Categorical Data Analyses Continued: Using the Fabulous 2x2 Table to Calculate Odds Ratio, Relative Risks, Sensitivity, Specificity, ROC…

Kathleen Torkko, PhD, MSPH, MS
Associate Professor
Department of Pathology (SOM)
Department of Biostatistics and Bioinformatics (SOPH)

November 6, 2019

Benjamini-Hochberg FDR

p-values sorted	rank
0.0002	1
0.0004	2
0.0016	3
0.0057	4
0.0091	5
0.0187	6
0.0225	7
0.0364	8
0.0441	9
0.0473	10
0.0536	11
0.0779	12
0.0862	13
0.1081	14
0.2341	15
0.357	16
0.4682	17
0.642	18
0.6833	19
0.8248	20

Calculate each individual p-value's Benjamini-Hochberg critical value, using the formula (i/m)Q, where:

i = the individual p-value's rank,
m = total number of tests
in our case = 20
Q = the false discovery rate (a percentage, chosen by you)
we will use 0.1

(i/m)Q
(1/20) x 0.10 = 0.005
(2/20) x 0.10 = 0.010
(3/20) x 0.10 = 0.015
(4/20) x 0.10 = 0.020
(5/20) x 0.10 = 0.025
(6/20) x 0.10 = 0.030
(7/20) x 0.10 = 0.035
(8/20) x 0.10 = 0.040
(9/20) x 0.10 = 0.045
(10/20) x 0.10 = 0.050
(11/20) x 0.10 = 0.055 …

p-values sorted	rank	(i/m)Q
0.0002	1	0.005
0.0004	2	0.010
0.0016	3	0.015
0.0057	4	0.020
0.0091	5	0.025
0.0187	6	0.030
0.0225	7	0.035
0.0364	8	0.040
0.0441	9	0.045
0.0473	10	0.050
0.0536	11	0.055
0.0779	12	0.060
0.0862	13	0.065
0.1081	14	0.070
0.2341	15	0.075
0.357	16	0.080
0.4682	17	0.085
0.642	18	0.090
0.6833	19	0.095
0.8248	20	0.100

Find the highest p-value that is also smaller than (i/m)Q

All p-values below 0.0536 are considered significant at Q=0.10

Even if the individual p-values might be greater than their individual calculated (i/m)Q value

Dietary variable	*P* value	Rank	(i/m)Q
Total calories	<0.001	1	0.010
Olive oil	0.008	2	0.020
Whole milk	0.039	3	0.030
White meat	0.041	4	0.040
Proteins	0.042	5	0.050
Nuts	0.060	6	0.060
Cereals and pasta	0.074	7	0.070
White fish	0.205	8	0.080

		Q=0.10				Q=0.20				Q=0.05
p-values sorted	rank	(i/m)Q		p-values sorted	rank	(i/m)Q		p-values sorted	rank	(i/m)Q
0.0002	1	0.005		0.0002	1	0.010		0.0002	1	0.003
0.0004	2	0.010		0.0004	2	0.020		0.0004	2	0.005
0.0016	3	0.015		0.0016	3	0.030		0.0016	3	0.008
0.0057	4	0.020		0.0057	4	0.040		0.0057	4	0.010
0.0091	5	0.025		0.0091	5	0.050		0.0091	5	0.013
0.0187	6	0.030		0.0187	6	0.060		0.0187	6	0.015
0.0225	7	0.035		0.0225	7	0.070		0.0225	7	0.018
0.0364	8	0.040		0.0364	8	0.080		0.0364	8	0.020
0.0441	9	0.045		0.0441	9	0.090		0.0441	9	0.023
0.0473	10	0.050		0.0473	10	0.100		0.0473	10	0.025
0.0536	11	0.055		0.0536	11	0.110		0.0536	11	0.028
0.0779	12	0.060		0.0779	12	0.120		0.0779	12	0.030
0.0862	13	0.065		0.0862	13	0.130		0.0862	13	0.033
0.1081	14	0.070		0.1081	14	0.140		0.1081	14	0.035
0.2341	15	0.075		0.2341	15	0.150		0.2341	15	0.038
0.357	16	0.080		0.357	16	0.160		0.357	16	0.040
0.4682	17	0.085		0.4682	17	0.170		0.4682	17	0.043
0.642	18	0.090		0.642	18	0.180		0.642	18	0.045
0.6833	19	0.095		0.6833	19	0.190		0.6833	19	0.048
0.8248	20	0.100		0.8248	20	0.200		0.8248	20	0.050

Overview/Objectives

- How can we describe the association between categorical variables?
  - What is the direction of the effect? How big is the effect?
    - Learn to calculate and interpret a odds ratio (OR), relative risk (RR)
- Diagnostic tests
  - How well does a test diagnose disease?
    - Learn to calculate and interpret sensitivity, specificity
  - How can we develop good diagnostic tests?
    - Learn to calculate and interpret a ROC analysis - partial

We will use the fabulous 2x2 table to help answer these questions

Ways to Measure Effect Size (measure of effect)

Means or differences in means
Medians or differences in medians
Proportions, counts

Correlation coefficients
Differences in slopes

Odds ratio (OR)
Relative risk (RR)
Sensitivity/specificity/NPV/PPV
Area under the curve

But first, some epidemiology speak for RR, OR

Outcome

Disease Y/N, Dead Y/N

Exposure

Test Pos/Neg, Male/Female

		Outcome Yes	Outcome No
Exposure	Yes	a	b
	No	c	d

Main Study Types that use RR or OR

Prospective (cohort) study

Define by exposure follow to outcome

Case-Control study

Defined by outcome assess exposure

Type of study determines if you calculate a OR (Case-control) or RR (prospective)

Risk vs. Odds

Both are ways of representing probability

The risk of an outcome (event) is the number of times it occurs divided by the number of times it could potentially ("at-risk") occur

Ranges from 0 to 1 (0% to 100%)

Odds are the number of people who developed the outcome divided by the number of people who did not

Ranges from 0 to ∞

Risk vs. Odds – cont.

1 out 8 women will develop breast cancer in their lifetime

risk of developing breast cancer = 1/8 = 0.125
odds of developing breast cancer = 1/7 = 0.143

1 out 3 people will develop any cancer in their lifetime

risk of developing any cancer = 1/3 = 0.333
odds of developing any cancer = 1/2 = 0.500

Odds approximates risk when the prevalence of disease is less common

Relative Risk vs. Odds Ratio

Use in terms of exposure and outcome

A relative risk is the ratio of two risks
From a prospective study where people are categorized by exposure

$$\frac{\text{Risk of outcome (yes) in people with exposure (yes)}}{\text{Risk of outcome (yes) in people without exposure (no)}}$$

An odds ratio is the ratio of two odds
From a case-control study where people are categorized by outcome

$$\frac{\text{Odds of exposure (yes) in people with outcome (yes)}}{\text{Odds of exposure (yes) in people without outcome (no)}}$$

2x2 Table Example: Prospective study and the RR

Below are data on the relationship between lung cancer and heavy drinking (>2 drinks per day)

Subjects without lung cancer were asked about their drinking habits and then followed for 10 years to determine incidence of lung cancer

Incidence means the rate of new cases (*e.g.*, # per 100,000 per year)

This is a prospective study. Heavy drinkers and light or non-drinkers without cancer were recruited and followed over time to see who developed cancer

Measure of effect: relative risk (RR)

	Lung Cancer +	Lung Cancer -	TOTALS
Heavy Drinkers	33	1667	1700
Others	27	2273	2300
TOTALS	60	3940	4000

2x2 Table Example: Prospective study and the RR

In a previous lecture you learned about how you could test for a difference in proportions using the chi-square test

Here the chi-square statistic = 3.89 and p=0.048

We conclude there is a difference in proportion of lung cancer between heavy drinkers and light or non-drinkers

There is an "association between drinking and lung cancer"

But how do we describe the magnitude or strength of this association?

	Lung Cancer +	Lung Cancer -	TOTALS
Heavy Drinkers	33	1667	1700
Others	27	2273	2300
TOTALS	60	3940	4000

2x2 Table Example: Prospective study and the RR

The risk of cancer in heavy drinkers= a/n_1
(AKA the probability of cancer among heavy drinkers)

The risk of cancer in light and non-drinkers = c/n_2
(AKA the probability of cancer among light and non-drinkers)

	Lung Cancer +	Lung Cancer -	TOTALS
Heavy Drinkers	a=33	b=1667	n_1=1700
Others	c=27	d=2273	n_2=2300
TOTALS	m_1=60	m_2=3940	N=4000

2x2 Table Example: Prospective study and the RR

The risk of cancer in heavy drinkers = a/n_1 = 33/1700 = 0.0194 = 1.94%
1.94% of heavy drinkers developed lung cancer during the 10 year study

The risk of cancer in light and non-drinkers = c/n_2 = 27/2300 = 0.0117 = 1.17%
1.17% of light or non-drinkers developed lung cancer during the 10 year study

	Lung Cancer +	Lung Cancer -	TOTALS
Heavy Drinkers	a=33	b=1667	n_1=1700
Others	c=27	d=2273	n_2=2300
TOTALS	m_1=60	m_2=3940	N=4000

The Relative Risk (AKA Risk Ratio, RR)

$$RR = \frac{\text{Probability of outcome in the exposed}}{\text{Probability of outcome in the unexposed}}$$

A relative risk of 1.0 corresponds to no difference between the exposure groups (exposure is not associated with outcome)

RR > 1.0 indicates higher probability of outcome in the exposed
Increased risk of outcome with exposure
RR < 1.0 indicates lower probability of outcome in the exposed
Decreased risk of outcome with exposure (exposure "protective")

$H_0 : RR_{population} = 1.0$
$H_A : RR_{population} \neq 1.0$

	Lung Cancer +	Lung Cancer -	TOTALS
Heavy Drinkers	33	1667	1700
Others	27	2273	2300
TOTALS	60	3940	4000

RR in heavy drinkers = 0.0194/0.0117 = 1.66
The probability of developing lung cancer is 1.66 times greater (66% higher) for heavy drinkers compared to those who do not drink heavily.

