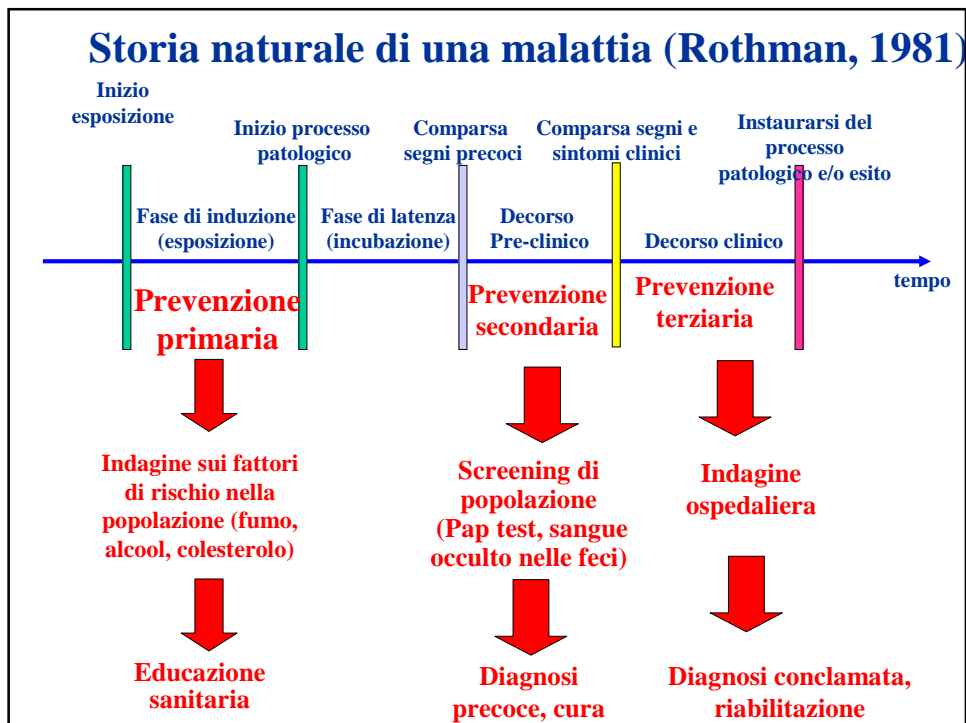


# Screening, sensibilità e specificità di un test diagnostico, curve R.O.C., teorema di Bayes

Prof. Giuseppe Verlato  
Sezione di Epidemiologia e Statistica Medica,  
Università di Verona



## **Prevenzione primaria, secondaria e terziaria**

### **Prima che si instauri la malattia:**

Prevenzione primaria = Rimozione dei fattori di rischio (ad esempio, campagne contro il fumo o contro l'alcoolismo).

### **La malattia si è instaurata, ma non è ancora evidente dal punto di vista clinico:**

Prevenzione secondaria = Individuazione precoce dei casi tramite uno **screening** (ad esempio, Pap test per il tumore dell'utero, mammografia per il tumore del seno, sangue occulto nelle feci per il tumore del colon).

### **La malattia si è manifestata clinicamente:**

Prevenzione terziaria = Terapia appropriata e riabilitazione per prevenire o ridurre le conseguenze negative della malattia stessa (ad esempio, assistenza agli infartuati e riabilitazione).

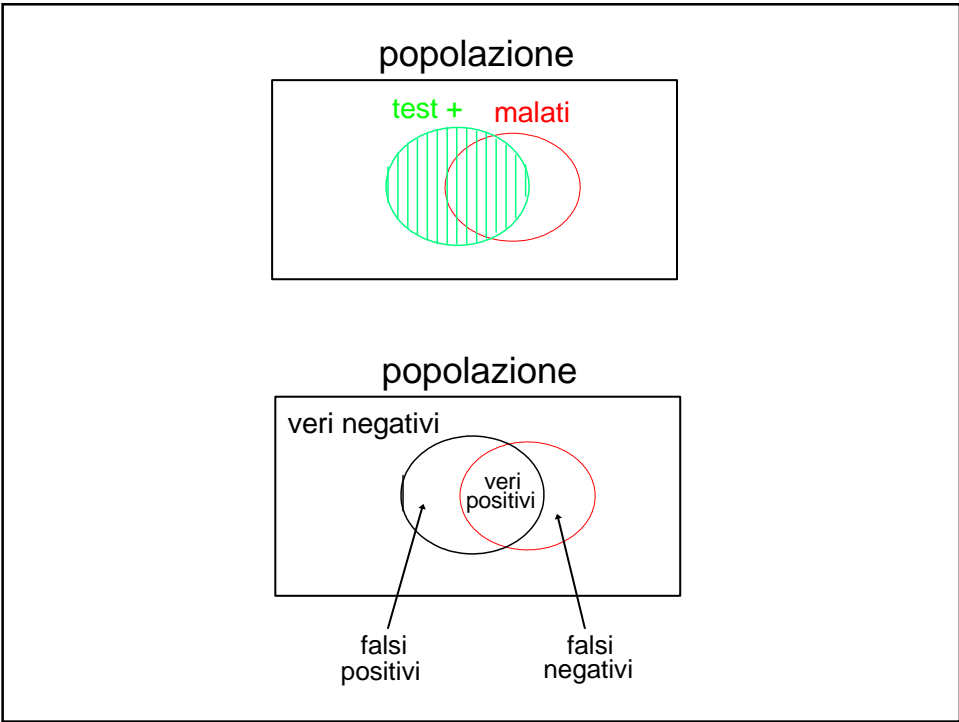
## **Screening**

1) Somministrazione di un test diagnostico poco costoso e poco invasivo

2) a larghi settori della popolazione a rischio per una determinata patologia

3) per identificare gli individui ammalati prima che la malattia si riveli dal punto di vista clinico.

Lo scopo dello screening è diagnosticare precocemente la malattia, quando è ancora curabile.



