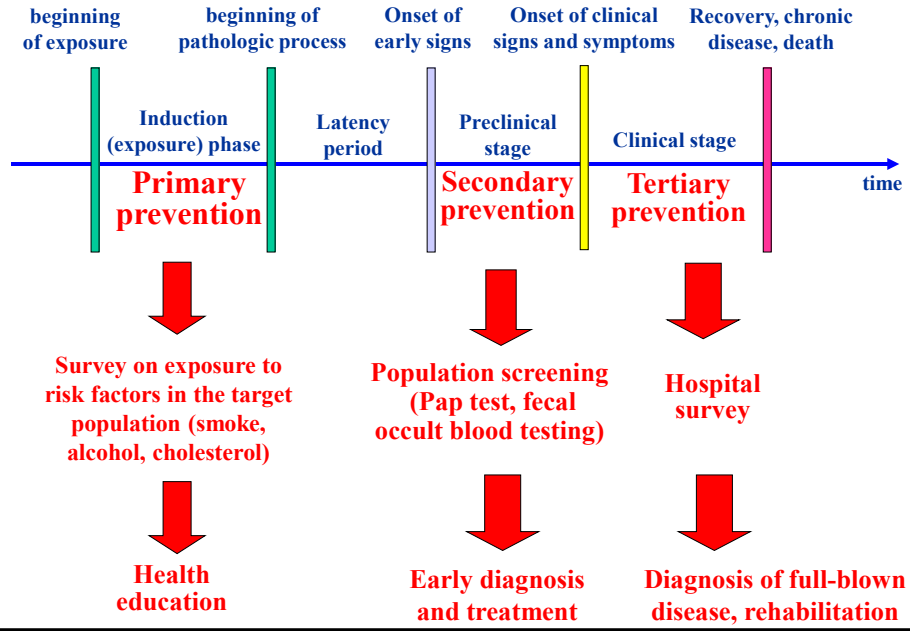
Screening, sensitivity and specificity of a diagnostic test, R.O.C. curves, Bayes' theorem

Prof. Giuseppe Verlato

Unit of Epidemiology & Medical Statistics,
Department of Diagnostics & Public Health,
University of Verona

Natural history of a disease (Rothman, 1981)



Primary, secondary and tertiary prevention

Before the disease onset:
Primary prevention = preventing or eliminating exposure to risk factors (for example, anti-smoking or anti-alcohol campaigns).

The disease has already established, but it is still at an early stage and clinically undetectable:
Secondary prevention = detecting disease cases at an early stage through a screening (for example, Pap smear test for cervical cancer, mammography for breast cancer, fecal occult blood testing for colon cancer).

The disease has become fully evident:
Tertiary prevention = Proper treatment and rehabilitation to prevent or soften the negative impact of the disease (for instance, care and rehabilitation of people with myocardial infarction).

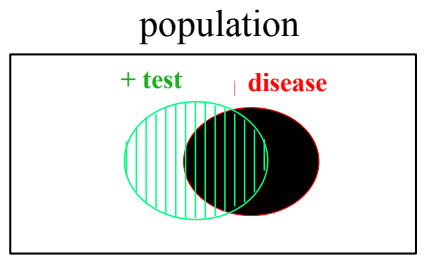
Screening

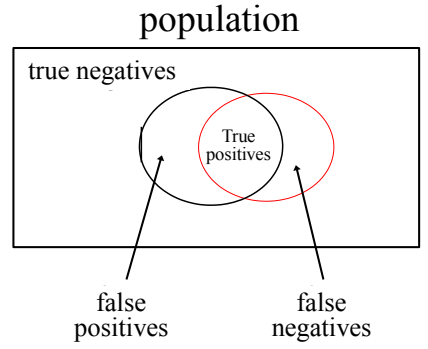
1) Administering a non-invasive and non-expensive test

2) to large population strata at risk for a certain disease

3) to detect affected individuals, before the disease itself becomes apparent from a clinical point of view.

The aim of a screening program is to detect the disease at an early stage, when chances of recovery are still high.





