
Introduzione agli studi di accuratezza diagnostica

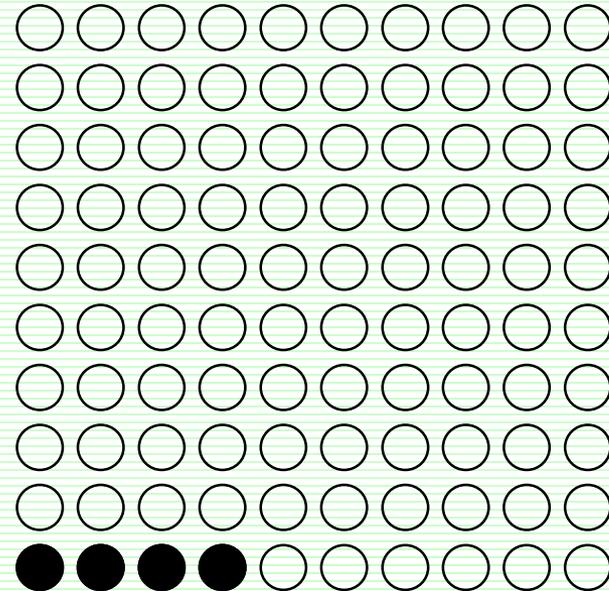
Di cosa ci occuperemo...

- **Le principali misure di accuratezza**
- **Gli studi di accuratezza: come valutarli?**

Misure di accuratezza: frequenze

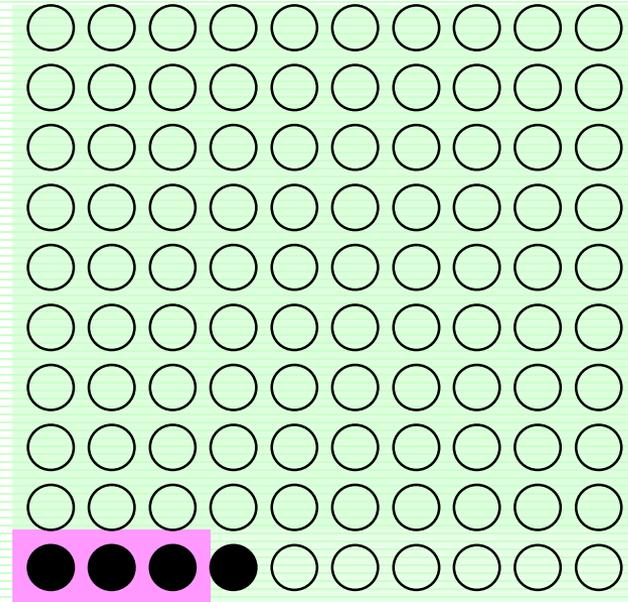
Persona senza la malattia	○
Persona con la malattia	●

Assumiamo che la prevalenza di malattia sia del 4%



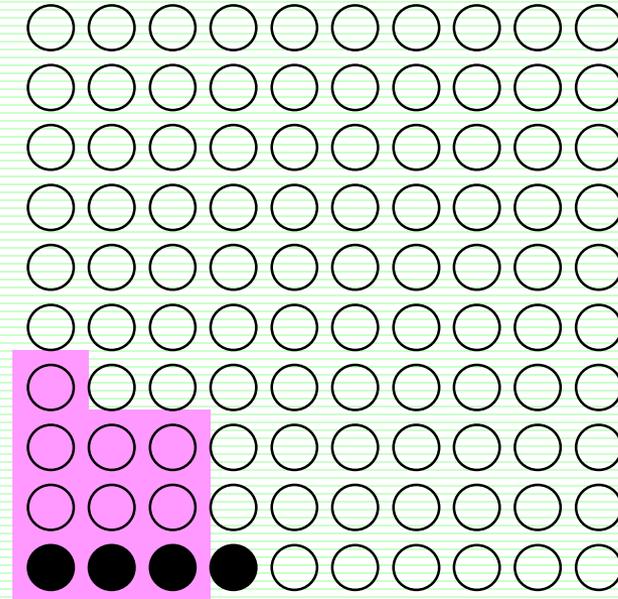
Assumiamo che dei 4 malati, 3 vengano individuati dal test

Persona senza la malattia	○
Persona con la malattia	●
Persona con test positivo	■
Persona con test negativo	□



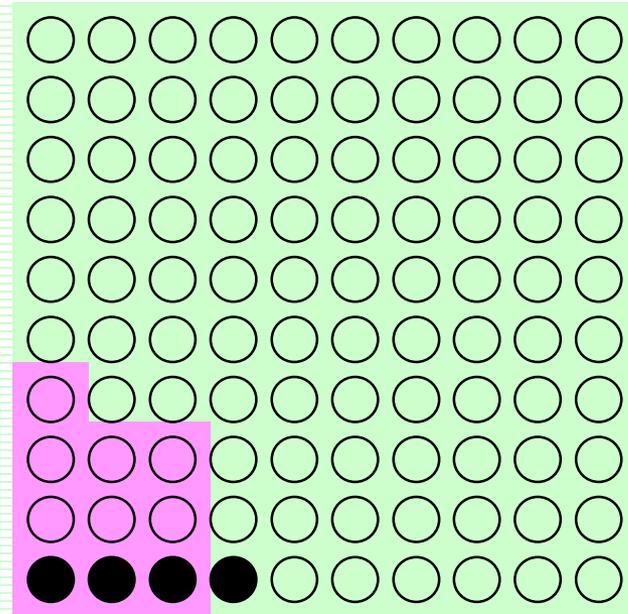
Assumiamo che il test sia positivo per altre 7 persone che in realtà sono sane

Persona senza la malattia	○
Persona con la malattia	●
Persona con test positivo	■
Persona con test negativo	□
Vero positivo al test	●■
Falso positivo al test	○■



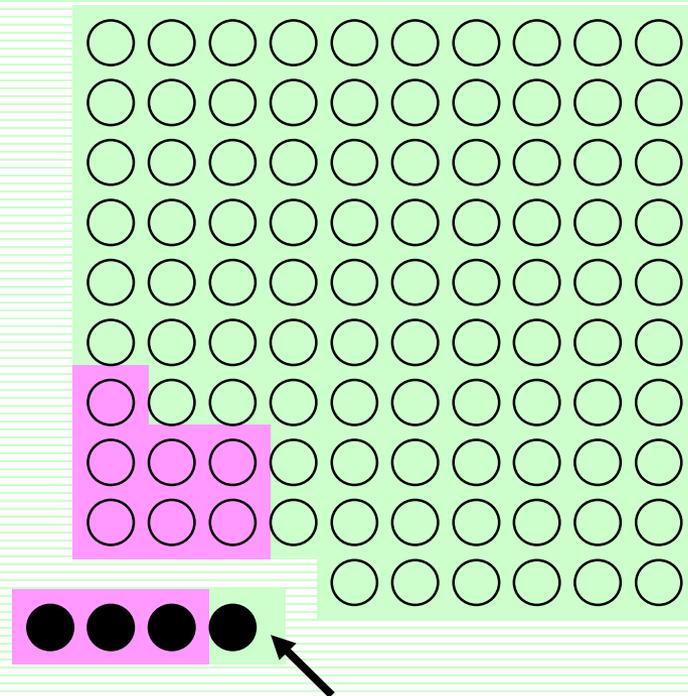
Il resto del campione è negativo al test

Persona senza la malattia	○
Persona con la malattia	●
Persona con test positivo	■
Persona con test negativo	□
Vero positivo al test	●■
Falso positivo al test	○■
Vero negativo al test	○□
Falso negativo al test	●□



Sensibilità

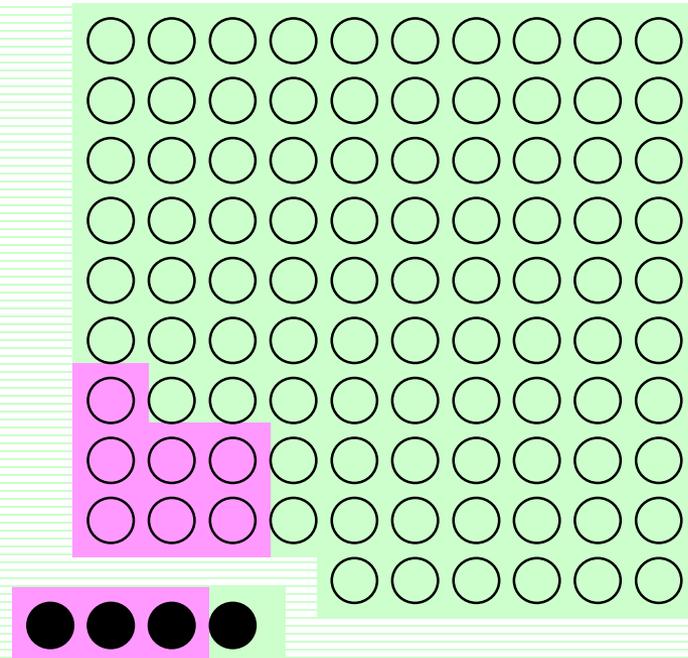
Persona senza la malattia	○
Persona con la malattia	●
Persona con test positivo	■
Persona con test negativo	□
Vero positivo al test	●
Falso positivo al test	○
Vero negativo al test	○
Falso negativo al test	●



- La Sensibilità è la proporzione di persone con la malattia correttamente identificate dal test
- $(1-Se)$ proporzione di falsi negativi

Sensibilità

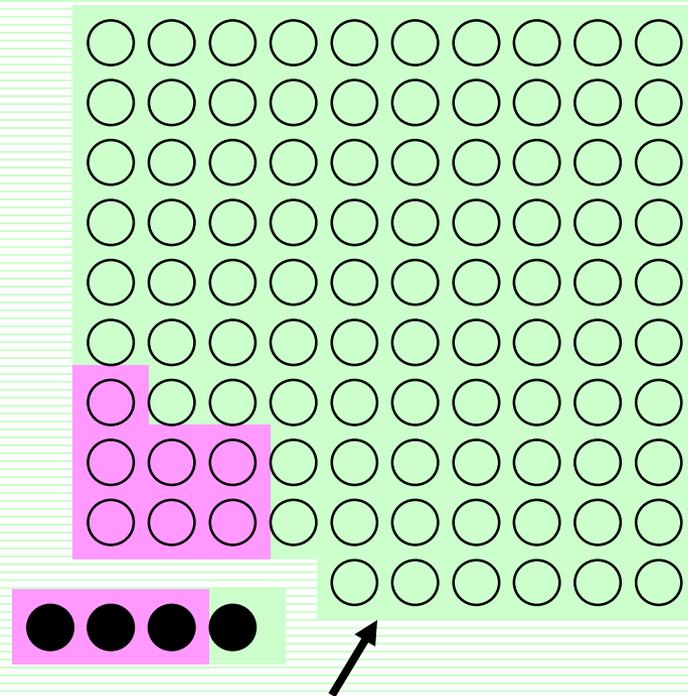
Persona senza la malattia	○
Persona con la malattia	●
Persona con test positivo	■
Persona con test negativo	□
Vero positivo al test	●■
Falso positivo al test	○■
Vero negativo al test	○□
Falso negativo al test	●□



In questo caso la sensibilità è del 75%

Specificità

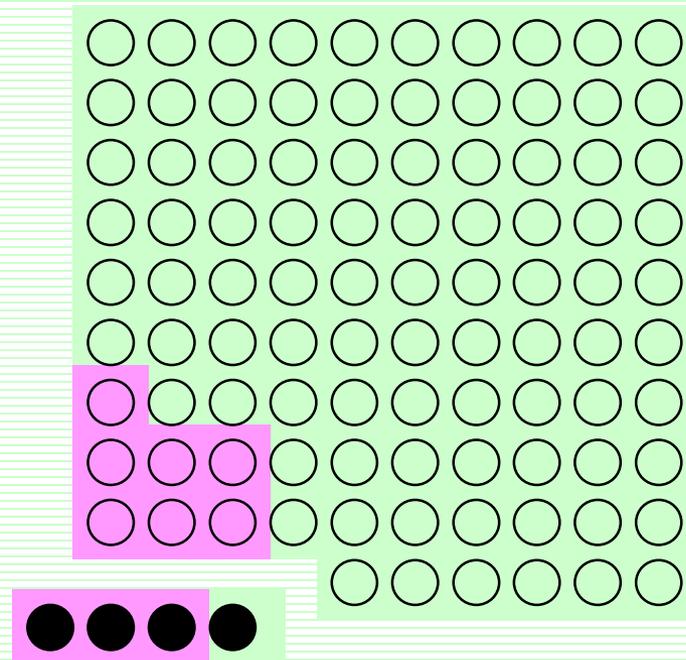
Persona senza la malattia	○
Persona con la malattia	●
Persona con test positivo	■
Persona con test negativo	□
Vero positivo al test	●■
Falso positivo al test	○■
Vero negativo al test	○□
Falso negativo al test	●□



- **Specificità è la proporzione di persone senza la malattia correttamente identificate dal test**
- **(1-Sp) proporzione di falsi positivi**

Specificità

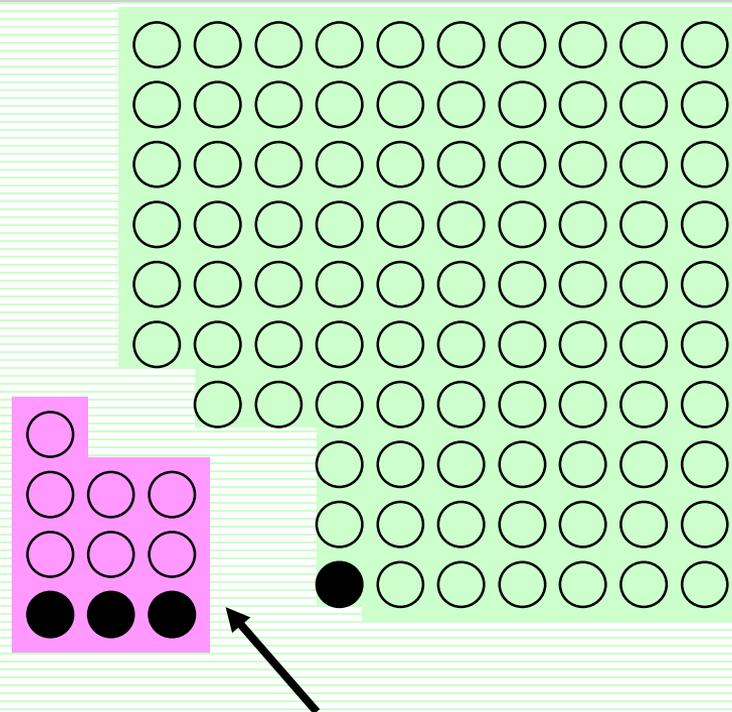
Persona senza la malattia	○
Persona con la malattia	●
Persona con test positivo	■
Persona con test negativo	□
Vero positivo al test	●
Falso positivo al test	○
Vero negativo al test	○
Falso negativo al test	●



In questo caso la specificità è pari a $(96-7)/96$ o 93%

Se qualcuno è positivo al test, che probabilità ha di avere effettivamente la malattia?

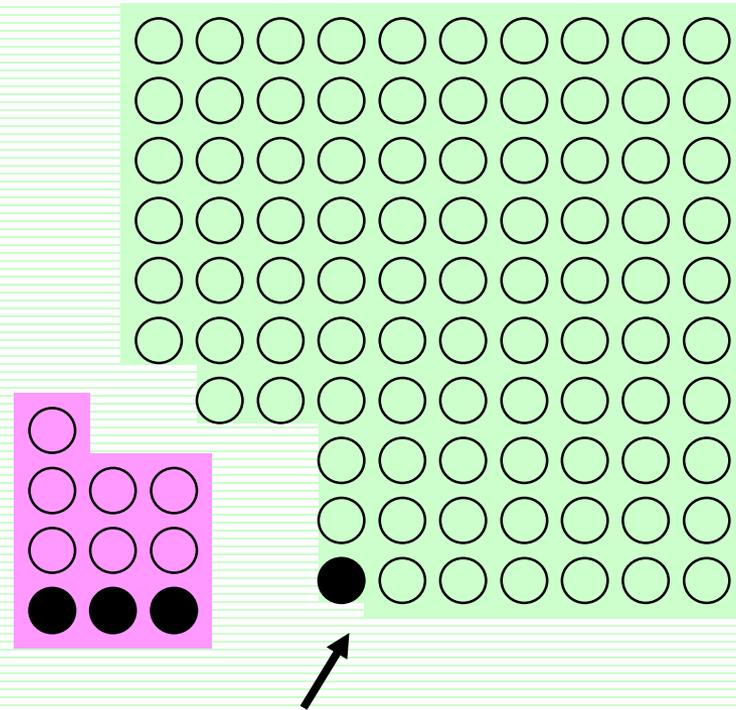
Persona senza la malattia	○
Persona con la malattia	●
Persona con test positivo	■
Persona con test negativo	□
Vero positivo al test	●■
Falso positivo al test	○■
Vero negativo al test	○□
Falso negativo al test	●□



- **Probabilità = $3/10 = 30\%$**
- **Questo è il valore predittivo positivo (il valore del test nel predire un risultato positivo)**

Se qualcuno è negativo al test, che probabilità ha di non avere la malattia?