RR in others = 0.0117/0.0117 = 1.00
RR needs a “referent” or “reference” group
RR in referent groups is always = 1.0

95% CI for a RR

- The standard error for the natural log of the RR is:

$$SE[\ln(RR)] = \sqrt{\left(\frac{b}{an_1} + \frac{d}{cn_2}\right)}$$

- We can use this to get a 95% confidence interval for the RR, although it takes a couple steps

  - First find the CI for ln(RR):

$$\begin{aligned} \text{CI for } \ln(RR) &= \ln(RR) \pm 1.96 SE[\ln(RR)] \\ &= (LL, UL) \end{aligned}$$

  - Then back-transform (ie exponentiate) the limits of the CI to get back to the regular, RR scale

$$\text{CI for } RR = (e^{LL}, e^{UL})$$

Because a relative risk of 1.0 corresponds to no difference between the exposure groups, if the 95%CI spans (includes) 1.0, then p>0.05

BRITISH MEDICAL JOURNAL

LONDON SATURDAY JUNE 26 1954

THE MORTALITY OF DOCTORS IN RELATION TO THEIR SMOKING HABITS

A PRELIMINARY REPORT

BY

RICHARD DOLL, M.D., M.R.C.P.
Member of the Statistical Research Unit of the Medical Research Council

AND

A. BRADFORD HILL, C.B.E., F.R.S.
Professor of Medical Statistics, London School of Hygiene and Tropical Medicine; Honorary Director of the Statistical Research Unit of the Medical Research Council

Sent questionnaire to licensed physicians in the UK in 1951 asking about smoking habits. After 2.5 years of follow-up, they had the following data for their preliminary report.

	Lung cancer	No lung cancer
>1 cigarette/day	27	12841
1 or less cigarette/day	11	11512

Percentage of row total	Lung Cancer	No lung cancer
>1 cigarette/day	0.21%	99.79%
1 cigarette/day	0.10%	99.90%

P value and statistical significance	
Test	Fisher's exact test
P value	0.0331
P value summary	*
One- or two-sided	Two-sided
Statistically significant (P < 0.05)?	Yes

Effect size	Value	95% CI
Relative Risk	2.198	1.106 to 4.369
Reciprocal of relative risk	0.4550	0.2289 to 0.9043

Compared to people who smoke 1 cigarette or less each day, people who smoke >1 cigarette each day are at 2.2 times more likely to develop lung cancer in a 2.5 year period.

The 95%CI do not include 1.0 so $p<0.05$.

2x2 Table Example: Case-Control Study and the OR

Suppose we decided to study drinking and lung cancer using a case control study, we enroll:

100 subjects with lung cancer and
100 subjects without lung cancer

Then ask them about their drinking habits over the previous 10 years

Measure of effect: odds ratio (OR)

	Lung Cancer +	Lung Cancer -	TOTALS
Heavy Drinkers	55	42	97
Others	45	58	103
TOTALS	100	100	200

	Lung Cancer +	Lung Cancer -	TOTALS
Heavy Drinkers	a 55	b 42	97
Others	c 45	d 58	103
TOTALS	100	100	200

The odds of heavy drinking in lung cancer patients = a/c
= 55/45 = 1.22

The odds of heavy drinking in patients without lung cancer = b/d
= 42/58 = 0.72

$$OR = \frac{(a/c)}{(b/d)} = \frac{ad}{bc}$$

$H_0 : OR_{population} = 1.0$

$H_A : OR_{population} \neq 1.0$

The Odds Ratio (OR)

$$OR = \frac{\text{Odds of exposure in those with outcome}}{\text{Odds of exposure in those without outcome}}$$

An odds ratio of 1.0 corresponds to no difference between the outcome groups (outcome is not associated with exposure)

- OR > 1.0 indicates higher odds of exposure in outcome –yes group
  - Higher odds of exposure with outcome
- OR < 1.0 indicates lower odds of exposure in outcome –yes group
  - Lower odds of exposure with outcome (exposure "protective")

OR = 1.22/0.72 = 1.69

People with lung cancer are 1.69 times (69%) more likely to have been heavy drinkers compared to those without lung cancer (OR=1.0)

BRITISH MEDICAL JOURNAL

LONDON SATURDAY SEPTEMBER 30 1950

SMOKING AND CARCINOMA OF THE LUNG

PRELIMINARY REPORT

BY

RICHARD DOLL, M.D., M.R.C.P.

Member of the Statistical Research Unit of the Medical Research Council

AND

A. BRADFORD HILL, Ph.D., D.Sc.

Professor of Medical Statistics, London School of Hygiene and Tropical Medicine; Honorary Director of the Statistical Research Unit of the Medical Research Council

In England and Wales the phenomenal increase in the number of deaths attributed to cancer of the lung provides one of the most striking changes in the pattern of mortality recorded by the Registrar-General. For example, in the quarter of a century between 1922 and 1947 the annual number of deaths recorded increased from 612 to 9,287, or roughly fifteenfold. This remarkable increase is, of course, out of all proportion to the increase of population—both in total and, particularly, in its older age groups. Stocks (1947), using standardized death rates to allow for

whole explanation, although no one would deny that it may well have been contributory. As a corollary, it is right and proper to seek for other causes.

Possible Causes of the Increase

Two main causes have from time to time been put forward: (1) a general atmospheric pollution from the exhaust fumes of cars, from the surface dust of tarred roads, and from gas-works, industrial plants, and coal fires; and (2) the smoking of tobacco. Some characteristics of the

What type of study is this?

TABLE IV.—*Proportion of Smokers and Non-smokers in Lung-carcinoma Patients and in Control Patients with Diseases Other Than Cancer*

Disease Group	No. of Non-smokers	No. of Smokers	Probability Test
Males:			
Lung-carcinoma patients (649)	2 (0·3%)	647	
Control patients with diseases other than cancer (649) ..	27 (4·2%)	622	
Females:			
Lung-carcinoma patients (60)	19 (31·7%)	41	
Control patients with diseases other than cancer (60) ..	32 (53·3%)	28	

Doll & Hill Study on Smoking & Lung Cancer: Odds Ratio Calculation

	Lung cancer	Controls
Smokers	647	622
Non-Smokers	2	27

OR = (647 x 27) / (622 x 2) = 17469/1244 = 14.0

Odds of smoking
Lung Ca = 647/2 = 323.5
No Ca = 622/27 = 23.0

OR = 323.5/23.0 = 14.0

OR=14.0 means people with lung cancer are 14 times more likely to smoke compared to people without lung cancer; positive association between smoking and lung cancer

P value and statistical significance		
Test	Fisher's exact test	
P value	<0.0001	
P value summary	****	
One- or two-sided	Two-sided	
Statistically significant (P < 0.05)?	Yes	
Effect size	**Value**	**95% CI**
Odds ratio	14.04	3.621 to 60.16
Reciprocal of odds ratio	0.07121	0.01662 to 0.2762

Why not do a RR for a Case-Control Study

We created an artificial population with a 50% prevalence* of lung cancer

100/200 = 50%

We can’t use relative risk

The relative risk depends on the prevalence of lung cancer in the population

We did not collect data that would allow us to estimate the prevalence in the general population

With a case-control study we are not measuring risk

*Prevalence: all cases of disease = incident (new) + existing

Effect of Prevalence

Prospective study

	Lung Cancer +	Lung Cancer -	TOTALS
Heavy Drinkers	33	1667	1700
Others	27	2273	2300
TOTALS	60	3940	4000

Prevalence =
60/4000 = 0.015
15 per 1000

Case-Control study

	Lung Cancer +	Lung Cancer -	TOTALS
Heavy Drinkers	55	42	97
Others	45	58	103
TOTALS	100	100	200

Prevalence =
100/200 = 0.500
5 per 10 or 50% which we set by choosing 100 lung cancers and 100 non-cancer

	Lung cancer Prevalence (per 1,000)	RR	OR
Prospective study data	15	1.65	1.67
Case-Control study data	500	1.30	1.69

For more common diseases the OR will be higher than the RR

Prospective study: prevalence of lung cancer is 0.16% (1.6 per 1,000)

Effect size	**Value**	**95% CI**
Relative Risk	2.198	1.106 to 4.369
Reciprocal of relative risk	0.4550	0.2289 to 0.9043
Odds ratio	2.201	1.108 to 4.421
Reciprocal of odds ratio	0.4544	0.2262 to 0.9025

Case-control study: prevalence of lung cancer is 50% (50 per 100)

Effect size	**Value**	**95% CI**
Relative Risk	7.393	2.318 to 26.69
Reciprocal of relative risk	0.1353	0.03747 to 0.4314
Odds ratio	14.04	3.621 to 60.16
Reciprocal of odds ratio	0.07121	0.01662 to 0.2762

Getting RR and OR Measures of Association in Prism

While it does not matter how you enter your data for a chi-square or Fisher's exact tests, it does matter for the calculation of the measures of association

Make sure your columns are for the outcome status and rows are for the exposure

	Table format: Contingency	Outcome A Lung Cancer +	Outcome B Lung Cancer -
		Y	Y
1	Heavy Drinker	33	1667
2	Not Heavy Drinker	27	2273
3	Title		

	Table format: Contingency	Outcome A Heavy Drinker	Outcome B Non Drinker
		Y	Y
1	Lung Cancer +	33	27
2	Lung Cancer -	1667	2273
3	Title		