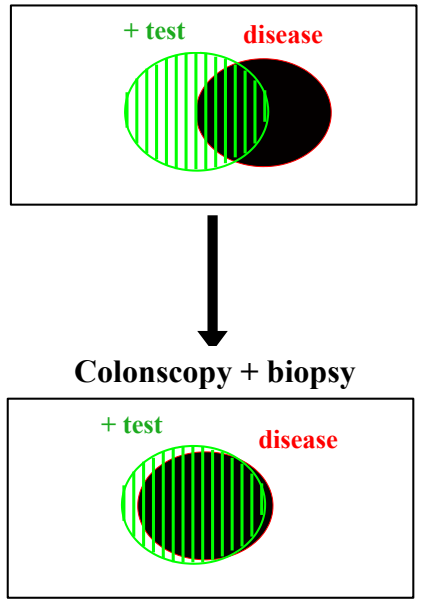
Ideal situation for a screening test

	diseased	healthy	
+ test	a	-----	
- test	-----	d	

In the real world

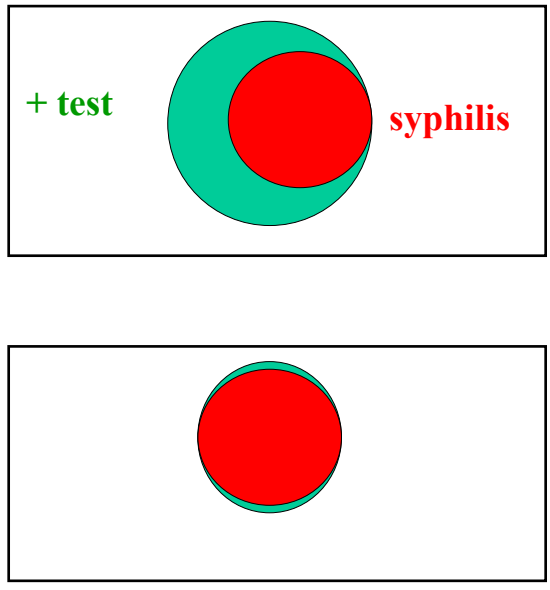
	diseased	healthy	
+ test	a	False positives	
- test	False negatives	d	

Occult blood testing in stools



Verlato, 1998

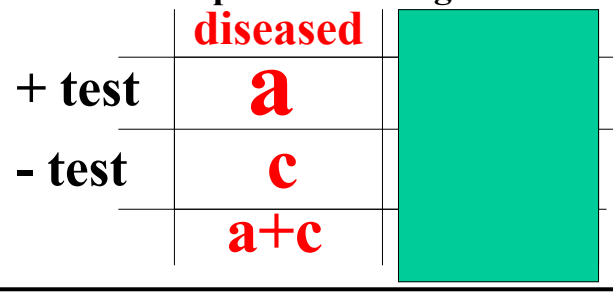
Screening for syphilis infection



1st step: false negatives are removed by VDRL test.

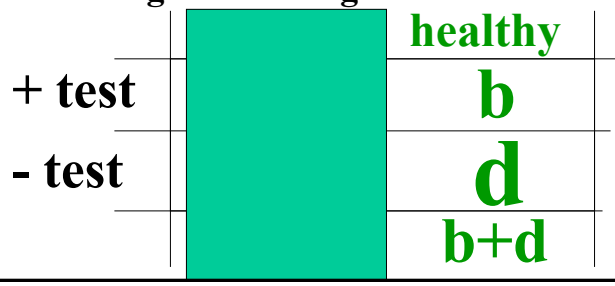
2nd step: false positives are removed by Nelson test.

Sensitivity = p(T+/M+) = probability that the test will be positive among those who are diseased



Sens = a/(a+c)

Specificity = p(T-/M-) = probability that the screening test will be negative among those who do not have the disease



Spec = d/(b+d)

SCREENING

Population at high risk

	M+	M-	
T+	291	7	298
T-	9	693	702
	300	700	1000

General population

	M+	M-	
T+	2910	9970	12880
T-	90	987030	987120
	3000	997000	1000000

Prevalenc : P(M+) =	300/1000= 0.30	_____
sensitivity : P(T+/M+) =	291/300 = 0.97	_____
specificity : P(T-/M-) =	693/700 = 0.99	_____
V+ = P(M+/T+) =	291/298 = 0.977	_____
V- = P(M-/T-) =	693/702 = 0.987	_____

V+ = positive predictive value
V- = negative predictive value

SCREENING

Population at high risk	M+	M-	
T+	291	7	298
T-	9	693	702
	300	700	1000

General population	M+	M-	
T+	2910	9970	12880
T-	90	987030	987120
	3000	997000	1000000

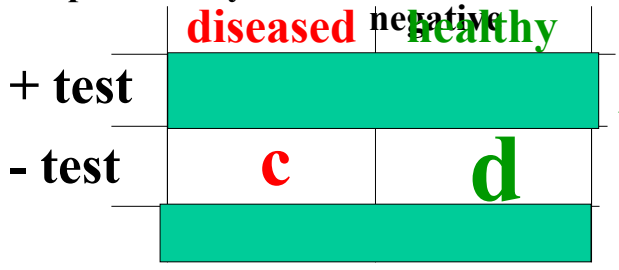
Prevalenc : P(M+) =	300/1000= 0.30	**3000 / 1 000 000 = 0.003 = 0.3%**
sensitivity = P(T+/M+) =	291/300 = 0.97	**2910 / 3000 = 0.97 = 97%**
specificity : P(T-/M-) =	693/700 = 0.99	**987030 / 997000 = 0.99 = 99%**
V+ = P(M+/T+) =	291/298 = 0.977	**2910 / 12880 = 0.226 = 22.6%**
V- = P(M-/T-) =	693/702 = 0.987	**987030 / 987120 = 0.9999=99.99%**

