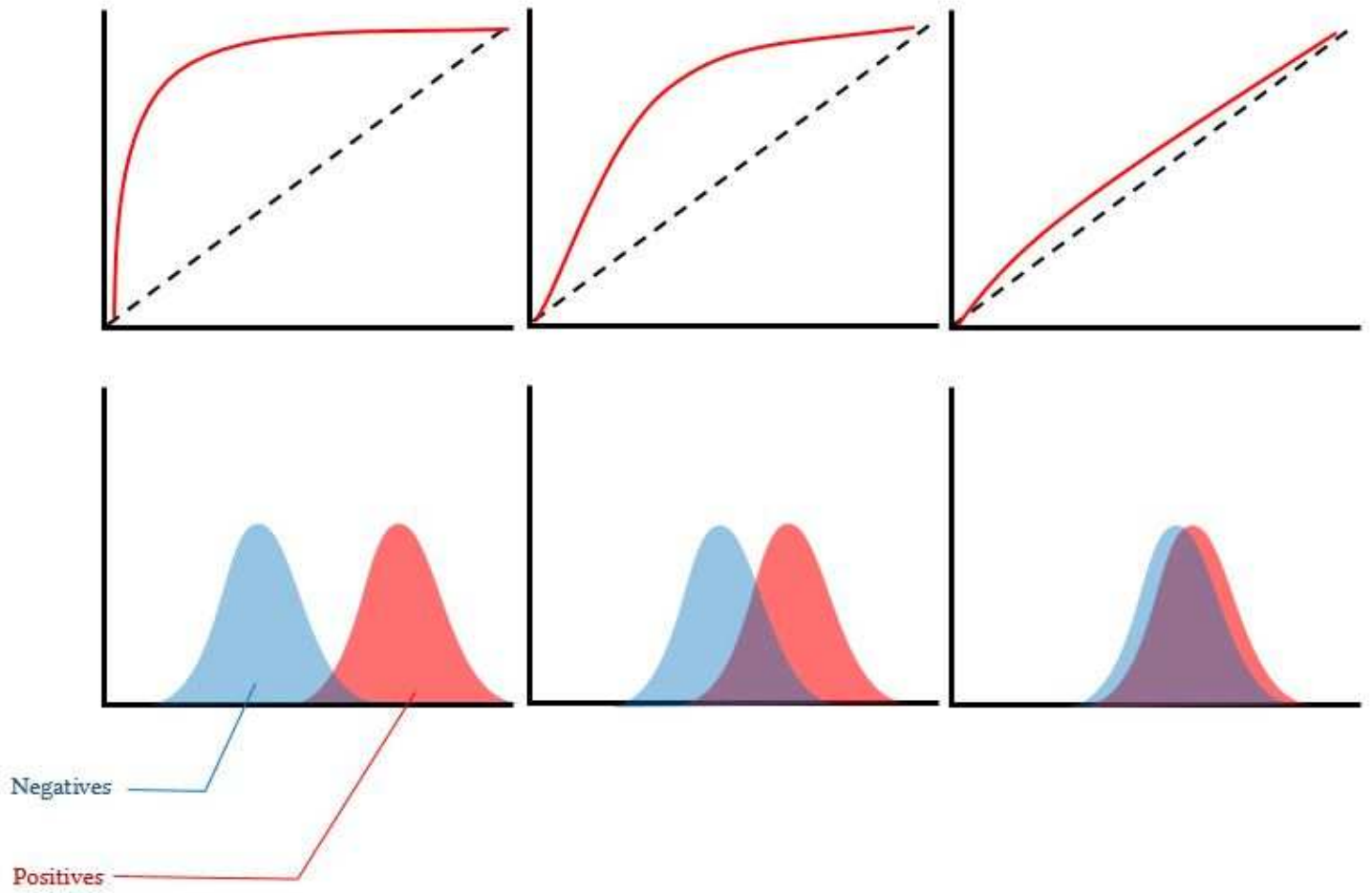


Sensitivity vs false positive rate

Find best threshold for discrimination

(c-index) is the fraction of pairs (i,j) in which the predicted risk was
higher for the subject who had the event:
C<0.5 No classification
C>0.5 Successful classification
C =1 perfect classification