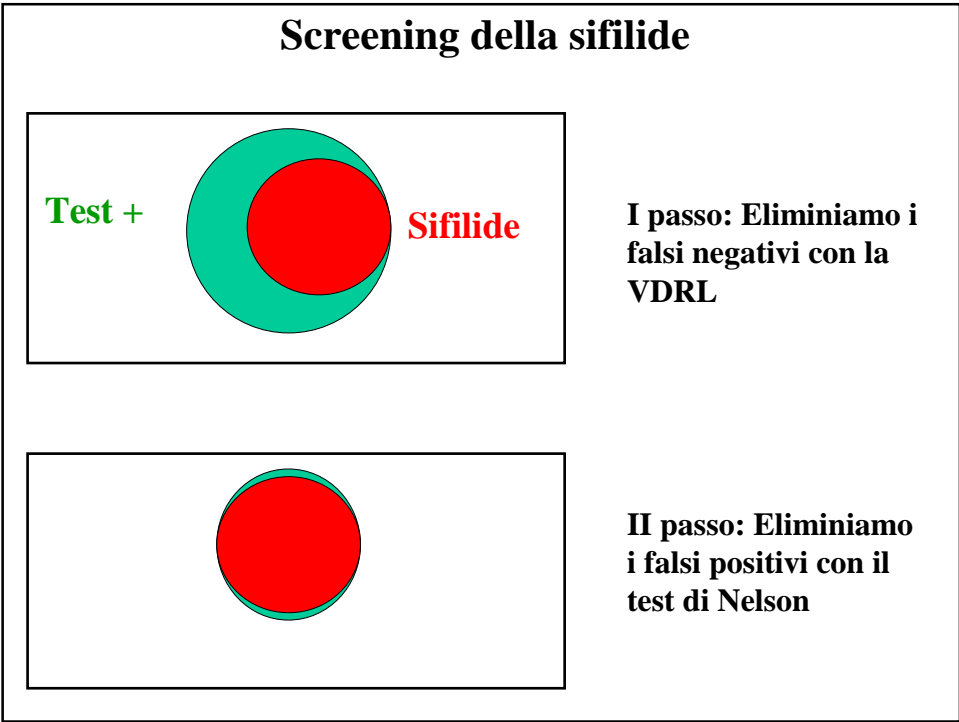
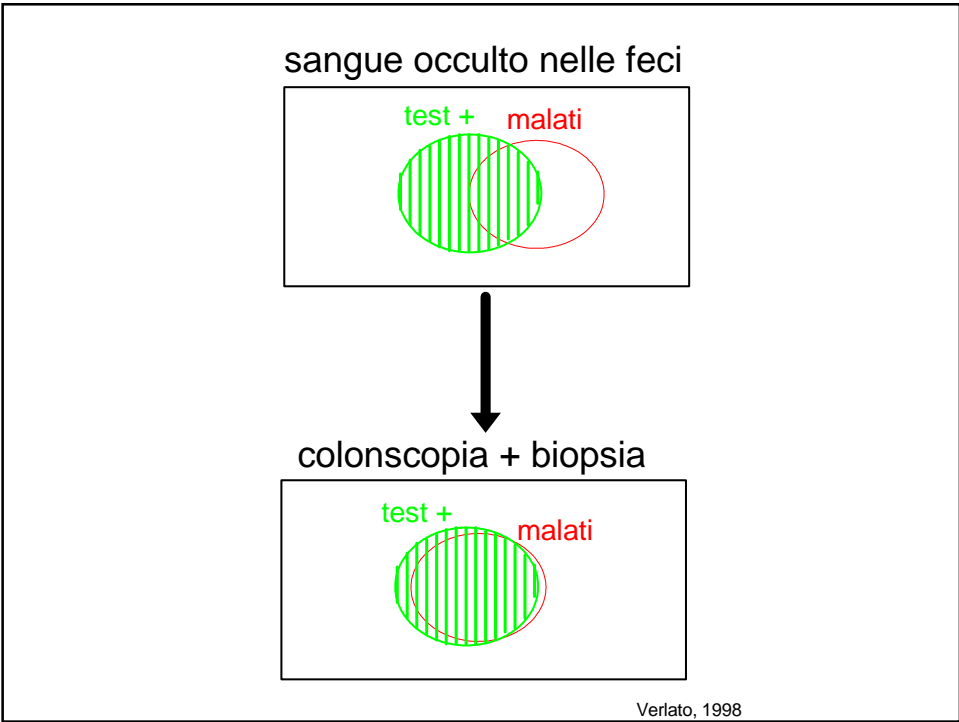
**Situazione ideale in un test di screening**

	malati	sani	
Test +	<b>a</b>	-----	
Test -	-----	<b>d</b>	

**Nella cruda realtà**

	malati	sani	
Test +	<b>a</b>	Falsi positivi	
Test -	Falsi negativi	<b>d</b>	



<b>Sensibilità = <math>p(T+/M+) =</math> probabilità che il test sia positivo nei malati</b>		
	<b>malati</b>	
<b>Test +</b>	<b>a</b>	
<b>Test -</b>	<b>c</b>	
	<b>a+c</b>	
		<b>Sens = <math>a/(a+c)</math></b>
<b>Specificità = <math>p(T-/M-) =</math> probabilità che il test sia negativo nei sani</b>		
		<b>sani</b>
<b>Test +</b>		<b>b</b>
<b>Test -</b>		<b>d</b>
		<b>b+d</b>
		<b>Spec = <math>d/(b+d)</math></b>