V+ = positive predictive value
V- = negative predictive value

Positive predictive value (V+) = p(M+/T+) = probability to actually have the disease if the test is positive

V+ = a/(a+b)

Negative predictive value (V-) = p(M-/T-) = probability not to have the disease if the test is negative



V- = d/(c+d)

EXAMPLE: SCREENING for BREAST CANCER

Giorgi et al [2006] summarized the results of screening programs for breast cancer, performed in Italy in 2003-04: 7.8% of women undergoing their 1st mammography were referred for further examinations, while the percentage of women diagnosed with breast cancer was 0.65% in the overall population participating in screening programs [Giorgi et al, 2006].

Hence the positive predictive value of mammography was 0.65% / 7.8% = 0.083: in other words 1 in 12 women, referred for invasive diagnostic procedures, did have a malignancy. Positive predictive value is always rather low in screening programs on the general population .

Of course, it is fully acceptable that 11 healthy women could uselessly undergo invasive procedures, in order to detect and eliminate a malignancy at an early stage. However, "*this value needs to be reasonably low, in order to limit the negative psychological impact (anxiety), the invasive procedure (cytology, core, or surgical biopsies), which may be required, as well as costs*" (*questo valore deve essere ragionevolmente basso, per limitare l'impatto psicologico negativo (ansietà), le procedure invasive indicate (citologia, prelievo dal centro del nodulo, o biopsie chirurgiche), come pure i costi*) [Giorgi et al, 2006].

Giorgi D, Giordano L, Ventura L, Puliti D, Piccini P, Paci E (2006) Mammography screening in Italy: 2003-2004 survey. Epidemiologia e Prevenzione, 30(1) supplemento 3: 7-16.

Other measures of diagnostic accuracy, mainly used in the clinical setting

Positive likelihood ratio (LR+)

Ratio between the probability of a POSITIVE test given the PRESENCE of the disease and the probability of a POSITIVE test given the ABSENCE of the disease:

$$LR+ = \frac{P(T+/M+)}{P(T+/M-)} = \frac{\text{sensitivity}}{\text{1-specificity}}$$

Negative likelihood ratio (LR-)

Ratio between the probability of a NEGATIVE test given the PRESENCE of the disease and the probability of a NEGATIVE test given the ABSENCE of the disease:

$$LR- = \frac{P(T-/M+)}{P(T-/M-)} = \frac{\text{1-sensitivity}}{\text{specificity}}$$

Cut-off for LR+ and LR-

When LR+ is greater than 5, a positive test will confidently confirm the presence of the disease

When LR- is lower than 0.2, a negative test will confidently exclude the disease

Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. JAMA 1994 Mar 2;271(9):703-7.

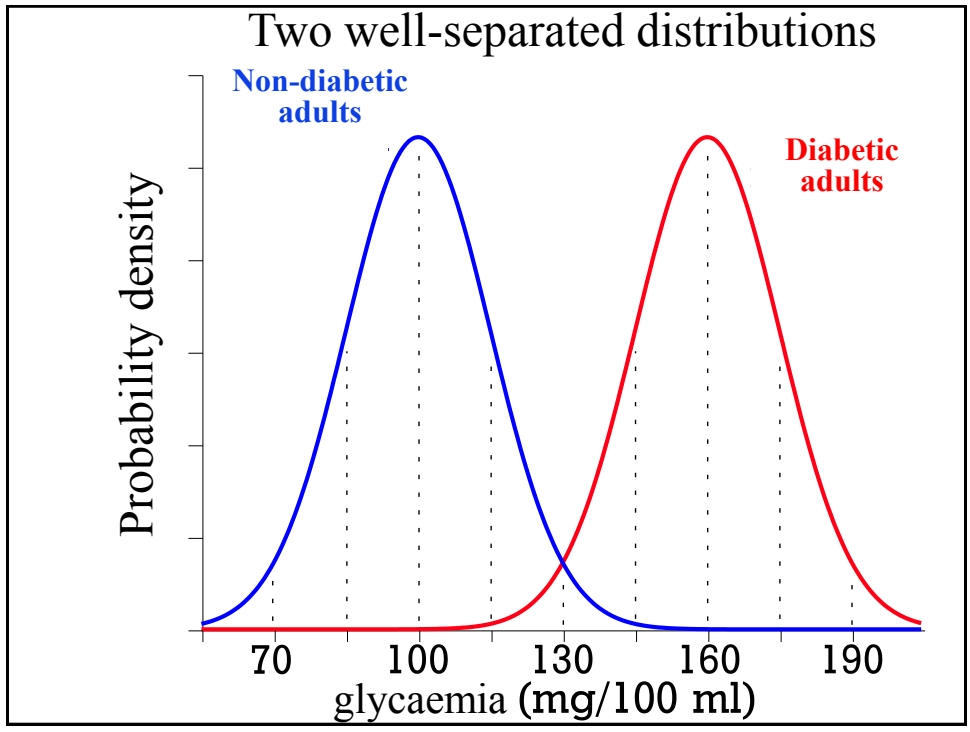
Sometimes the diagnostic test is based on a CONTINUOUS variable. For instance:

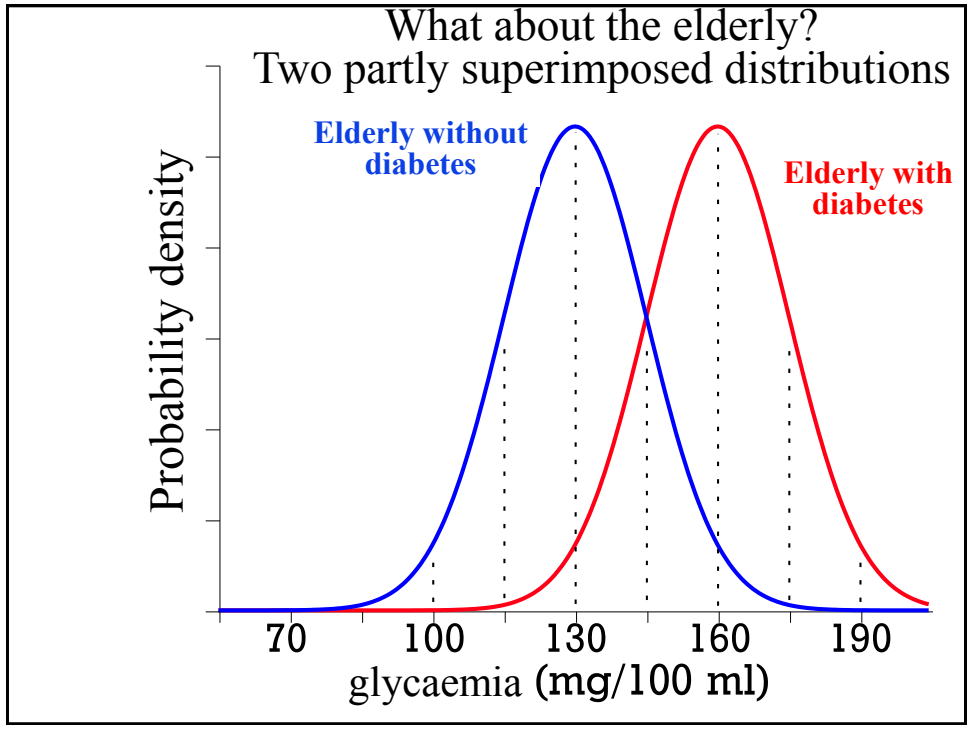
The fasting plasma glucose diagnostic threshold for diabetes is 7.0 mmol/l (126 mg/dl).

The blood pressure threshold for defining hypertension is 140/90 mmHg.

DECISION LEVEL PLOT → to choose the optimal cut-off

R.O.C. CURVE → to evaluate the overall performance of the test over the entire range of possible cut-offs


Two well-separated distributions
Non-diabetic adults
Diabetic adults
Probability density
70
100
130
160
190
glycaemia (mg/100 ml)

