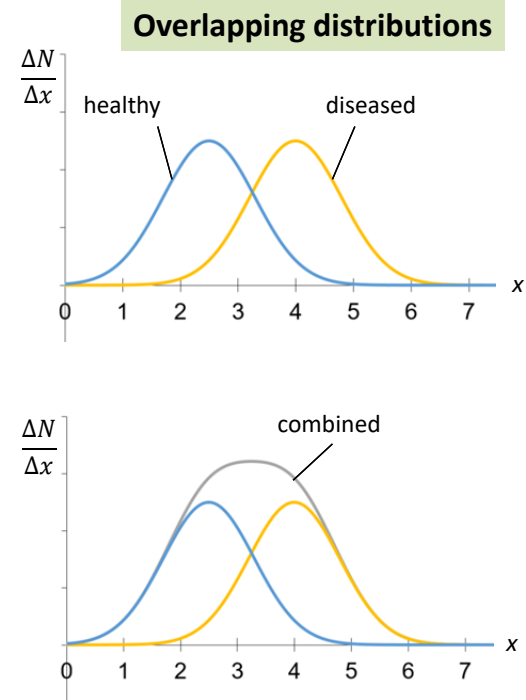


KAD 2021.11.12



everything has a **distribution** (eye color, height, cholesterol level,...)

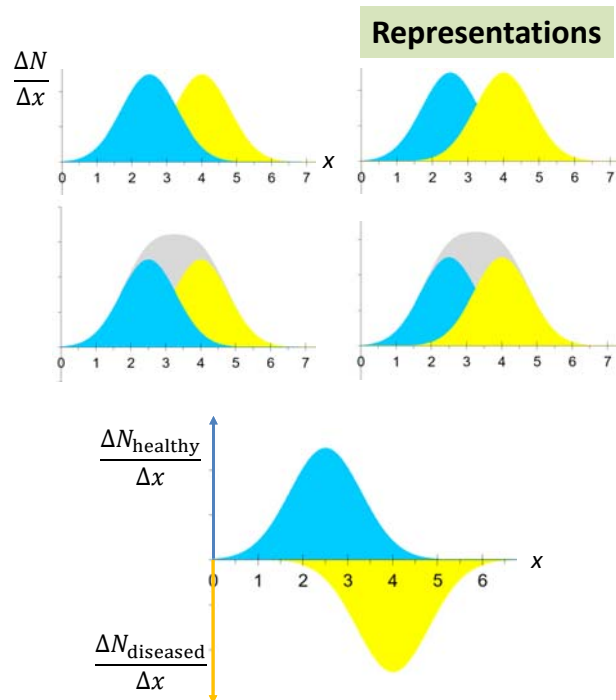
consider a continuous variable with a different distribution in the diseased population than in the healthy one

suppose that the measurable parameter is typically **larger** in the **diseased** population than in the healthy one (if smaller, the reciprocal of the original parameter can be used)

the figure shows two such density functions; the area under the curve corresponds to the number of individuals

in the present example the number of healthy and diseased is the same (the standard deviation of the parameter is the same)