SCREENING						
	Popolazione a rischio			Popolazione generale		
	M+	M-		M+	M-	
T+	291	7		2910	9970	
T-	9	693		90	987030	
	300	700		3000	997000	
	1000			1000000		
Prevalenza = $P(M+) =$	$300/1000 = 0.30$			_____		
Sensibilità = $P(T+/M+) =$	$291/300 = 0.97$			_____		
Specificità = $P(T-/M-) =$	$693/700 = 0.99$			_____		
$V+ = P(M+/T+) =$	$291/298 = 0.977$			_____		
$V- = P(M-/T-) =$	$693/702 = 0.987$			_____		
V+ = Valore predittivo dei positivi V- = Valore predittivo dei negativi						

## SCREENING

Popolazione a rischio

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

Popolazione generale

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

Prevalenza = P(M+) =	300/1000 = 0.30	3000 / 1 000 000 = 0,003 = 0,3%
Sensibilità = P(T+/M+) =	291/300 = 0.97	2910 / 3000 = 0,97 = 97%
Specificità = 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+ = Valore predittivo dei positivi  
V- = Valore predittivo dei negativi

**Valore predittivo dei positivi (V+) =  $p(M+/T+) =$**   
**probabilità che chi ha il test positivo sia malato**

	malati	sani	
<b>Test +</b>	<b>a</b>	<b>b</b>	<b>V+ = a/(a+b)</b>
<b>Test -</b>			

**Valore predittivo dei negativi (V-) =  $p(M-/T-) =$**   
**probabilità che chi ha il test negativo sia sano**

	malati	sani	
<b>Test +</b>			<b>V- = d/(c+d)</b>
<b>Test -</b>	<b>c</b>	<b>d</b>	

## ESEMPIO: SCREENING DEL CANCRO MAMMARIO

Negli screening del cancro mammario, effettuati in Italia nel periodo 2003-04, alla prima mammografia il 7.8% delle donne veniva sottoposto ad ulteriori accertamenti e la percentuale di cancri scoperti sull'intero campione risultava pari allo 0.65% [Giorgi et al, 2006].

Pertanto il valore predittivo positivo della mammografia risultava pari a  $0.65\% / 7.8\% = 0.083$ , ovvero tra le donne che si sottoponevano a procedure invasive 1 su 12 avevano effettivamente un tumore maligno. Negli screening di popolazione il valore predittivo dei positivi è sempre piuttosto basso.

Nessuno contesta che, per trovare un tumore maligno in fase iniziale, 11 donne vengano sottoposte inutilmente ad accertamenti invasivi, anche se *"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.

### Un altro modo per valutare il test diagnostico, utilizzato soprattutto in ambito clinico

#### Rapporto di verosimiglianza per esito positivo (positive likelihood ratio, LR+)

Rapporto tra la proporzione di test positivi nei malati e la proporzione di test positivi nei sani:

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

#### Rapporto di verosimiglianza per esito negativo (negative likelihood ratio, LR-)

Rapporto tra la proporzione di test negativi nei malati e la proporzione di test negativi nei sani:

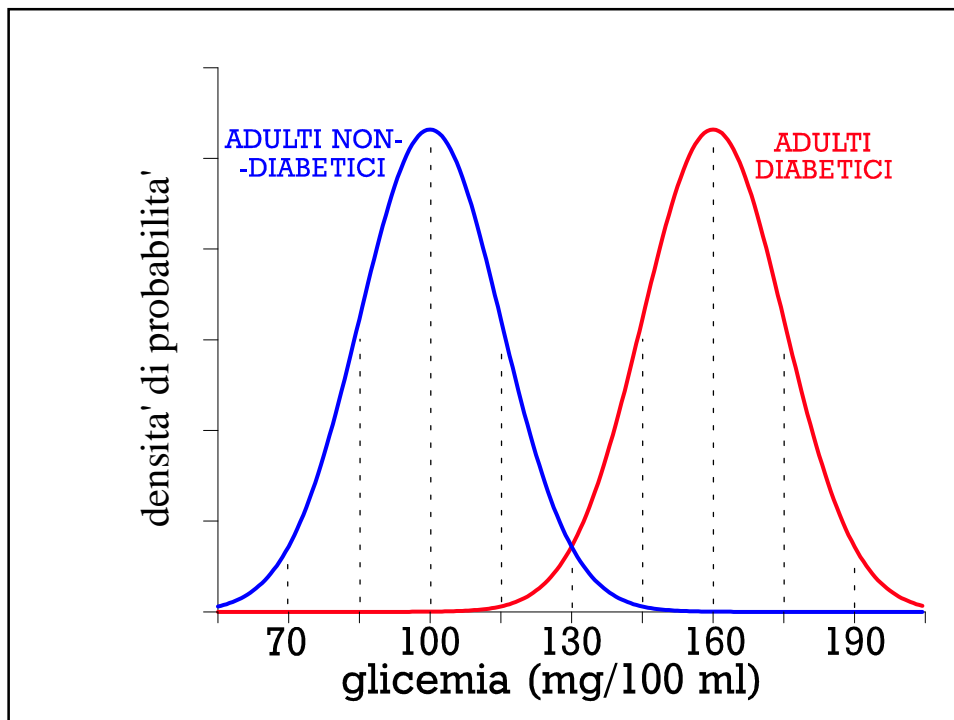
$$LR - = \frac{P(T-/M+)}{P(T-/M-)} = \frac{1-\text{sensibilità}}{\text{specificità}}$$