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

Example

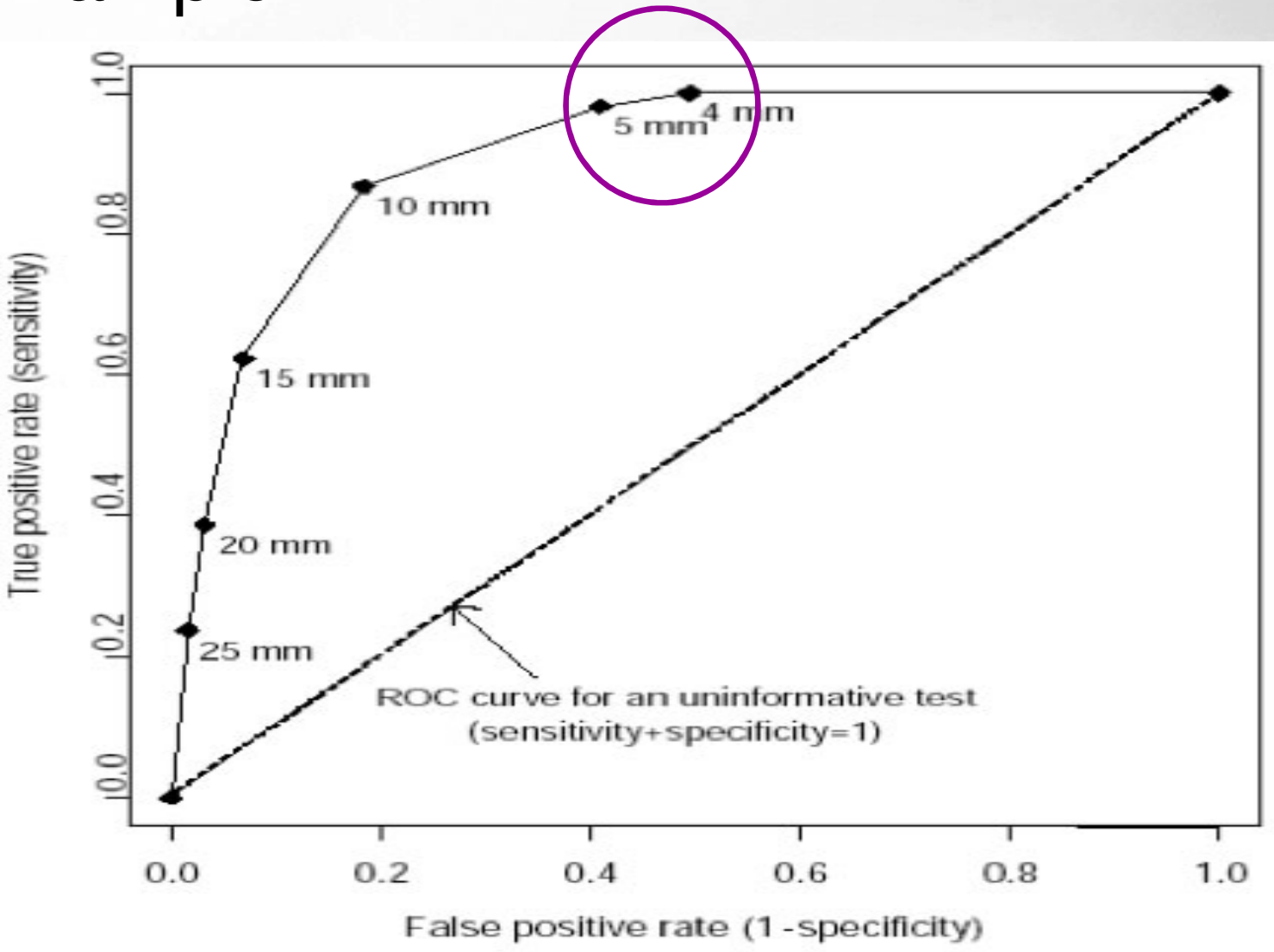
Ultrasounds can be used to detect thinning of the uterus Wall (indicative of posible tumor). If the result is positive a biopsy is required

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)

Receiver operator characteristic curves (ROC)

- Example.