Persona senza la malattia	○
Persona con la malattia	●
Persona con test positivo	■
Persona con test negativo	□
Vero positivo al test	●■
Falso positivo al test	○■
Vero negativo al test	○□
Falso negativo al test	●□



- **Probabilità = $89/90 = 99\%$**
- **Questo è il valore predittivo negativo (il valore del test nel predire un risultato negativo)**

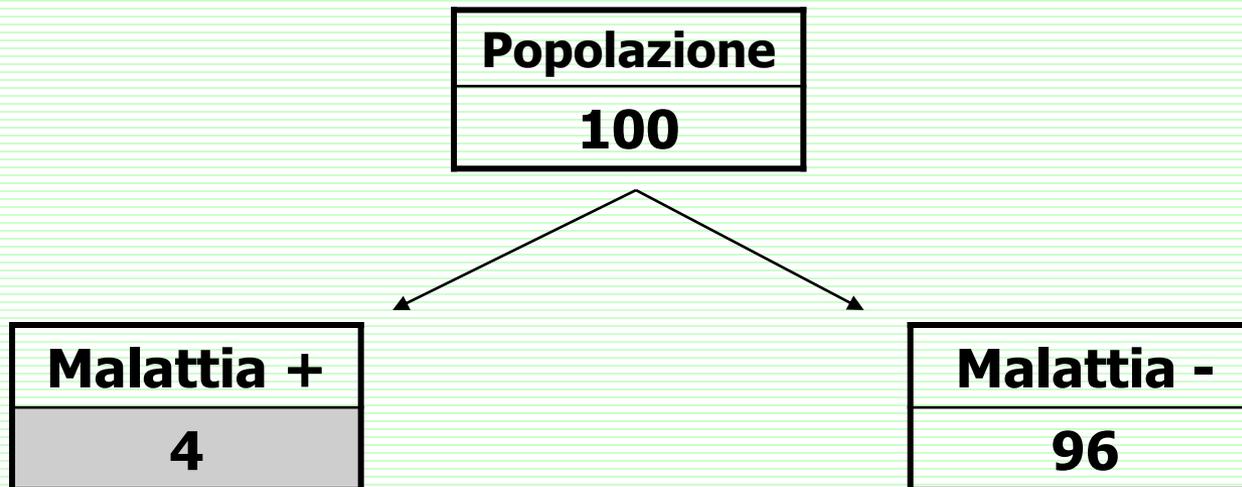
Sensibilità, Specificità e Valori Predittivi

- **Per sensibilità e specificità, il “denominatore” è la malattia**
- **Per i valori predittivi, il “denominatore” è il test**

Metodo 2: Albero

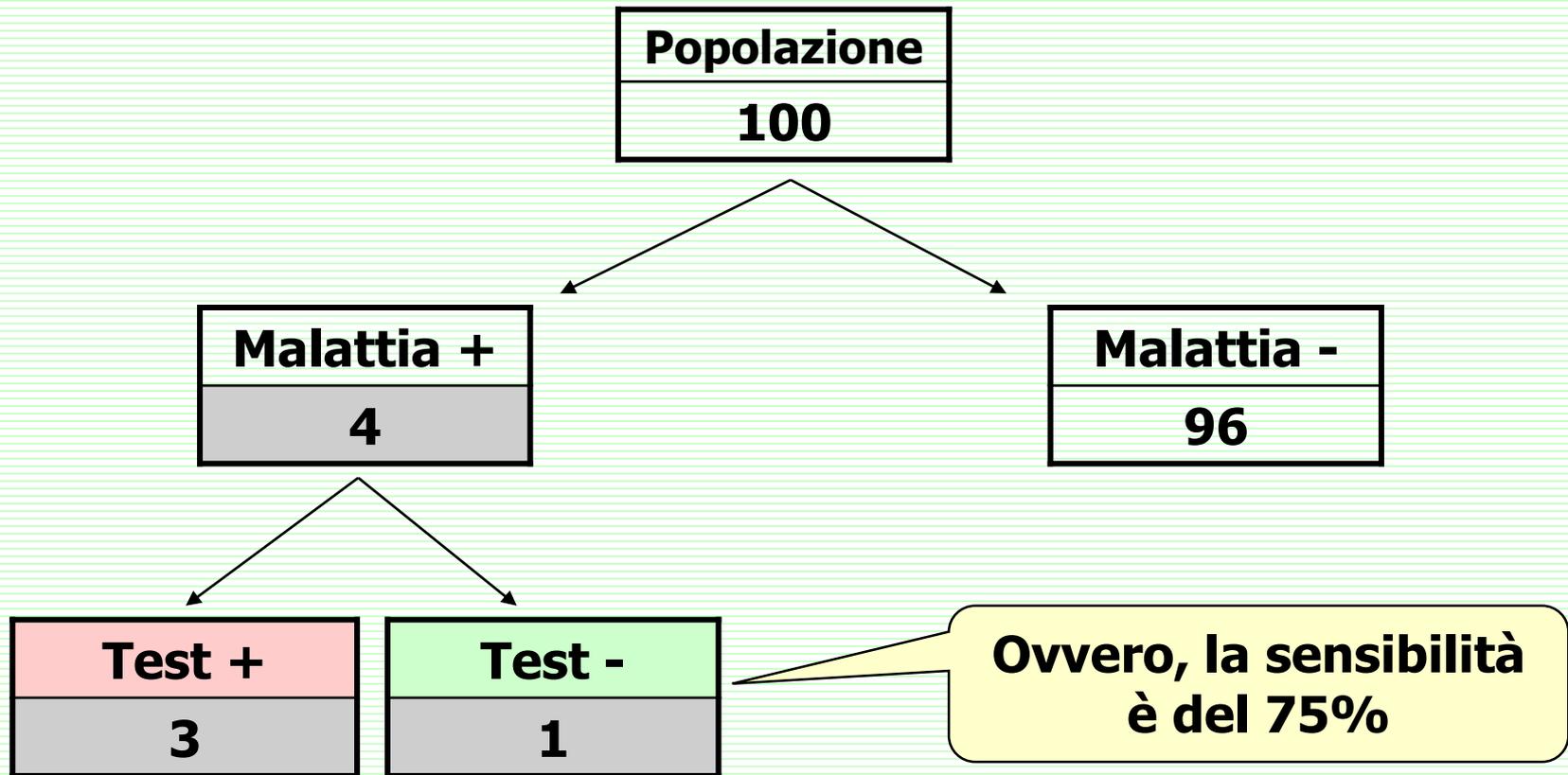
Popolazione
100

Per ogni 100 persone, 4 saranno malati

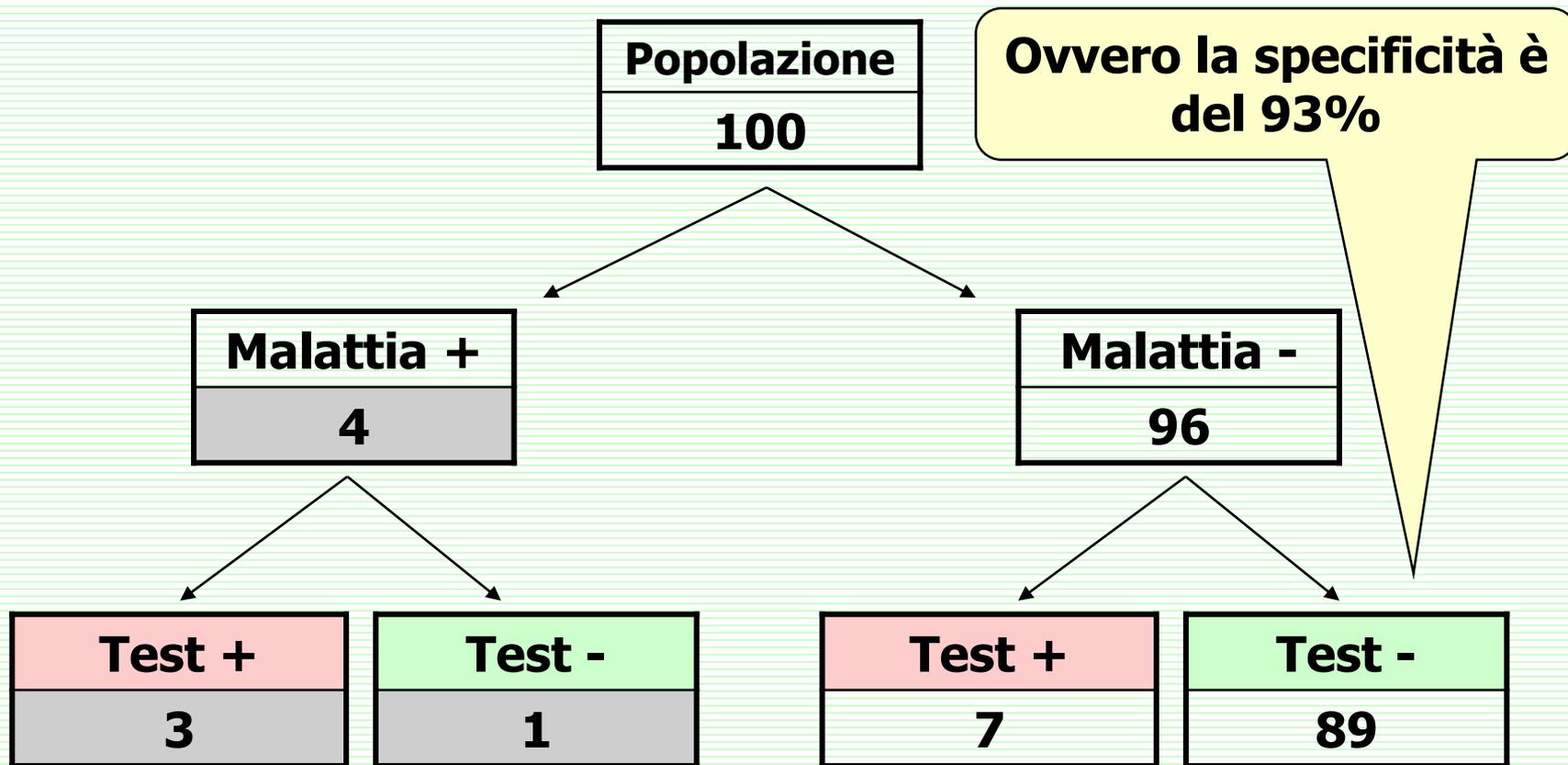


Se queste 100 persone sono rappresentative della popolazione a rischio, la proporzione di quelle con la malattia (4%) rappresenta la prevalenza della malattia e può essere considerata la probabilità pre-test di essere malato

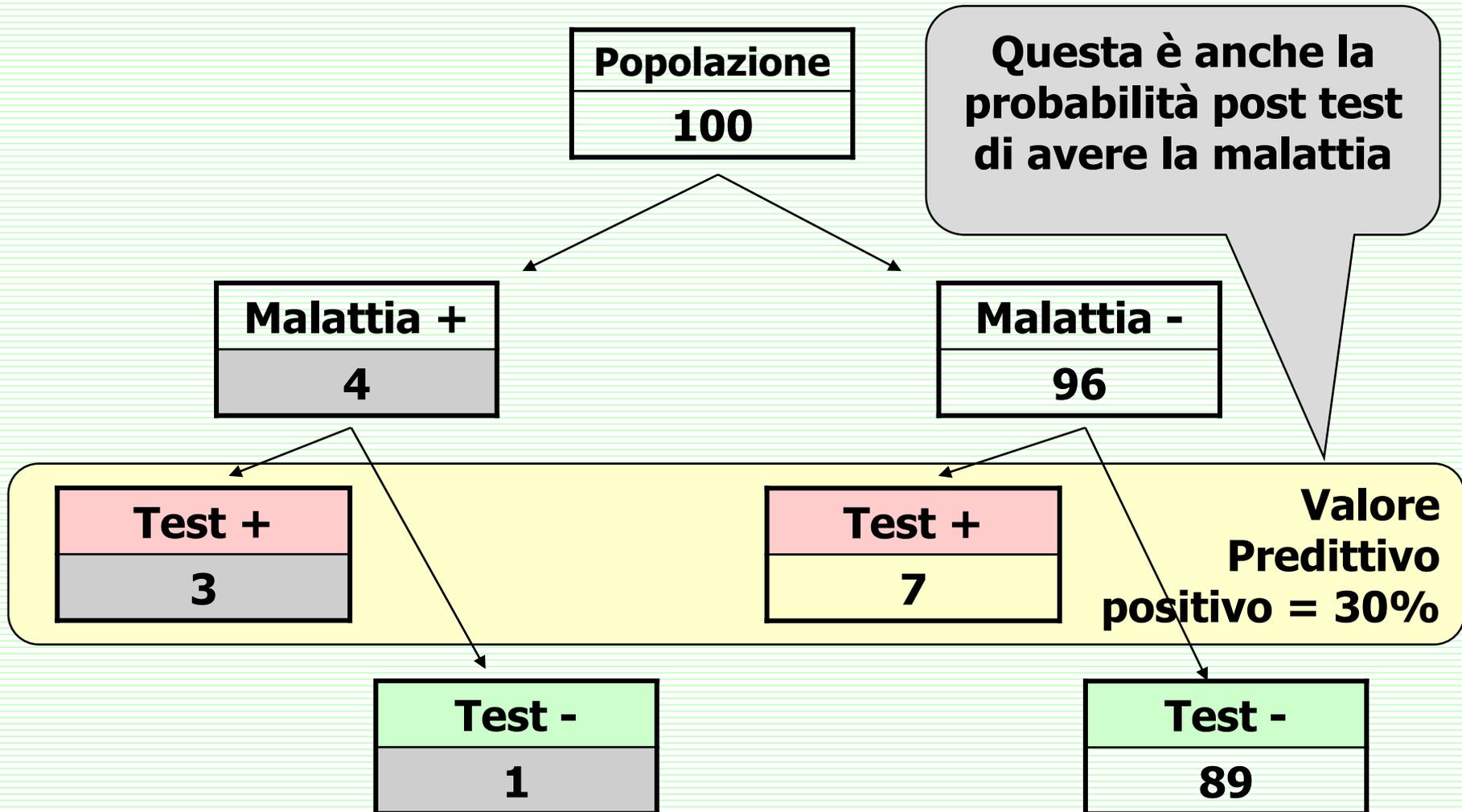
Delle 4 persone con la malattia, il test ne individuerà 3



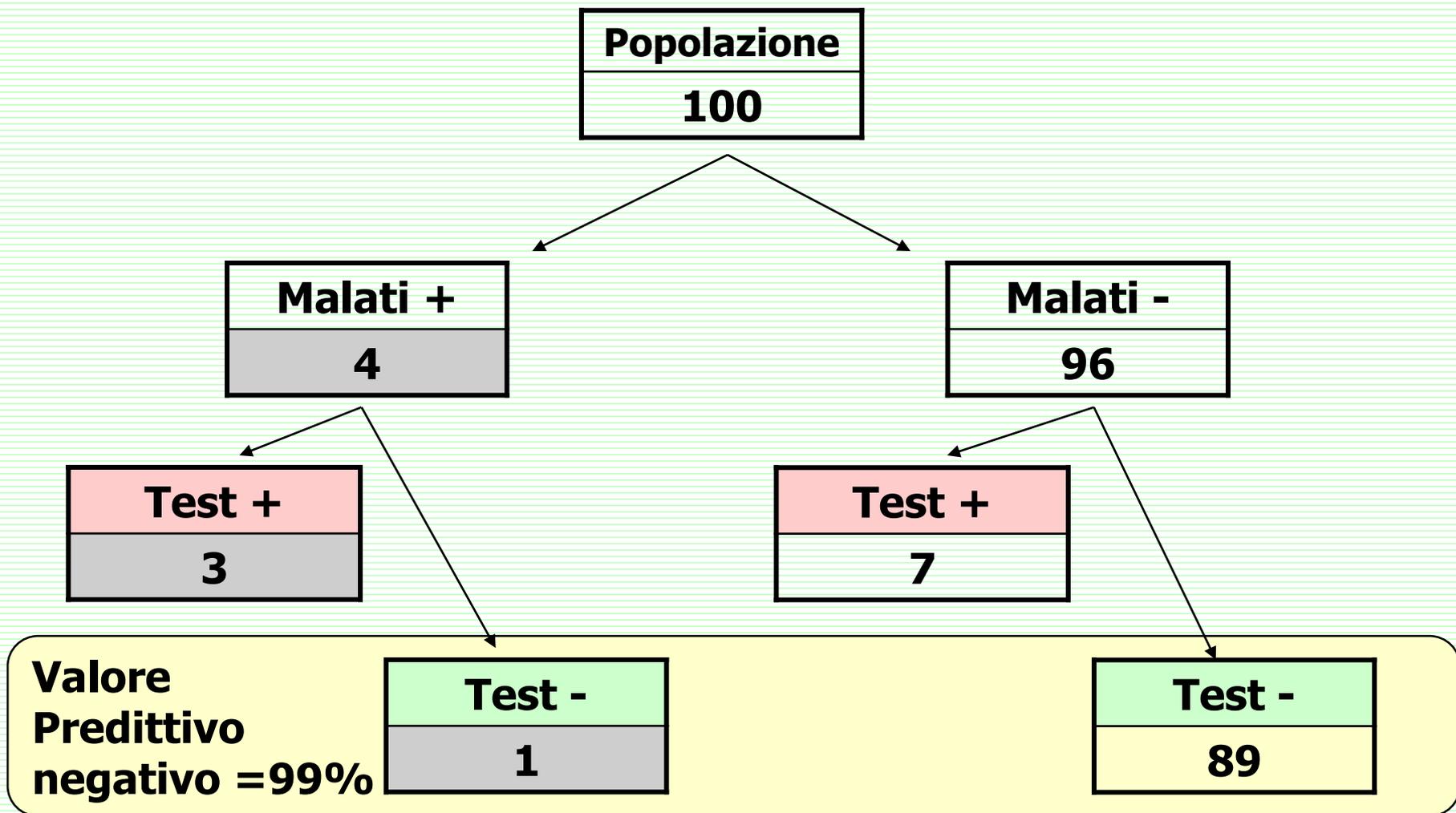
Tra i 96 sani, 7 saranno comunque positivi