## Cut-off per LR+ ed LR-

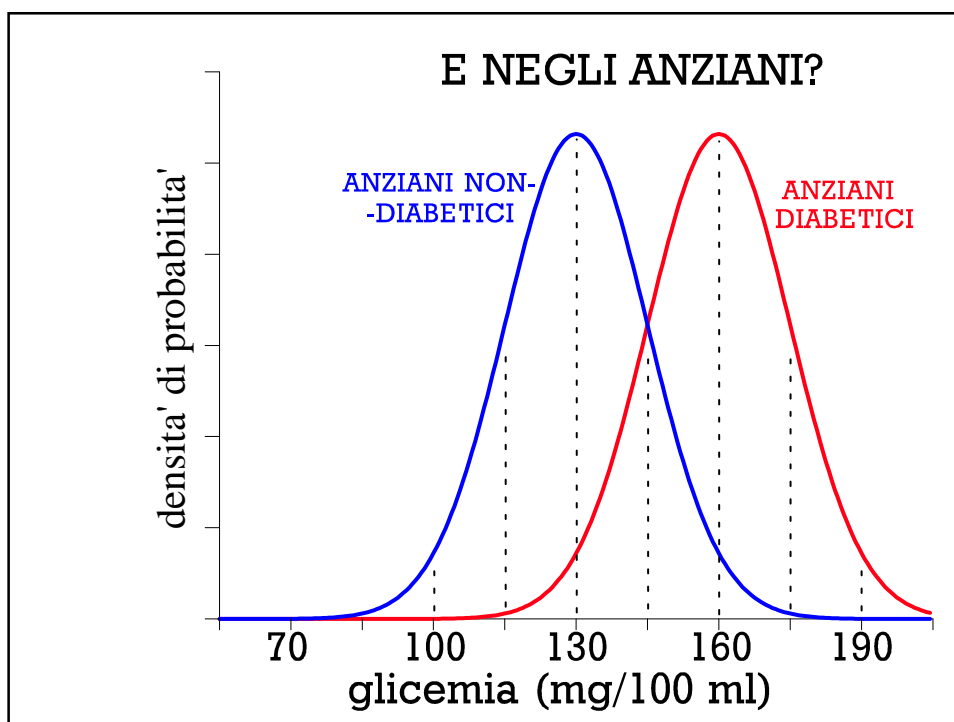
Quando LR+ è maggiore di 5, un test positivo consente di confermare la presenza della malattia con un buon livello di confidenza

Quando LR- è minore di 0,2, un test negativo consente di escludere la malattia con un buon livello di confidenza

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.



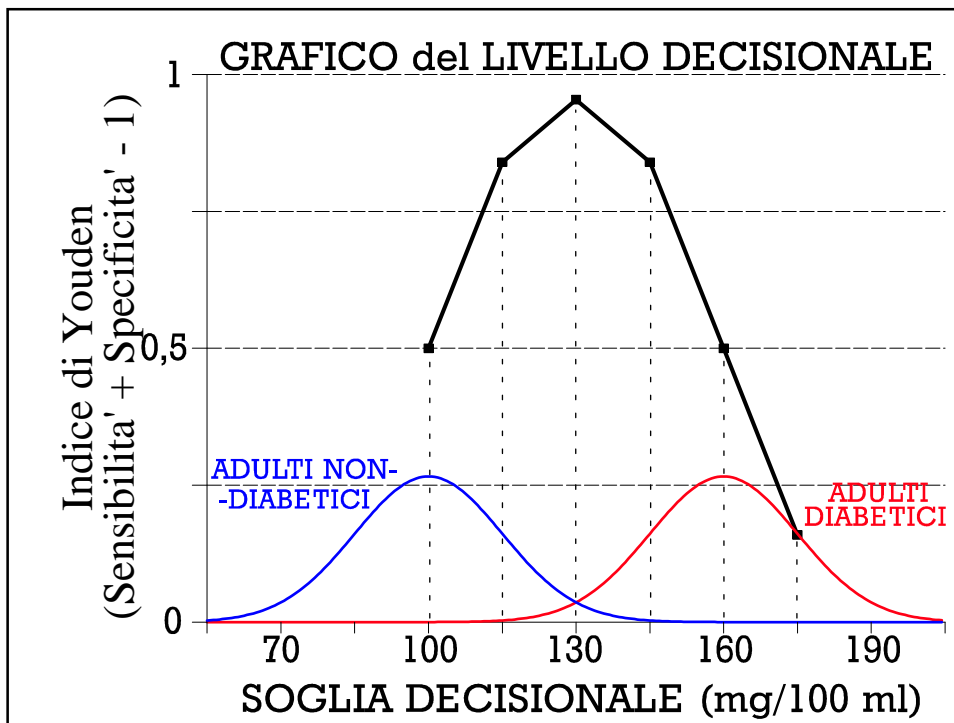
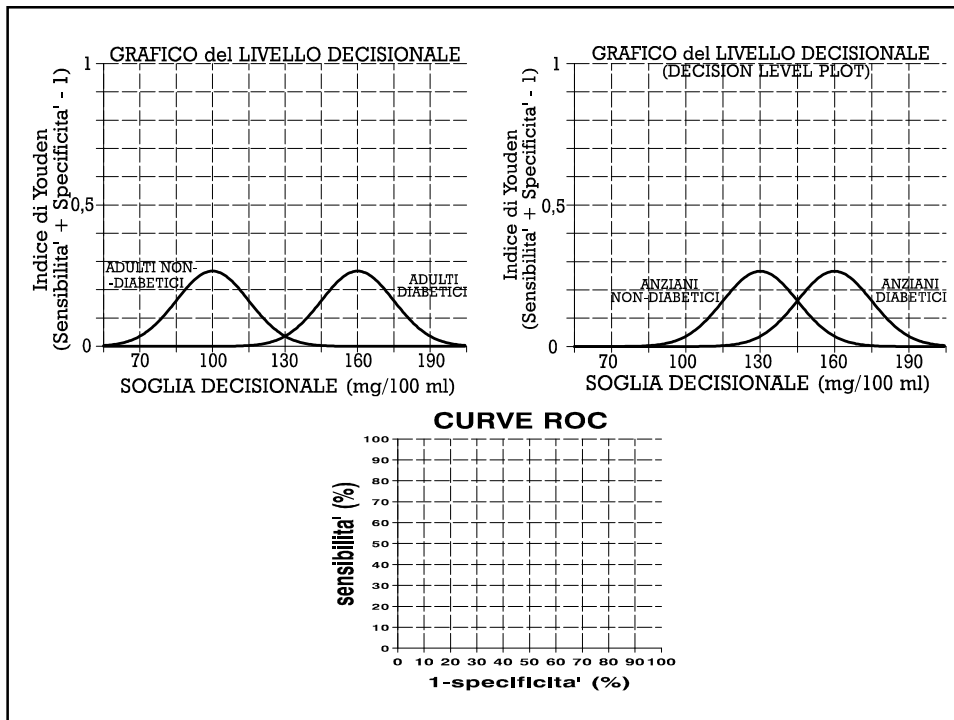