2

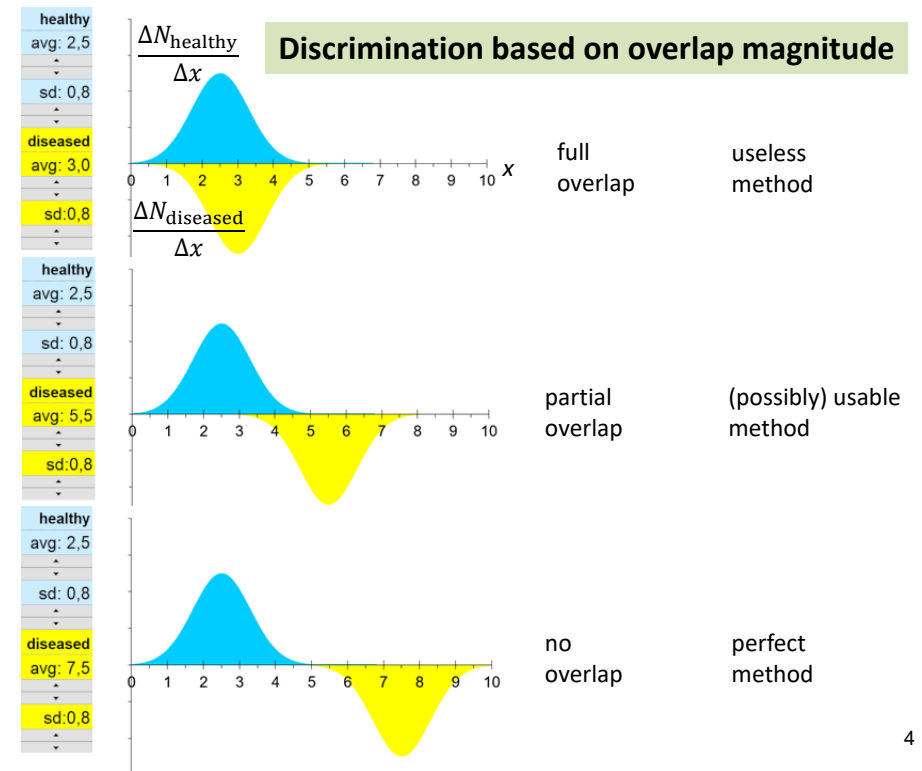


due to the great importance of the areas under the curve, we prefer an image that colors the areas instead of a line drawing

due to the overlaps, correct coloring is difficult (or impossible) in the usual representation

proposed new representation: instead of the negative axis, another positive axis, for the diseased

3



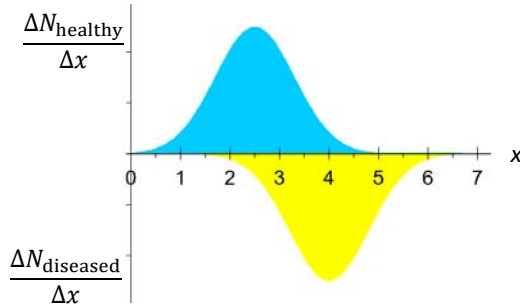
4

## Prevalence

frequency of diseased in examined population

measure of how common the disease is

= probability prior to test  
= a-priori-probability

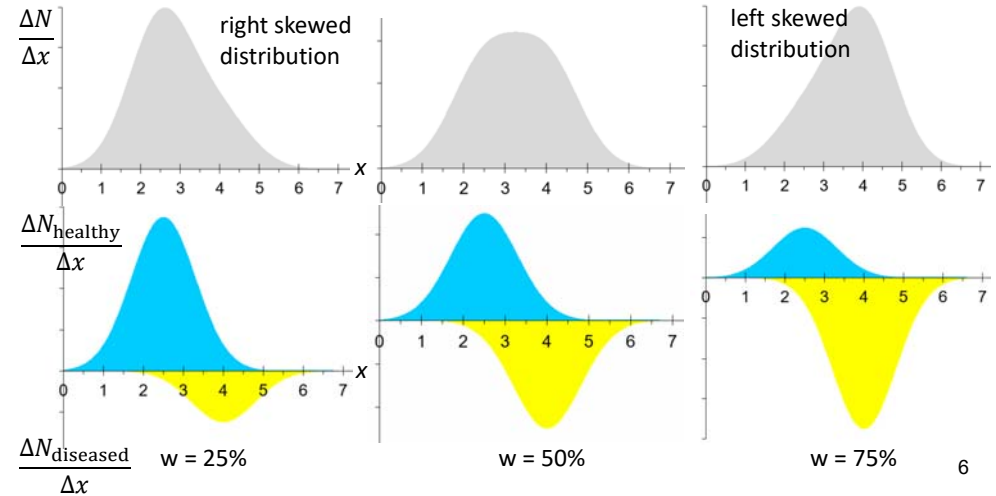


$$w = \frac{\text{diseased}}{\text{total}} = \frac{\text{diseased}}{\text{diseased} + \text{healthy}} = \frac{de - sp}{se - sp}$$

cf: incidence = the number of new cases in a given period and in a given number of population, e.g. 29 per year, per 10 000 people

5

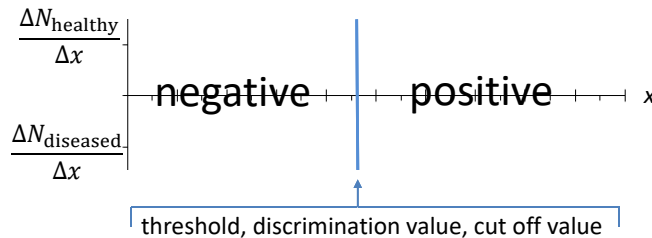
## Effect of prevalence on combined distributions