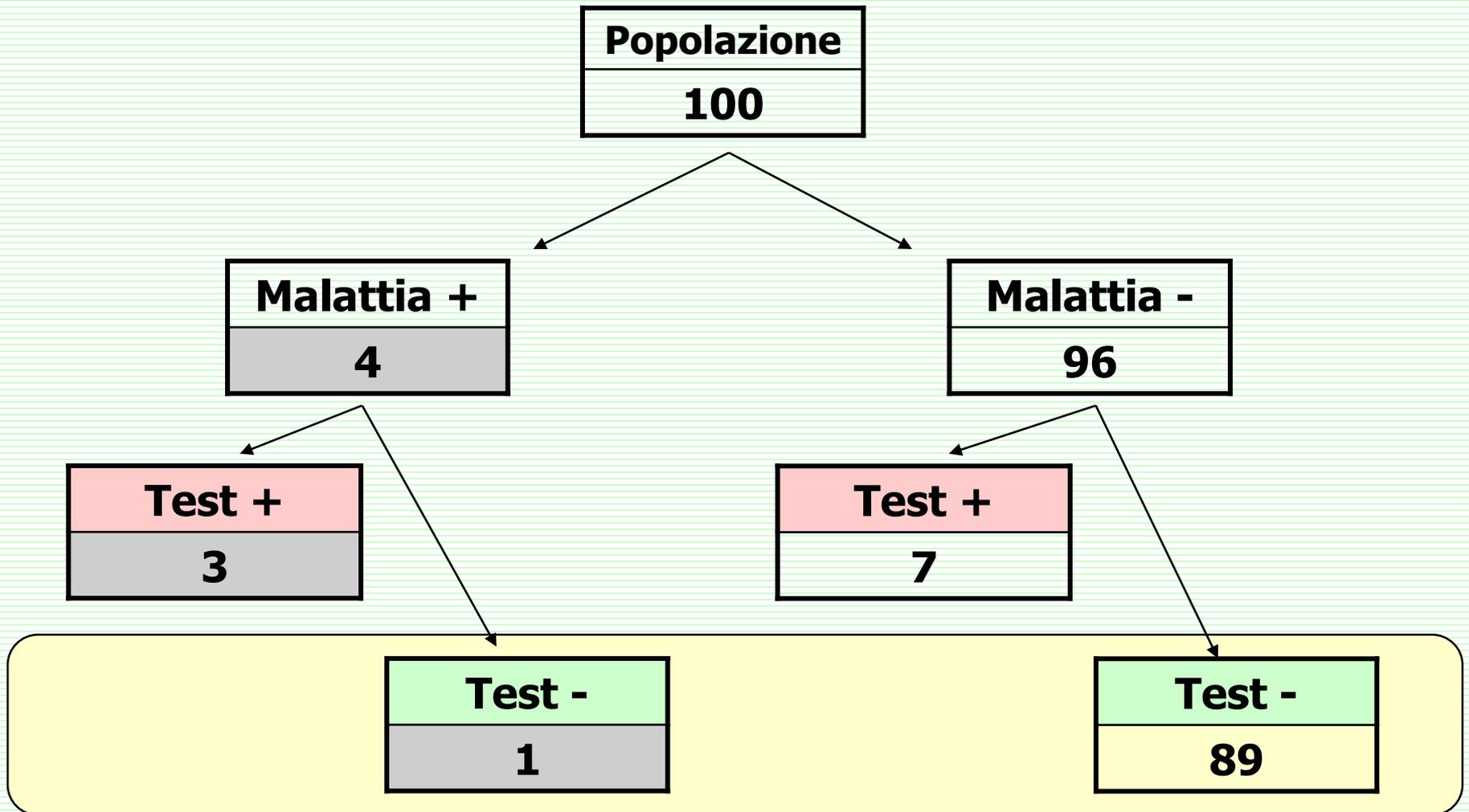
Tra quelli con il test positivo, 3 su 10 sono davvero malati



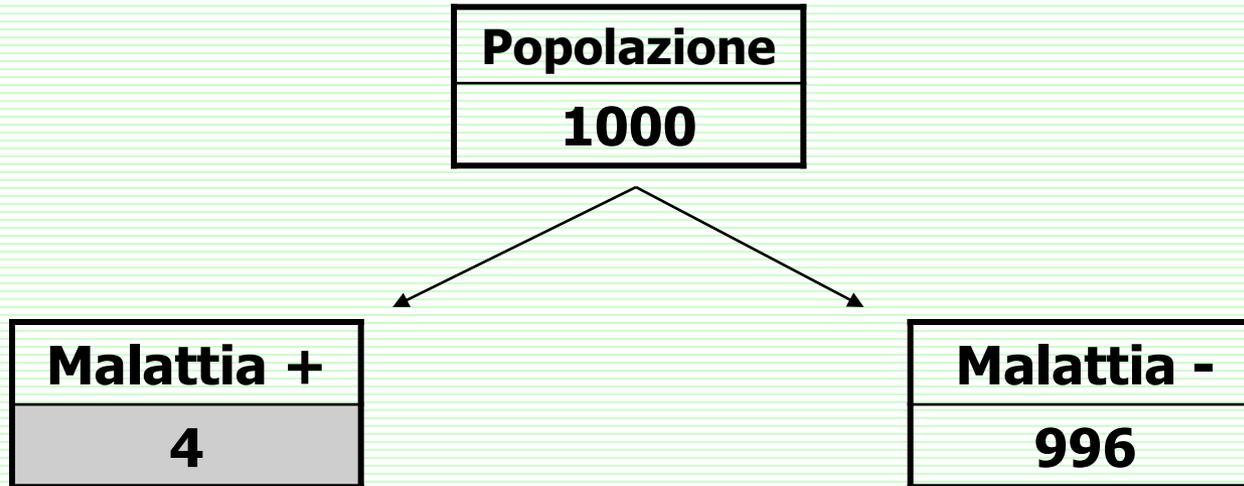
Tra quelli con il test negativo, 89 su 90 sono sani



Invece, sempre per quelli col test negativo, la probabilità di avere la malattia è solo 1 su 90

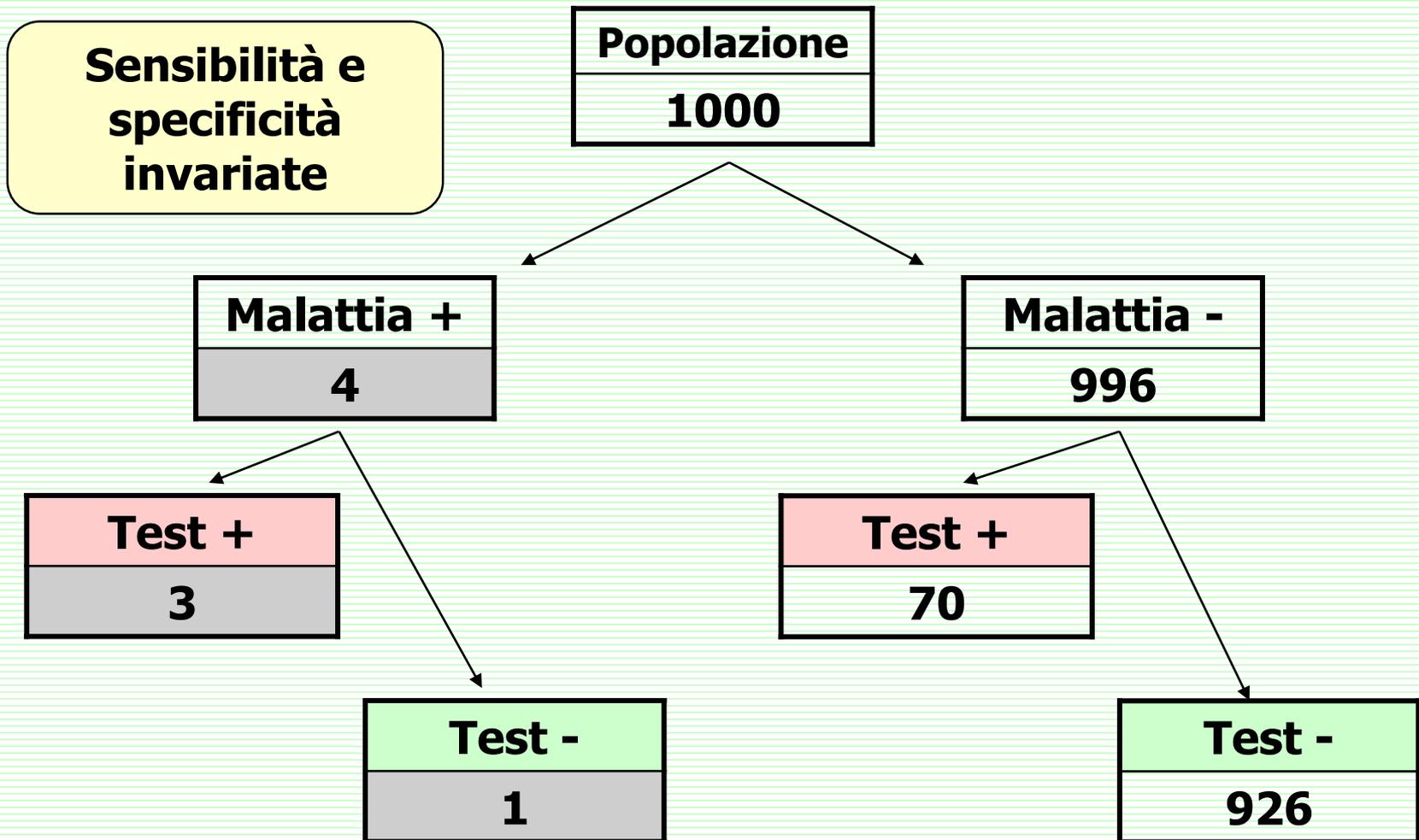


Valori predittivi e cambiamenti nella prevalenza

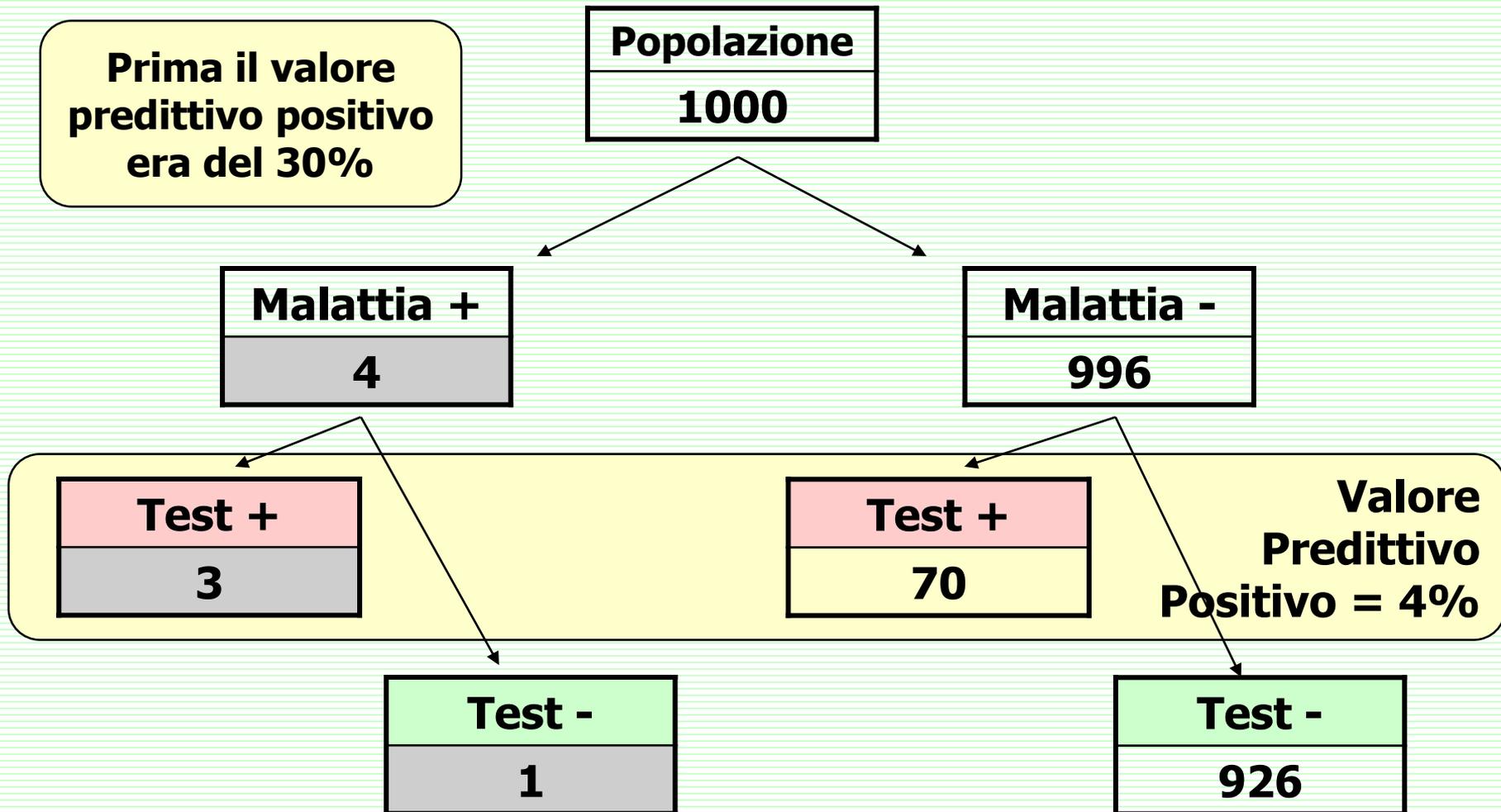


Prevalenza ridotta dal 4% allo 0.4%

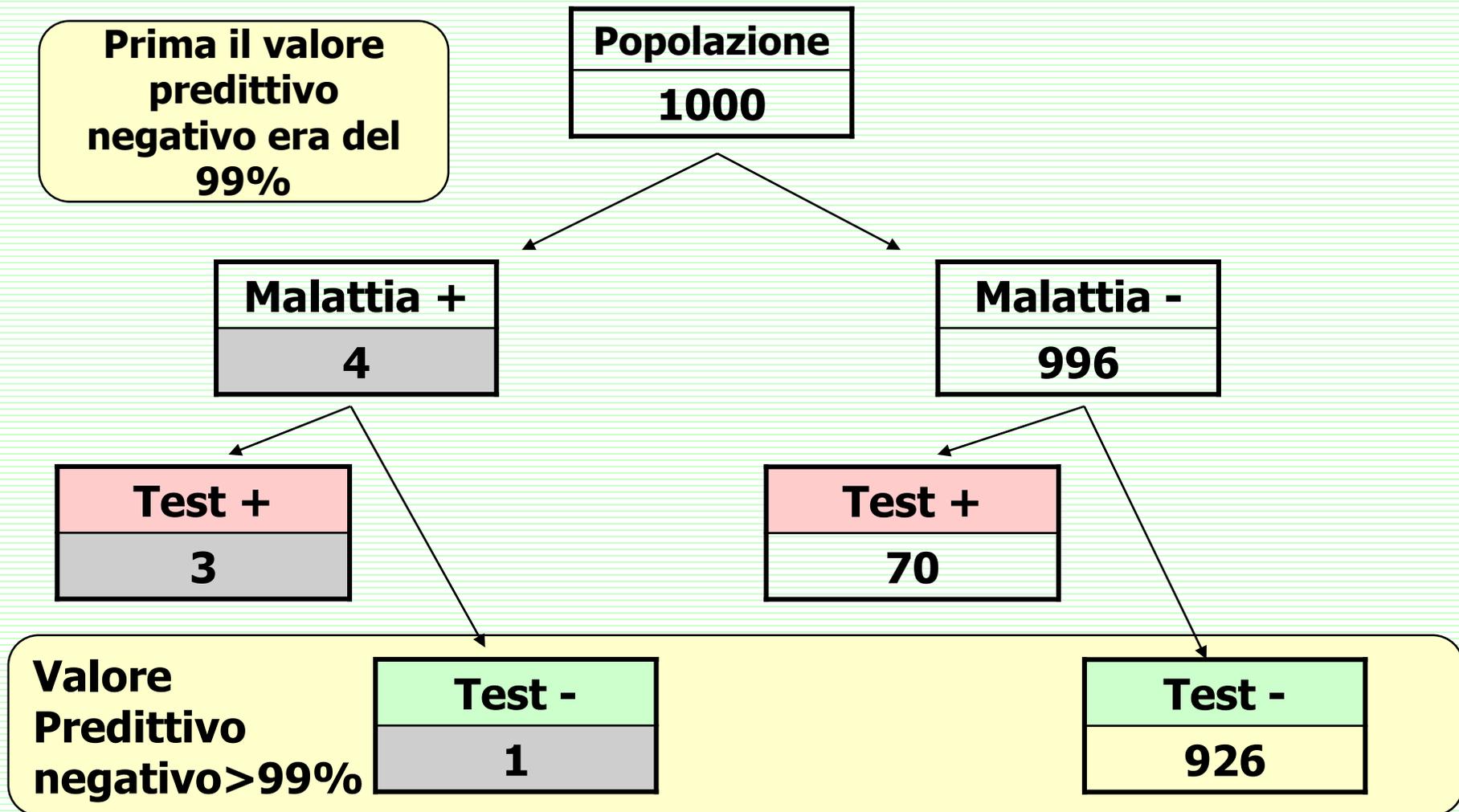
Valori predittivi e cambiamenti nella prevalenza