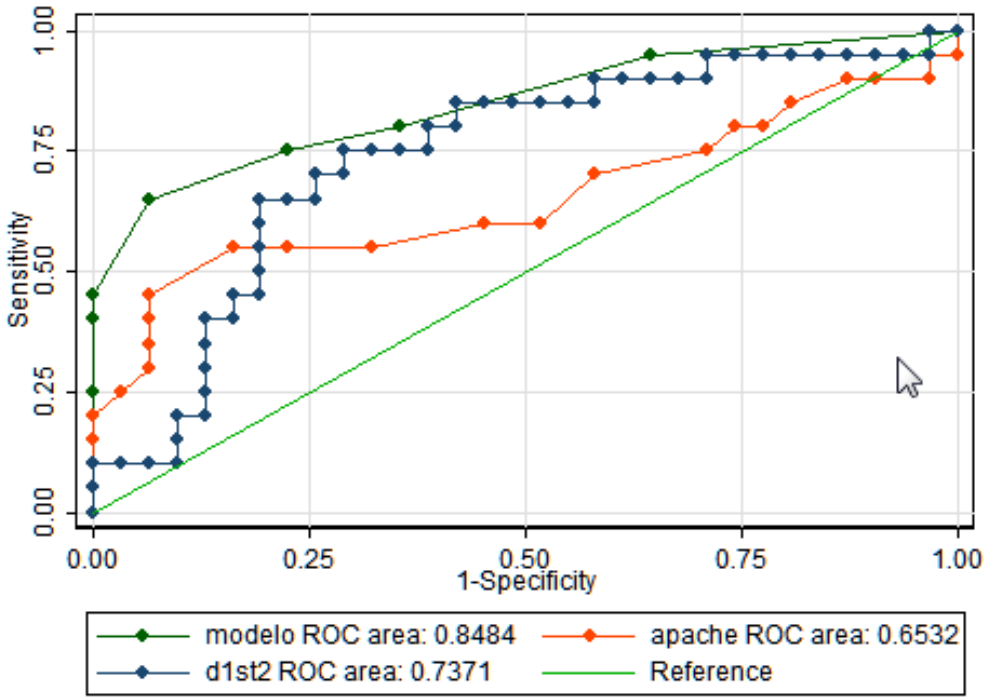


Regressio Multivariant logit per a mort_60

Number of obs = 51

VARIABLE		OR	(95%CI)	p-value
Día 1 ST-2 >4153.60	No(<4153.60)	1		0.0095
	Si(>4153.60)	8.12	(1.67;39.60)	
APACHE II >26.00	No(<26.00)	1		0.0062
	Si(>26.00)	15.88	(2.19;114.85)	
Inmunosupresión previa	No	1		0.0254
	Sí	7.90	(1.29;48.43)	

