Problemi cognitivi nell'inferenza diagnostica

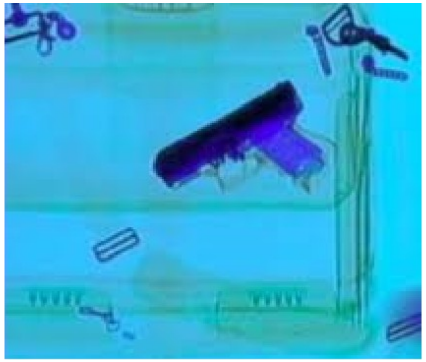
TSD (Teoria della rilevazione del segnale)

TSD (Teoria della rilevazione del segnale)


TSA

IGNORA

SELEZIONA



TSD (Theory of Signal Detection, Teoria della rilevazione del segnale)

Anomalia presente / assente?

Tono acustico udito / non udito

Diagnosi su un gruppo di pazienti

Sommario

Falsi positivi e Falsi negativi

Prevalenza

Effetti della prevalenza sulla decisione medica

Sensibilità

Specificità

Uso delle informazioni sulla sensibilità e specificità nel processo diagnostico

Un problema tipico

Spiegazione grafica

Il teorema di Bayes

Errori di ragionamento

Generalizzazione

Sommario

Falsi positivi e Falsi negativi

Prevalenza

Effetti della prevalenza sulla decisione medica

Sensibilità

Specificità

Uso delle informazioni sulla sensibilità e specificità nel processo diagnostico

Un problema tipico

Spiegazione grafica

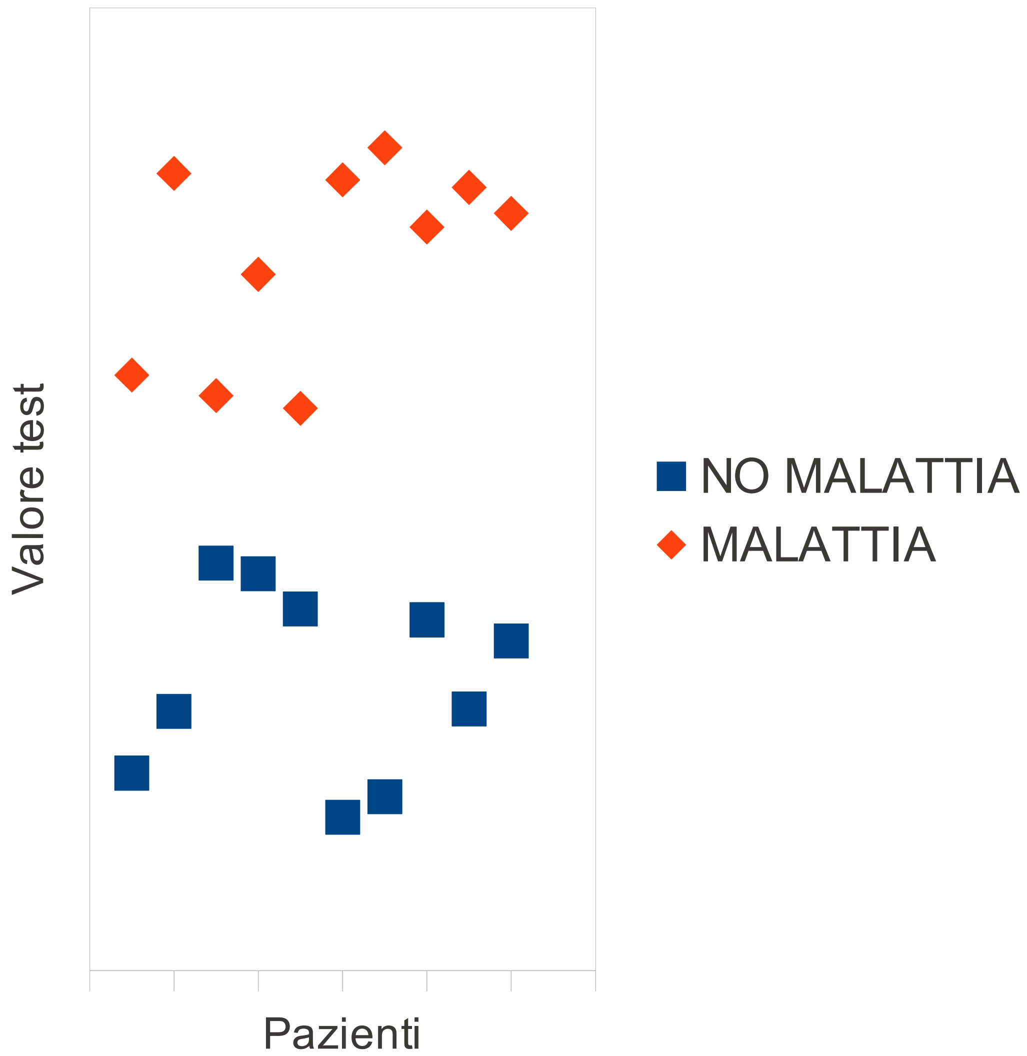
Il teorema di Bayes