Use the Contingency Table

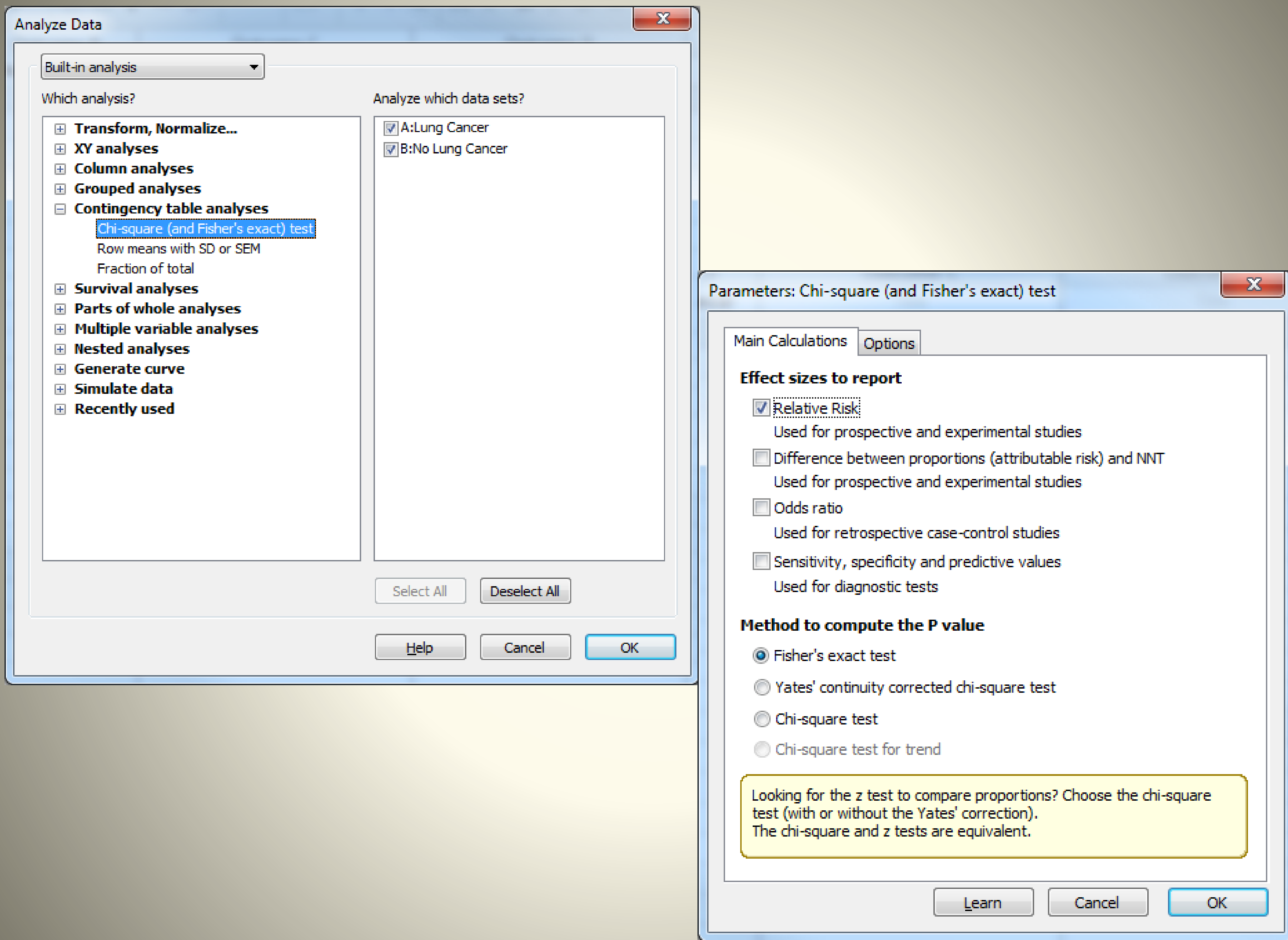
Prospective study

Table format: Contingency	Outcome A	Outcome B	C
	Lung Cancer	No Lung Cancer	
	Y	Y	
1 Heavy Drinkers	33	1667	
2 Others	27	2273	
3 Title			

Case-Control study

Table format: Contingency	Outcome A	Outcome B	Ou
	Lung Cancer	No Lung Cancer	
	Y	Y	
1 Heavy Drinkers	55	42	
2 Others	45	58	
3 Title			

Go to Analyses> Contingency table analyses>Chi-square (and Fisher's Exact) test


Analyze Data
Built-in analysis
Which analysis?
Transform, Normalize...
XY analyses
Column analyses
Grouped analyses
Contingency table analyses
Chi-square (and Fisher's exact) test
Row means with SD or SEM
Fraction of total
Survival analyses
Parts of whole analyses
Multiple variable analyses
Nested analyses
Generate curve
Simulate data
Recently used
Analyze which data sets?
A:Lung Cancer
B:No Lung Cancer
Select All
Deselect All
Help
Cancel
OK
Parameters: Chi-square (and Fisher's exact) test
Main Calculations
Options
Effect sizes to report
Relative Risk
Used for prospective and experimental studies
Difference between proportions (attributable risk) and NNT
Used for prospective and experimental studies
Odds ratio
Used for retrospective case-control studies
Sensitivity, specificity and predictive values
Used for diagnostic tests
Method to compute the P value
Fisher's exact test
Yates' continuity corrected chi-square test
Chi-square test
Chi-square test for trend
Looking for the z test to compare proportions? Choose the chi-square test (with or without the Yates' correction).
The chi-square and z tests are equivalent.
Learn
Cancel
OK

Contingency	A	B	C
Table Analyzed	Prospective		
P value and statistical significance			
Test	Fisher's exact test		
P value	0.0646		
P value summary	ns		
One- or two-sided	Two-sided		
Statistically significant (P < 0.05)?	No		
Effect size	**Value**	**95% CI**	
Relative Risk	1.654	1.003 to 2.725	
Reciprocal of relative risk	0.6047	0.3670 to 0.9966	
Methods used to compute CIs			
Relative Risk	Koopman asymptotic score		
Data analyzed	**Lung Cancer**	**No Lung Cancer**	**Total**
Heavy Drinkers	33	1667	1700
Others	27	2273	2300
Total	60	3940	4000
Percentage of row total	**Lung Cancer**	**No Lung Cancer**	
Heavy Drinkers	1.94%	98.06%	
Others	1.17%	98.83%	
Percentage of column total	**Lung Cancer**	**No Lung Cancer**	
Heavy Drinkers	55.00%	42.31%	
Others	45.00%	57.69%	

Prospective study

Prevalence = 1.5%

10	Effect size	Value	95% CI
11	Relative Risk	1.654	1.003 to 2.725
12	Reciprocal of relative risk	0.6047	0.3670 to 0.9966
13			
14	Odds ratio	1.667	0.9985 to 2.754
15	Reciprocal of odds ratio	0.6000	0.3631 to 1.001

Case-Control study

Prevalence = 50%

10	Effect size	Value	95% CI
11	Relative Risk	1.298	0.9831 to 1.725
12	Reciprocal of relative risk	0.7705	0.5796 to 1.017
13			
14	Odds ratio	1.688	0.9609 to 2.889
15	Reciprocal of odds ratio	0.5925	0.3461 to 1.041

What we calculated by hand

	RR	OR
Prospective study data	1.65	1.67
Case-Control study data	1.30	1.69

Data format changes RR but not OR

Table format: Contingency		Outcome A	Outcome B
		Lung Cancer	No cancer
		Y	Y
1	Heavy Drinkers	55	42
2	Others	45	58

Relative Risk	1.298
Reciprocal of relative risk	0.7705
Odds ratio	1.688
Reciprocal of odds ratio	0.5925

Table format: Contingency		Outcome A	Outcome B
		Heavy Drinkers	Others
		Y	Y
1	Lung Cancer	55	45
2	No cancer	42	58

Relative Risk	1.31
Reciprocal of relative risk	0.7636
Odds ratio	1.688
Reciprocal of odds ratio	0.5925

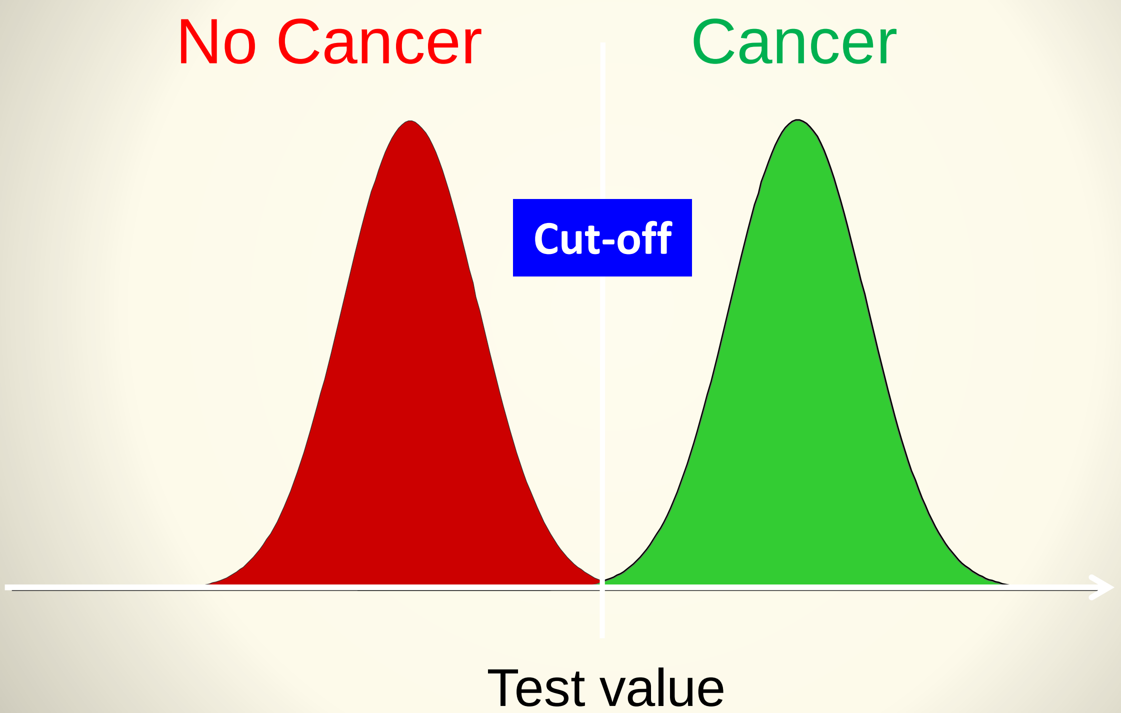
Diagnostic Tests

Common to use 2x2 tables to develop and evaluate how well a diagnostic test works