6

## A negative test result below the threshold and a positive test result above it

among the possible measurement parameter values, by designating a **threshold** value, we decide which will be the positive values and which are the negative ones according to the test method



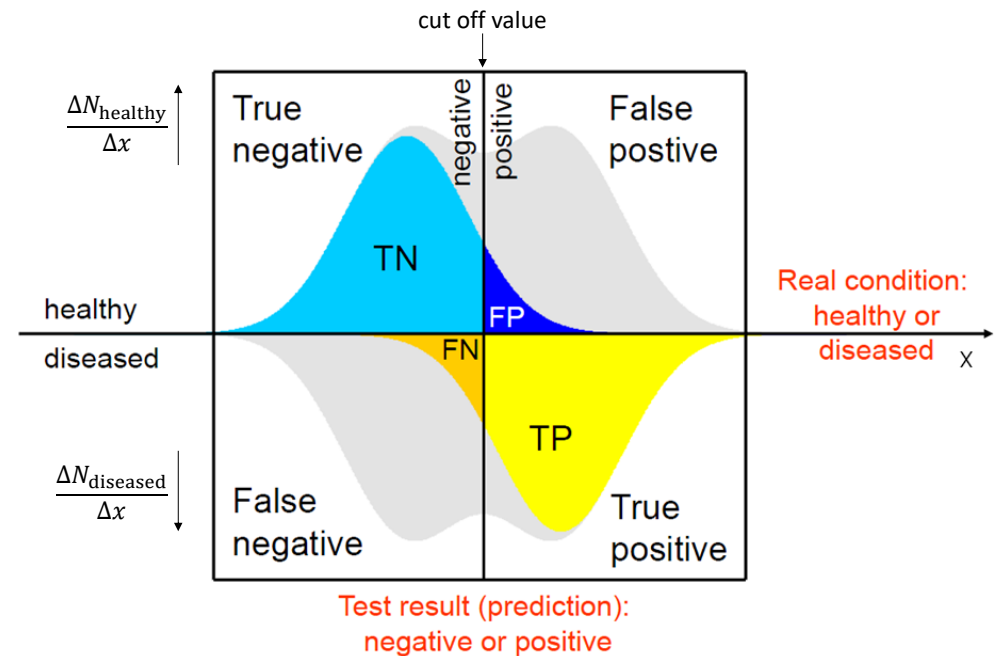
the wish/**desire**/request that the diseased and the positive, respectively healthy and negative match each other as much as possible

however, the **classification** is almost never perfect:

- there will be diseased who are positive: true positive, TP ✓
- there will be diseased who are negative: false negative, FN ●
- there will be healthy who are negative: true negative, TN ✓
- there will be healthy who are positive: false positive, FP ●

7

## Confusion matrix



8

## Parameters of diagnostic „goodness”

based on one (or more) measured parameters diagnostic tests divide the examined into (test) **positive** and (test) **negative** groups

the “goodness” of grouping **cannot** be characterized by a single number

(a) how well does it catch those **to be caught**?

e.g. the probability of a COVID infected stating/determining to be positive

(b) how well does it leave those **to be left alone**?

e.g. the probability of claiming to be negative for a person not infected with a COVID

(c) how reliable is a **positive test result**?

in the case of a positive test result, how certain the patient is diseased

e.g. in the case of a positive COVID test, how certain it is that the person is infected with COVID

(d) how reliable the **negative test result** is?

in the case of a negative test result, how certain the person is healthy

e.g. in the case of a negative COVID test, how certain it is that the person is not infected with COVID

9

The goodness of a test can be described in terms of the following diagnostic parameters

Sensitivity

Specificity

only 3 independent!

PPV, relevance

NPV, segregation

Every method must be compared with a reference-method: **gold standard**  
method known to always work  
(sometimes only the result of an autopsy)



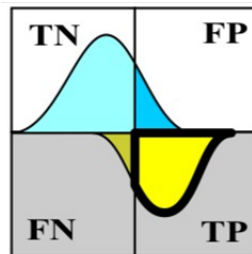
10

## Diagnostic sensitivity

= positive within diseased

= true positive rate

= recall rate



probability that the test finds the diseased positive

$$\frac{\text{TP}}{\text{TP} + \text{FN}} = \text{se} = \frac{\text{true positive}}{\text{diseased}} = p(\text{positive}|\text{diseased})$$

discr. threshold ↓ sens. ↑

Large-sensitivity tests are required:

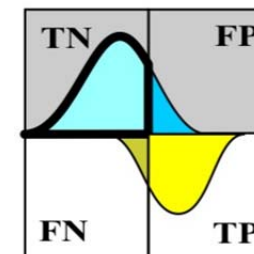
In early diagnosis (screening) so that few patients remain unrecognized.  
If the risk of disease is higher than the risk of treatment.