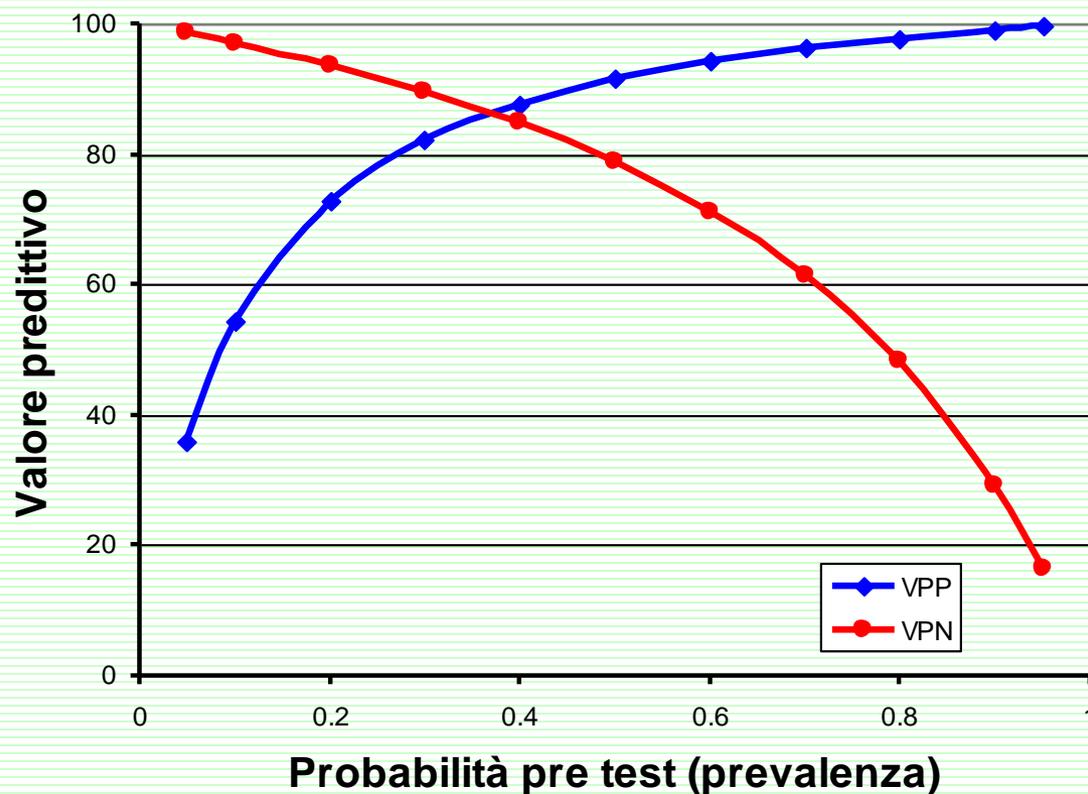
Valore predittivo positivo a basse prevalenze



Valore predittivo positivo a basse prevalenze

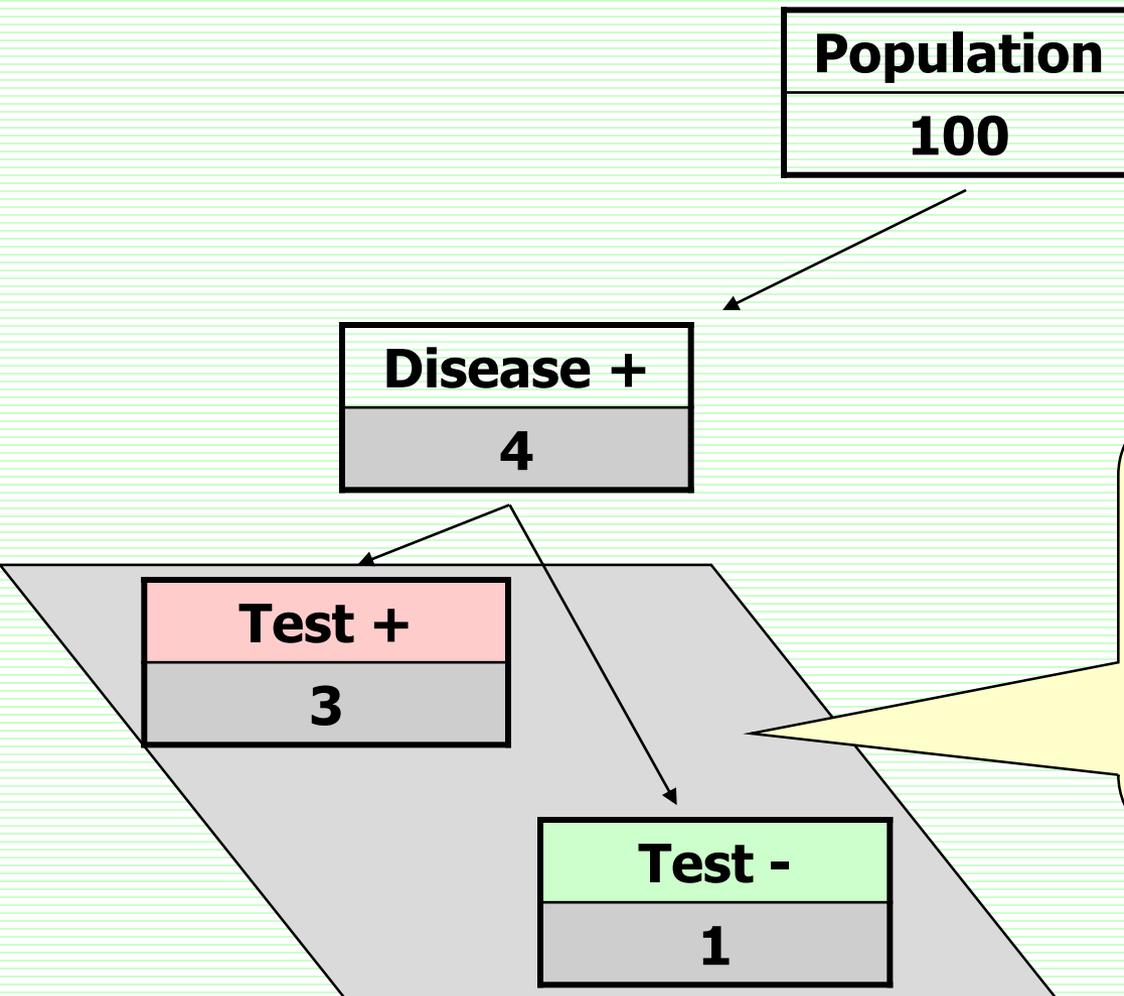


Relazione tra prevalenza e valori predittivi



Basato su un test con sensibilità del 75% e specificità del 93%

Likelihood (Verosimiglianza)



La likelihood che qualcuno con la malattia avrà un test positivo è del 75% (= sensibilità!)

Likelihood (2)

Population
100

Disease -
96

Test +
7

Test -
89

**La likelihood che
qualcuno sano avrà
un test positivo è
7/96 o 7%**

= 1-specificità!

Likelihood Ratio (rapporto di verosimiglianza)

$$\begin{aligned} \text{Likelihood ratio +} &= \frac{\text{Likelihood di un test positivo data la malattia}}{\text{Likelihood di un test positivo data l'assenza di malattia}} \\ LR_+ &= \frac{\text{Pr}(T+ | D+)}{\text{Pr}(T+ | D-)} \\ &= \frac{\text{Sensibilità}}{1 - \text{Specificità}} = \frac{0.75}{0.07} = 10.7 \end{aligned}$$

Una Likelihood Ratio pari a 1.0 indica un test non informativo (ad es. capita quando sensibilità e specificità sono entrambi del 50%)

Più alta la Likelihood Ratio +, migliore il test

Problema clinico

- **Il test è sul valore diagnostico del natriuretico di tipo B (BNP)**
- **il BNP distingue tra gli anziani quelli con e senza disfunzione ventricolare sinistra?**

$$LR_+ = \frac{\Pr(T+ | D+)}{\Pr(T+ | D-)} = \frac{11/12}{50/143} = 2.6$$

		LVSD		
		Presente	Assente	
BNP	≥ 18.7	11	50	61
	< 18.7	1	93	94
		12	143	155

Likelihood Ratio di un test positivo

Quanto è più probabile avere un risultato positivo del test in soggetto malato rispetto ad un soggetto non malato

$$LR_+ = \frac{\Pr(T+ | D+)}{\Pr(T+ | D-)}$$

Likelihood Ratio di un test negativo

$$LR_{-} = \frac{\Pr(T- | D+)}{\Pr(T- | D-)}$$

Quanto è più probabile avere un risultato negativo del test in soggetto malato rispetto ad un soggetto non malato

Likelihood ratio

$$LR_{-} = \frac{\Pr(T- | D+)}{\Pr(T- | D-)} = \frac{1/12}{93/143} = 0.13$$

		LVSD		
		Presente	Assente	
BNP	>=18.7	11	50	61
	<18.7	1	93	94
		12	143	155

Si possono calcolare gli LR dalle percentuali di colonna

		LVSD		LR
		Presente	Assente	
BNP	≥ 18.7	91,67%	34,97%	2,62
	< 18.7	8,33%	65,03%	0,13
		100%	100%	

Interpretazione degli LR

LR è la probabilità che un test sia positivo in un soggetto malato rispetto ad un soggetto non malato

Significato:

LR=1 la probabilità post test è uguale alla probabilità pre test

LR >1 incrementa la probabilità di avere la patologia in esame

LR <1 diminuisce la probabilità di avere la patologia in esame

Dimensione degli LRs

- LRs >10 o < 0.1 genera grandi, e spesso conclusivi cambiamenti da pre- a post-test probabilità;
- LRs di 5-10 e 0.1-0.2 genera moderati cambiamenti da pre- a post-test probabilità;
- LRs di 2-5 e 0.5-0.2 genera piccoli(ma talvolta importanti) cambiamenti nelle probabilità;
- LRs di 1-2 e 0.5-1 altera le probabilità di un grado molto piccolo

Tabelle 2x2

		Malattia		
		Si	No	Totale
TEST	Si	3	7	10
	No	1	89	90
	Totale	4	96	100
		a+c	b+d	a+b+c+d

Labels for cells:

- Top-left cell (3): a
- Top-right cell (7): b
- Bottom-left cell (1): c
- Bottom-right cell (89): d
- Row 1 Total (10): a+b
- Row 2 Total (90): c+d
- Column 1 Total (4): a+c
- Column 2 Total (96): b+d
- Grand Total (100): a+b+c+d

Sensibilità

		Malattia		
		Si	No	Totale
TEST	Si	3	7	10
	No	1	89	90
	Totale	4	96	100
		a+c	b+d	a+b+c+d

a *b*
c *d*

c+d *a+b*



Sensibilità

La proporzione di persone con la diagnosi (N=4) che vengono correttamente identificate (N=3)

$$\text{Sensibilità} = a/(a+c) = 3/4 = 75\%$$

Specificità

		Malattia		
		Si	No	Totale
TEST	Si	3	7	10
	No	1	89	90
	Totale	4	96	100
		a+c	b+d	a+b+c+d

a b
c d

a+b
c+d

Falsi
Positivi

Specificità

La proporzione di persone senza la diagnosi (N=96) che vengono correttamente identificate (N=89)

$$\text{Specificità} = d / (b+d) = 89 / 96 = 93\%$$

Odds Pre-test [odds= $p/(1-p)$]

		Malattia		
		Si	No	Totale
TEST	Si	3 <small>a</small>	7 <small>b</small>	10 <small>a+b</small>
	No	1 <small>c</small>	89 <small>d</small>	90 <small>c+d</small>
	Totale	4 <small>a+c</small>	96 <small>b+d</small>	100 <small>a+b+c+d</small>

Nel campione, l'odds di avere la malattia è 4 su 96 o 4% (odds pre-test)

Odds post-test

		Malattia		
		Si	No	Totale
TEST	Si	3	7	10
	No	1	89	90
	Totale	4	96	100
		a+c	b+d	a+b+c+d

Note: In the original image, the cells containing 3, 7, 4, and 96 are highlighted with light green circles. The labels 'a', 'b', 'c', and 'd' are placed near the cells containing 3, 7, 1, and 89 respectively.

Nel campione, l'odds di avere la malattia è 4 su 96 o 4% (odds pre-test)

In quelli positivi al test, l'odds di avere la malattia è 3 su 7 o 43% (odds post-test)

Odds post-test

		Malattia		
		Si	No	Totale
TEST	Si	3	7	10
	No	1	89	90
	Totale	4	96	100
		a+c	b+d	a+b+c+d

Labels in the table: 'a' and 'b' are above the cells containing 3 and 7 respectively; 'c' and 'd' are to the left of the cells containing 1 and 89 respectively. Marginal totals are labeled 'a+b' on the right of the first row, 'c+d' on the right of the second row, and 'a+c', 'b+d', and 'a+b+c+d' at the bottom of the columns.

Nel campione, l'odds di avere la malattia è 4 su 96 o 4%
(odds pre-test)

In quelli positivi al test, l'odds di avere la malattia è 3 su 7
o 43% (odds post-test)

In quelli negativi al test, l'odds di avere la malattia è 1 su
89 o approssimativamente 1%

Odds Ratio Diagnostico

		Malattia			
		Si	No	Totale	
TEST	Si	3	7	10	a+b
	No	1	89	90	c+d
	Totale	4	96	100	a+b+c+d
		a+c	b+d		

Potenzialmente utile come misura di sintesi, ma solo congiuntamente ad altre misure (LR, sensibilità, specificità)

$$\begin{aligned} \text{DOR} &= \frac{3/7}{1/89} \\ &= \frac{0.429}{0.011} = 38.2 \end{aligned}$$

L' Odds Ratio Diagnostico è pari al rapporto tra l'odds di avere la diagnosi dato un test positivo e l'odds di avere la diagnosi dato un test negativo

Teorema di Bayes

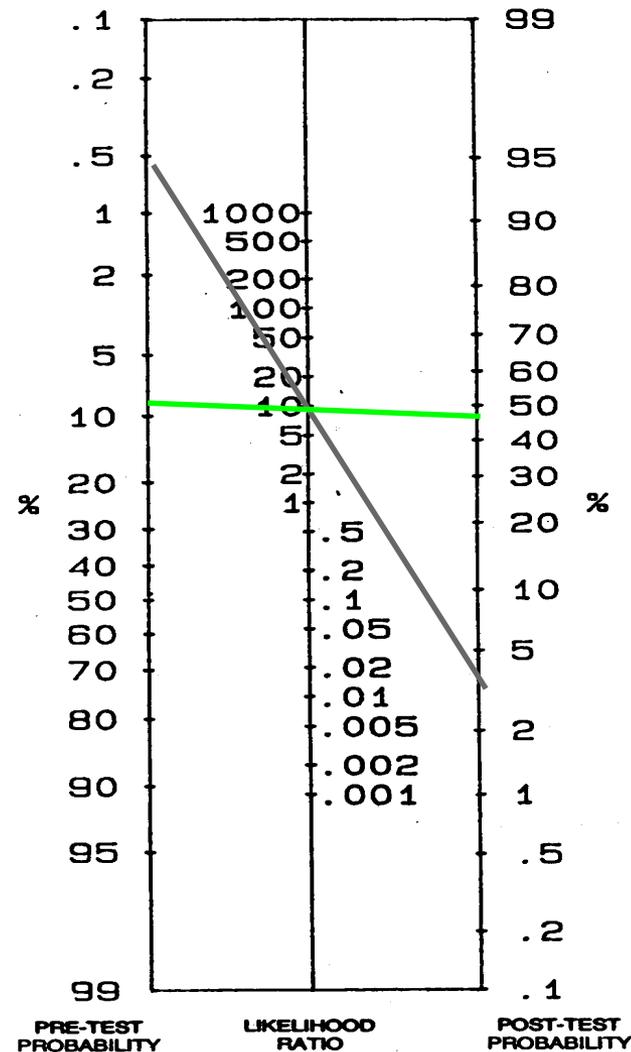


Odds post-test =
Likelihood Ratio x Odds pre-test



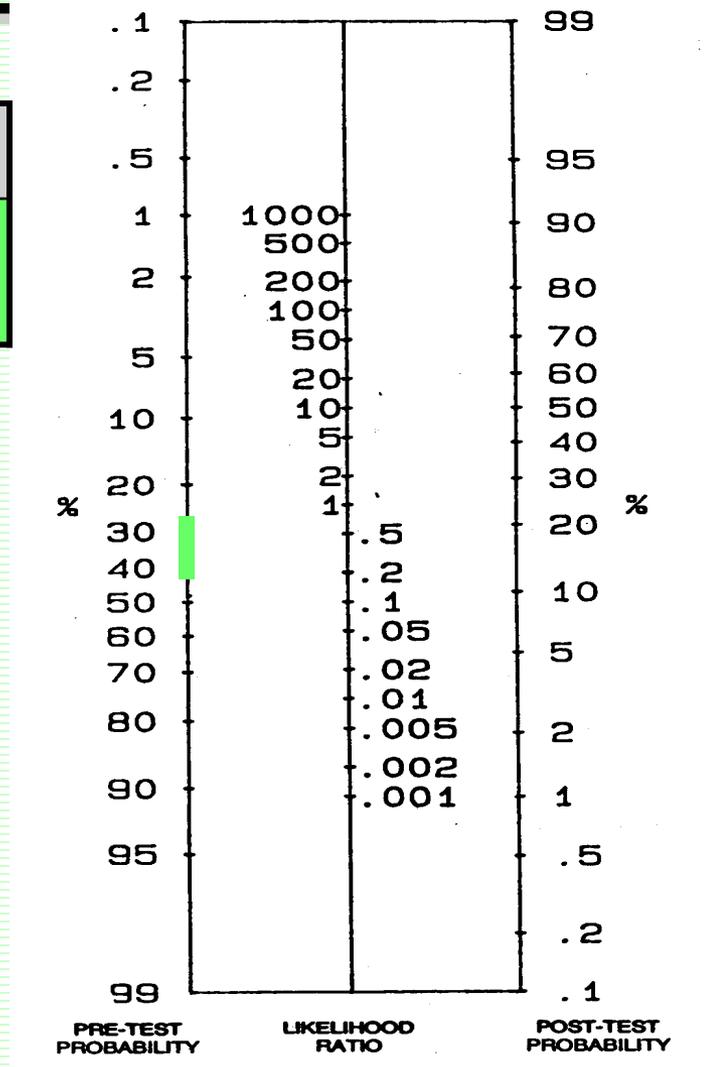
Likelihood Ratio e probabilità pre- e post- test