LL model= -22.01 ; AIC model= 52.03 ; BIC model= 59.75



Detailed report of Sensitivity and Specificity

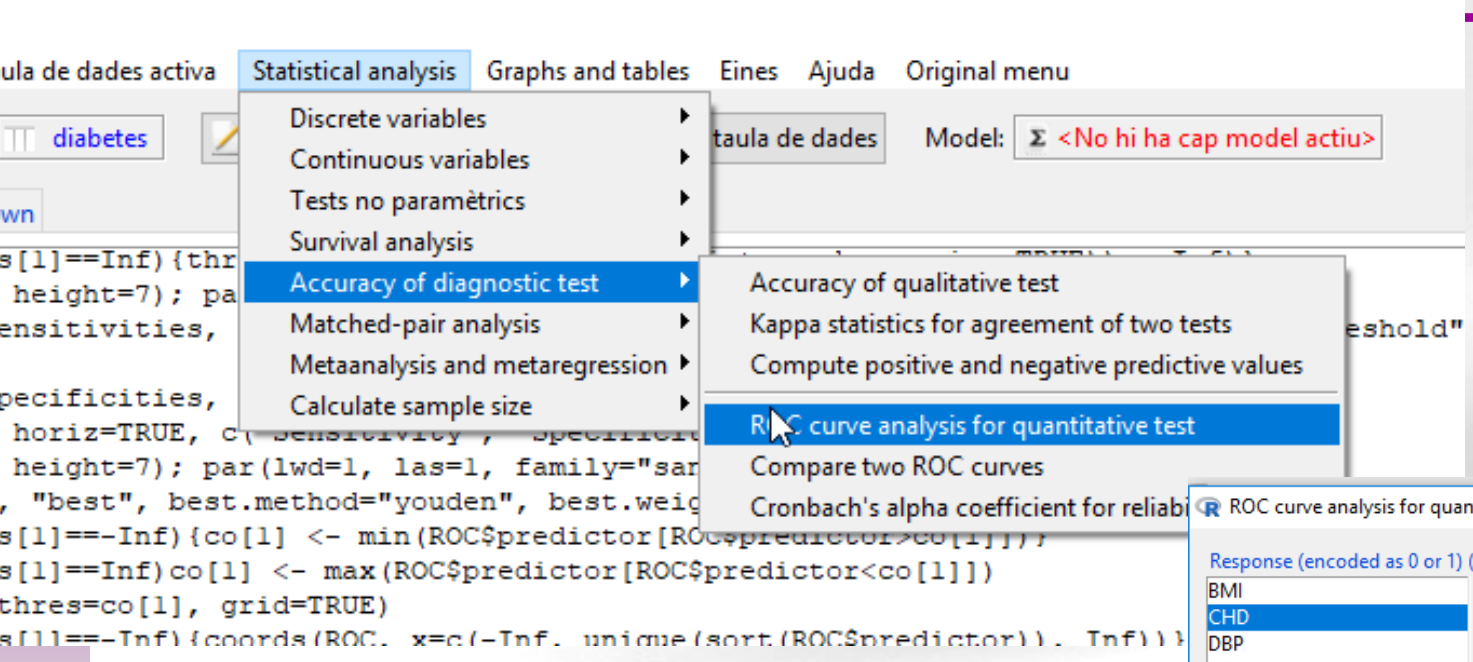
CUTPOINT	SENSITIVITY	SPECIFICITY	CORRECTED CLASSIFIED	LR+	LR-
(>= .0393311)	100.00%	0.00%	39.22%	1.0000	
(>= .244451)	95.00%	35.48%	58.82%	1.4725	0.1409
(>= .2496142)	80.00%	64.52%	70.59%	2.2545	0.3100
(>= .3939386)	75.00%	77.42%	76.47%	3.3214	0.3229
(>= .7244239)	65.00%	93.55%	82.35%	10.0750	0.3741
(>= .8370442)	45.00%	100.00%	78.43%		0.5500
(>= .8407952)	40.00%	100.00%	76.47%		0.6000
(>= .9766)	25.00%	100.00%	70.59%		0.7500
(> .9766)	0.00%	100.00%	60.78%		1.0000

Observaciones	ROC AREA	Std. Error	Asymptotic Normal [95% Conf. Interval]	Indice Youden	Punto corte según Youden	Mejor Punto corte AUC
51	0.8484	0.0585	0.7338 ; 0.9630	0.5855	0.72	0.39