11

## Diagnostic specificity

= negative among healthy

= true negative rate



probability that the test finds a healthy negative

$$\frac{\text{TN}}{\text{TN} + \text{FP}} = \text{sp} = \frac{\text{true negative}}{\text{healthy}} = p(\text{negative}|\text{healthy})$$

discr. threshold ↑ spec. ↑

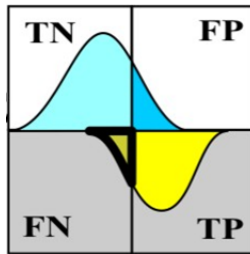
High-specificity tests are important:

When the false positive values have severe consequences (e.g. surgery).  
When the risk of treatment is higher than the risk of disease.

12

## Diagnostic False Negative Rate

Type-II error



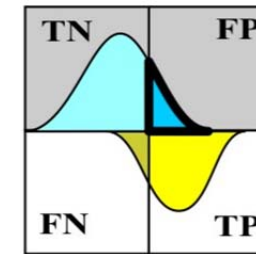
the probability that the test will find a diseased negative  
negative among diseased

$$\frac{\text{FN}}{\text{diseased}} = 1 - \text{se} = \frac{\text{FN}}{\text{FN} + \text{TP}} = \underline{p(\text{negative}|\text{diseased})}$$

13

## Diagnostic False Positive Rate

Type-I error



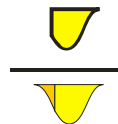
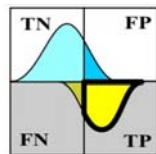
the probability that the test will find a healthy positive  
positive among the healthy

$$\frac{\text{FP}}{\text{healthy}} = 1 - \text{sp} = \frac{\text{FP}}{\text{TN} + \text{FP}} = \underline{p(\text{positive}|\text{healthy})}$$

14

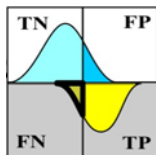
## Horizontal rates are independent of prevalence

sensitivity  
(se)



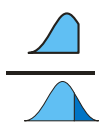
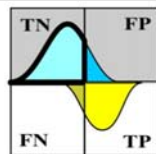
$$\text{se} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

false negative rate  
(1-se)



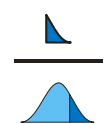
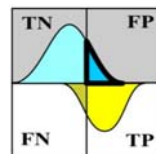
$$1 - \text{se} = \frac{\text{FN}}{\text{FN} + \text{TP}}$$

specificity  
(sp)



$$\text{sp} = \frac{\text{TN}}{\text{TN} + \text{FP}}$$

false positive rate  
(1-sp)



$$1 - \text{sp} = \frac{\text{FP}}{\text{TN} + \text{FP}}$$

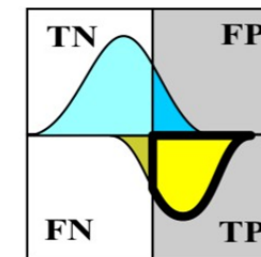
15

## Predictive values (vertical rates)

a-posteriori-probabilities; they depend strongly on prevalence

### Positive predictive value

= PPV  
= predictive value positive  
= PVP  
= diagnostic relevance



probability of disease if test is positive  
diseased among positives

$$\frac{\text{TP}}{\text{positive}} = \text{PPV} = \frac{\text{TP}}{\text{TP} + \text{FP}} = \frac{\text{se} \cdot w}{\text{se} \cdot w + (1 - \text{sp}) \cdot (1 - w)} = \underline{p(\text{diseased}|\text{positive})}$$