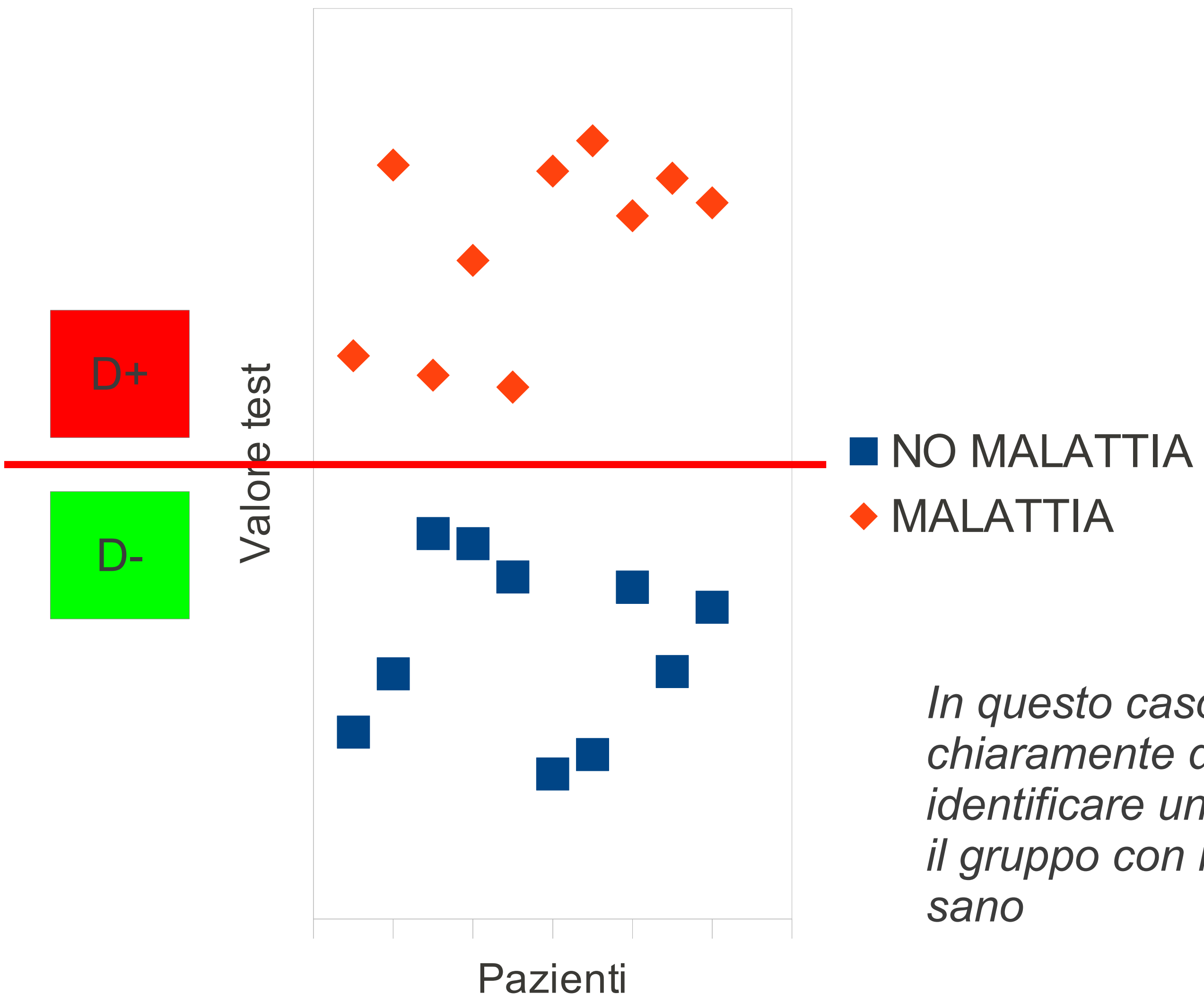
Errori di ragionamento

Generalizzazione

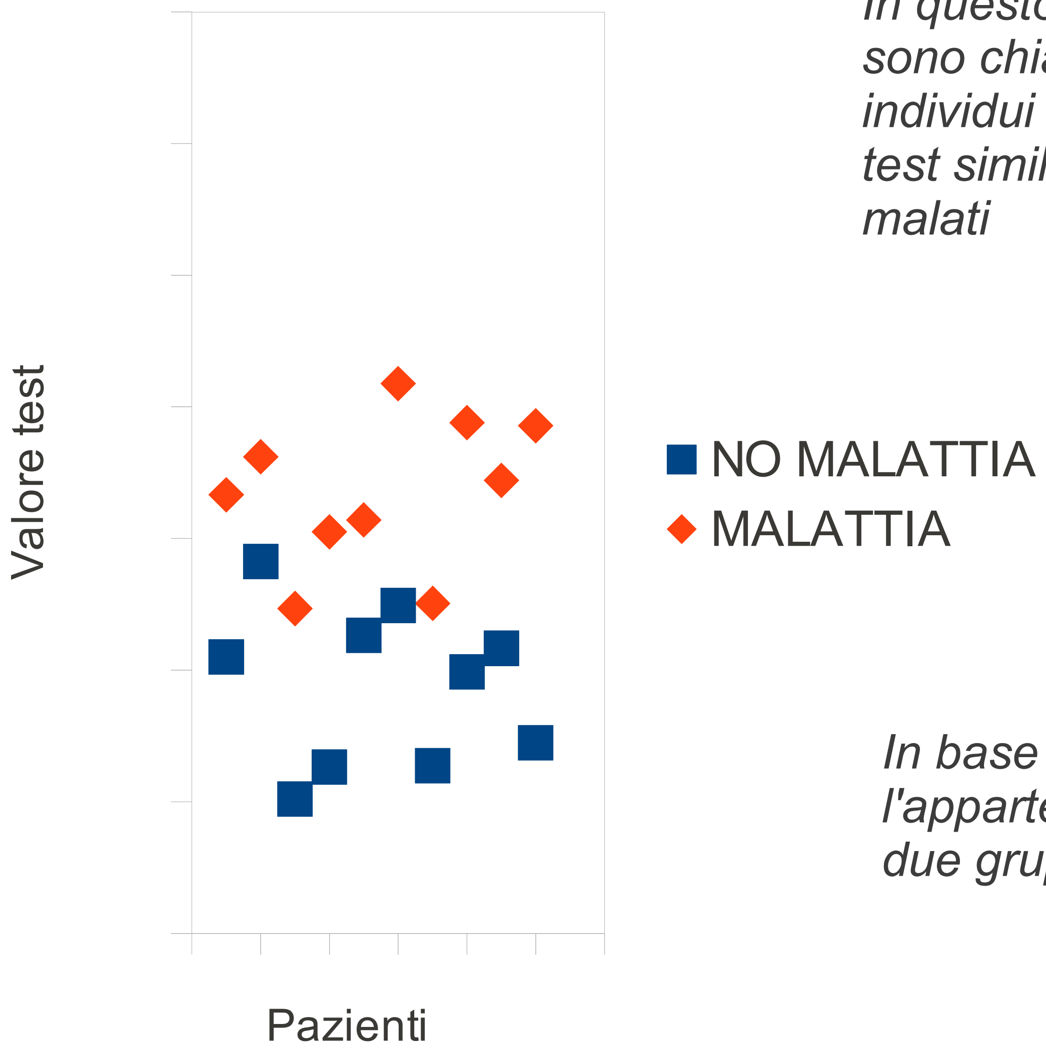
Venti pazienti si sottopongono ad un test; tra questi, è noto lo stato di malattia di dieci pazienti e di assenza di malattia degli altri dieci.

Sulla base del risultato di un test, si richiede di distinguere il gruppo con patologia da quello sano





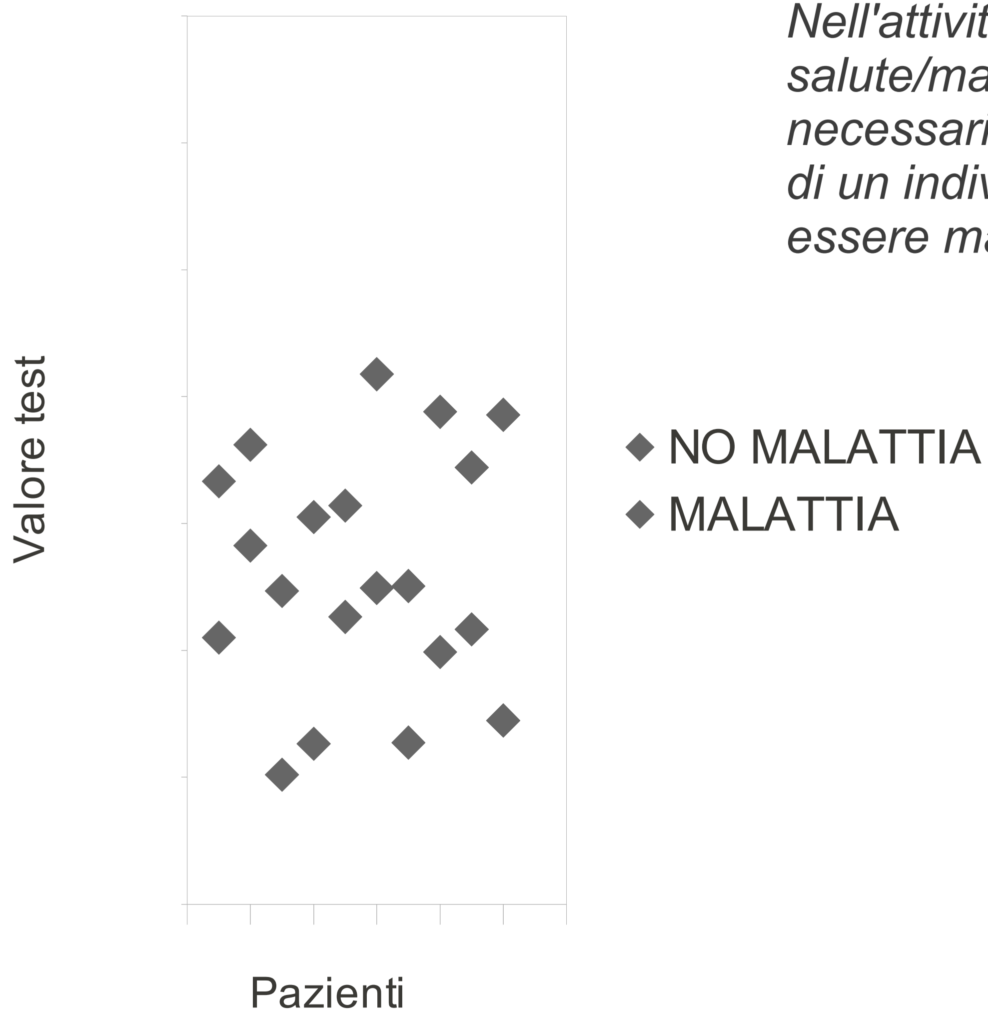
In questo caso, i due gruppi sono chiaramente distinti ed è possibile identificare un valore che distingua il gruppo con malattia da quello sano



In questo caso, i due gruppi non sono chiaramente distinti, e alcuni individui sani hanno risultati del test simili a quelli degli individui malati

In base a cosa si può decidere l'appartenenza di un individuo ai due gruppi?

Lavorare in condizioni di incertezza

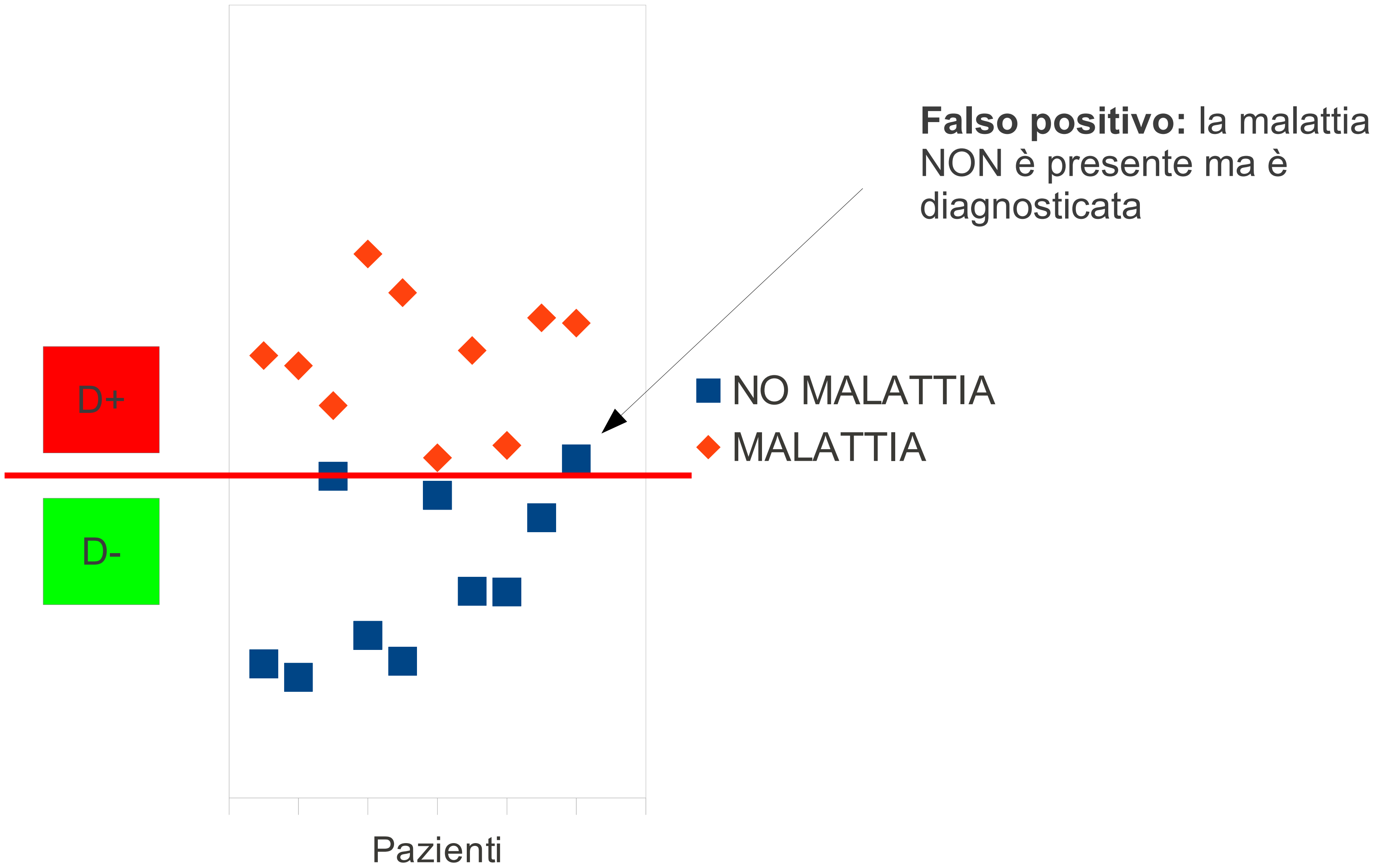


Nell'attività reale lo stato di salute/malattia NON è noto ed è necessario valutare la probabilità di un individuo di essere / non essere malato