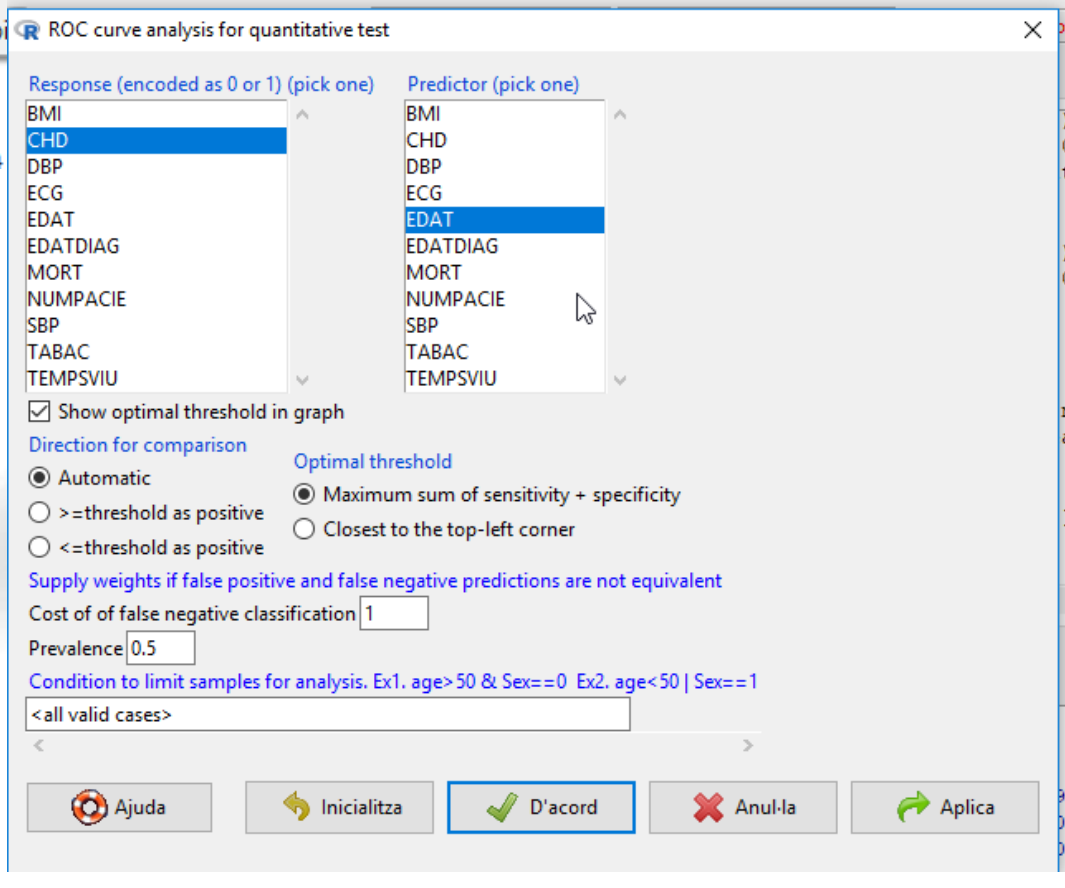
Vall d'Hebron
Institut de Recerca

UNITAT
D'ESTADÍSTICA I
BIOINFORMÀTICA

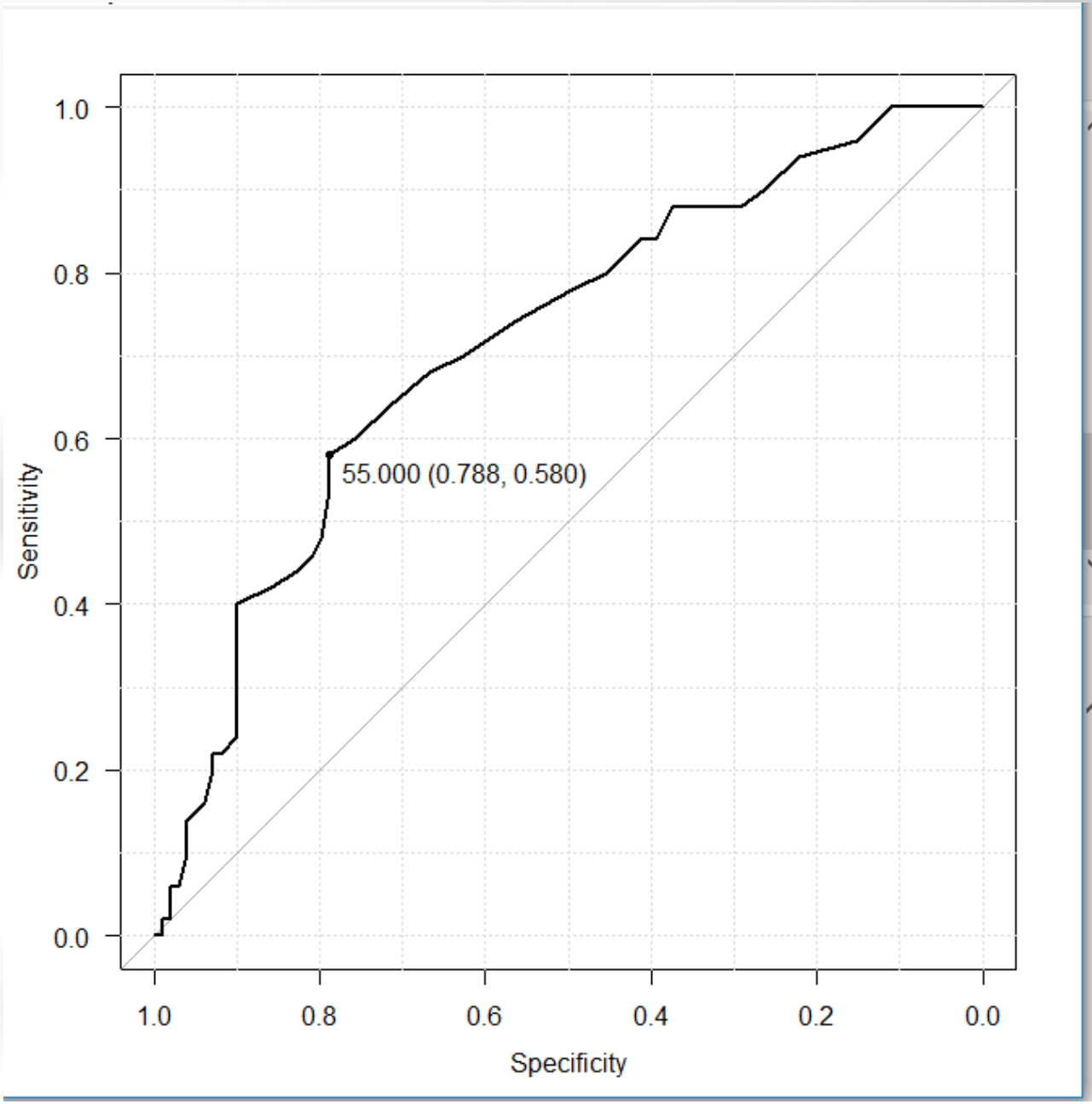
ROC curves in R Commander



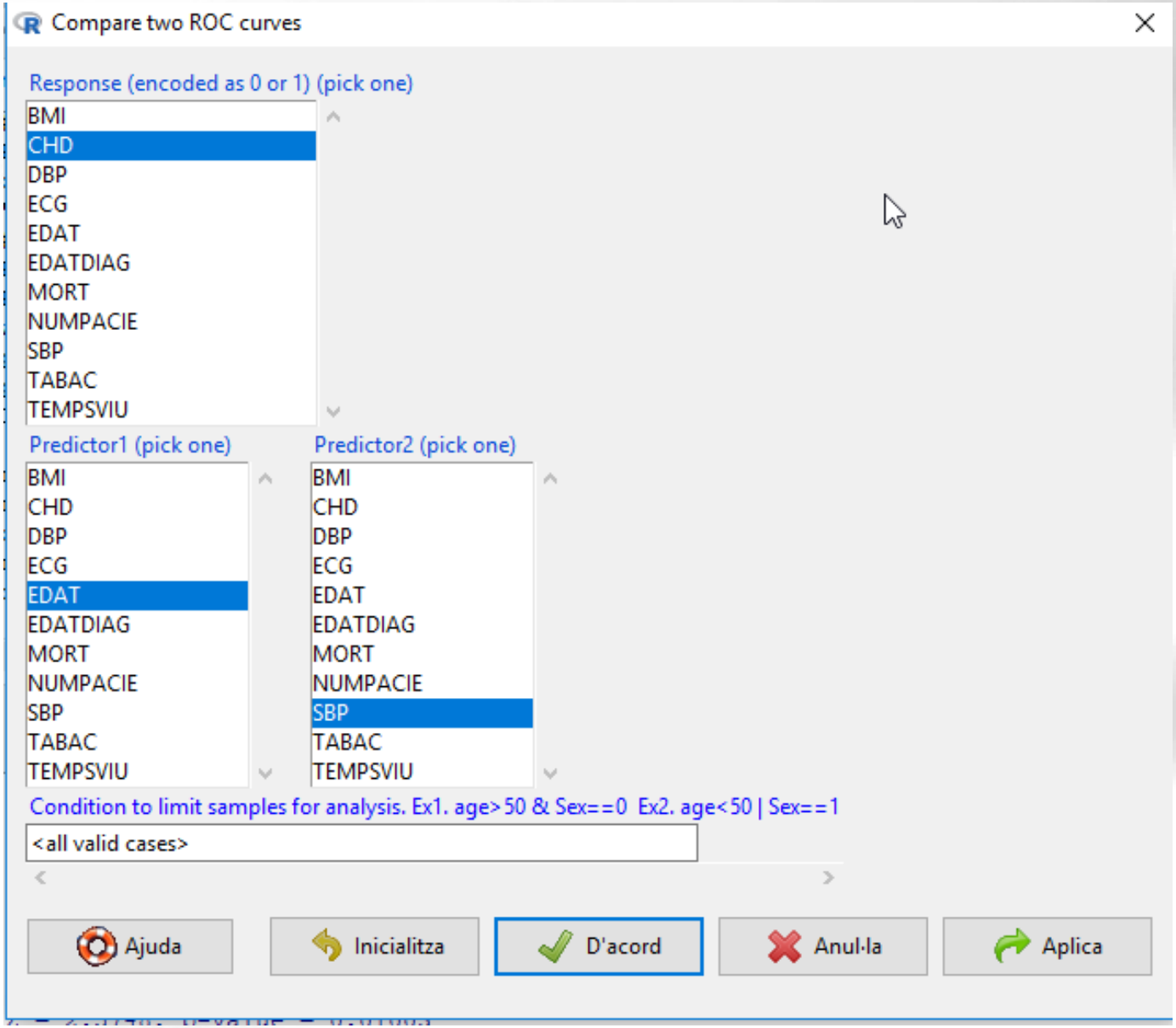

```
s[1]==Inf){thr
 height=7); pa
ensitivities,
                                                    eshold"
pecificities,
 horiz=TRUE, c(
 height=7); par(lwd=1, las=1, family="sar
, "best", best.method="youden", best.weig
s[1]==-Inf){co[1] <- min(ROC$predictor[ROC$predictor>co[1]])}
s[1]==Inf)co[1] <- max(ROC$predictor[ROC$predictor<co[1]])
thres=co[1], grid=TRUE)
s[1]==-Inf){coords(ROC, x=c(-Inf, unique(sort(ROC$predictor)), Inf))}
```



ROC curve for edat



Compare ROC curve for edat and SBP



Compare ROC curve for edat and SBP