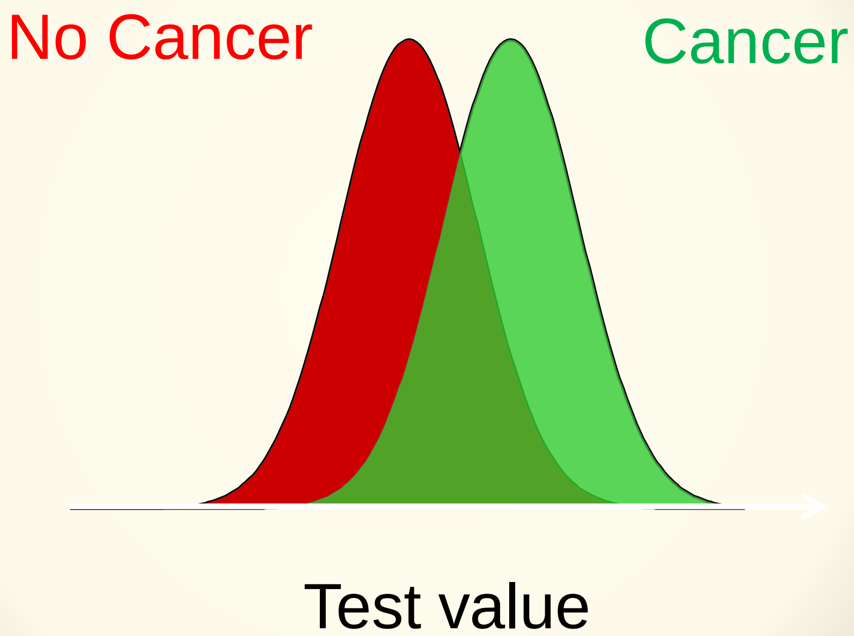
When developing a diagnostic test, case control study designs are often used to ensure that there is a sufficient number of people with disease included in the study

Purpose: See how well a test can classify subjects known to be disease positive and negative

Ideal Situation



The Reality



Diagnostic Tests: The goal is to correctly classify subjects

A good diagnostic test should be able to discriminate between those with and without disease

We can summarize this data in a 2 x 2 table

You can see this table has many similarities to hypothesis testing and the idea of type I and type II errors

	Disease +	Disease -
Test +	True Positive (Correct!)	False Positive (Mistake)
Test -	False Negative (Mistake)	True Negative (Correct!)

Diagnostic Tests: Types of errors

As in our hypothesis testing framework, we can make 2 kinds of errors:

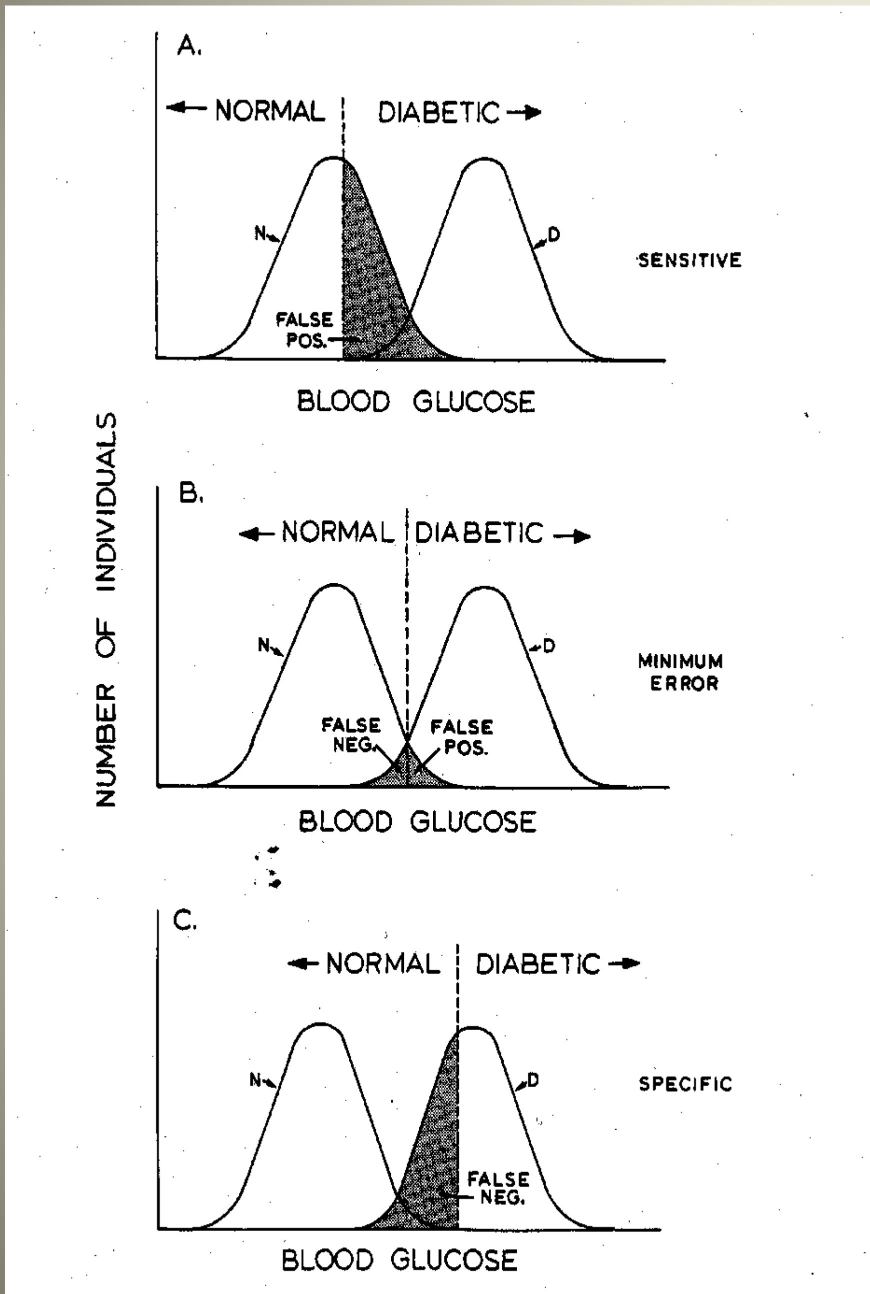
- False positive: Someone without disease has a positive test
  - Similar to a type I error in hypothesis testing
- False negative: Someone with disease has a negative test
  - Similar to a type II error in hypothesis testing

We want our test to have:

- High numbers of true positives and true negatives
- Low numbers of false positives and false negatives

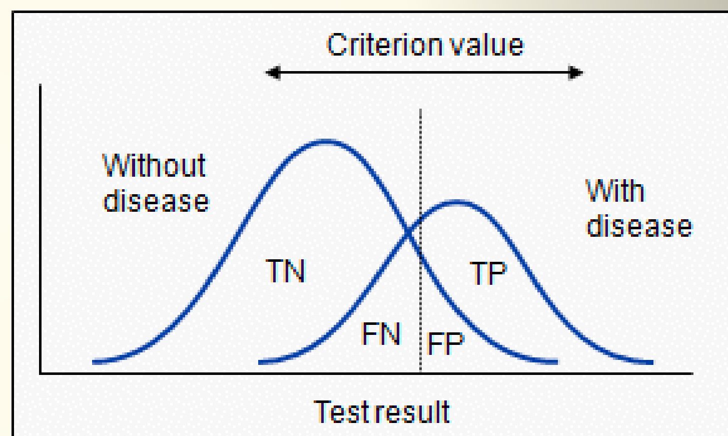
	Disease +	Disease -
Test +	True Positive	False Positive
Test -	False Negative	True Negative

Where Do We Set The Cut-off For A Screening Test?



The impact of false positives:

Anxiety, cost of further testing



Impact of false negatives:

With serious disease, do not want to miss a case

Describing the Performance of Diagnostic Tests: Sensitivity

The probability of a positive test result given a subject is disease positive

P(T+|D+)

1- sensitivity is sometimes called the “false negative rate”

	Disease +	Disease -
Test +	TP	FP
Test -	FN	TN

Sensitivity = TP/(TP+FN)
False negative rate= FN/(TP+FN)

	Breast Cancer +	Breast Cancer -
Mammogram +	90	5
Mammogram -	10	95

Sensitivity = 90/(90+10) = 90%

If a woman has breast cancer, she has a 90% chance of having a positive mammogram and a 10% chance of having a negative mammogram.