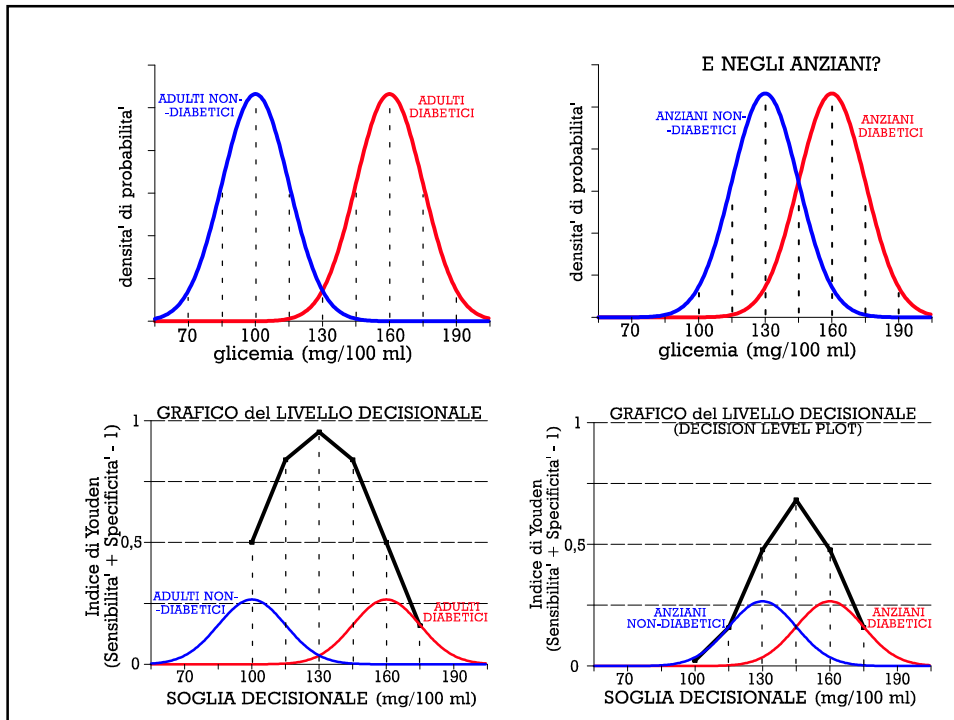
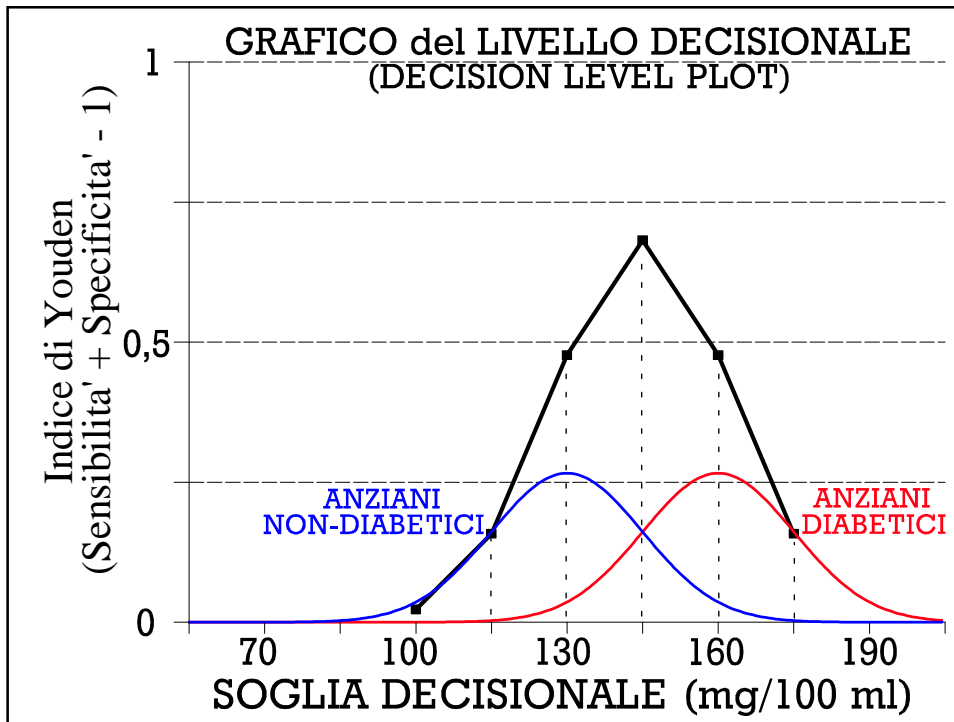