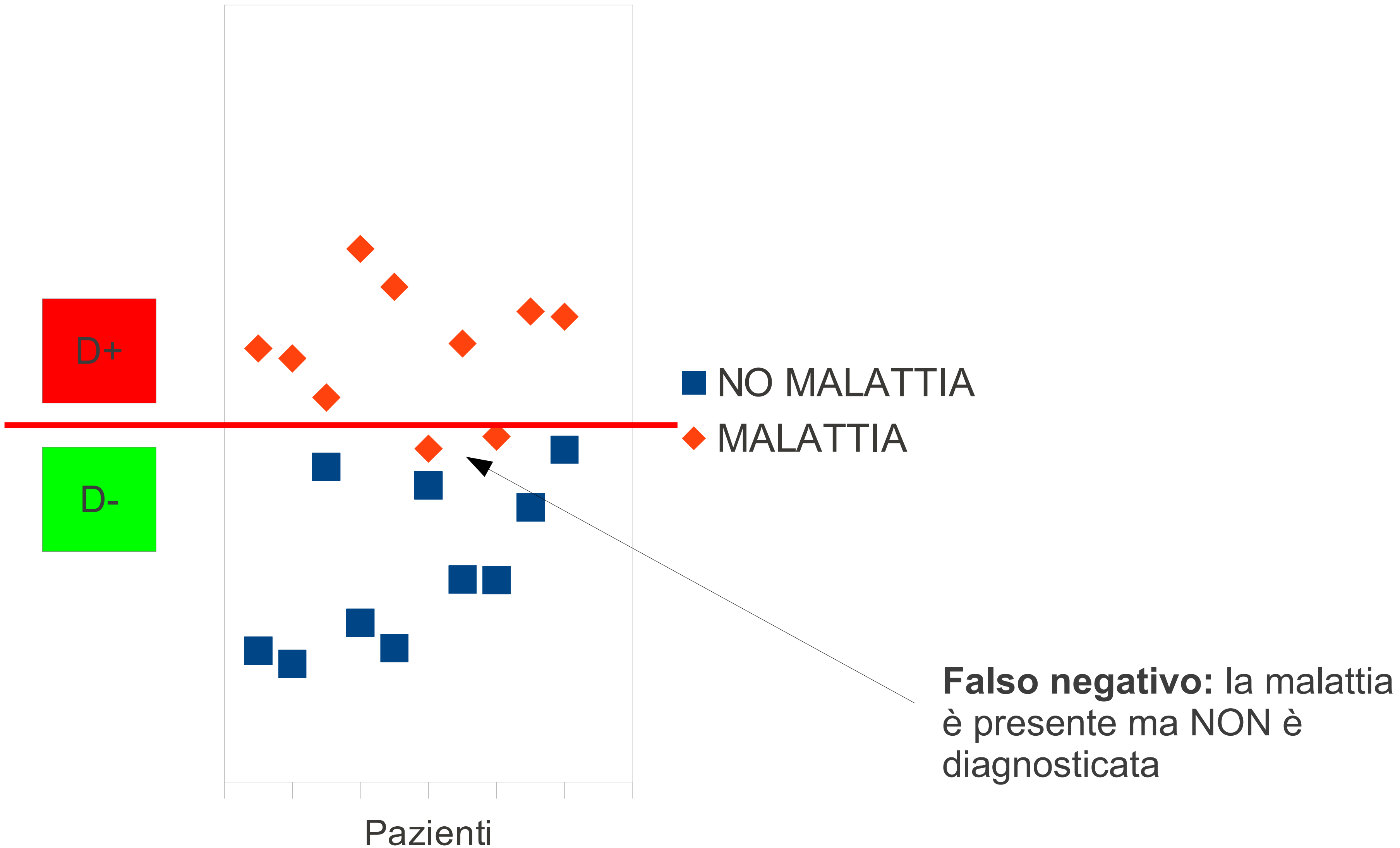
Se i risultati di un test in due popolazioni (sani e malati) si sovrappongono, non è possibile individuare un valore critico che individui i due gruppi senza commettere due tipi di errori:

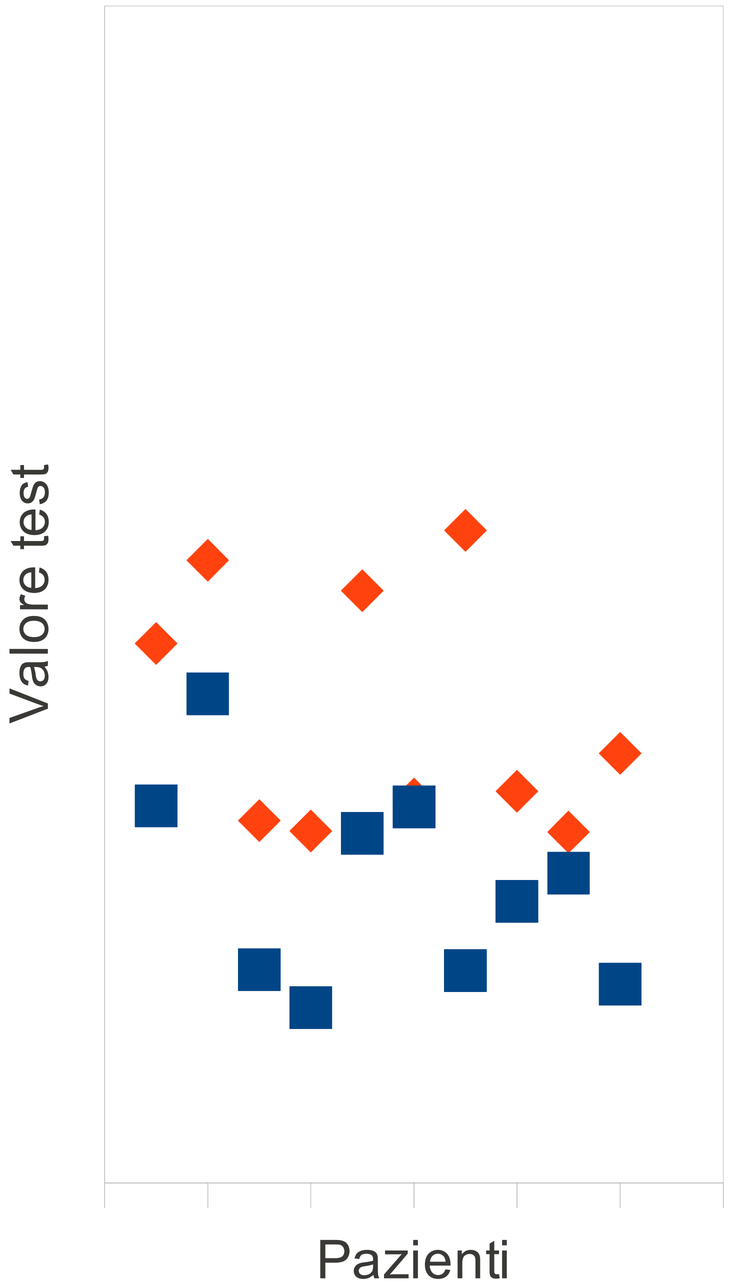
Individui sani indicati come con patologia (falsi positivi)

Individui con patologia indicati come sani (falsi negativi)