Per un dato test ed una particolare likelihood ratio, la probabilità post-test dipenderà dalla probabilità pre-test (la prevalenza della condizione nel campione studiato)



Analisi di sensitività di un test diagnostico

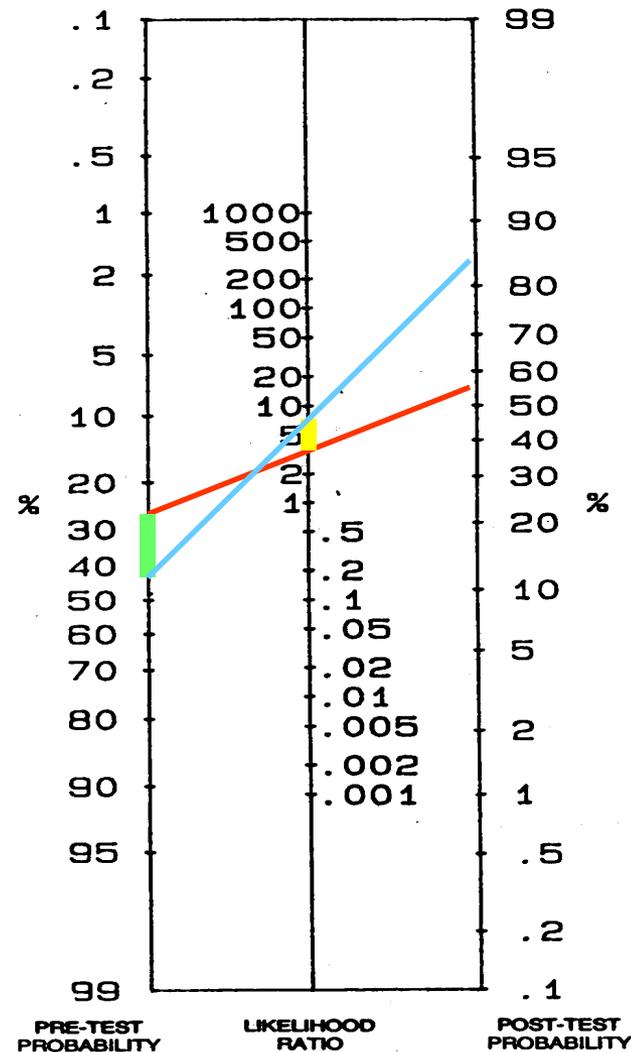
	Valore	95% IdC
Probabilità pre-test	35%	26% - 44%



Analisi di sensitività di un test diagnostico

	Valore	95% IdC
Probabilità pre-test	35%	26% - 44%
Likelihood ratio	5.0	3.0 - 8.5

Applicando gli intervalli di confidenza al 95% sul nomogramma, la probabilità post-test risulterà all'interno dell'intervallo 55-85%



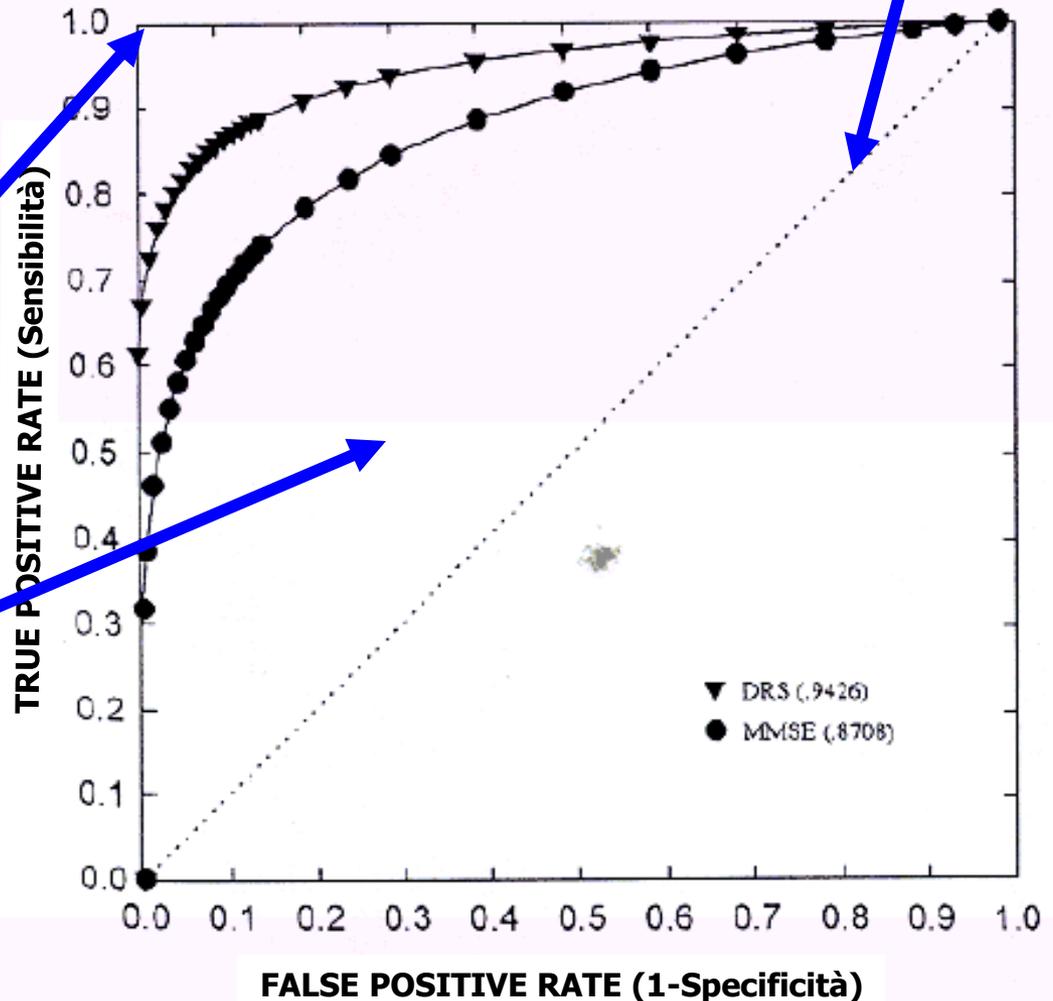
La linea diagonale (Sensibilità=0.5 e Specificità=0.5) rappresenta una performance "casuale"

Curva ROC (Receiver Operating Characteristic Curve)

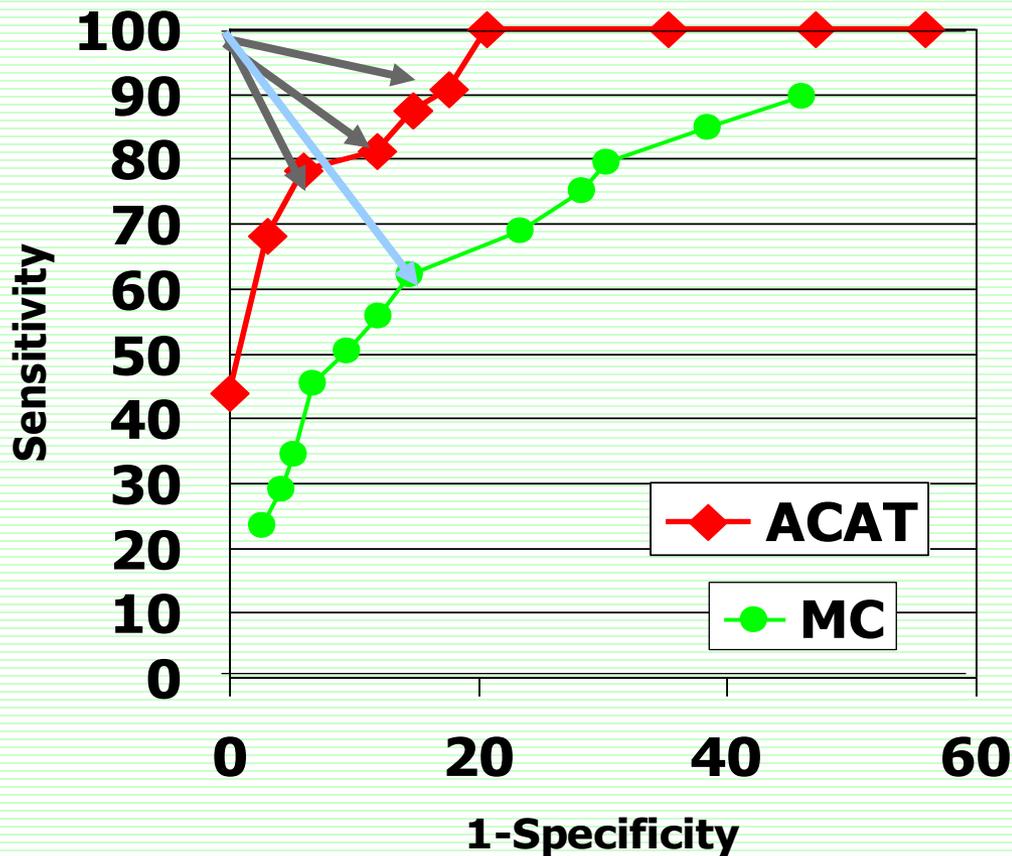
La forma complessiva dipende dalla relazione tra sensibilità e specificità

Più la curva si avvicina al punto Sensibilità=1 and Specificità=1, migliore la performance del test

L' "area sotto la curva" fornisce una misura della performance del test



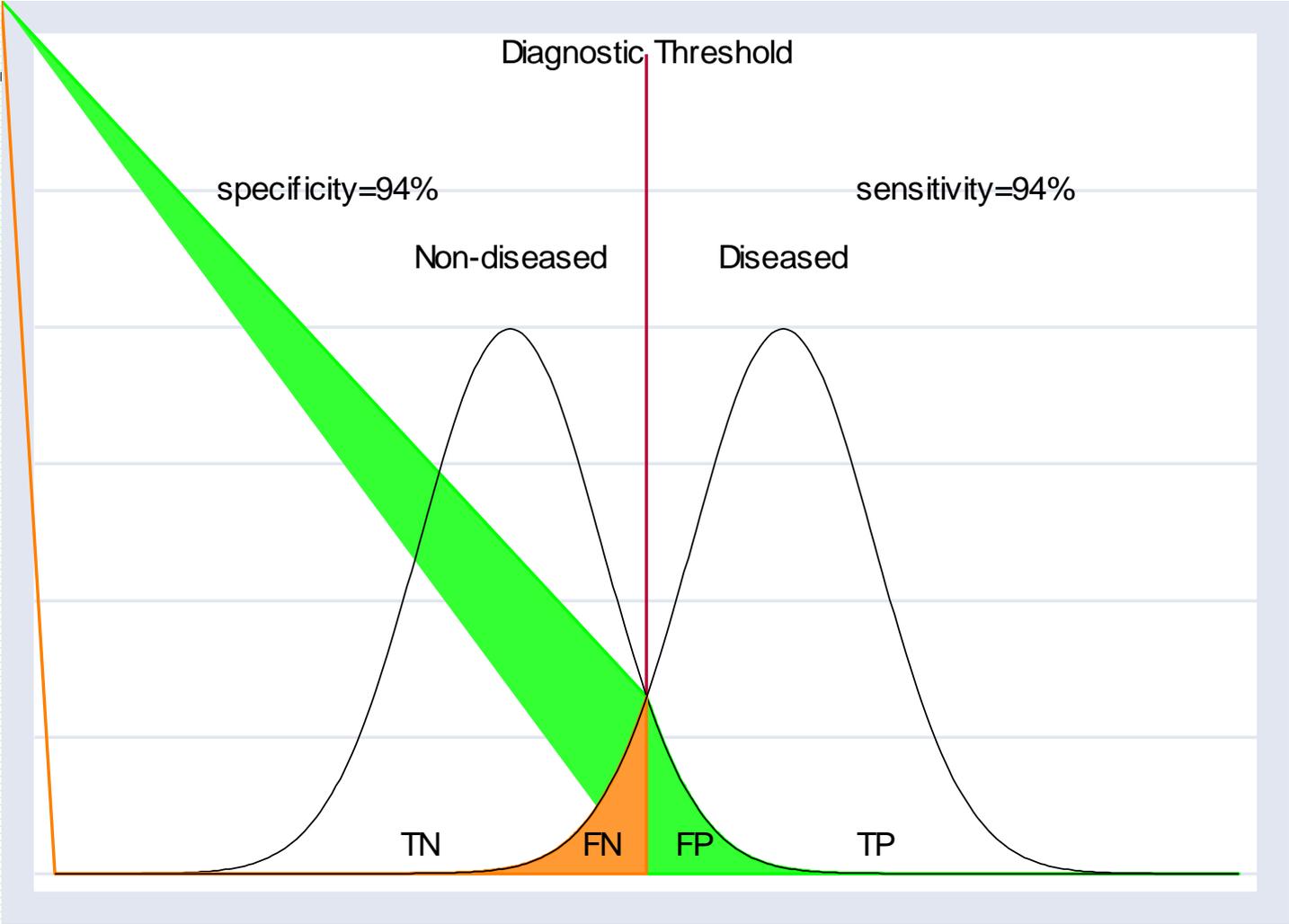
ROC CURVE



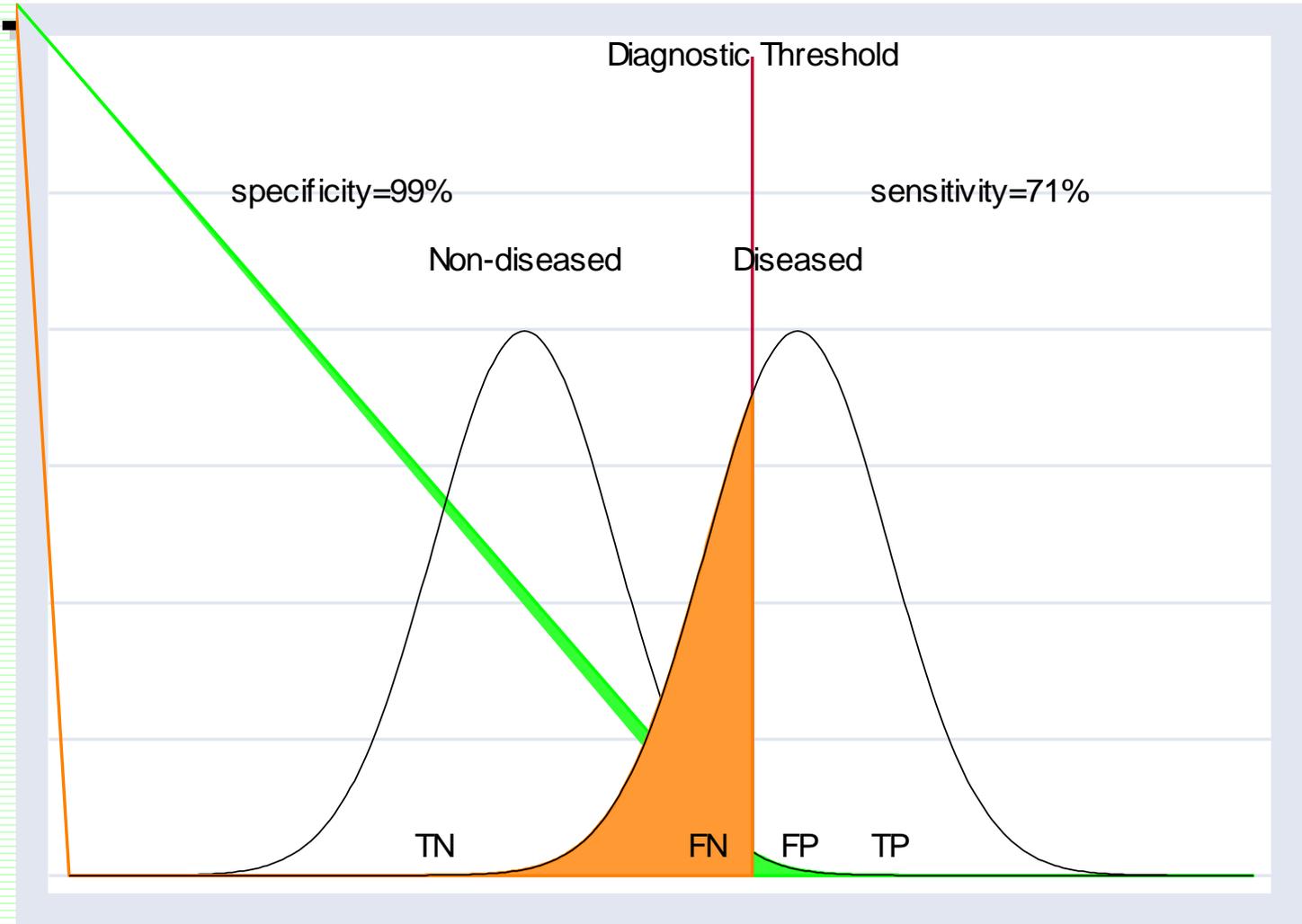
Questo studio confronta la performance del test di screening per la demenza in un campione di comunità (ACAT) e in un campione clinico (MC)

Flicker L, Loguidice D, Carlin JB, Ames D. The predictive value of dementia screening instruments in clinical populations. *International Journal of Geriatric Psychiatry* 1997 ; 12 : 203-209