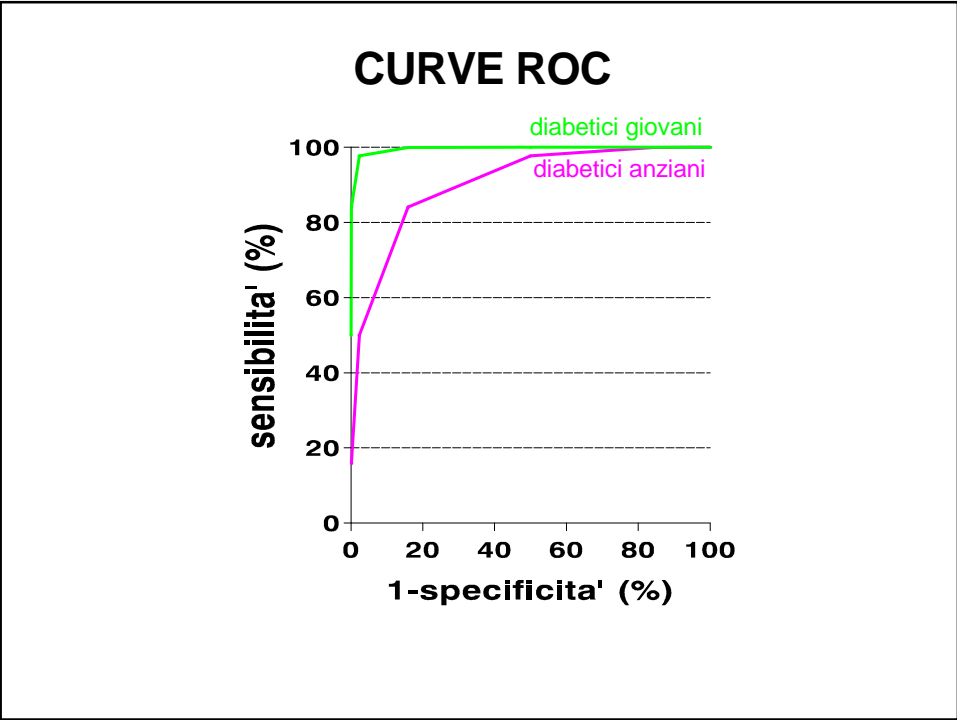
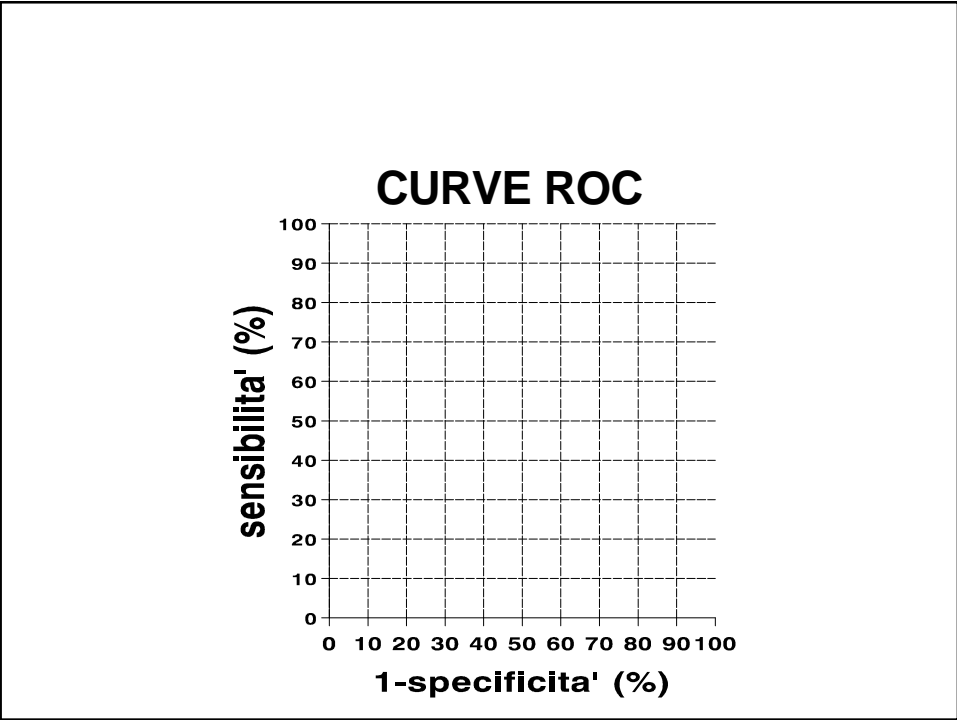
I ESEMPIO (PAZIENTI DIABETICI)			LIVELLO DECISIONALE	II ESEMPIO (PAZIENTI DIABETICI ANZIANI)		
specificità	1-specificità	sensibilità		specificità	1-specificità	sensibilità
50.0 %	50.0 %	99.997 %	100 mg/dl	2.3 %	97.7 %	99.997 %
84.1 %	15.9 %	99.9 %	115 mg/dl	15.9 %	84.1 %	99.9 %
97.7 %	2.3 %	97.7 %	130 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 %