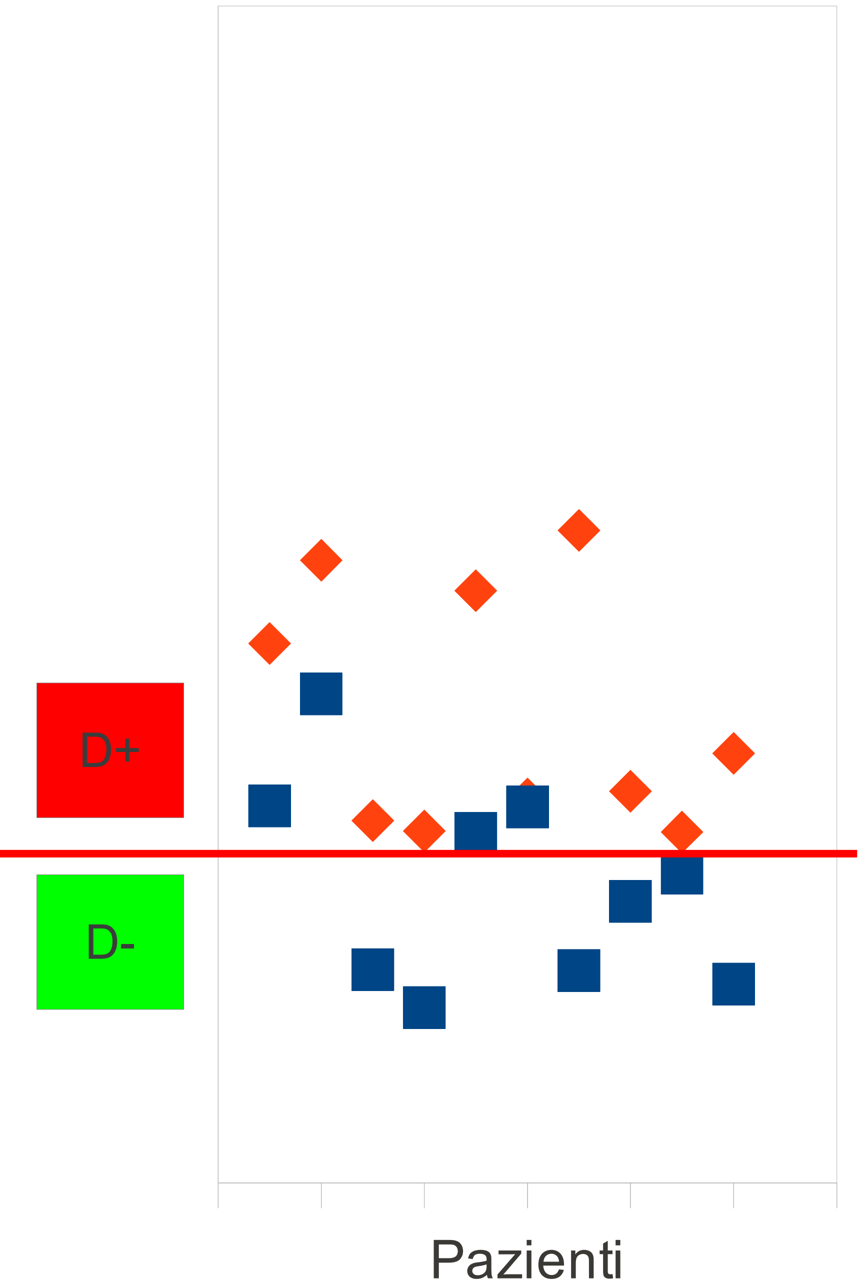
Falso positivo: la malattia NON è presente ma è diagnosticata
D+
D-
NO MALATTIA
MALATTIA
Pazienti


D+
D-
NO MALATTIA
MALATTIA
Pazienti
Falso negativo: la malattia è presente ma NON è diagnosticata



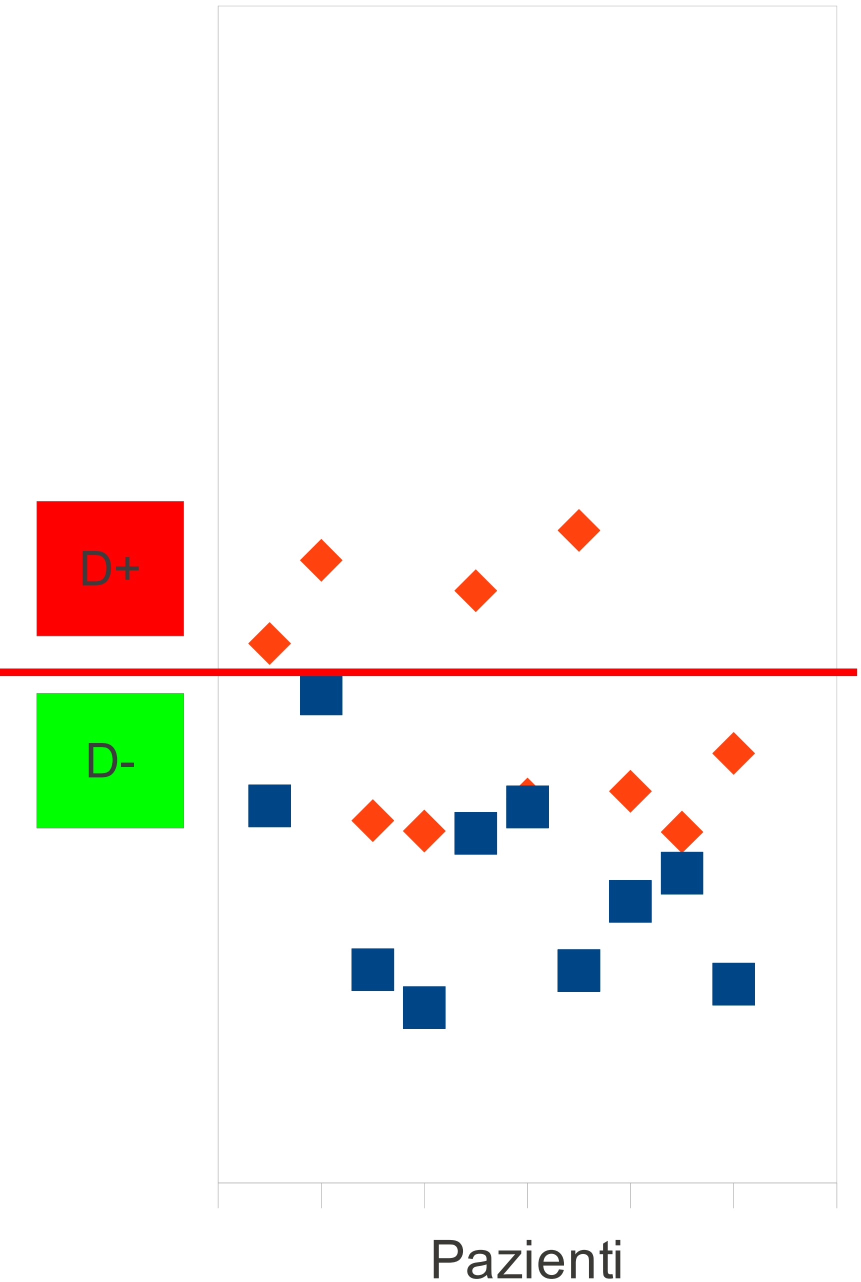
I possibili esiti di una decisione possono essere descritti nella seguente tabella:

	Malattia presente	Malattia assente
Malattia diagnosticata	CORRETTO (vero positivo)	ERRORE (falso positivo)
Malattia non diagnosticata	ERRORE (falso negativo)	CORRETTO (vero negativo)



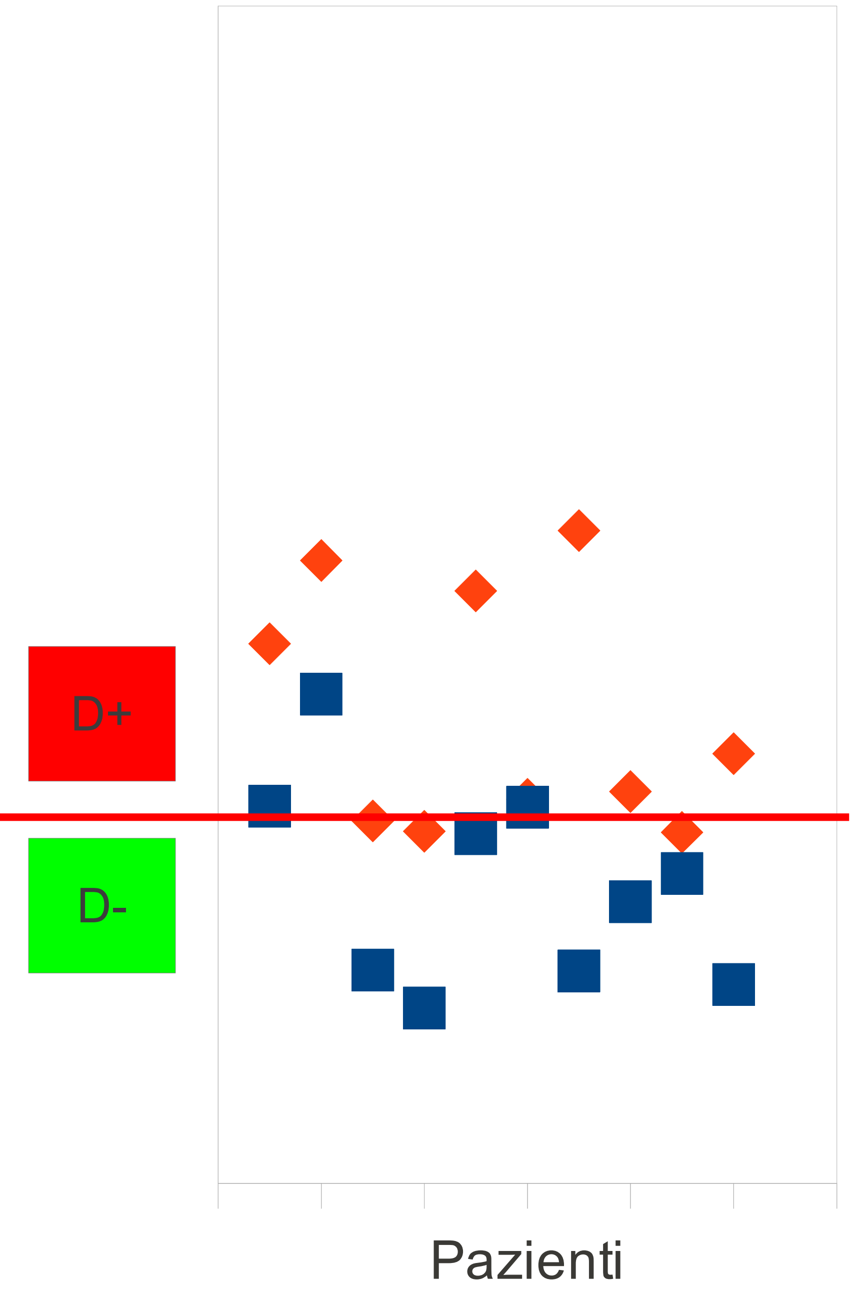
Falsi... ?