16

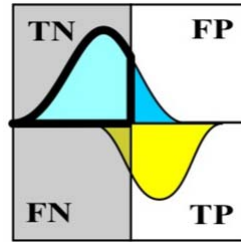
## Negative predictive value

= NPV

= predictive value negative

= PVN

= diagnostic **segregation**



probability of health  
if test is negative

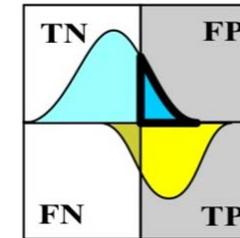
healthy among  
negatives

$$\frac{\text{TN}}{\text{TN} + \text{FN}} = \text{NPV} = \frac{\text{TN}}{\text{negative}} = \frac{\text{TN}}{\text{TN} + \text{FN}} = \frac{\text{sp} \cdot (1 - w)}{\text{sp} \cdot (1 - w) + (1 - \text{se}) \cdot w} = p(\text{healthy}|\text{negative})$$

17

## False alarm rate

= 1 - PPV



the probability of the  
absence of the disease  
if the test is positive

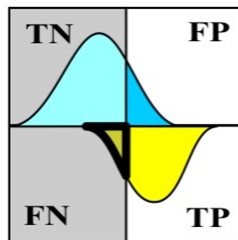
healthy among  
positives

$$\frac{\text{FP}}{\text{FP} + \text{TP}} = 1 - \text{PPV} = \frac{\text{FP}}{\text{positive}} = \frac{\text{FP}}{\text{FP} + \text{TP}} = p(\text{healthy}|\text{positive})$$

18

## False reassurance rate

= 1 - NPV



the probability of the  
presence of the disease  
if the test is negative

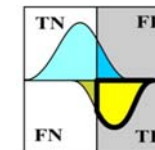
diseased among  
negatives

$$\frac{\text{FN}}{\text{FN} + \text{TN}} = 1 - \text{NPV} = \frac{\text{FN}}{\text{negative}} = \frac{\text{FN}}{\text{FN} + \text{TN}} = p(\text{diseased}|\text{negative})$$

19

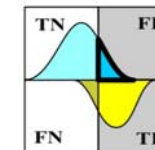
## Vertical rates are dependent of prevalence

positive predictive  
value (PPV)



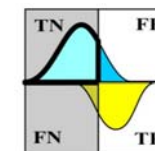
$$\text{PPV} = \frac{\text{TP}}{\text{FP} + \text{TP}}$$

false alarm rate  
(1 - PPV)



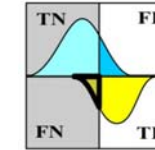
$$1 - \text{PPV} = \frac{\text{FP}}{\text{FP} + \text{TP}}$$

negative  
predictive value  
(NPV)



$$\text{NPV} = \frac{\text{TN}}{\text{TN} + \text{FN}}$$

false reassurance rate  
(1 - NPV)



$$1 - \text{NPV} = \frac{\text{FN}}{\text{TN} + \text{FN}}$$

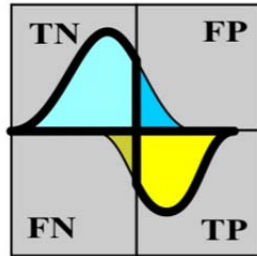
20

## Diagnostic accuracy

= da=de

= efficacy/efficiency

= correct classification rate



probability of correct diagnosis

$$\frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FN} + \text{FP}} = \text{de} = \frac{\text{TP} + \text{TN}}{\text{total}} = \frac{\text{TP} + \text{TN}}{\text{TN} + \text{FP} + \text{FN} + \text{TP}} = \frac{\text{se} \cdot w + \text{sp} \cdot (1 - w)}{1}$$

often: discrimination threshold is chosen so that accuracy is maximized

21

## Effect of prevalence

NPV = 90%

case1:  $w = 50\%$

sp = 90%

		Test	
		negative	positive
Gold-standard	healthy	90	10
	diseased	10	90

se = 90%

(de = 90%)

PPV = 90%

NPV = 99%

Case 2:  $w = 10\%$

sp = 90%

		Test	
		negative	positive
Gold-standard	healthy	810	90
	diseased	10	90

se = 90%

(de = 90%)