Describing the Performance of Diagnostic Tests: Specificity

The probability of a negative test result given the subject is disease negative

P(T-|D-)

1-specificity is called the "false positive rate"

	Disease +	Disease -
Test +	TP	FP
Test -	FN	TN

Specificity = TN/(TN+FP)
False positive rate = FP/(TN+FP)

	Breast Cancer +	Breast Cancer -
Mammogram +	90	5
Mammogram -	10	95

Specificity = 95/(95+5) = 95%

If a woman does not have breast cancer, she has a 95% chance of having a negative mammogram and a 5% chance of having a positive mammogram.

Sensitivity, Specificity, and Predictive Values

Sensitivity and specificity are characteristics of the diagnostic test

Tell us how well the test works

Aid in comparing tests and developing a good tests

When using the test in the population at large, we are usually more interested in predictive value

Sensitivity answers the question, “If I have the disease, what are the chances my test will be positive?”

$$P(T+|D+)$$

Generally people taking the test want to know “What is the probability that I have the disease if I have a positive test?”

$$P(D+|T+)$$

Positive predictive value (PPV)

The probability of being disease positive given a positive test result

$PPV = P(D+|T+) = TP/(TP+FP)$

Don't confuse this with sensitivity, $P(T+|D+) = TP/(TP+FN)$

PPV depends on the prevalence of the disease in the population

If the prevalence is higher, PPV will be higher as well

Breast cancer in a representative sample of 100,000 women

	Breast Cancer +	Breast Cancer -
Mammogram +	90	4995
Mammogram -	10	94,905

PPV = TP/(TP+FP) = 90/(90+4995) = 1.8%

Sensitivity = TP/(TP+FN) = 90/(90+10) = 90%

If a woman has a positive mammogram, there is only a 1.8% chance she actually has breast cancer.

Since the prevalence of breast cancer is low, the PPV is low even though sensitivity is 90%.

Negative predictive value (NPV)

The probability of being disease negative given a negative test result

$P(D-|T-) = TN/(TN+FN)$
Don't confuse this with specificity, $P(T-|D-) = TN/(TN+FP)$

The negative predictive value also depends on the prevalence of the disease in the population

If the prevalence is higher, NPV will be lower

Breast cancer in a Representative sample of 100,000 women

	Breast Cancer +	Breast Cancer -
Mammogram +	90	4995
Mammogram -	10	94,905

NPV = TN/(TN+FN) =94905/(10+94905) = 99.999%

Specificity = TN/(TN+FP) =94905/(94905+4995) = 95%

If a women has a negative mammogram, there is a 99.999% chance she truly does not have breast cancer.

Predictive Value in a Higher Risk Population

Breast cancer in a sample of 100,000 women with BRCA mutation

	Breast Cancer +	Breast Cancer -
Mammogram +	900	4950
Mammogram -	100	94,050

PPV = 900/(900+4950) = 15.4% (*before it was 1.8%)*
NPV = 94050/(100+94050) = 99.89% (*very similar to before, 99.999%*)
Sensitivity = 900/1000 = 90%
Specificity = 94050/(94050+4950) = 95%

Note the sensitivity and specificity of the test remain the same as before.

	Breast Cancer +	Breast Cancer -
Mammogram +	900	4950
Mammogram -	100	94,050

PPV= 15.4%, NPV= 99.89%, Sensitivity= 90%, Specificity = 95%

If a woman with the BRCA mutation has a positive mammogram, there is a 15.4% chance she has breast cancer.

If a woman with the BRCA mutation has a negative mammogram, there is a 99.89% chance she does not have breast cancer.

The likelihood ratio (LR)

The LR is the ratio of sensitivity/(1-specificity)

	Breast Cancer +	Breast Cancer -
Mammogram +	900	4950
Mammogram -	100	94,050

In the breast cancer examples: 0.9/(1-0.95)=18

A person with breast cancer is 18 times more likely to have a positive test than a person without breast cancer

Sample 1: Low Prevalence of Carotid Atherosclerotic Disease

	Disease (Carotid Atherosclerotic Disease)	No Disease (No Carotid Atherosclerotic Disease)	Total
Test positive (positive MR angiography)	20	10	30
Test negative (negative MR angiography)	4	120	124
Total	24	130	154

Sample 3: High Prevalence of Carotid Atherosclerotic Disease

	Disease (Carotid Atherosclerotic Disease)	No Disease (No Carotid Atherosclerotic Disease)	Total
Test positive (positive MR angiography)	500	10	510
Test negative (negative MR angiography)	100	120	220
Total	600	130	730

Sample 2: Intermediate Prevalence of Carotid Atherosclerotic Disease

	Disease (Carotid Atherosclerotic Disease)	No Disease (No Carotid Atherosclerotic Disease)	Total
Test positive (positive MR angiography)	100	10	110
Test negative (negative MR angiography)	20	120	140
Total	120	130	250

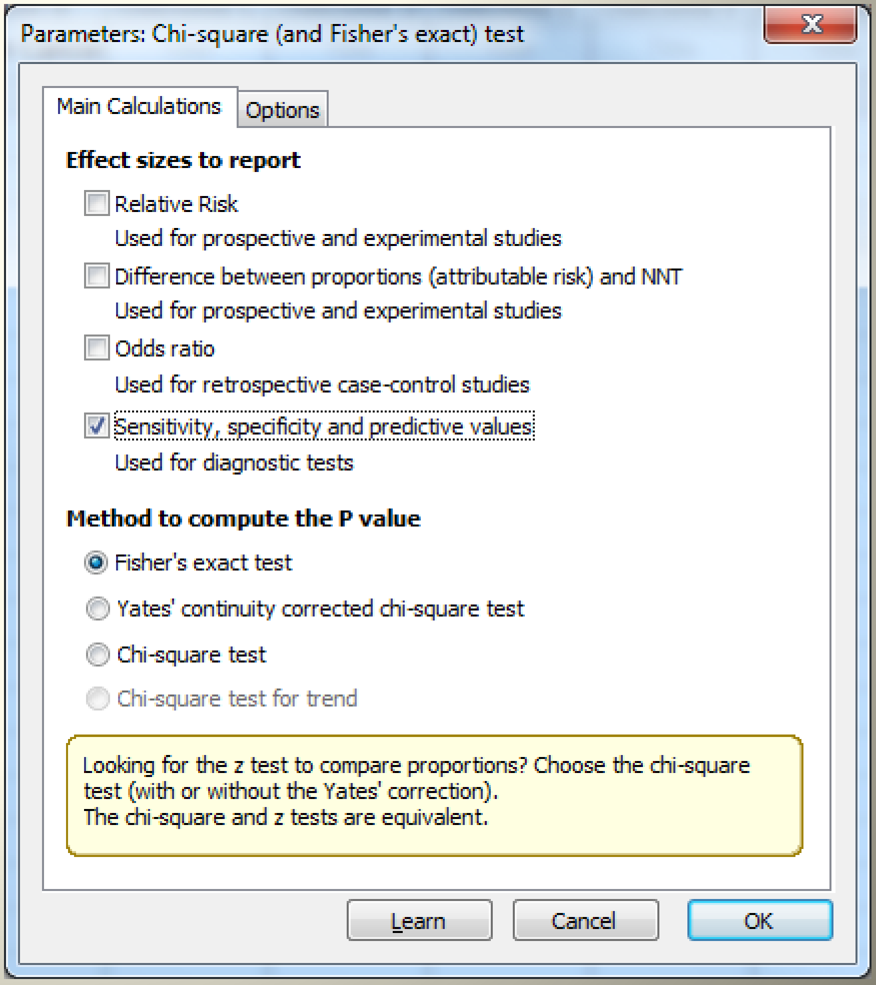
	Low	Mid	High
Prevalence	16%	48%	82%
Sens	83%	83%	83%
Spec	92%	92%	92%
PPV	67%	91%	98%
NPV	98%	86%	55%

AJNR Am J Neuroradiol 20:1584–1596, October 1999

Calculations in PRISM

Table format: Contingency	Outcome A	Outcome B
	Breast Cancer	No Breast Cancer
	Y	Y
1 Mammogram +	900	4950
2 Mammogram -	100	94050
3 Title		

Much like calculating RR and OR, go to Analyze> Contingency table Analyses> Chi-square (and Fisher's exact) test



Under main calculations, choose "sensitivity, specificity, and predictive values"

We get the same results as our calculations by hand

Effect size	Value	95% CI
Sensitivity	0.9000	0.8798 to 0.9171
Specificity	0.9500	0.9486 to 0.9513
Positive Predictive Value	0.1538	0.1448 to 0.1633
Negative Predictive Value	0.9989	0.9987 to 0.9991
Likelihood Ratio	18.00	

Introduction to ROC

Determining the best cut-off for a diagnostic test

ARTICLE IN PRESS

Jan 2019

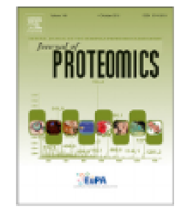
Journal of Proteomics xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Journal of Proteomics

journal homepage: www.elsevier.com/locate/jprot

ELSEVIER



Putative salivary biomarkers useful to differentiate patients with fibromyalgia

Federica Ciregia[a,b,*], Camillo Giacomelli[c], Laura Giusti[a], Claudia Boldrini[b], Isabella Piga[d], Pasquale Pepe[c], Arianna Consensi[c], Sara Gori[e], Antonio Lucacchini[f], Maria R. Mazzoni[b], Laura Bazzichi[c]

[a] Department of Clinical and Experimental Medicine, University of Pisa, Via Savi 10, 56126 Pisa, Italy
[b] Department of Pharmacy, University of Pisa, Via Bonanno 6, 56126 Pisa, Italy
[c] Rheumatology Operative Unit, Department of Clinical and Experimental Medicine, University of Pisa, Via Roma 67, 56126 Pisa, Italy
[d] School of Medicine and Surgery, Università degli Studi di Milano-Bicocca, Via Cadore 48, 20900 Monza, Italy
[e] Neurology Unit, Department of Clinical and Experimental Medicine, University of Pisa, Via Roma 67, 56126 Pisa, Italy
[f] Fondazione ItPA, Via dei Vestini 23, 66100 Chieti, Italy

Table 4
Comparison between sensitivities and specificities of single biomarkers and their combination. For each proposed biomarker, and for their combination, we indicated sensitivity and specificity corresponding to different cut-off in order to identify the better cut-off. s.i.: signal intensity.