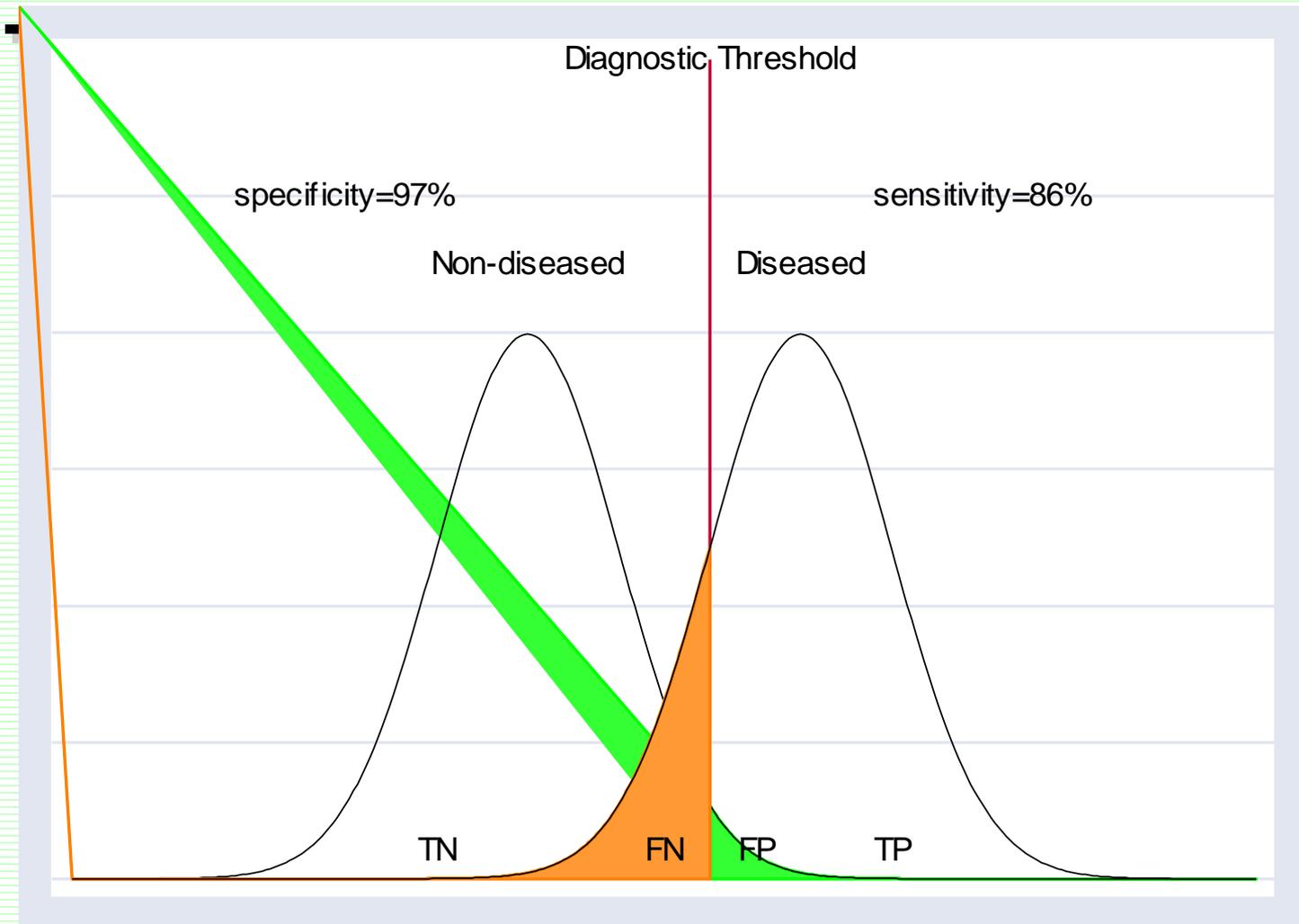
Continuous diagnostic test results



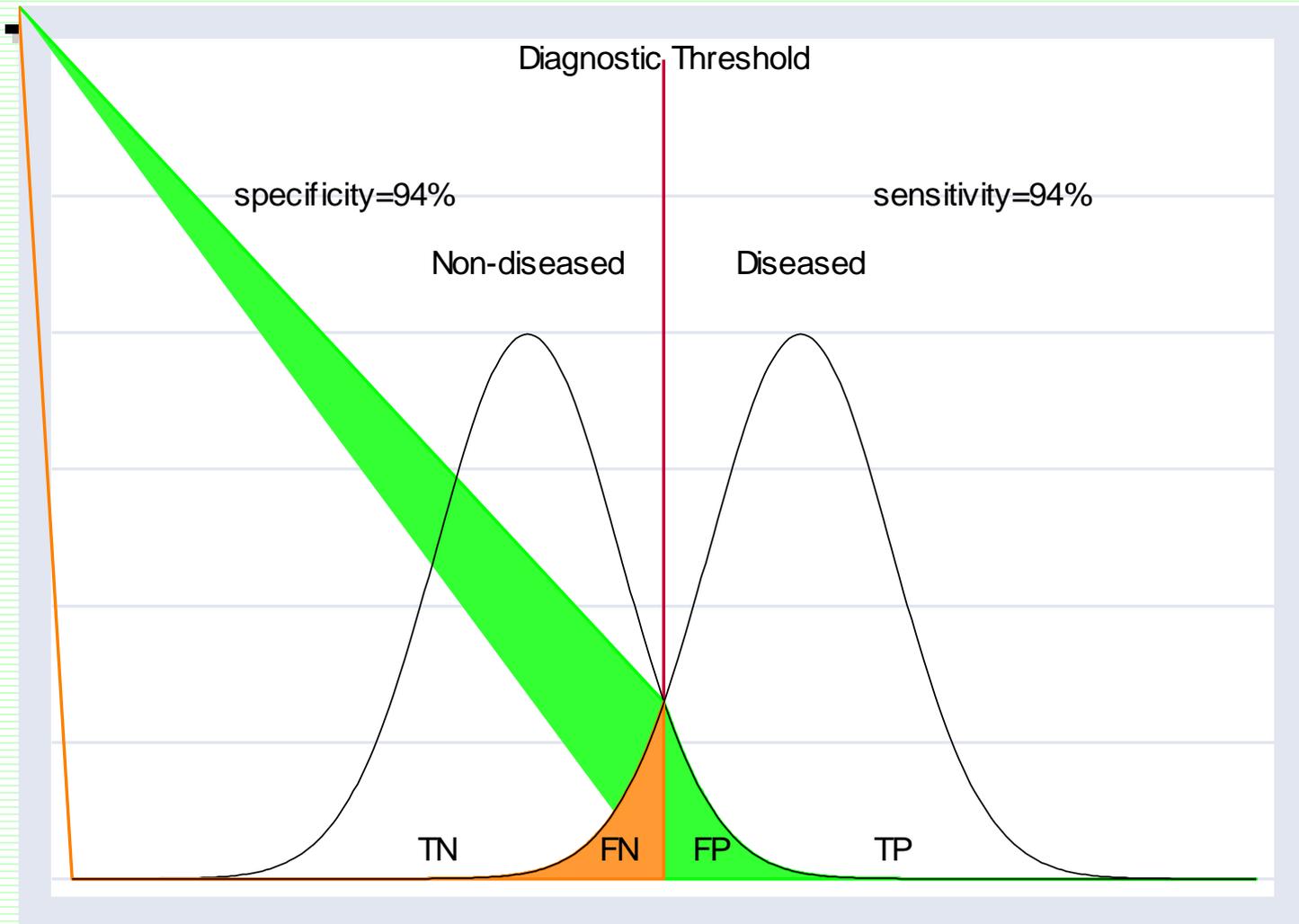
Continuous diagnostic test results



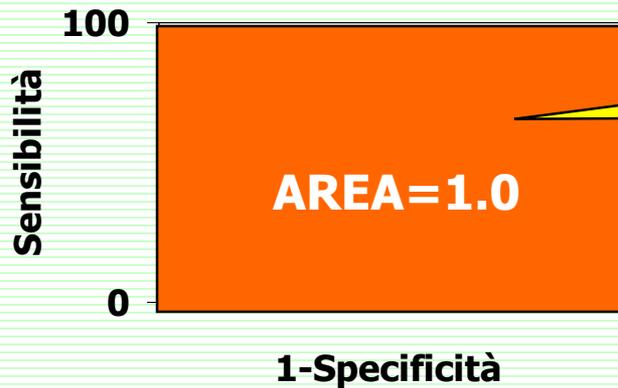
Continuous diagnostic test results



Continuous diagnostic test results

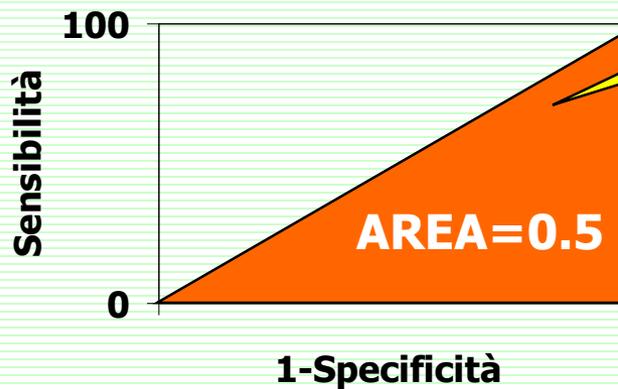


Area sotto le curve ROC



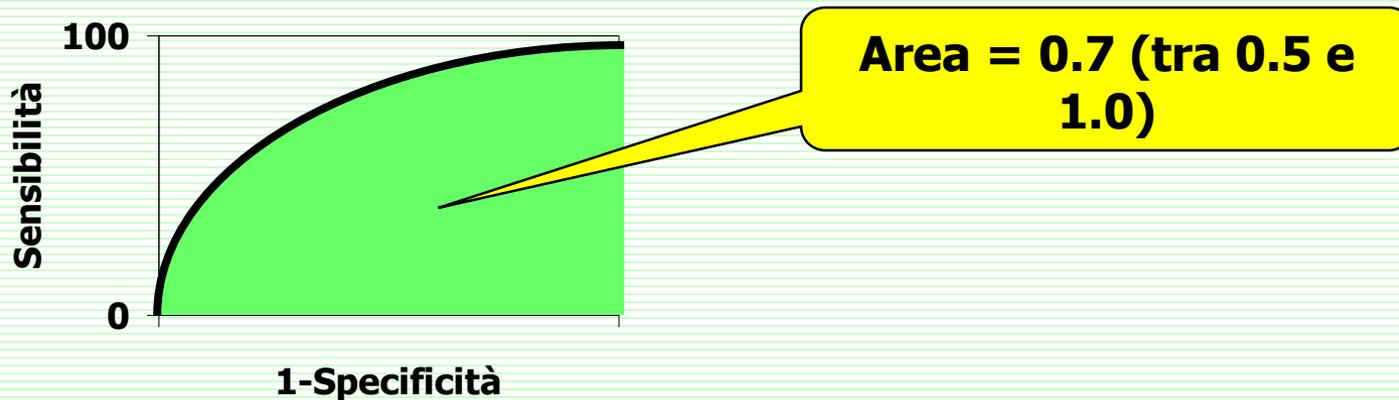
Sensibilità e Specificità entrambi del 100% - TEST PERFETTO

Sensibilità e Specificità entrambi del 50% - TEST INUTILE



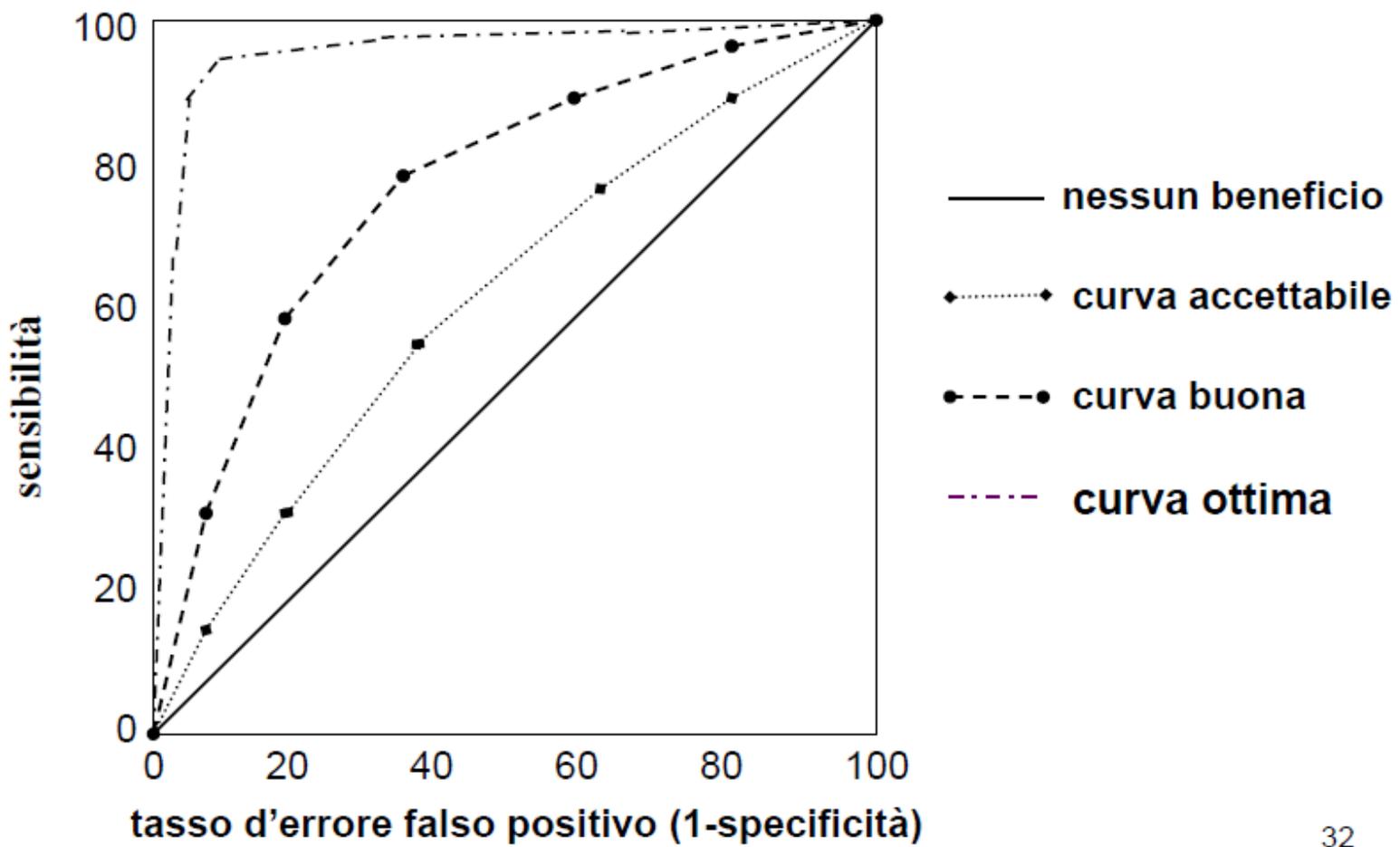
L'area sotto la curva ROC sarà compresa tra 0.5 e 1.0

Area sotto le curve ROC



- **Consideriamo due pazienti scelti a caso tra il gruppo di malati ed il gruppo di sani**
- **Se il test viene utilizzato per individuare quale paziente proviene dal gruppo dei malati, ci darà una risposta corretta il 70% delle volte**

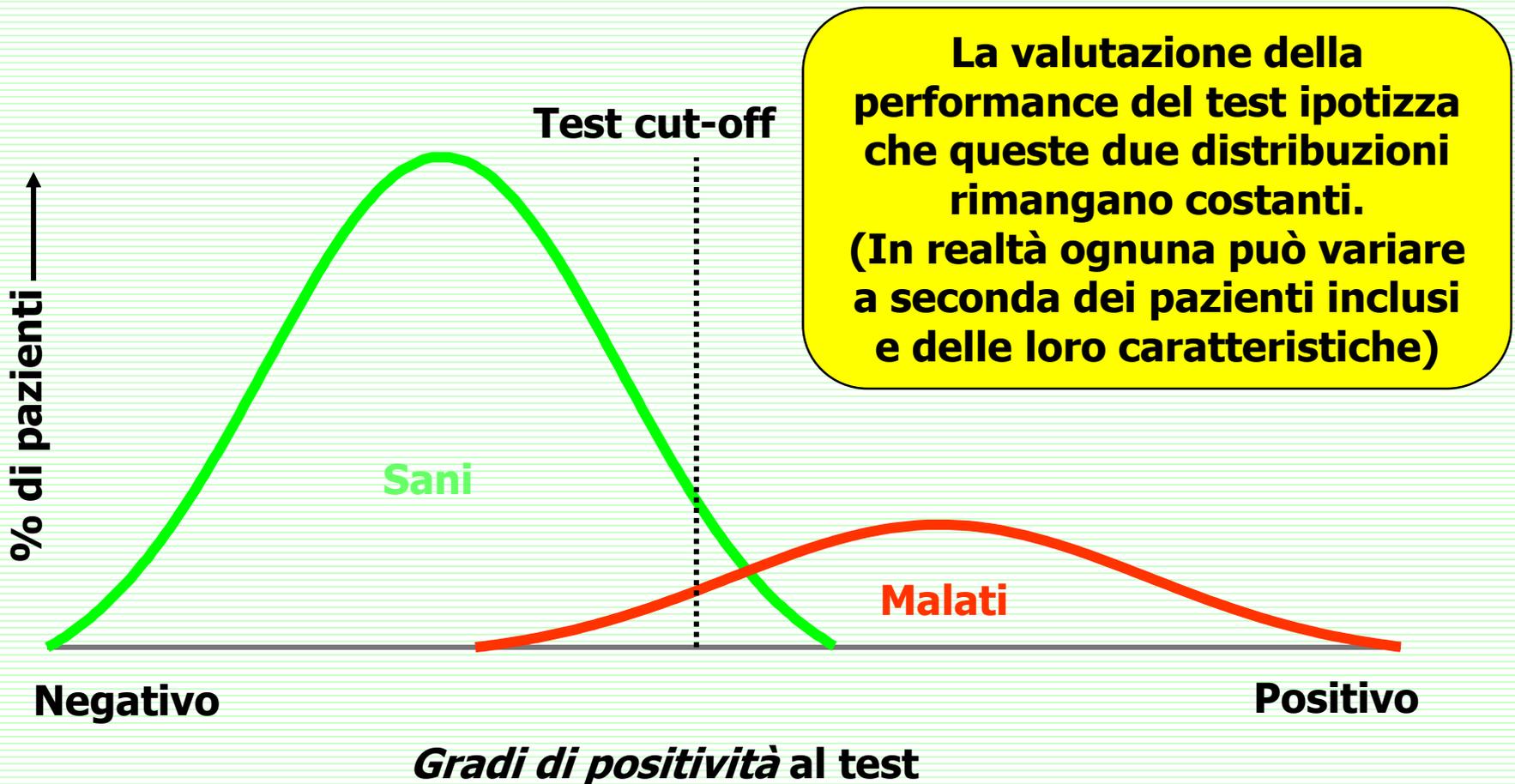
Curve ROC (Receiver Operating Characteristics)