PPV = 50%

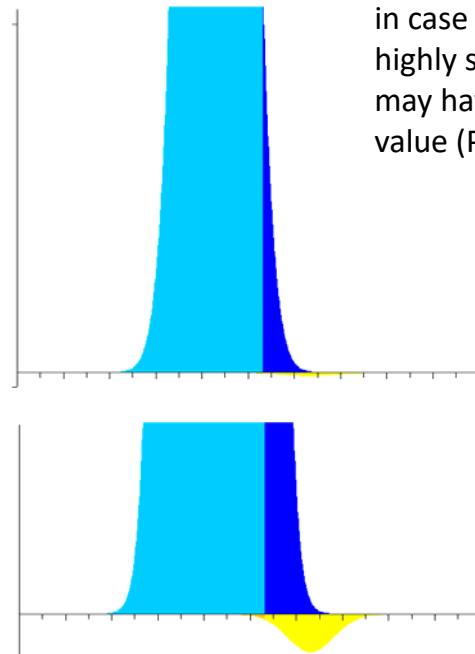
22

## Overview

<b>Sensitivity</b>	se	$\frac{\text{TP}}{\text{TP} + \text{FN}}$	$p(P D)$	positive within diseased	True Positive Rate	prevalence-independent
<b>Specificity</b>	sp	$\frac{\text{TN}}{\text{TN} + \text{FP}}$	$p(N H)$	negative among healthy	True Negative Rate	
<b>False Negative Rate</b>	1-se	$\frac{\text{FN}}{\text{TP} + \text{FN}}$	$p(N D)$	negative among diseased		
<b>False Positive Rate</b>	1-sp	$\frac{\text{FP}}{\text{TN} + \text{FP}}$	$p(P H)$	positive among the healthy		
<b>Positive Predictive Value</b>	PPV	$\frac{\text{TP}}{\text{TP} + \text{FP}}$	$p(D P)$	diseased among positives	Relevance	prevalence-dependent
<b>Negative Predictive Value</b>	NPV	$\frac{\text{TN}}{\text{TN} + \text{FN}}$	$p(H N)$	healthy among negatives	Segregation	
<b>False alarm rate</b>	1-PPV	$\frac{\text{FP}}{\text{TP} + \text{FP}}$	$p(H P)$	healthy among positives		
<b>False reassurance rate</b>	1-NPV	$\frac{\text{FN}}{\text{TN} + \text{FN}}$	$p(D N)$	diseased among negatives		

conditional probability (Bayes)

24



in case of very small prevalence a highly sensitive and specific test may have low positive predictive value (PPV)