TRFE AUC 0.699					
Cut-off s.i.	*1812*	*1889*	*2065*	*2238*	*2562*
Sensitivity %	71	68	68	65	55
Specificity %	57	60	63	63	73
ENOA AUC 0.738					
Cut-off s.i.	*0.95*	*1.01*	*1.05*	*1.25*	*1.36*
Sensitivity %	81	81	77	71	68
Specificity %	61	63	63	68	71
PGAM1 AUC 0.683					
Cut-off s.i.	*1.40*	*1.56*	*1.62*	*1.79*	*1.93*
Sensitivity %	48	45	45	42	39
Specificity %	80	87	90	90	90
Combining TRFE, ENOA, PGAM1 AUC 0.792					
Cut-off combined risk index	*0.36*	*0.41*	*0.43*	*0.44*	*0.52*
Sensitivity %	84	81	74	71	52
Specificity %	63	67	73	77	93

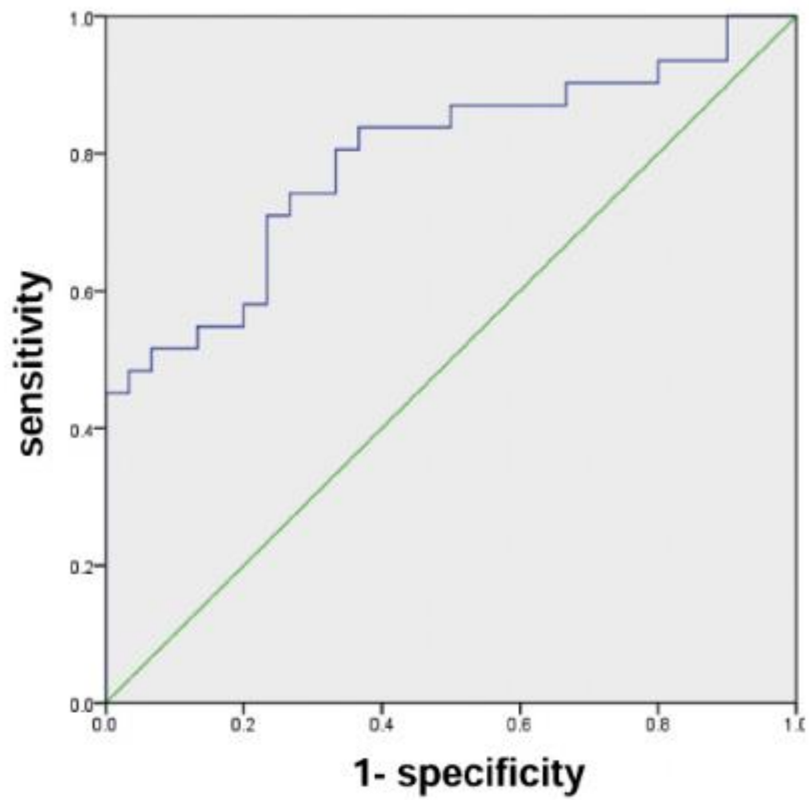
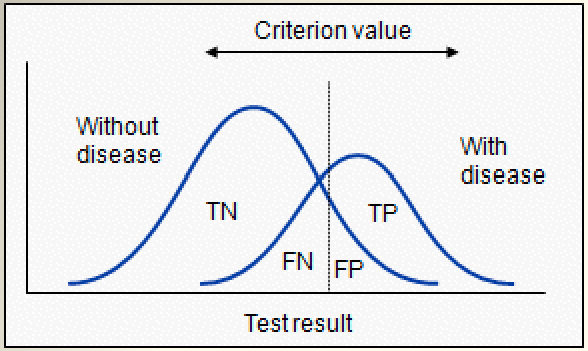


Fig. 5. ROC curve. Receiver operating characteristic curve (ROC) obtained from the combination of ENOA, TRFE, PGAM1, in WS. ROC curve was calculated to assess the clinical potential of these selected proteins to distinguish FM from control (healthy subjects plus migraine patients).

Changing Threshold of a Test Changes Sensitivity and Specificity



Numbers of false positives and false negatives also change

If the threshold is too high, then some individuals with low test results or mild forms of disease will be missed

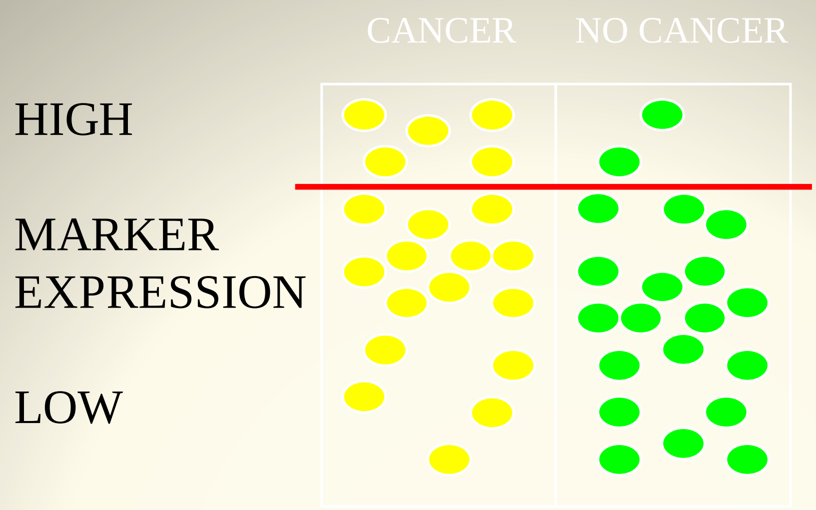
low sensitivity, high specificity

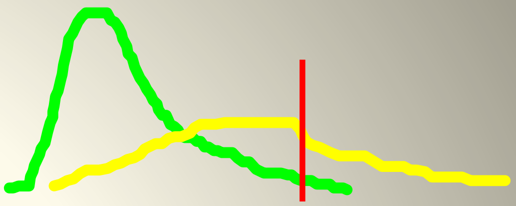
missing cases – false negatives

If the threshold is too low, most individuals with disease will be detected, but many people without the disease will also have positive tests.

high sensitivity, low specificity

more non cases picked up – false positives

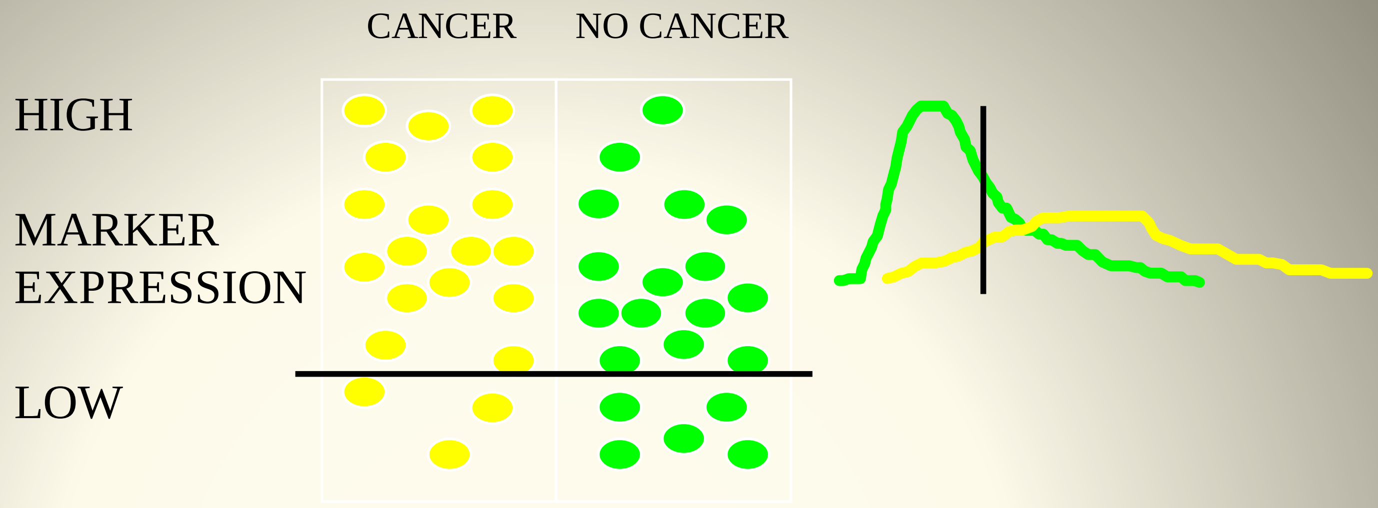




	CANCER	NO CANCER
HIGH	5	2
LOW	15	18
	20	20

Sensitivity = 5/20 = 25%

Specificity = 18/20 = 90%



	CANCER	NO CANCER
HIGH	17	15
LOW	3	5
	20	20

Sensitivity = 17/20 = 85%

Specificity = 5/20 = 25%

Receiver Operating Characteristic (ROC) Curve

Helps visualize the trade off between sensitivity and specificity when choosing a threshold for a diagnostic test

The ROC Curve is a plot of the sensitivity vs. (1-specificity) for each potential threshold (cut-off)

Why the funny name?