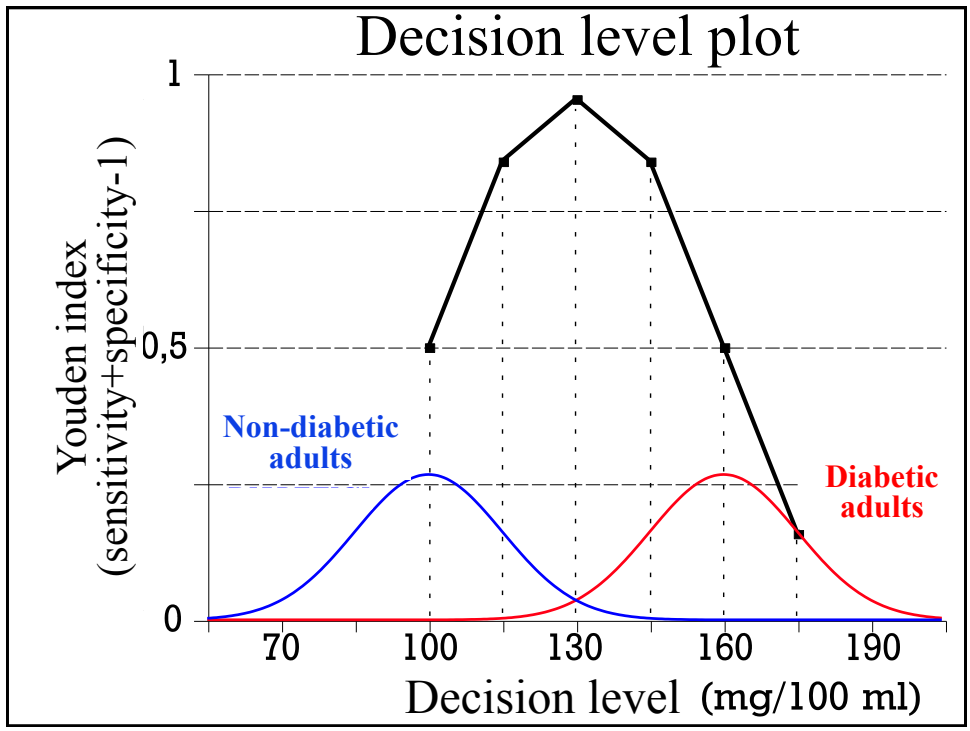

What about the elderly?
Two partly superimposed distributions
Elderly without diabetes
Elderly with diabetes
Probability density
70
100
130
160
190
glycaemia (mg/100 ml)

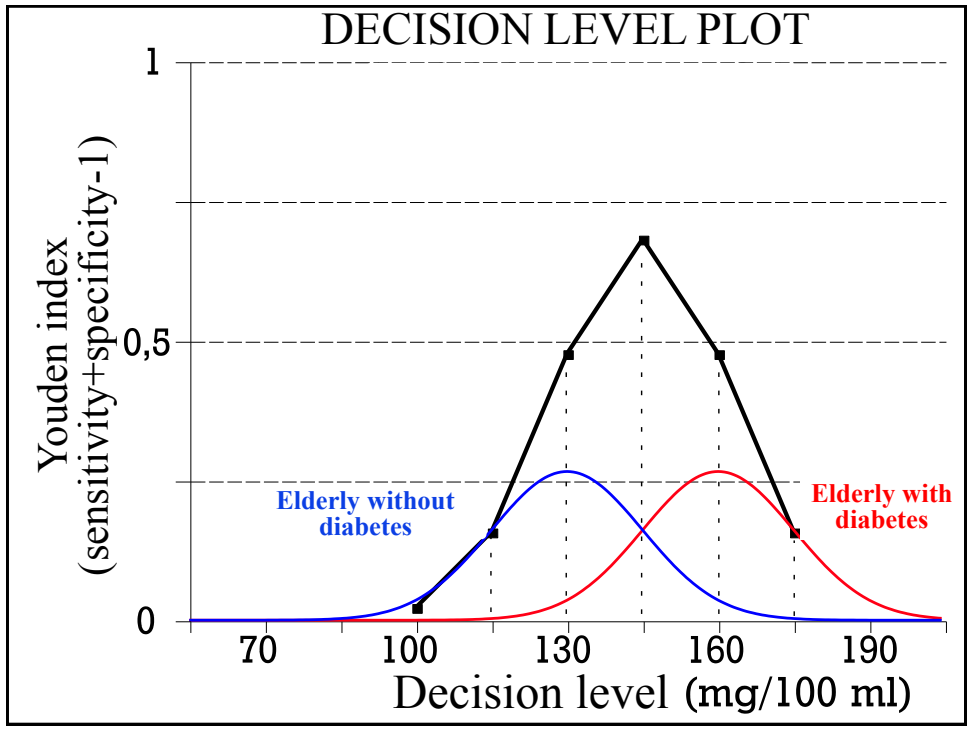
1st EXAMPLE: ADULTS with or without diabetes				2nd EXAMPLE: ELDERLY with or without diabetes		
specificity	1-specificity	sensitivity	CUT-OFF	specificity	1-specificity	sensitivity
50.0 %	50.0 %	99.997 %	100 mg/dl	2.3 %	97.7 %	99.997 %
84.1 %	15.9 %	99.9 %	115 mg/dl	15.9 %	84.1 %	99.9 %
97.7 %	2.3 %	97.7 %	130 mg/dl	50.0 %	50.0 %	97.7 %
99.9 %	0.1 %	84.1 %	145 mg/dl	84.1 %	15.9 %	84.1 %
99.997 %	0.003 %	50.0 %	160 mg/dl	97.7 %	2.3 %	50.0 %
---	---	---	175 mg/dl	99.9 %	0.1 %	15.9 %