prevalence = 0.1 %

sensitivity = 98 %

specificity = 98 %

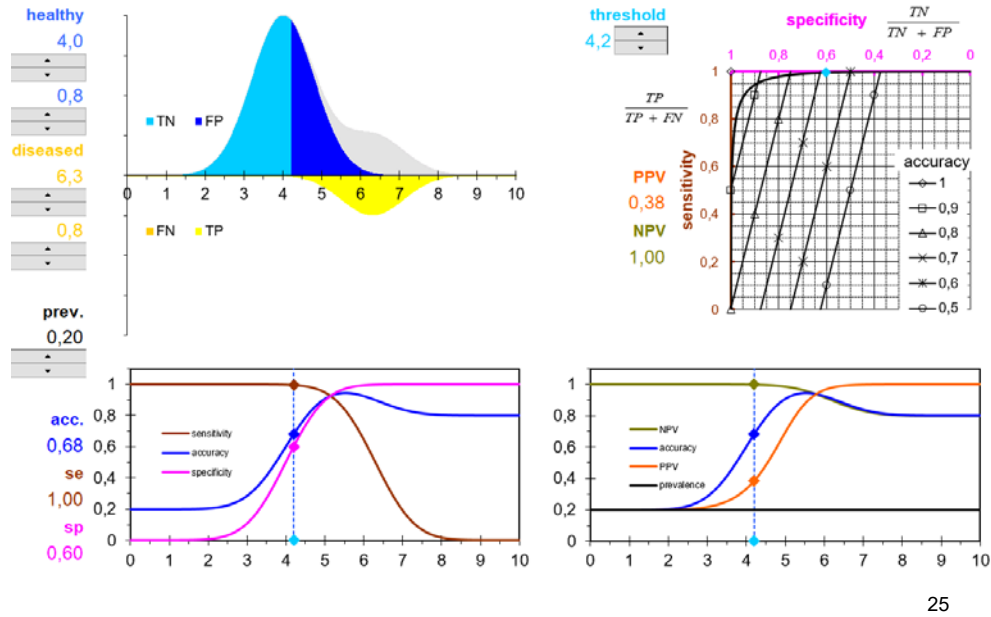
PPV = 4 %

23



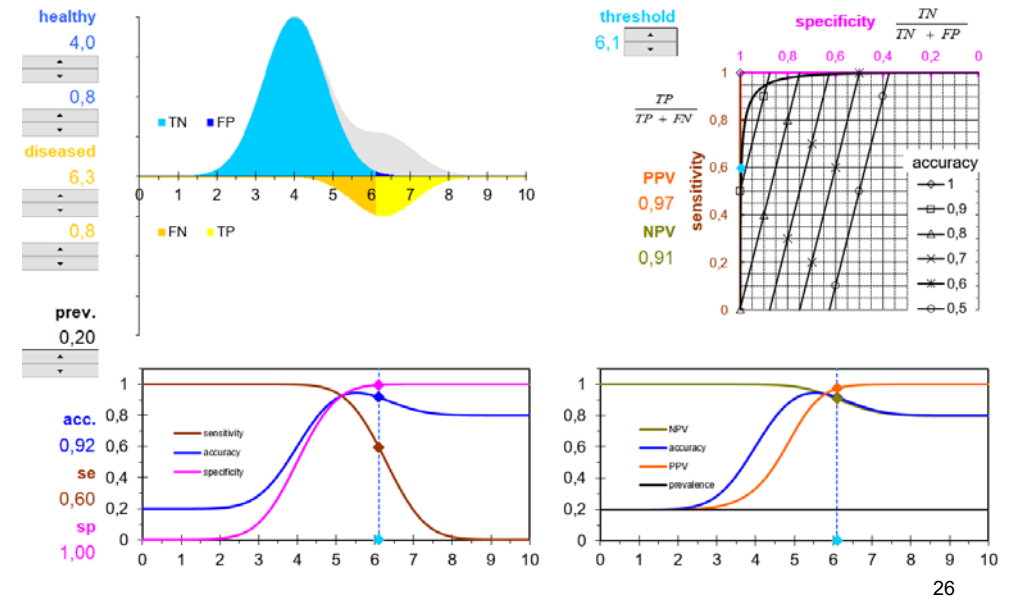
how well does it catch those **to be caught**?

## Maximize diagnostic sensitivity



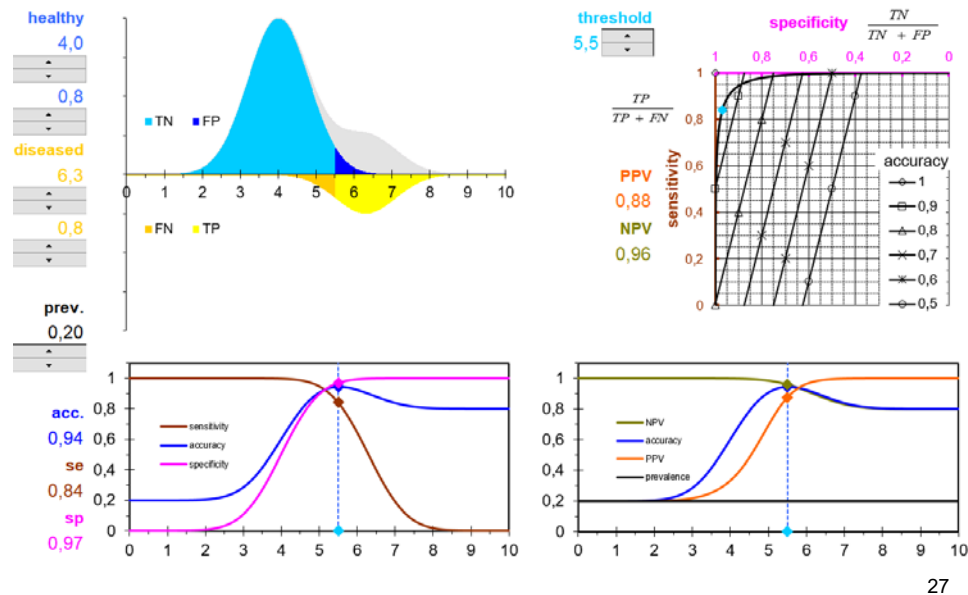
how well does it leave those **to be left alone**?

## Maximize diagnostic specificity



it is equally important to catch those **to be caught** and to leave those **to be left alone**?

## Maximize diagnostic accuracy



everything has a **distribution**; the distribution of sick and healthy values **overlap**  
whether it is possible to decide which is more important :

## Take-home message

to detect the disease in as many patients as possible in order to receive treatment (**maximizing sensitivity**), or

to assume a false positive value (minimizing false-positive ratio or **maximizing specificity**)  
in as few healthy people as possible so that they do not receive unnecessary therapy

if they cannot be decided, they are equally important: **maximizing accuracy**