These plots were originally used in WWII to determine whether blips on a radar screen were ships, planes, or birds or fish




Receiver Operating Characteristic (ROC) Curve

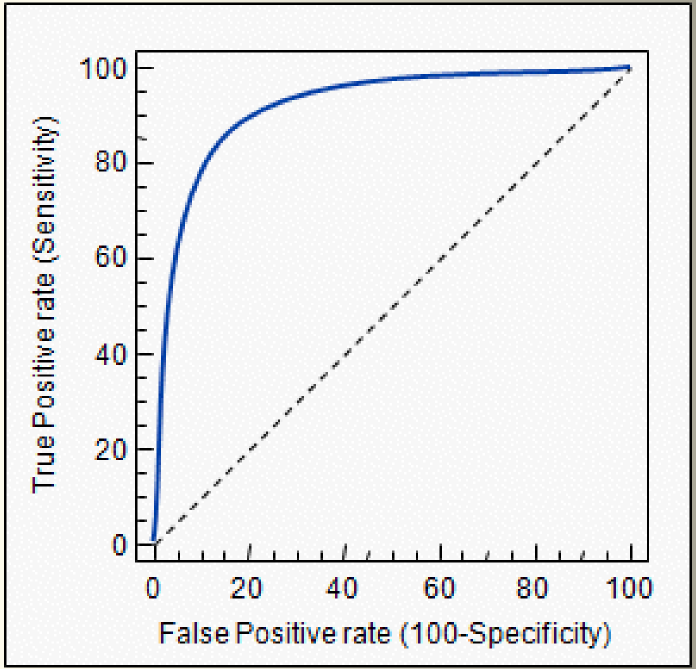
The bottom left of the curve is one silly extreme when the test will never return a diagnosis of disease

0% sensitivity, 100% specificity

At the top right, everyone is diagnosed with the disease

100% sensitivity, 0% specificity

The dotted line represents a test that has no ability to discriminate between those with and without disease

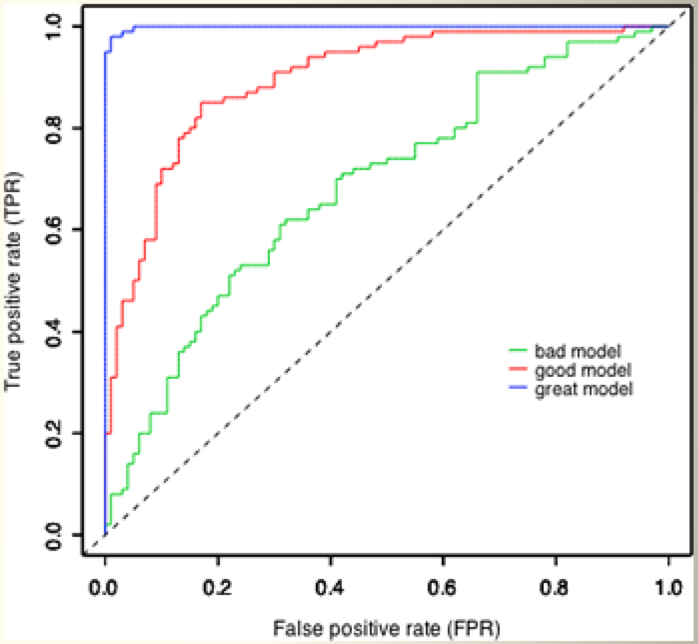


Receiver Operating Characteristic (ROC) Curve

ROC curves can be used to compare tests

Curves closer to the top left and further from the diagonal line indicate more accurate tests

This is often described numerically with the area under the curve (AUC).



ROC: Area Under the Curve

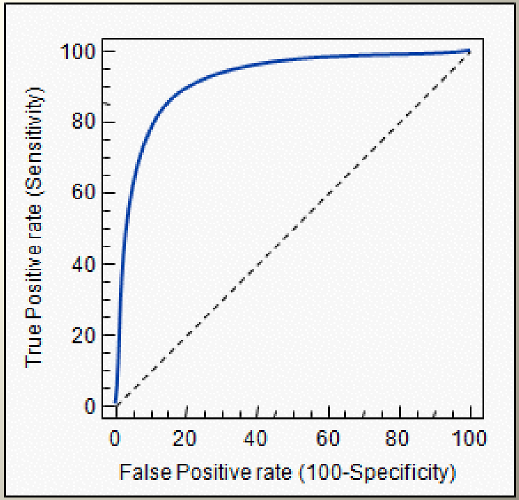
The area under the curve (AUC) is a reflection of how good the test is at distinguishing between patients with disease and those without disease.

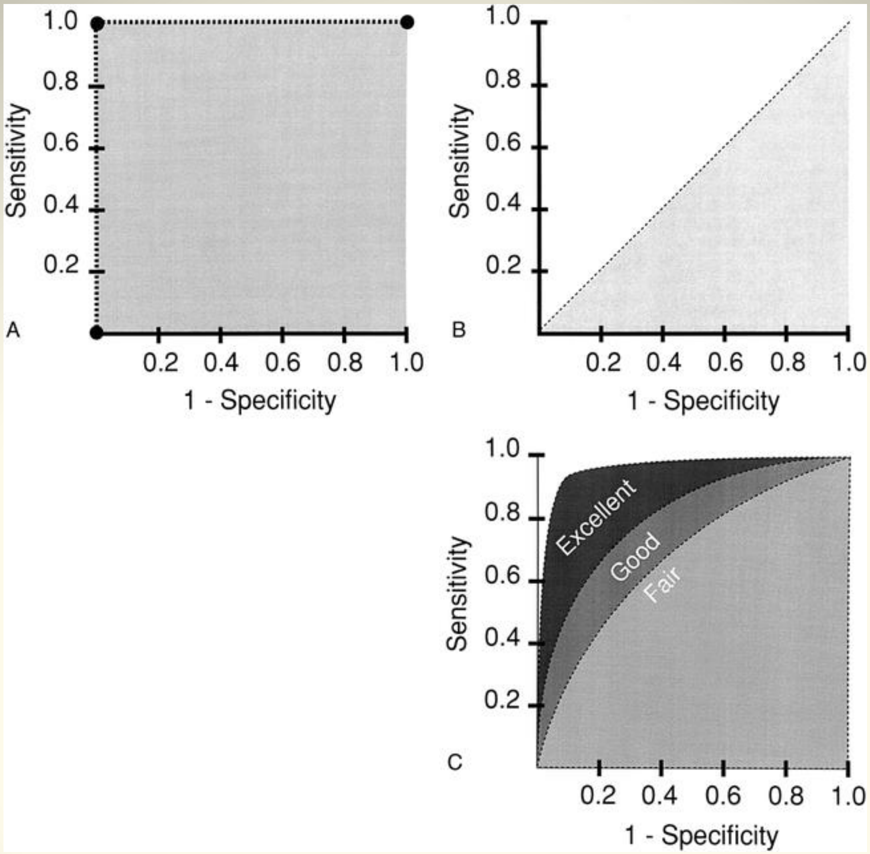
The AUC summarizes the discriminative ability of a test across the full range of cut-offs .

The greater the AUC, the better the test.

A perfect test will have an AUC of 1.0, while a completely useless test (one whose curve falls on the diagonal line) has an AUC of 0.5.

Excellent test	AUC 0.9-1.0
Good	AUC 0.8-0.9
Fair	AUC 0.7-0.8
Poor	AUC 0.6-0.7
Fail	AUC 0.5-0.6





The perfect test (A) has an area under the curve (AUC) of 1.
The useless test (B) has an AUC of 0.5. The typical test (*C*) has an AUC between 0.5 and 1.
The greater the AUC (i.e,, excellent > good > fair), the better the diagnostic performance

L. Santiago Medina AJNR Am J Neuroradiol 1999;20:1584-1596

The ROC curve has three purposes

Determine the cut-off point at which optimal sensitivity and specificity are achieved

Assess the diagnostic accuracy of a test
- Area under the curve (AUC)

Compare the usefulness of two or more diagnostic tests
- Different AUC

www.impactjournals.com/oncotarget/ Oncotarget, 2017, Vol. 8, (No. 29), pp: 47110-47120

Research Paper

Influence of serum cholesterol level and statin treatment on prostate cancer aggressiveness

Thomas J. Schnoeller[1], **Florian Jentzmik**[1,2], **Andres J. Schrader**[3] **and Julie Steinestel**[3]