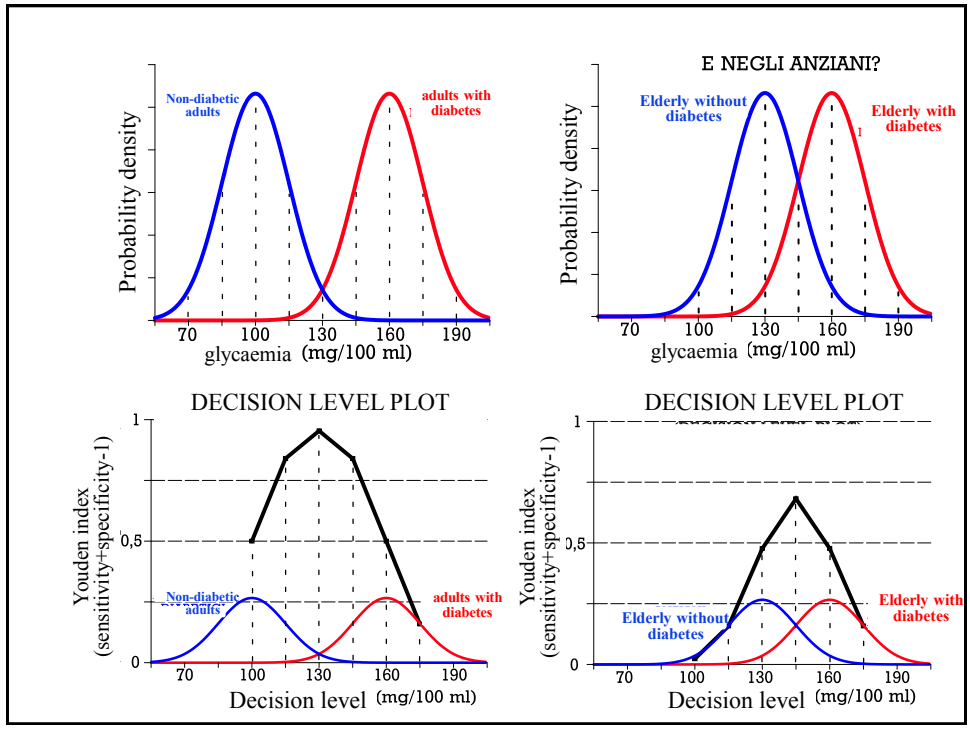
Used to create DECISION LEVEL PLOTS and R.O.C. CURVES

Used to create DECISION LEVEL PLOTS and R.O.C. CURVES

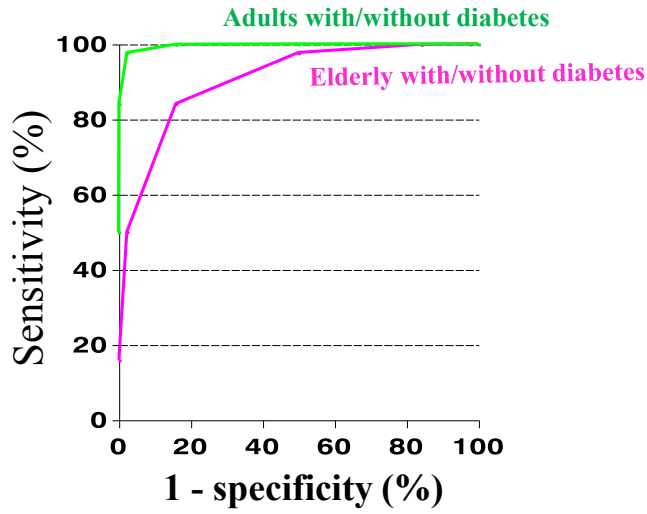
R.O.C. = *Receiver Operating Characteristic*

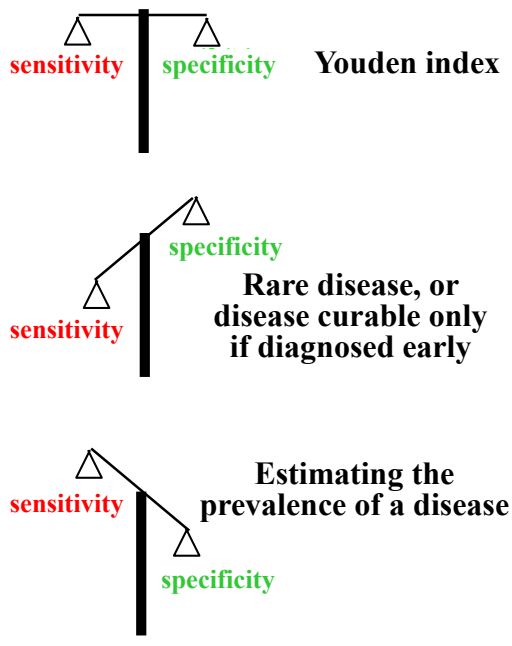



DECISION LEVEL PLOT
Youden index
(sensitivity+specificity-1)
1
0,5
0
Elderly without
diabetes
Elderly with
diabetes
70
100
130
160
190
Decision level (mg/100 ml)