	Malattia presente	Malattia assente
Malattia diagnosticata	10	4
Malattia non diagnosticata	0	6



Falsi... ?

	Malattia presente	Malattia assente
Malattia diagnosticata	4	0
Malattia non diagnosticata	6	10

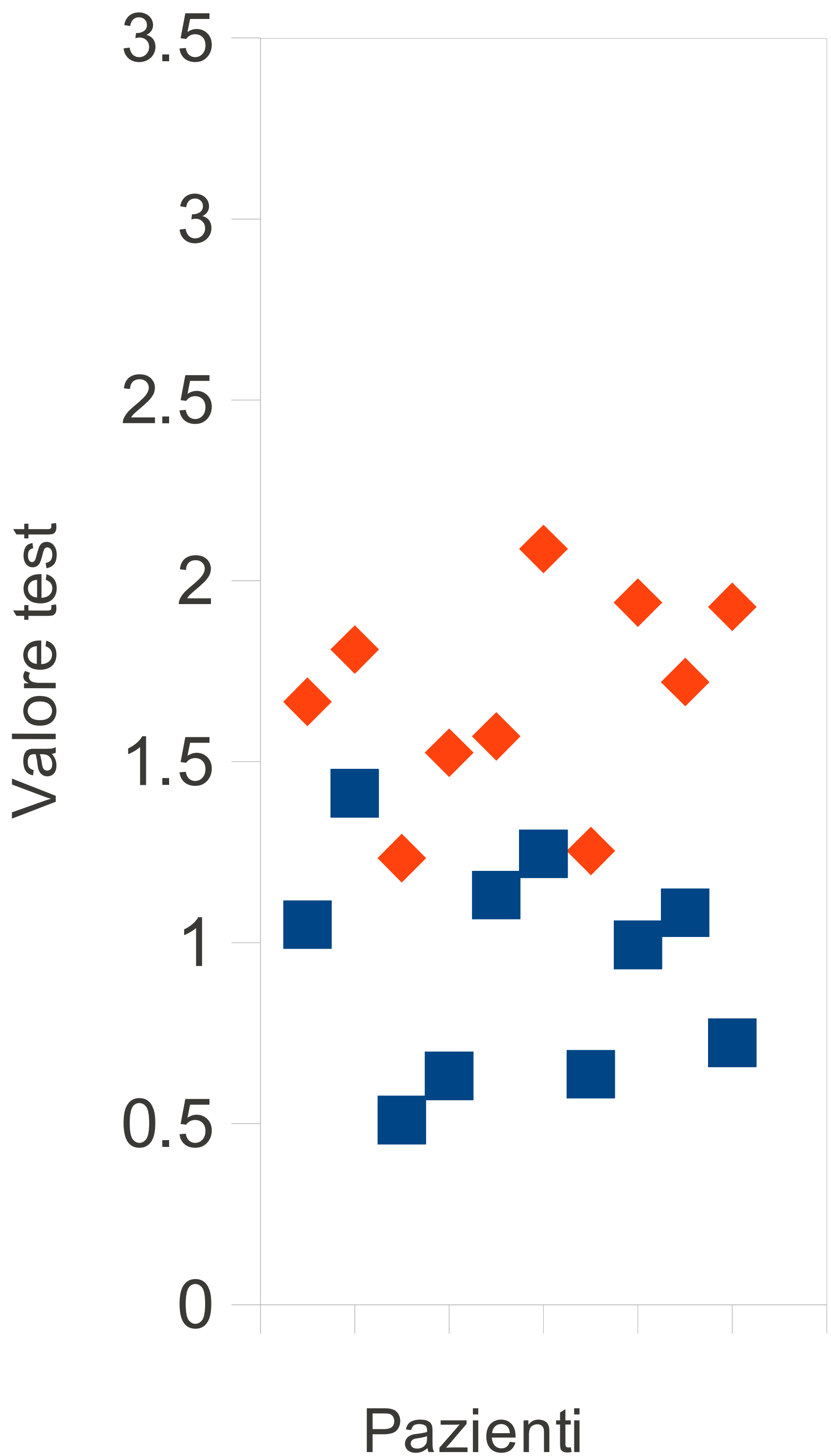


	Malattia presente	Malattia assente
Malattia diagnosticata	7	3
Malattia non diagnosticata	3	7

Cosa è <u>meglio</u> scegliere: FP o FN?

Cosa è meglio scegliere: FP o FN?

Benefici associati a VP	Benefici associati a VN	Costi associati a FP	Costi associati a FN
Una volta identificata la patologia, il paziente può ricevere la terapia Il beneficio è proporzionale alla gravità della malattia	Il paziente non soffre della patologia indagata Si indaga sull'esistenza di altre possibili patologie Si evitano gli effetti collaterali di una terapia errata Si risparmia il costo di una terapia non necessaria	Costo della terapia Effetti collaterali di una terapia	Il paziente soffre di una malattia e non riceve cure Importanza di una diagnosi precoce



Dove fissereste la soglia se...

Malattia poco grave e senza conseguenze nella popolazione di cui il paziente fa parte

Specificare il numero di VP, VN, FP, FN

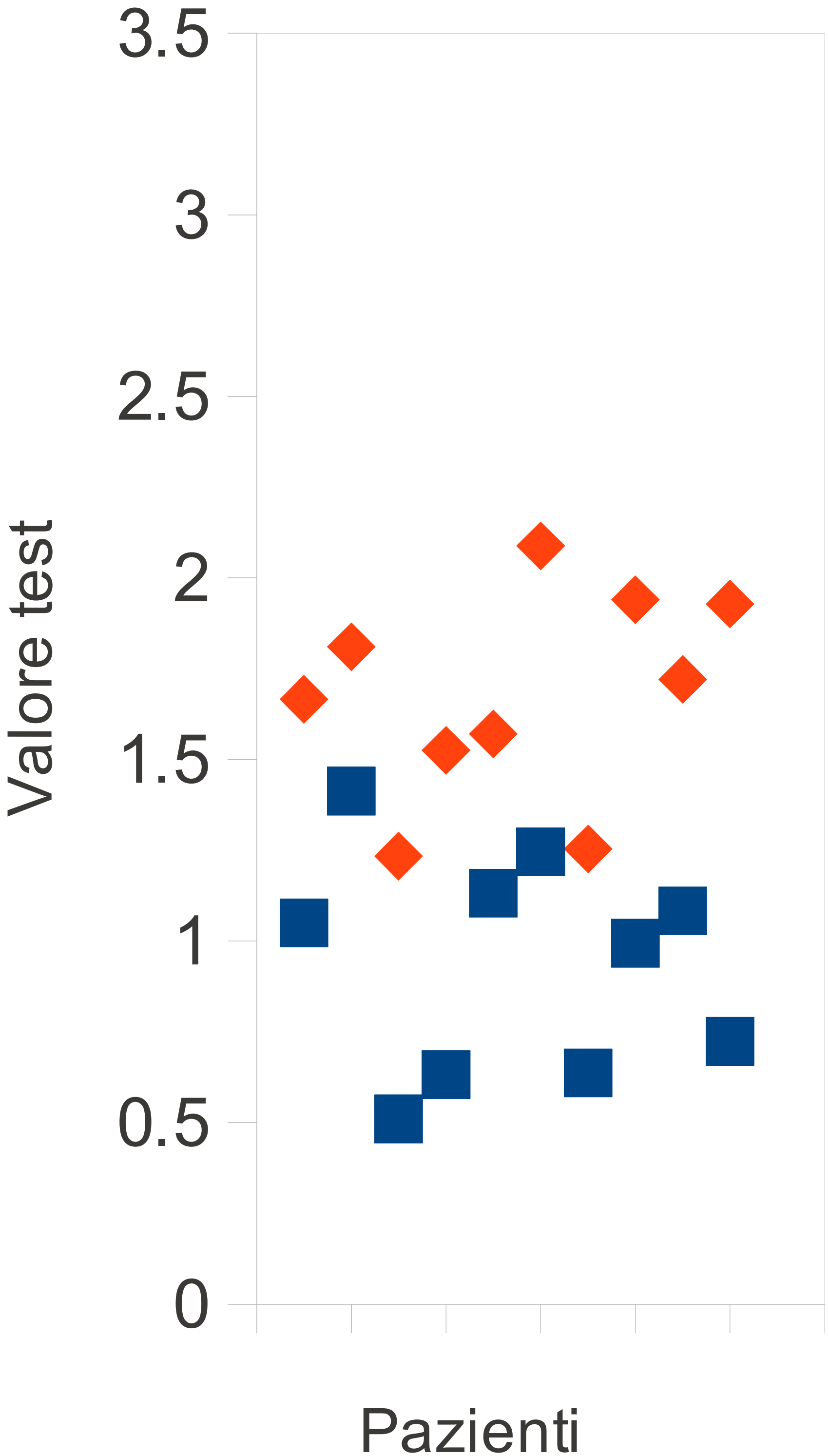
Soglia........................

VP............................

VN............................

FP............................

FN............................



Dove fissereste la soglia se...

Malattia ad alto rischio e potenzialmente contagiosa e mortale

Specificare il numero di VP, VN, FP, FN

Soglia.........................

VP.............................