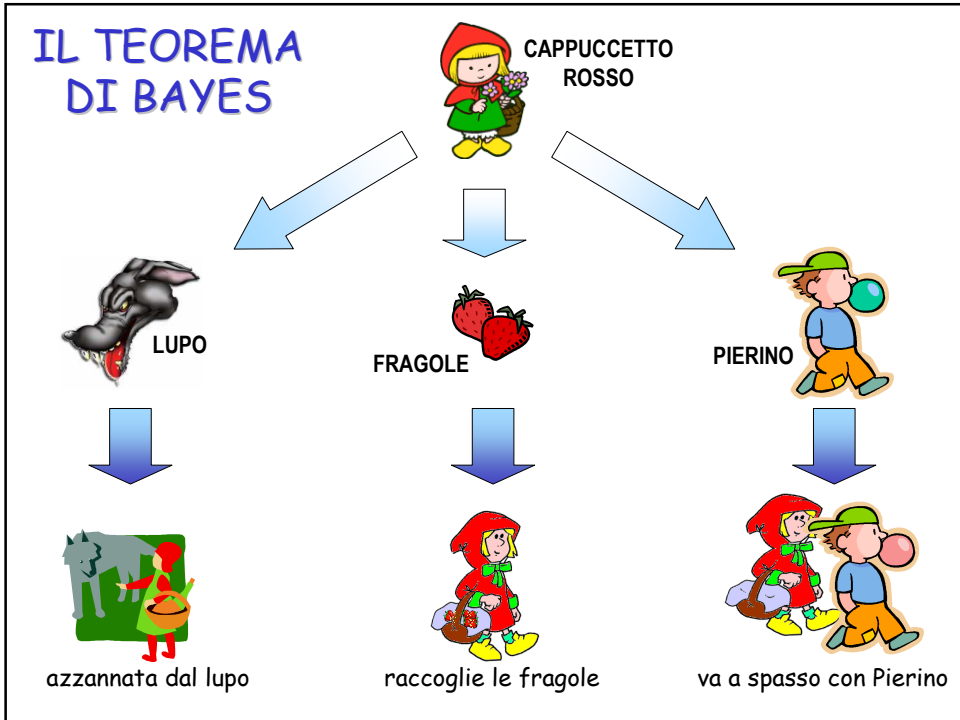
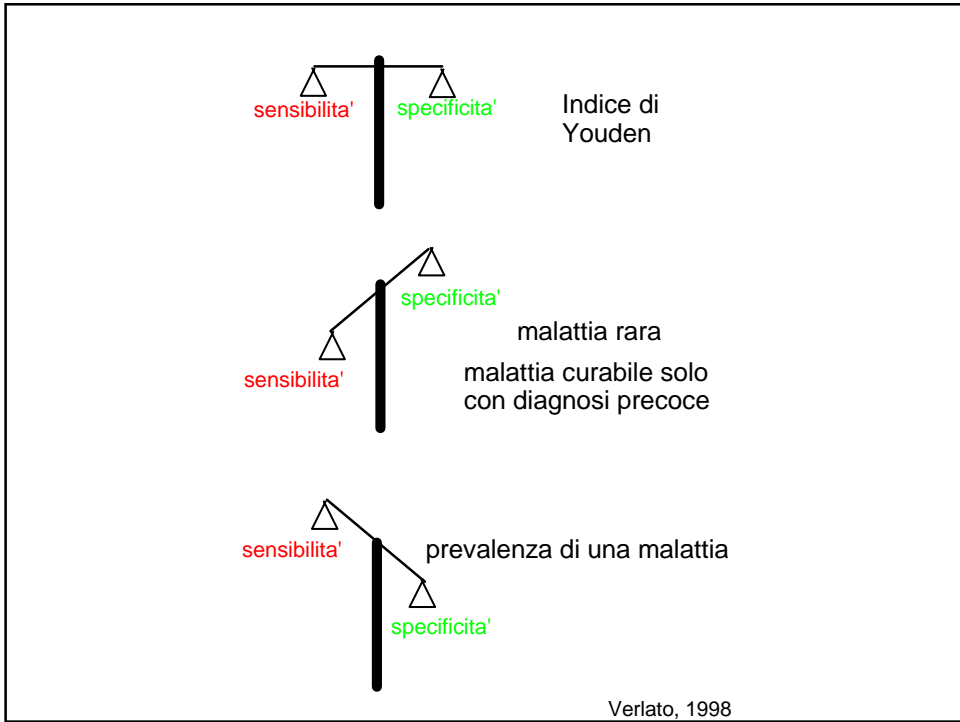
utilizzati per rappresentare il  
LIVELLO DECISIONALE e  
le CURVE ROC

utilizzati per rappresentare il  
LIVELLO DECISIONALE e  
le CURVE ROC





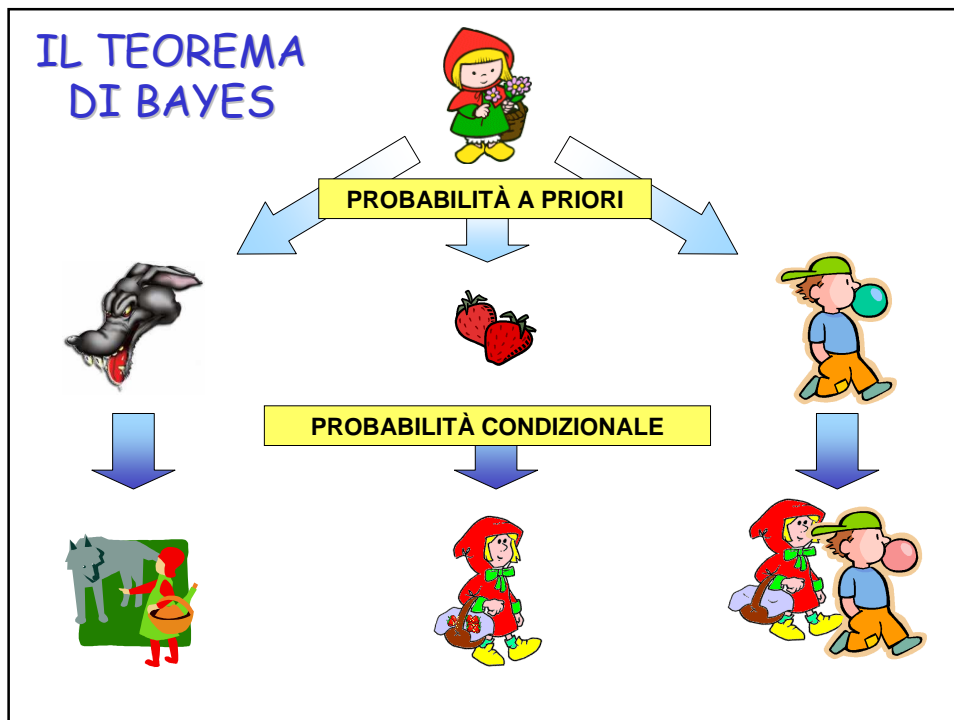




## Assunzioni:

1.  $P(\text{Lupo} \cap \text{Fragole}) = P(\text{Lupo} \cap \text{Pierino}) = P(\text{Pierino} \cap \text{Fragole}) = 0$

2.  $P(\text{azzannata dal lupo}) + P(\text{raccolta fragole}) + P(\text{a spasso con Pier.}) = 1$



## Teorema di Bayes

(Thomas Bayes 1702 - 1761)