Probability density
Non-diabetic
adults
adults with
diabetes
70
100
130
160
190
glycaemia (mg/100 ml)
E NEGLI ANZIANI?
Elderly without
diabetes
Elderly with
diabetes
DECISION LEVEL PLOT
Youden index
(sensitivity+specificity-1)
1
0,5
0
Decision level (mg/100 ml)

R.O.C. (*Receiver Operating Characteristic)* Curves



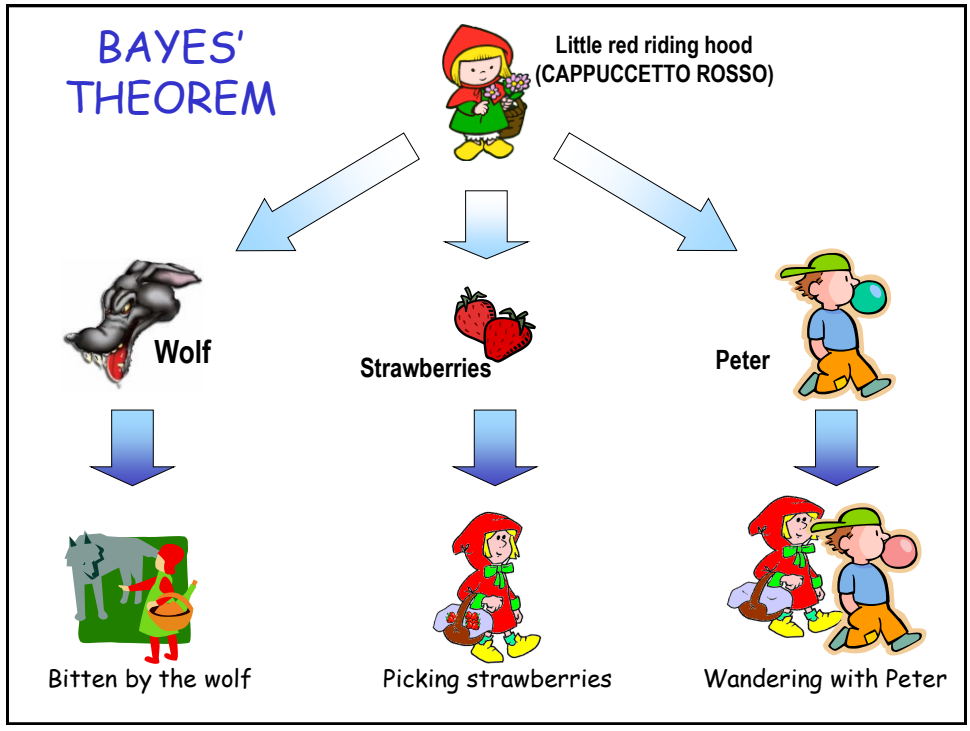


Verlato, 1998

BAYES' THEOREM and its application to DIFFERENTIAL DIAGNOSIS

1st EXAMPLE: from the world of fairy tales

2nd EXAMPLE: clinical examples

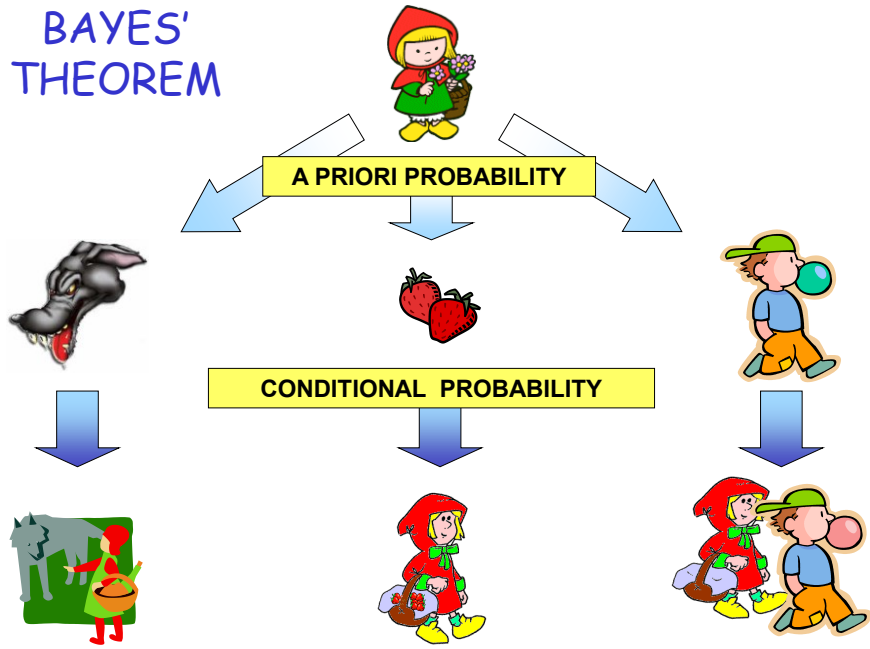


Assumptions:

1. P(Wolf ∩ Strawberries) =P(Wolf ∩ Peter) =P(Peter ∩ Strawberries) =0

2. P(bitten by wolf) + P(picking strawberries) + P(wandering with Peter) =1

BAYES' THEOREM



Bayes' theorem

(*Thomas Bayes 1702 - 1761*)

We know the **effect**, we have a list of **possible causes** and we want to assign to each cause the probability to have produced the effect.

In medicine a patient reports a **symptom** to a doctor, who has to finding the **disease** causing this symptom among a list of possible diseases.

Clinical application of Bayes' theorem
Clinical case: Hematuria in 25 years-old man

	Kidney stone	Glomerulo-nephritis	Cancer	Total
a priori probability– p(D)	0.1%	0.5%	0.01%	-------
conditional prob. - p(S/D)	50%	80%	60%	-------
probability product	5/10,000	40/10,000	0.6/10,000	45.6/10,000
posterior probability (D/S)	5/45.6	40/45.6	0.6/45.6	45.6/45.6
	11.0%	87.7%	1.3%	100%

$$p(D_1/S) = \frac{p(D_1) * p(S/D_1)}{p(D_1) * p(S/D_1) + p(D_2) * p(S/D_2) + p(D_3) * p(S/D_3)}$$

ASSUMPTIONS

1) There are only three diseases (kidney stone, glomerulonephritis, cancer) causing hematuria

2) The three diseases are mutually exclusive

Bayes' formula

It conveniently displays the single steps of diagnostic procedure, showing how probabilities initially attributed to different causes (diseases) are subsequently modified by newly collected information (symptoms).

$$P(D_i \mid S) = \frac{P(D_i) \cdot P(S \mid D_i)}{\sum_{i=1}^{k} P(D_i) \cdot P(S \mid D_i)}$$

where:

$D_1 \ldots D_i \ldots D_k \Rightarrow$ ***possible causes of the symptom under study***

$S \Rightarrow$ ***symptom/sign under study***

$P(D_i)$ is the ***a priori probability*** of the cause D_i: it can be viewed as the *probability* that a physician assigns to a *given disease* BEFORE visiting the patient, according to disease occurrence (incidence/prevalence)

$P(S|D_i)$ is the ***conditional probability***: the probability of the symptom given the disease Di.

$P(D_i|S)$ is the ***posterior probability*****:** it measures the probability that the event S, already occurred, could be attributed to the cause D_i, among a finite set ***k*** of possible causes. In the clinical setting it represents the new probability that the physician assigns to the disease AFTER having visited the patient.

Computation of posterior probabilities $\Rightarrow$ DIFFERENTIAL DIAGNOSIS

Clinical application of Bayes' theorem
Clinical case: Hemoptysis in a 40 years-old man

	TBC	Lung cancer	Pneumonia	Total
a priori probability– p(D)	0.01%	0.1%	1%	-------
conditional prob. - p(S/D)	80%	40%	2%	-------
probability product				
posterior probability (D/S)				

ASSUMPTIONS

Clinical application of Bayes' theorem
Clinical case: Hemoptysis in a 40 years-old man

	TBC	Lung cancer	Pneumonia	Total
a priori probability– p(D)	0.01%	0.1%	1%	-------
conditional prob. - p(S/D)	80%	40%	2%	-------
probability product	0.8/10,000	4.0/10,000	2.0/10,000	6.8/10,000
posterior probability (D/S)	0.8/6.8	4/6.8	2/6.8	6.8/6.8
	11.8%	58.8%	29.4%	100%

$$p(D_1/S) = \frac{p(D_1) * p(S/D_1)}{p(D_1) * p(S/D_1) + p(D_2) * p(S/D_2) + p(D_3) * p(S/D_3)}$$

ASSUMPTIONS

1) There are only three diseases (TBC, lung cancer, pneumonia) causing hemoptysis

2) The three diseases are mutually exclusive

Application of the Bayes' theorem in clinical practice

Bayes' theorem has not been extensively applied in clinical practice, as its assumptions are not met.

1. It is seldom possible to identify all possible diseases which could cause a certain symptom/sign.

2. Diseases often simultaneously occur in the same subjects (comorbidities, multimorbidities, syndromes).