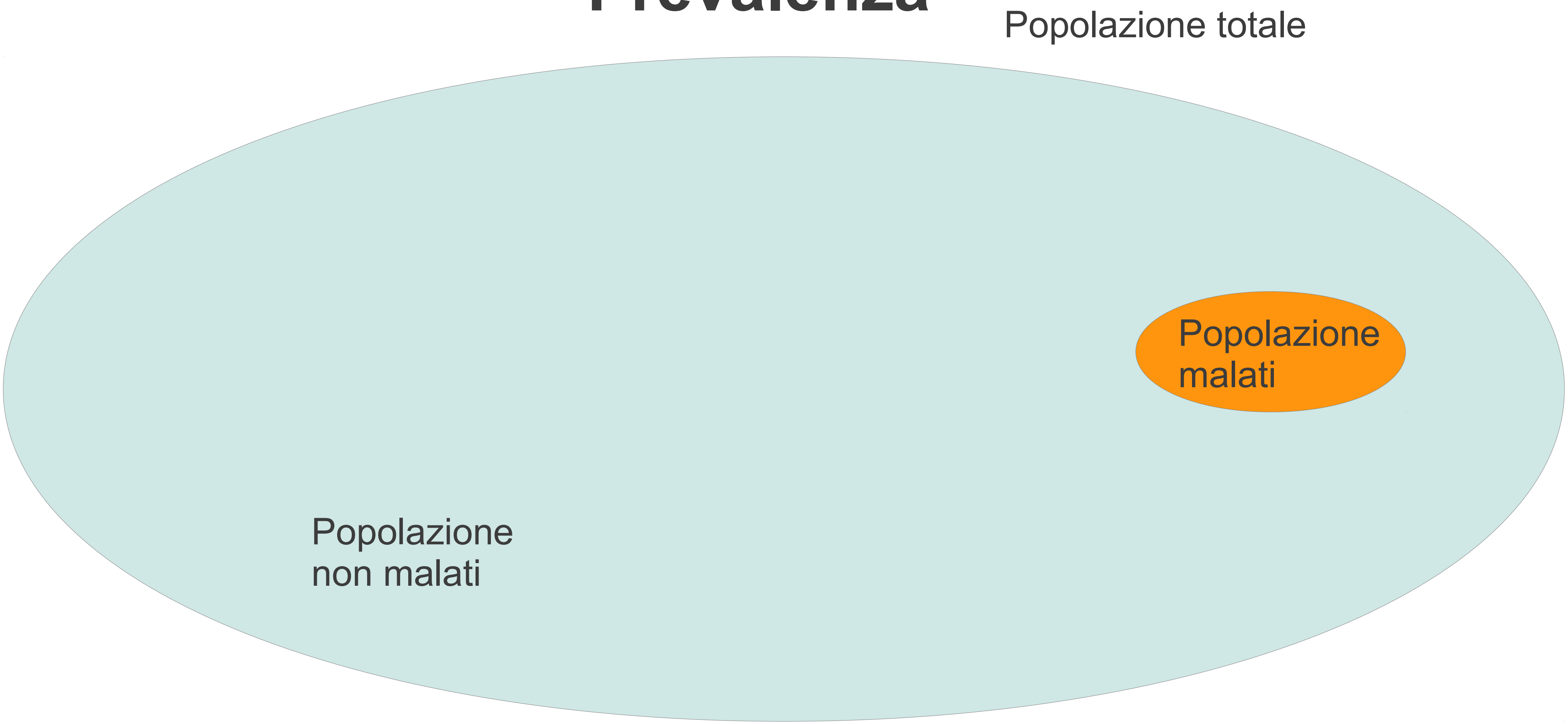
VN.............................

FP.............................

FN.............................

PREVALENZA

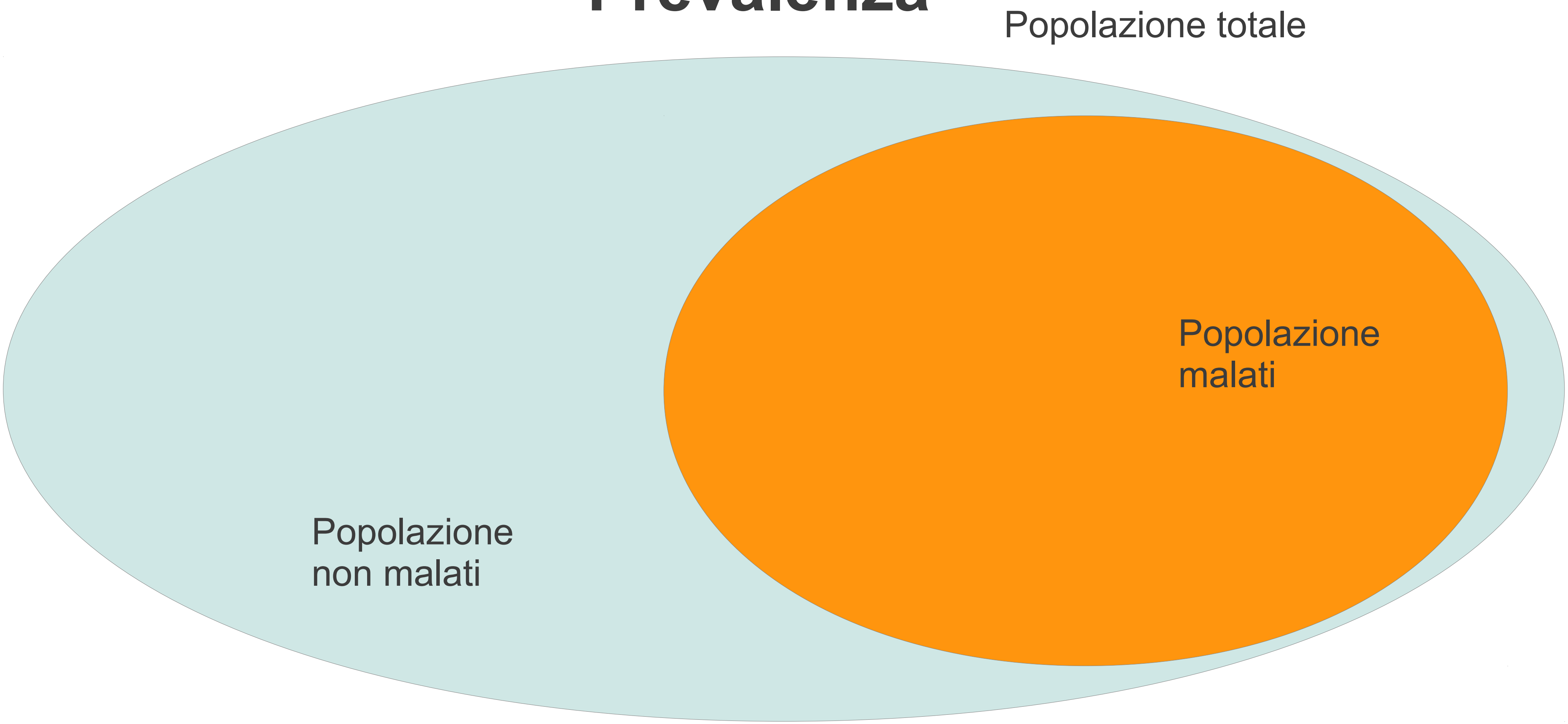
Prevalenza



Proporzione, nella popolazione totale testata, degli individui che hanno la malattia, ovvero:

$$P = \frac{\text{numero totale di malati}}{\text{popolazione totale (N)}}$$

Prevalenza



Proporzione, nella popolazione totale testata, degli individui che hanno la malattia, ovvero:

$$P = \frac{\text{numero totale di malati}}{\text{popolazione totale (N)}}$$

Proporzione, nella popolazione totale testata, degli individui che hanno la malattia

Quando si ha a che fare con la probabilità di una ipotesi diagnostica, è fondamentale individuare correttamente la popolazione a cui fare riferimento!

Ad esempio: esseri umani di sesso maschile, italiani, abitanti dell'Emilia Romagna, persone seguite da un centro, ...

Ad esempio:

probabilità di aver contratto il virus HIV nella popolazione totale = 0.1%

probabilità di aver contratto il virus HIV tra le persone seguite in un centro di disintossicazione = 18%

Alcune popolazioni possono avere valori di prevalenza diversi da altre, e non individuare la popolazione corretta porta a una stima errata delle probabilità della presenza di una malattia

Quando si considera la prevalenza?

paziente

test

decisione

Quando si considera la prevalenza?

paziente	Caratteristiche della popolazione a cui il paziente appartiene (es., prevalenza)
test	Risultato del test (positivo o negativo)
decisione	

Quando si considera la prevalenza?

paziente

Caratteristiche della popolazione a cui il paziente appartiene (es., prevalenza)

test

Risultato del test (positivo o negativo)

decisione

Corretto: la decisione è presa sulla base delle caratteristiche della popolazione a cui il paziente appartiene e al risultato del/dei test condotti

Quando si considera la prevalenza?

paziente

Caratteristiche della popolazione a cui il paziente appartiene (es., prevalenza)

test

Risultato del test (positivo o negativo)

decisione

È un grave errore utilizzare l'informazione di prevalenza nel determinare il risultato di un test!

EFFETTO DELLA PREVALENZA SULLA DECISIONE MEDICA

Evans, Tambouret, Evered, Wilbur, Wolfe, 2011

Effetto della prevalenza su FP / FN

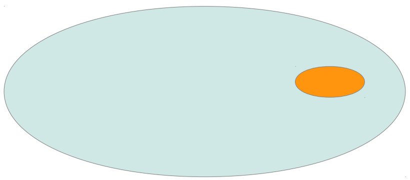
Campione di 10 esperti

Screening del cancro cervicale

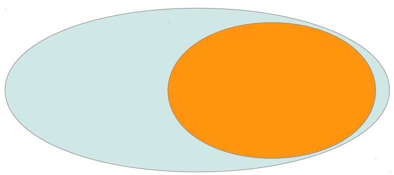
Stimoli: preparati citologici

Due condizioni: popolazioni di cellule con...