Si conosce il **risultato** di un *esperimento* e si vuole conoscere la probabilità che esso sia dovuto ad una determinata '**causa**'

In medicina raramente esiste una relazione univoca tra **causa** (patologia) ed **effetti** (sintomi - *outcome*)

Spesso si hanno difficoltà ad individuare tutte le *possibili cause* che contribuiscono a 'determinare' un *outcome* e quindi le probabilità *a priori* ad esse associate

## Formula di Bayes

Permette di formalizzare i singoli passaggi del *procedimento diagnostico* e di modificare l'attribuzione iniziale di probabilità a delle **cause** (malattie) sulla base delle **conoscenze acquisite** (sintomi)

$$P(H_i | A) = \frac{P(H_i) \cdot P(A | H_i)}{\sum_{i=1}^k P(H_i) \cdot P(A | H_i)}$$

*dove:*

$H_1 \dots H_i \dots H_k \Rightarrow$  *possibili cause del sintomo in questione*

$A \Rightarrow$  *sintomo/malessere in studio*

$P(H_i)$  è nota come la **probabilità a priori** della causa  $H_i$ : può essere intesa come la **probabilità** che un medico può attribuire ad una *determinata patologia*, PRIMA di guardare il malato, sulla base della diffusione della malattia (incidenza/prevalenza)

$P(A|H_i)$  è la **probabilità condizionale**: la probabilità del sintomo data la malattia  $H_i$ . Il sintomo è patognomonico della malattia?

$P(H_i|A)$  è la **probabilità a posteriori**: misura la probabilità che l'evento  $A$ , già realizzatosi, sia dovuto alla causa  $H_i$ , tra una serie finita  $k$  di cause possibili (i.e. rappresenta la probabilità che il medico attribuisce alla malattia, dopo aver acquisito le informazioni sui sintomi del paziente)

Calcolo delle probabilità a posteriori  $\Rightarrow$  DIAGNOSTICA DIFFERENZIALE

### Applicazione del teorema di Bayes Caso clinico: Ematuria in maschio di 25 anni

	Calcolo	Glomerulonefrite	Neoplasia	Totale
probabilità a priori – $p(M)$	0,1%	0,5%	0,01%	-----
prob. condizionale - $p(S/M)$	50%	80%	60%	-----
prodotto di probabilità	5 / 10000	40 / 10000	0,6 / 10000	45,6/100000
prob. a posteriori – $p(M/S)$	5 / 45,6 11,0%	40 / 45,6 87,7%	0,6 / 45,6 1,3%	45,6 / 45,6 100%

$$p(M_1/S) = \frac{p(M_1) * p(S/M_1)}{p(M_1) * p(S/M_1) + p(M_2) * p(S/M_2) + p(M_3) * p(S/M_3)}$$

**Assunzioni**

- 1) L'ematuria può avere solo 3 cause (calcolo, glomerulonefrite, neoplasia)
- 2) Le 3 cause sono mutuamente esclusive



### ***Esempio illustrativo***

*Un ginecologo riscontra anomalie all'utero di una paziente:*

- displasia al collo dell'utero ( $H_1$ )
  - stadio iniziale di tumore al collo dell'utero ( $H_2$ )
- } **Possibili cause**

*Dalla letteratura medica e per esperienza il medico sa che:*

$$P(H_1) = 0.85, P(H_2) = 0.15$$

*Il medico chiede se la paziente ha avuto perdite continue (**A**) nelle ultime settimane. Si sa che:*

$$P(A | H_1) = 0.10$$

$$P(A | H_2) = 0.80$$

$$P(H_1 | A) = \frac{0.85 \cdot 0.10}{0.85 \cdot 0.10 + 0.15 \cdot 0.80} = 0.40$$

$$P(H_2 | A) = \frac{0.15 \cdot 0.80}{0.85 \cdot 0.10 + 0.15 \cdot 0.80} = 0.60$$

1/31/2013

### **APPLICAZIONE del TEOREMA di BAYES CASO CLINICO: Emoftoe in maschio di 40 anni**

	TBC	Cancro polmonare	Bronco-polmonite	Totale
Prob. a priori Prevalenza= $p(M+)$	0.01 %	0.1 %	1.0 %	---
Prob. condizionale $p(S+/M+)$	80.0 %	40.0 %	2.0 %	---
Prodotto delle prob.				
Prob. a posteriori $p(M+/S+)$				
Prob. a posteriori $p(M+/S+)$				

ASSUNZIONI

## APPLICAZIONE del TEOREMA di BAYES CASO CLINICO: Emoftoe in maschio di 40 anni

	TBC	Cancro polmonare	Bronco-polmonite	Totale
Prob. a priori Prevalenza= $p(M+)$	0.01 %	0.1 %	1.0 %	---
Prob. condizionale $p(S+/M+)$	80.0 %	40.0 %	2.0 %	---
Prodotto delle prob.	0.8/10000	4.0/10000	2.0/10000	6.8/10000
Prob. a posteriori $p(M+/S+)$	0.8/6.8	4.0/6.8	2.0/6.8	6.8/6.8
Prob. a posteriori $p(M+/S+)$	11.8 %	58.8 %	29.4 %	100.0 %

### ASSUNZIONI

- 1) L'emoftoe puo' avere solo 3 cause (TBC, Ca polmonare, Broncopolmonite)
- 2) Le 3 cause sono mutuamente esclusive

## Perdite di sangue dal corpo umano

### **Sangue dall'apparato digerente:**

ematemesi = sangue vomitato

melena = sangue digerito (nero) nelle feci

rettorragia = sangue non digerito (rosso vivo) nelle feci

### **Sangue dall'albero respiratorio:**

emoftoe o emottisi

### **Sangue nell'urina:**

ematuria

### **Sangue dalle vie genitali femminili:**

menorragia = perdita mestruale abbondante

metrorragia = perdita non nel periodo delle mestruazioni