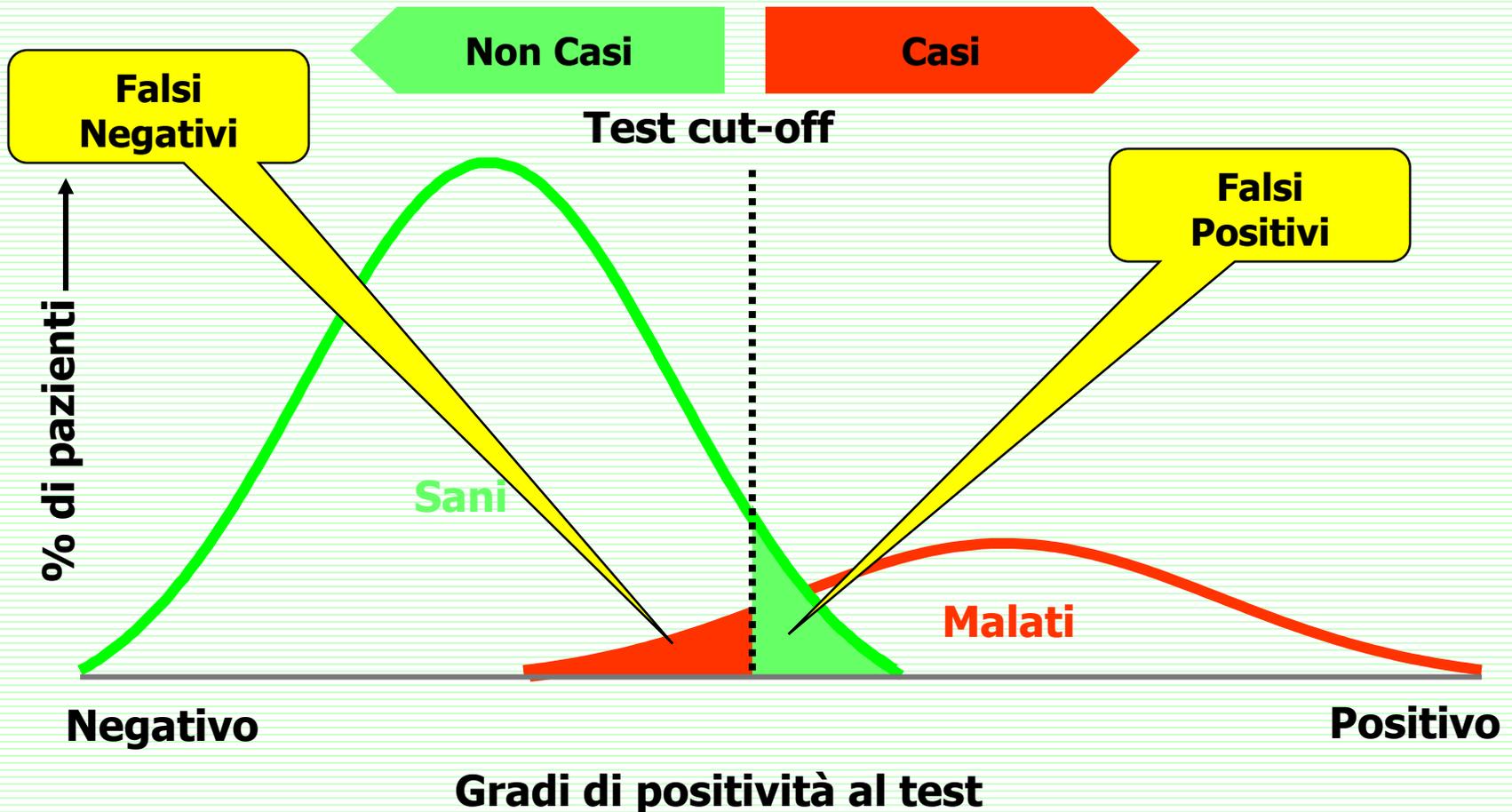
Applicare un test diagnostico in setting differenti

- **Il Valore Predittivo Positivo di un test cambia in base alla prevalenza della malattia nel setting di studio**
- **Sensibilità e specificità sono solitamente considerate proprietà del test piuttosto che del setting, e perciò vengono ritenute costanti**
- **In realtà, sensibilità e specificità sono influenzate dalla complessità di diagnosi differenziali e da una moltitudine di altri fattori (spettro dei pazienti, ...)**

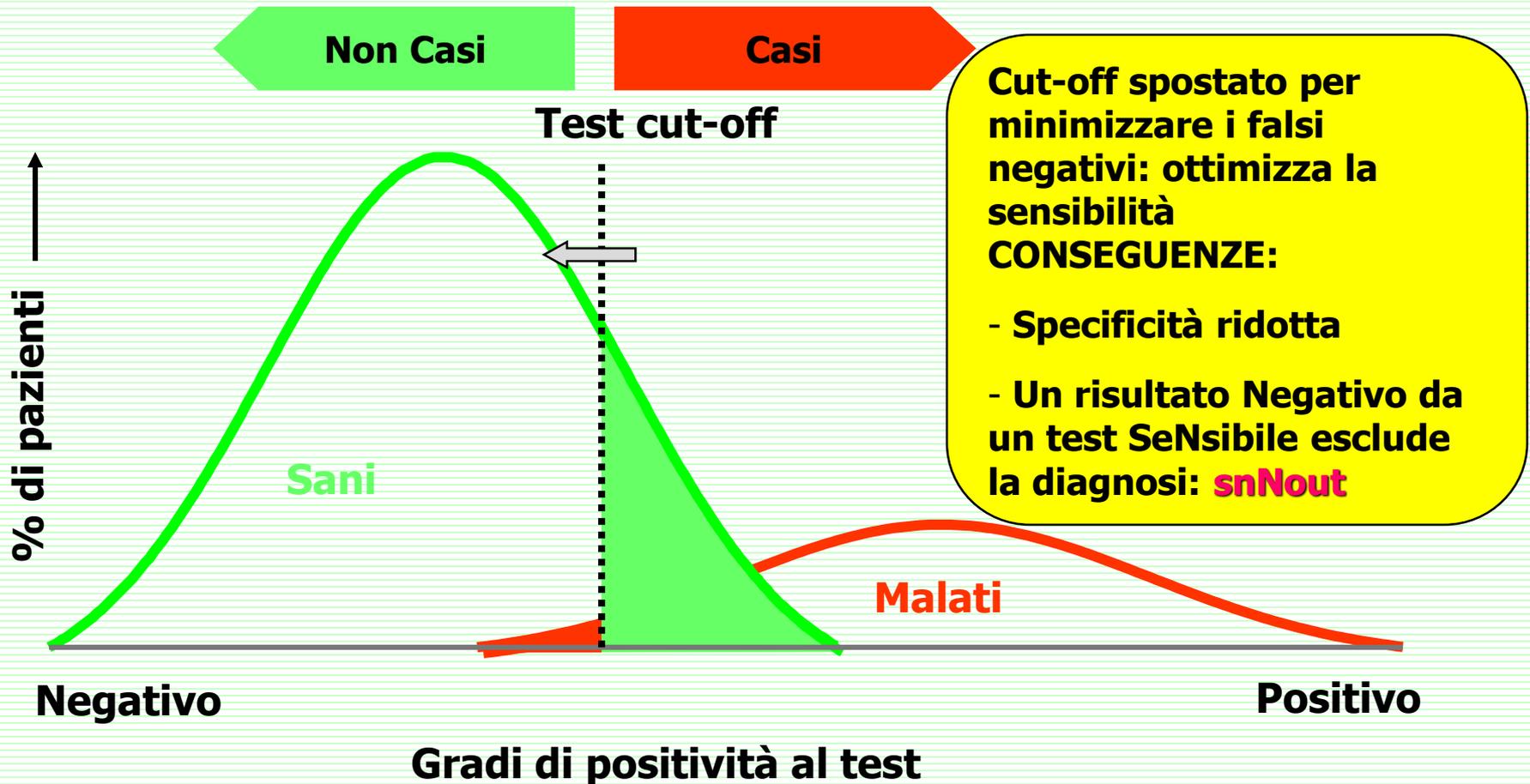
Un test con valori distribuiti secondo una curva normale



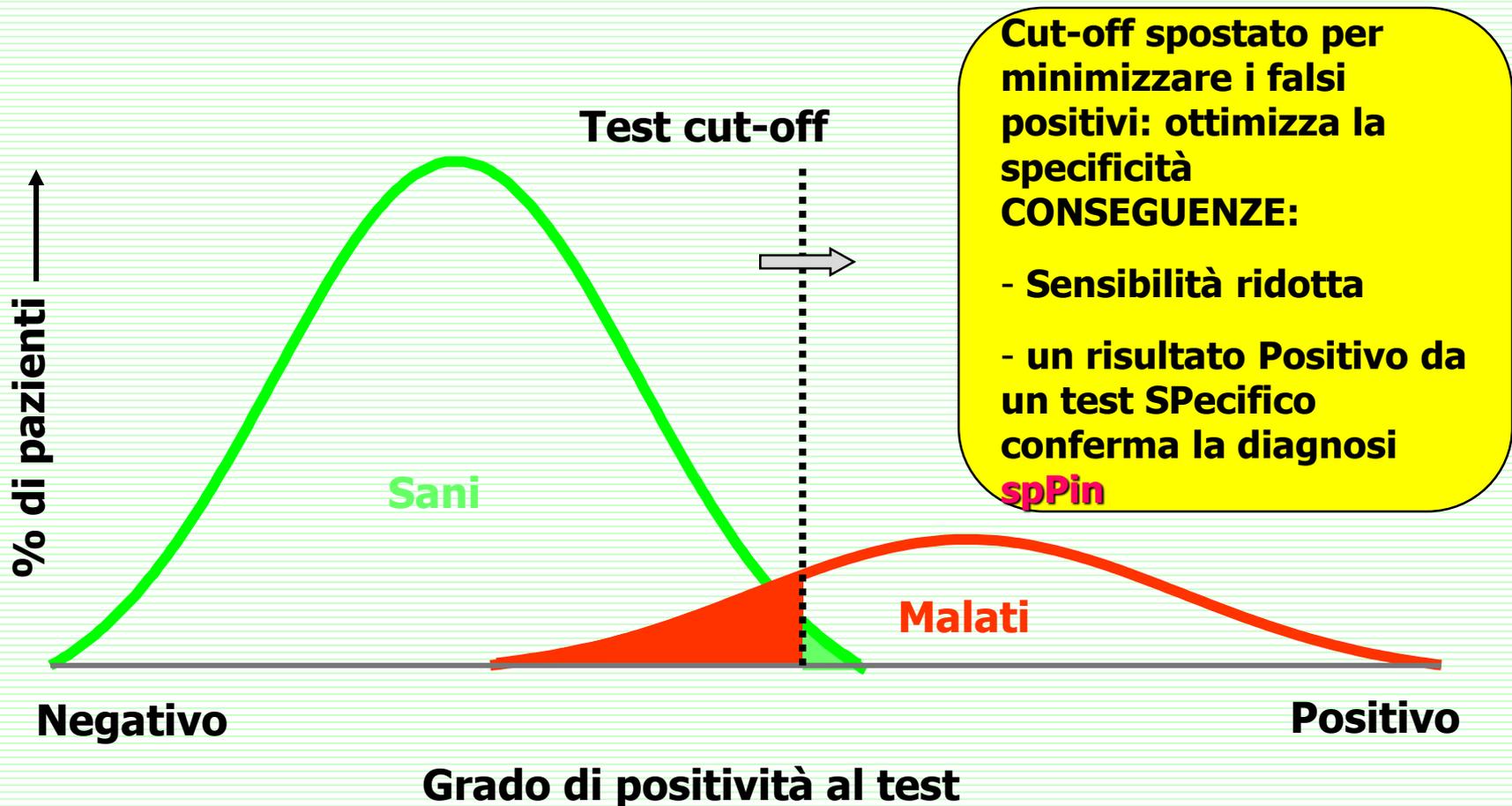
La performance di un test diagnostico



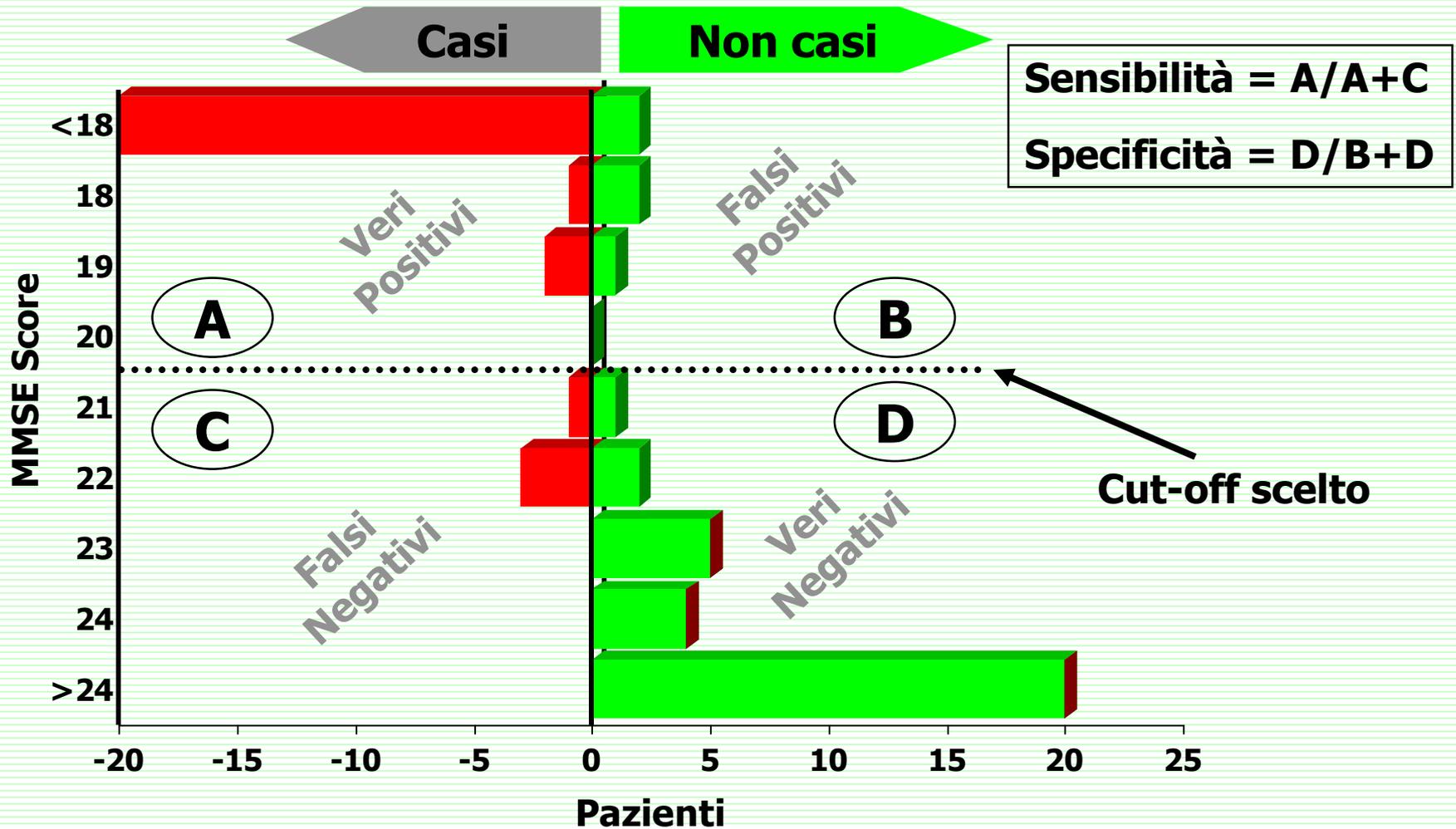
Minimizzare i falsi negativi: un test sensibile