...prevalenza BASSA (patologia rara)



...prevalenza ALTA   (patologia comune)

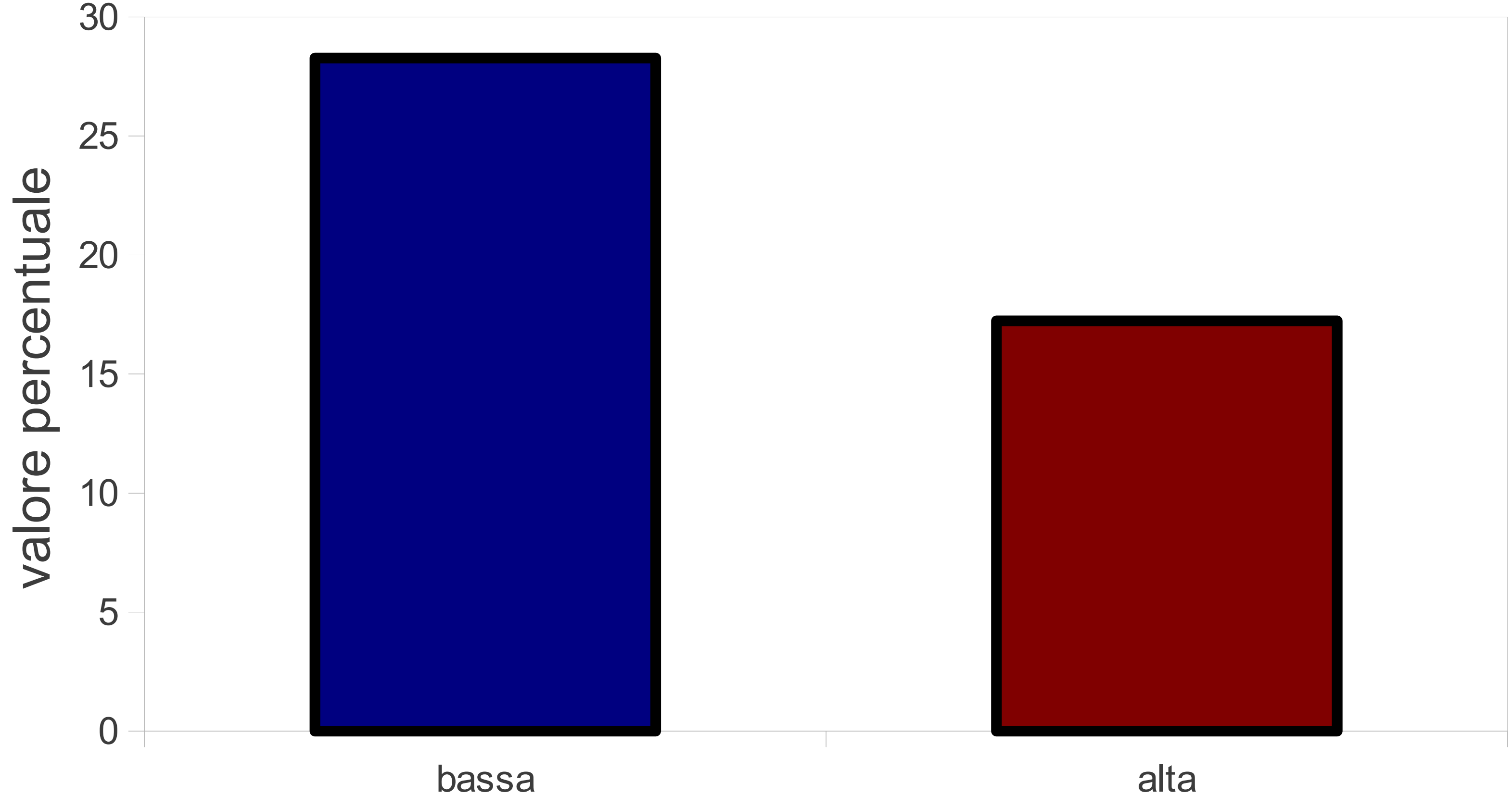


Falsi Negativi

Patologia presente ma NON diagnosticata

Falsi Negativi

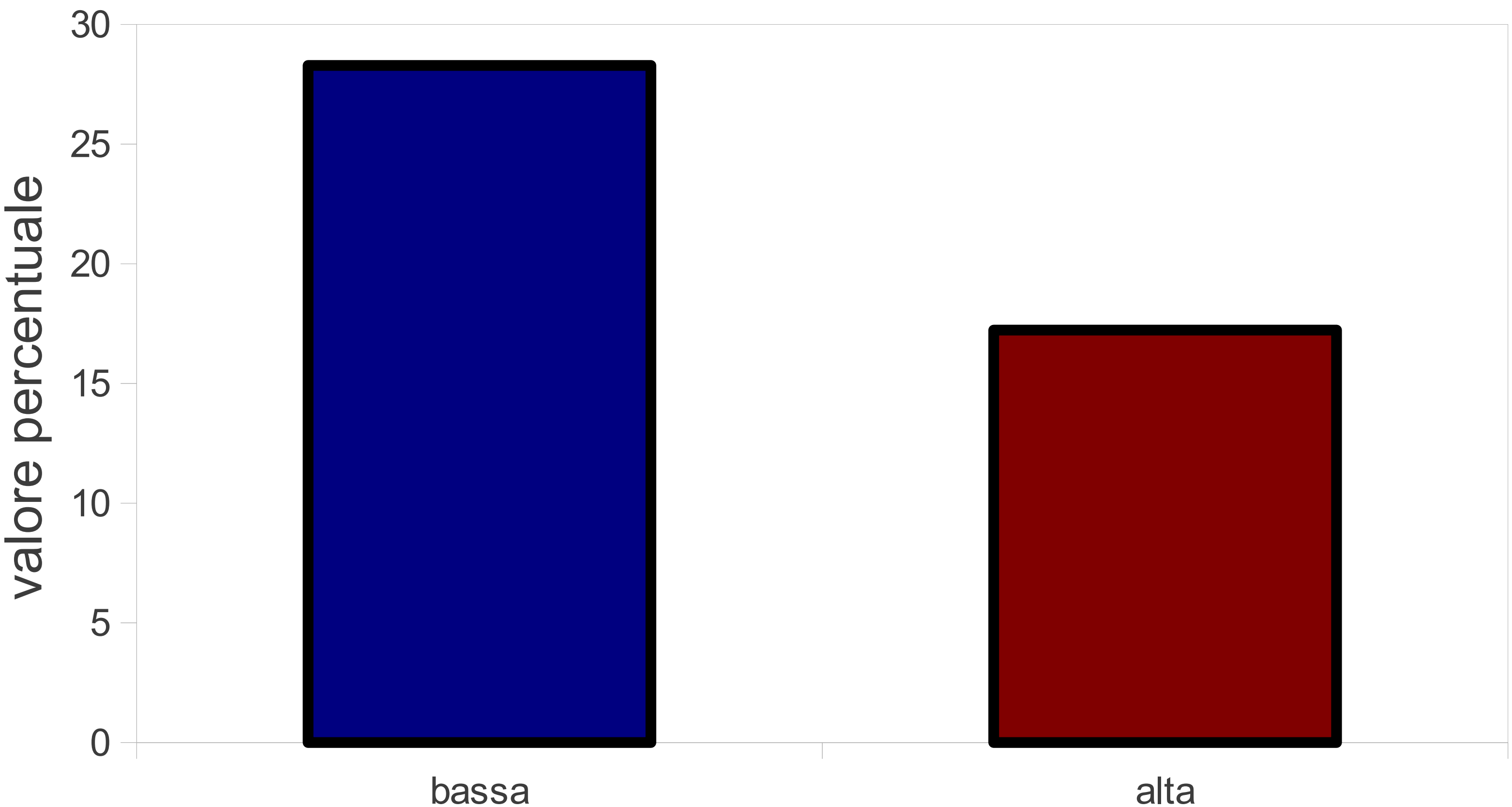
Patologia presente ma NON diagnosticata



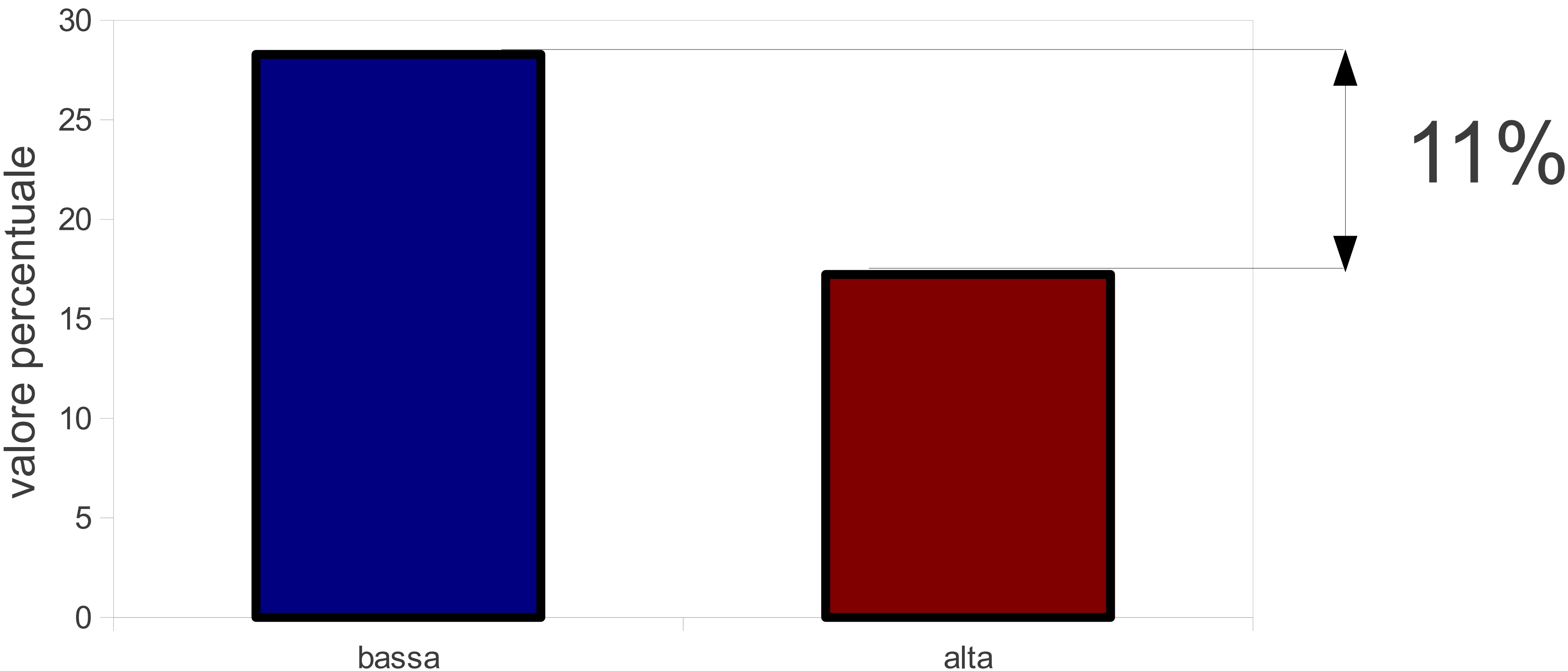
prevalenza BASSA (patologia rara)