[1]Department of Urology, Ulm University Medical Center, Ulm, Germany

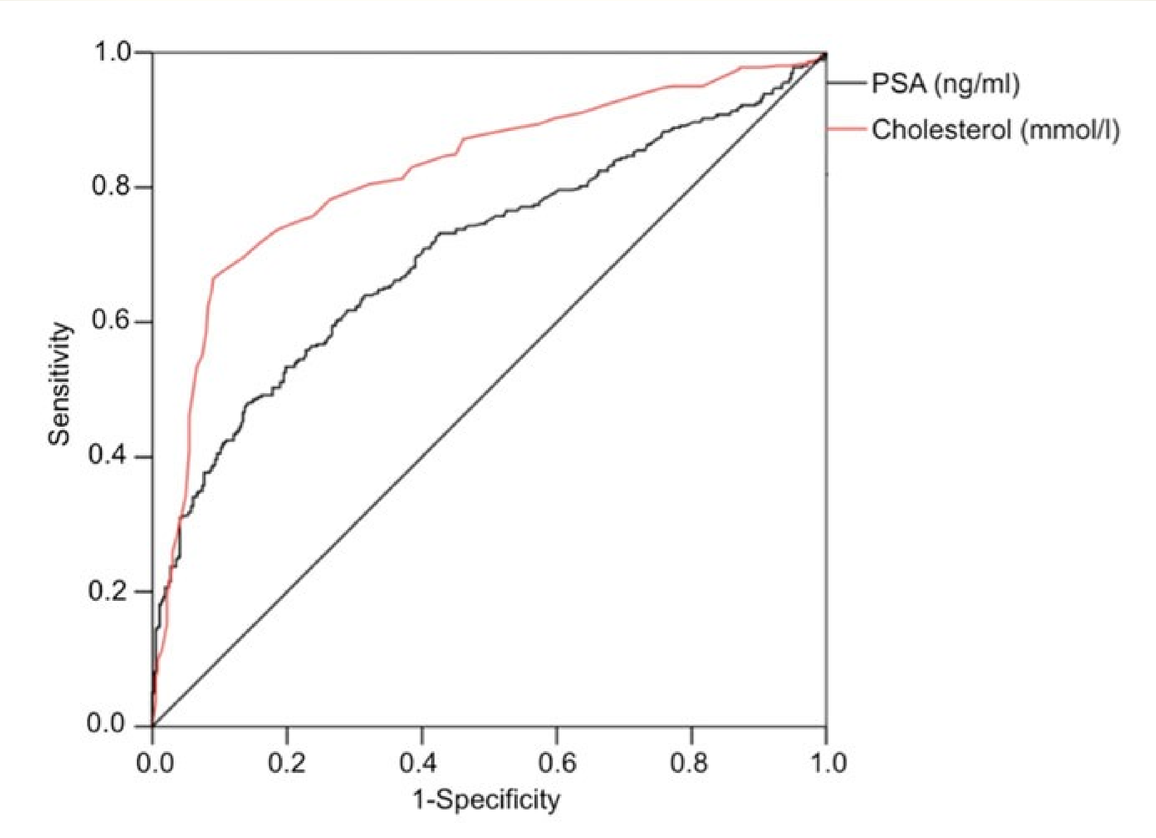
[2]Department of Urology, St. Elisabeth Hospital, Ravensburg, Germany

[3]Department of Urology, Muenster University Medical Center, Muenster, Germany

ABSTRACT

Both cholesterol levels and the use of statins have been described to influence the development and prognosis of prostate cancer (PC). In this retrospective, cross-sectional analysis of consecutive cases from a tertiary referral center we evaluated an association between hypercholesterolemia (≥5.0mmol/l), the use of statins, and advanced/aggressive PC in 767 men with histologically confirmed, clinically localized PC awaiting radical prostatectomy. We found that patients with HCE (n=287, 37.4%) had a significantly higher incidence of poorly differentiated PC (Gleason score ≥7b, 81.1% vs. 4.9%), advanced local tumor stage (≥pT3, 57.7% vs. 22.2%), and nodal involvement (19.8% vs. 1.6%). Multivariate logistic regression analysis identified hypercholesterolemia as a risk factor for aggressive and/or advanced PC (OR 2.01, p<0.001) whereas statin intake showed an odds ratio of 0.49 (p=0.005) indicating a negative association with high-risk PC. Despite a limited number of patients using statins (~9.5%), adjusted and weighed multivariate logistic regression models revealed that preoperative hypercholesterolemia is associated with a diagnosis of high-risk PC which is negatively influenced by statin intake.

Receiver operating characteristic (ROC) ROC analysis curve comparing preoperative total serum cholesterol values versus high-risk disease (pT≥3 and/or pN+ and/or Gleason score 7b) in patients with prostate cancer. AUC=0.82, 95% CI 0.79-0.85 (p<0.001). PSA served as positive control (AUC=0.71, 95% CI 0.67-0.74, p<0.001).

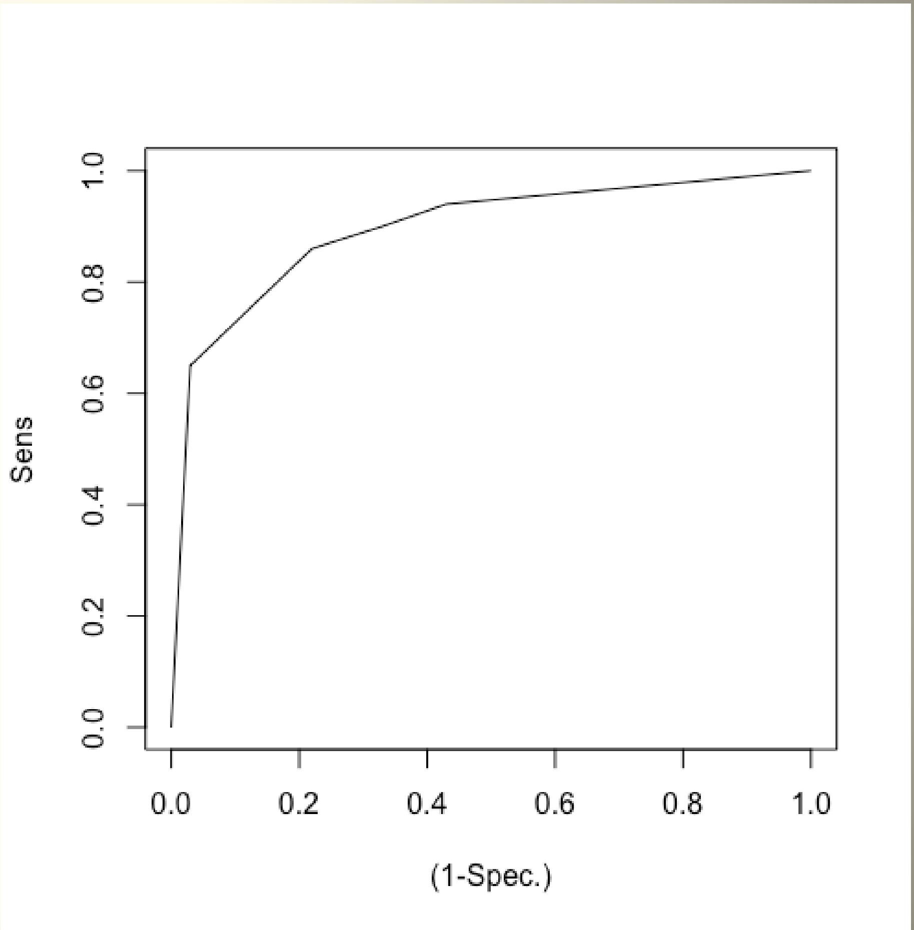


The ROC: Determining Test Threshold

Where we ultimately decide to put the threshold for a + test will depend on the costs of false positives vs. false negatives

If we think the costs are similar, then we would like to strike a balance between FP and FN

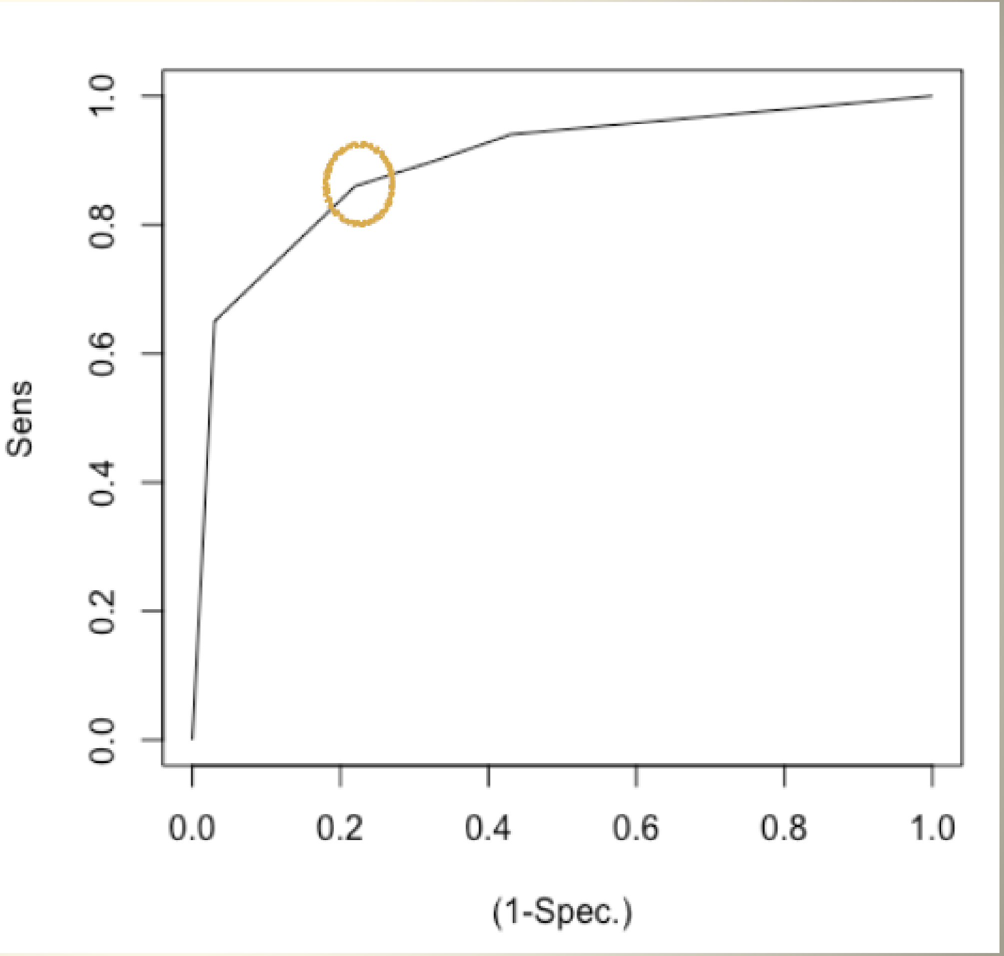
Choose the threshold closest to upper left corner of the plot



The ROC: Determining Test Threshold

This corresponds to giving positive test results to those with a score of 4 or higher

Sens:0.86
Spec: 0.78



Will show how to do an ROC in Prism next lecture