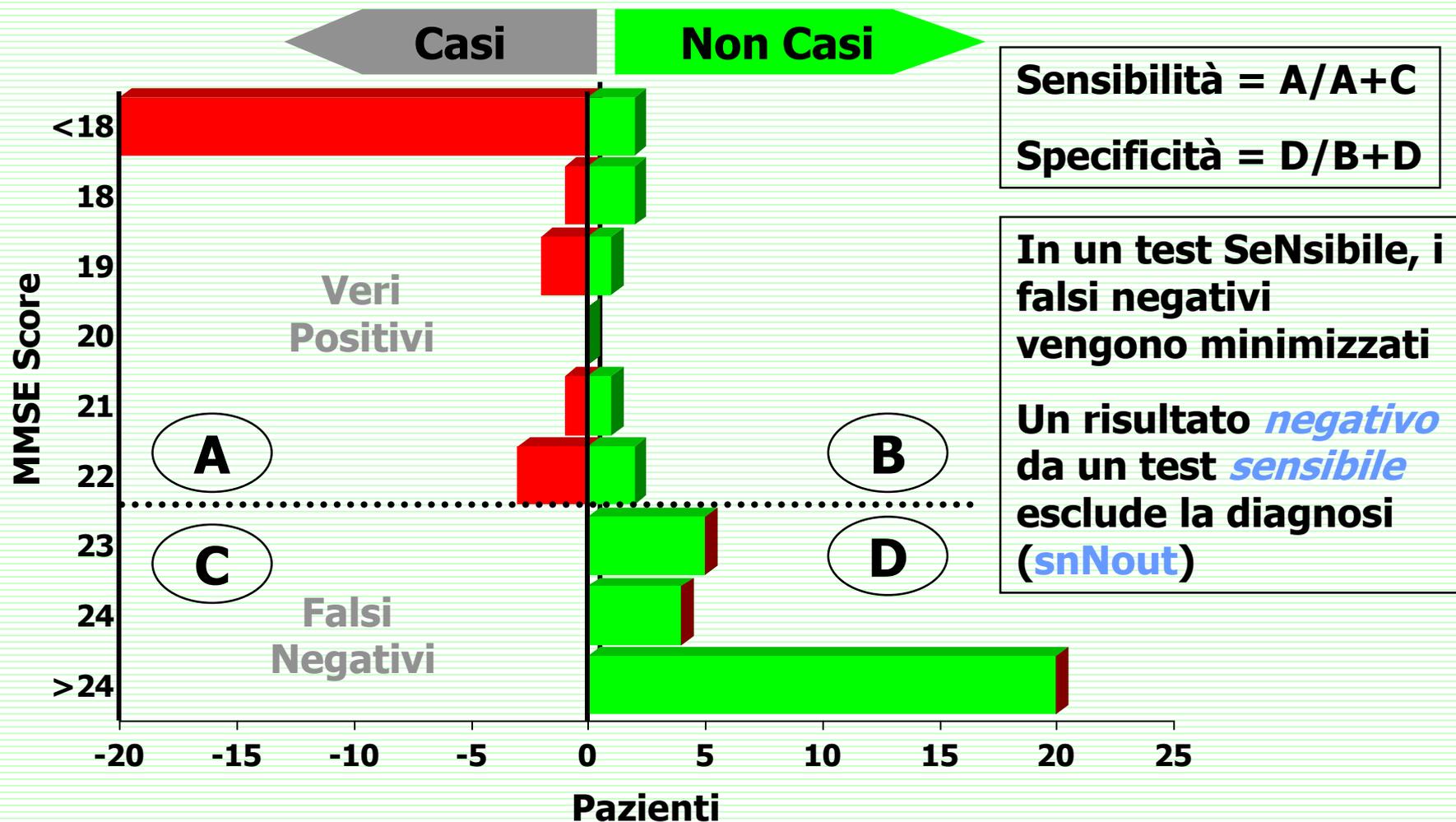
Minimizzare i falsi positivi: un test specifico



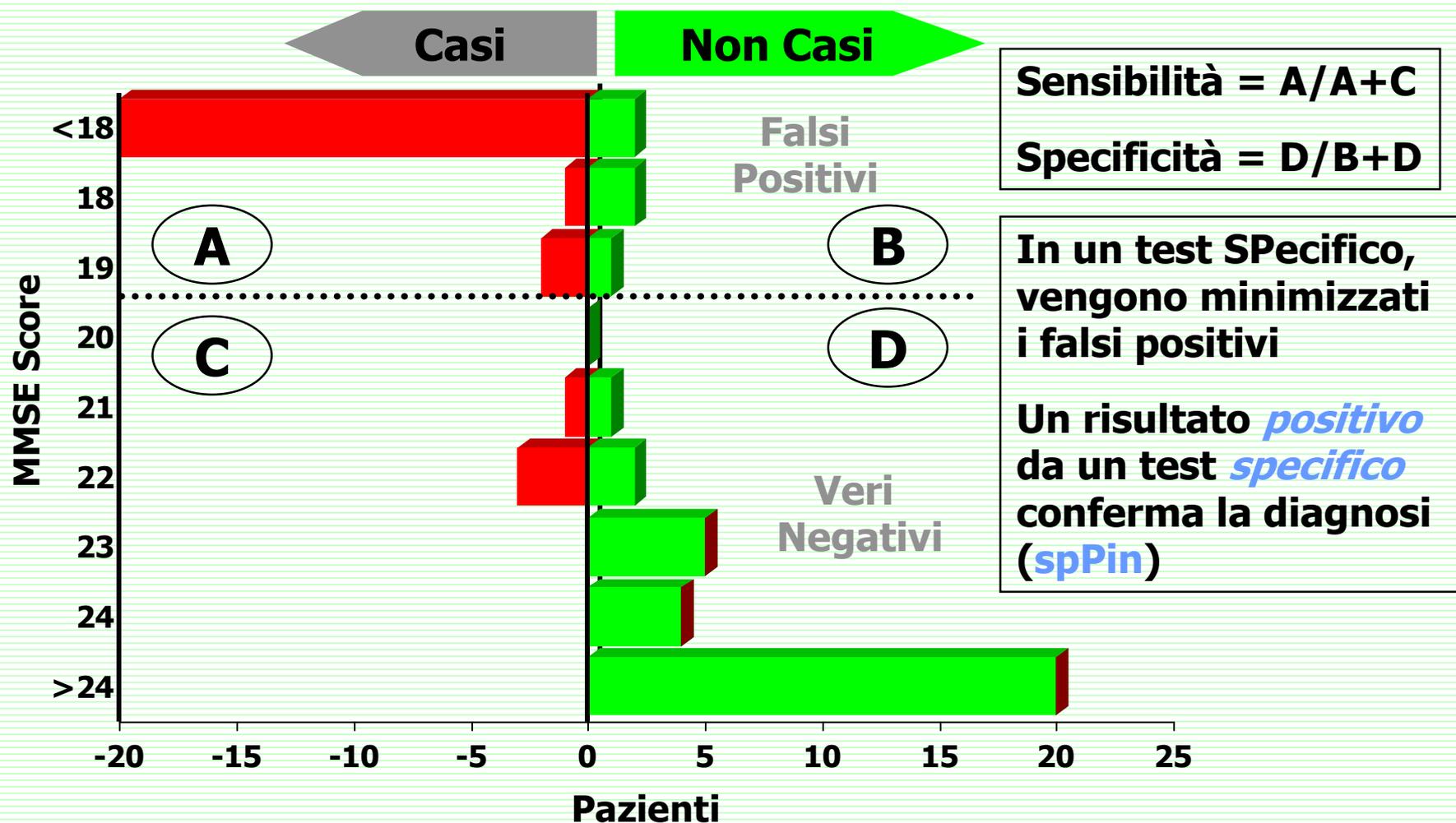
Utilizzare differenti Cut-offs



Aumentare la Sensibilità

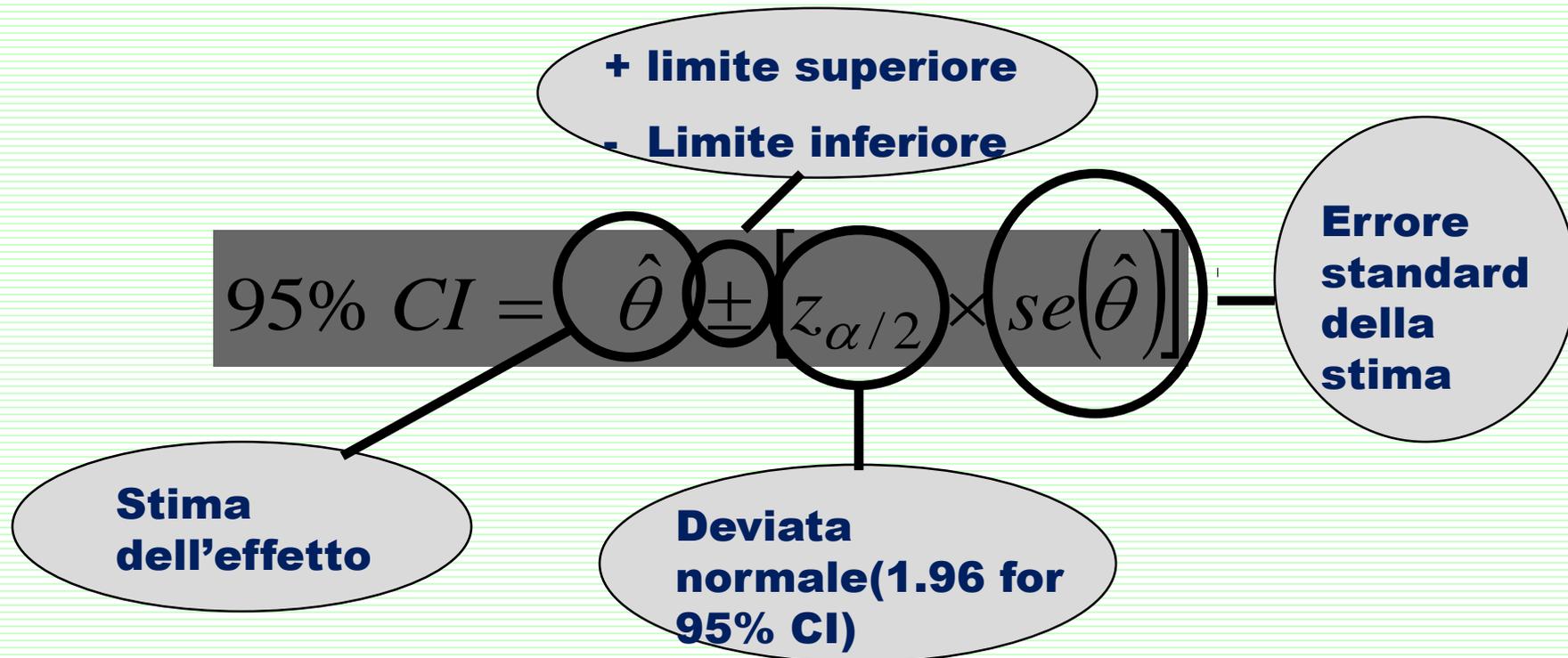


Aumentare la Specificità



INTERVALLO DI CONFIDENZA

- L'incertezza campionaria deve sempre essere descritti, per tutte le statistiche, attraverso l'intervallo di confidenza



1. Il test è potenzialmente rilevante nella mia pratica?

- **Anche se il test è valido al 100%, accurato, riproducibile è utile?**
- **Potrebbe identificare una patologia trattabile?**
- **Lo userei al posto del test che utilizzo ora?**
- **cambia la probabilità di diagnosi differenziale in modo da cambiare il trattamento?**

2. Il test è stato confrontato con un vero gold (reference) standard?

- **Per molte condizioni non esiste un reference standard**
- **Viene descritto il metodo con cui si sta confrontando?**

3. Questo studio include un adeguato spettro di soggetti?

- **Devono essere definiti i criteri di arruolamento dei soggetti: sesso, età, sintomi o gravità della malattia e specificati i criteri di esclusione**
- **Il test deve essere applicato ad un collettivo che include casi più o meno severi, trattati e non trattati**
- **Virtualmente Se e Sp sono costanti rispetto alla prevalenza**
- **I valori predittivi dipendono fortemente dalla prevalenza**
- **I medici di base sono scettici circa l'utilità di test sviluppati in ambito ospedaliero**
- **Un buon test diagnostico non è necessariamente un buon test di screening**

4. Sono stati evitati bias di progettazione?

- **Ogni soggetto che è stato sottoposto al nuovo test è stato sottoposto anche al reference standard e viceversa?**
- **Un bias potenziale è di valutare un reference standard nei soggetti che sono risultati positivi al test che deve essere validato**

5. Sono stati evitati gli errori sistemati?

- **L'interpretazione dei risultati (immagini diagnostiche, valori,...) da parte degli operatori sono inconsciamente influenzati dalla conoscenza**
- **Es. la presenza di dolore cardiaco quando si legge un ECG**
- **La validazione dovrebbe essere condotta in cieco**

6. Il test è riproducibile?

- **Concordanza fra test diversi applicati alla stessa casistica (tra test)**
- **Concordanza fra operatori diversi che interpretano il risultato dello stesso test sugli stessi individui (tra lettori)**
- **Concordanza tra le valutazioni di uno stesso operatore in tempi diversi sullo stesso test e gli stessi individui (entro lettori)**

7. Quali sono le caratteristiche del test che derivano dallo studio?

- **Se, Sp e Valori predittivi potrebbero essere troppo bassi. Il test non è valido**
- **La validità del test dipende però dalla condizione che si sta indagando.**
- **Es test di screening di Guthrie per l'ipotiroidismo congenito $Se=0.99$**
- **Ma $VPP=0.06$ (è capace di cogliere tutti i malati ma a spese di molti falsi positivi)**

8. Vengono presentati gli intervalli di confidenza?

- **Se il campione è grande l'intervallo di confidenza sarà stretto**
- **Particolare attenzione agli studi eseguiti su piccoli campioni**

9. Si possono derivare dei range di normalità?

- **Se la variabile è continua è importante conoscere quali valori sono da considerare anormali**
- **Questo si può fare con la curva ROC oppure con i LR**

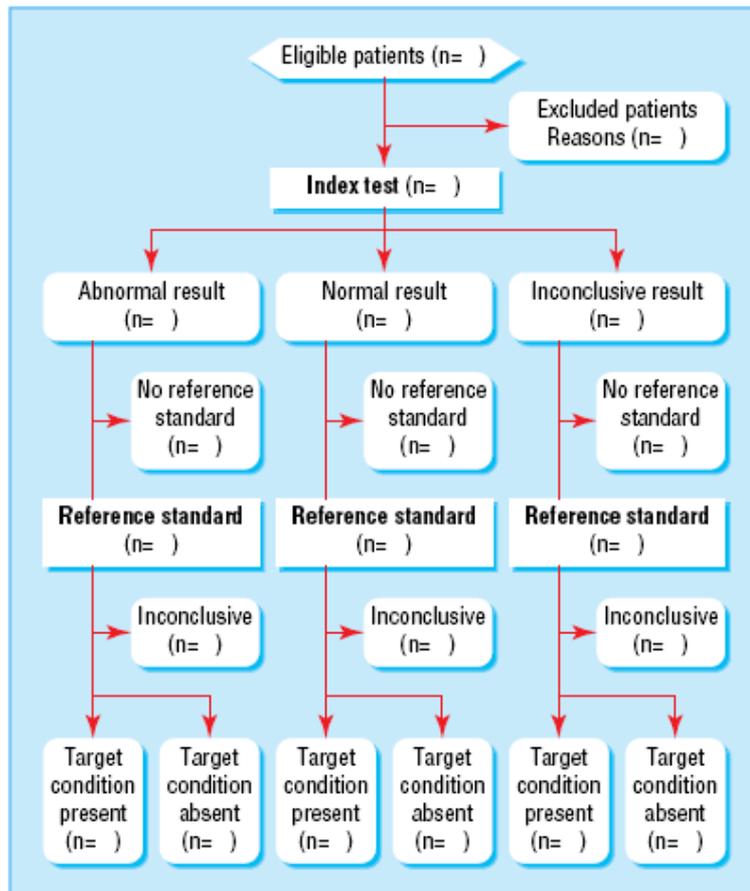
10. Il test può essere collocato al posto di un altro in una sequenza di test?

- **Ogni test diagnostico entra in un albero decisionale di test successivi**
- **Es non tutti i soggetti ipertesi devono sottoporsi a scintigrafia cardiaca, ma solo quelli per i quali si hanno positività di ECG sotto sforzo.**

Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative

Patrick M Bossuyt, Johannes B Reitsma, David E Bruns, Constantine A Gatsonis, Paul P Glasziou, Les M Irwig, Jeroen G Lijmer, David Moher, Drummond Rennie, Henrica C W de Vet for the STARD steering group

T
a
e



STARD) steering group aims to improve the diagnostic accuracy. The group describes and aims for authors of reports

Diagnostic accuracy studies

Description
Identify the article as a study of diagnostic accuracy (recommend MeSH heading "sensitivity and specificity")
State the research questions or aims, such as estimating diagnostic accuracy or comparing accuracy between tests or across participant groups
Describe the study population: the inclusion and exclusion criteria and the settings and locations where the data were collected
Describe participant recruitment: was this based on presenting symptoms, results from previous tests, or the fact that the participants had received the index tests or the reference standard?
Describe participant sampling: was this a consecutive series of participants defined by selection criteria in items 3 and 4? If not, specify how participants were further selected
Describe data collection: was data collection planned before the index tests and reference standard were performed (prospective study) or after (retrospective study)?
Describe the reference standard and its rationale
Describe technical specifications of material and methods involved, including how and when measurements

Bibliografia essenziale

Sedlmeier P and Gigerenzer G. Teaching Bayesian reasoning in less than two hours. *Journal of Experimental Psychology: General.* 130 (3):380-400, 2001.

Greenhalgh T. *How to read a paper. The basic of evidence based medicine.* BMJ Books, 2004.

Knotternus JA (ed). *The Evidence Base of Clinical Diagnosis.* London: BMJ Books, 2002.

Strauss SE, Richardson WS, Glasziou P, Haynes RB. *Evidence-based medicine: how to practice and teach EBM.* Elsevier Churchill Livingstone, 2005.

Loong TW. Understanding sensitivity and specificity with the right side of the brain. *BMJ* 2003; 327: 716-19.