prevalenza ALTA (patologia comune)

Quando la prevalenza di una patologia è BASSA aumenta la probabilità (rispetto alla condizione di ALTA prevalenza) che si commetta un FALSO NEGATIVO: la patologia **è presente** ma non viene rilevata.



Quando la prevalenza di una patologia è BASSA aumenta la probabilità (rispetto ala condizione di ALTA prevalenza) che si commetta un FALSO NEGATIVO: la patologia è presente ma non viene rilevata.



Sommario

Falsi positivi e Falsi negativi

Prevalenza

Effetti della prevalenza sulla decisione medica

Sensibilità

Specificità

Uso delle informazioni sulla sensibilità e specificità nel processo diagnostico

Un problema tipico

Spiegazione grafica

Il teorema di Bayes

Errori di ragionamento

Generalizzazione

Sommario

Falsi positivi e Falsi negativi

Prevalenza

Effetti della prevalenza sulla decisione medica

Sensibilità

Specificità

Uso delle informazioni sulla sensibilità e specificità nel processo diagnostico

Un problema tipico

Spiegazione grafica

Il teorema di Bayes

Errori di ragionamento

Generalizzazione

Cosa è <u>meglio</u> scegliere: FP o FN?

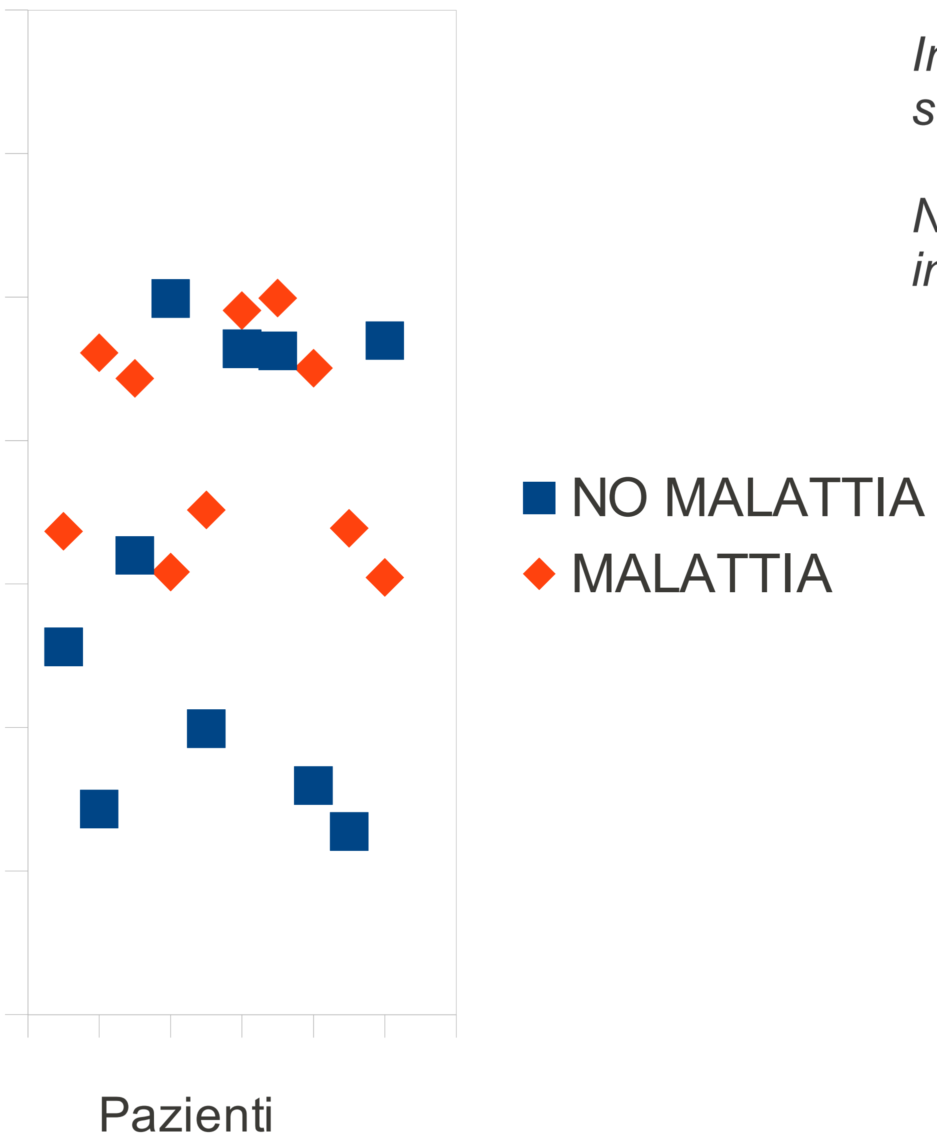
Benefici associati a VP
Benefici associatia VN
Costi associati a FP
Costi associati a FN

È <u>possibile</u> scegliere?

Sensibilità
Specificità

SENSIBILITÀ

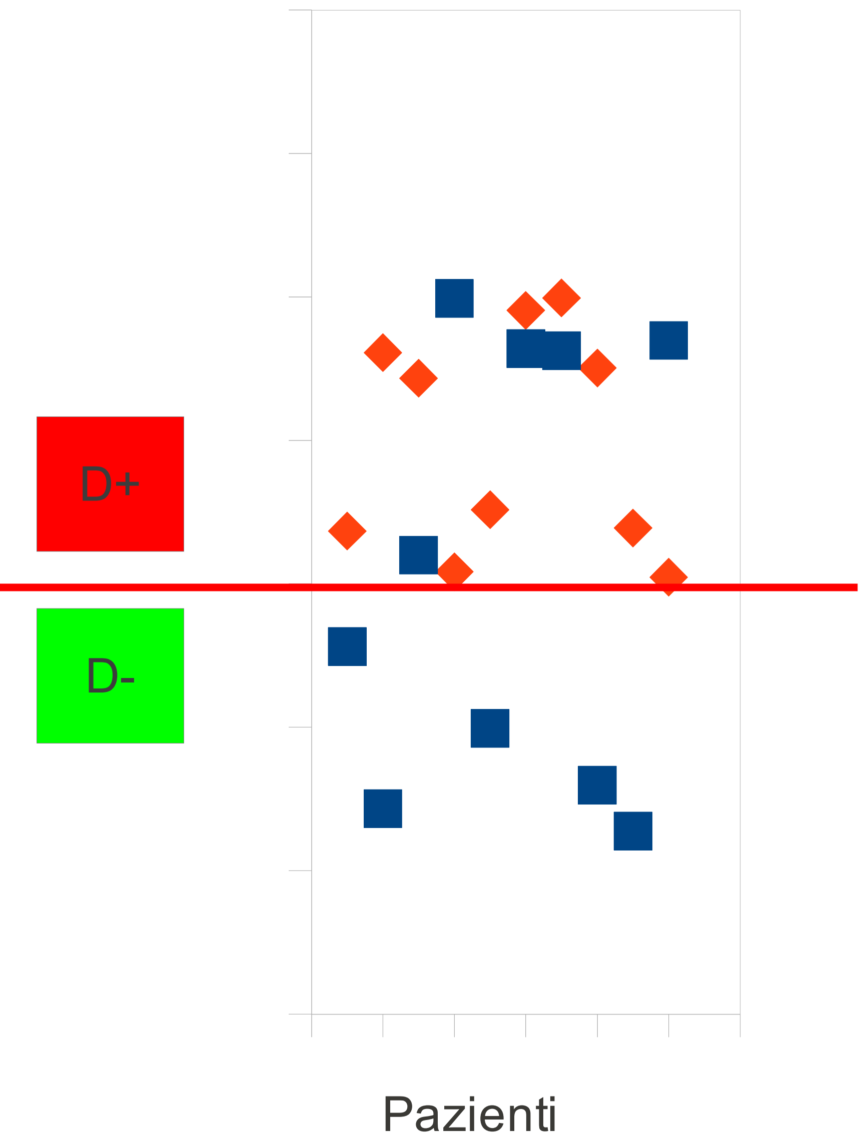
Sensibilità



In questo caso, i due gruppi non sono chiaramente distinti

Non è possibile evitare di individuare falsi positivi

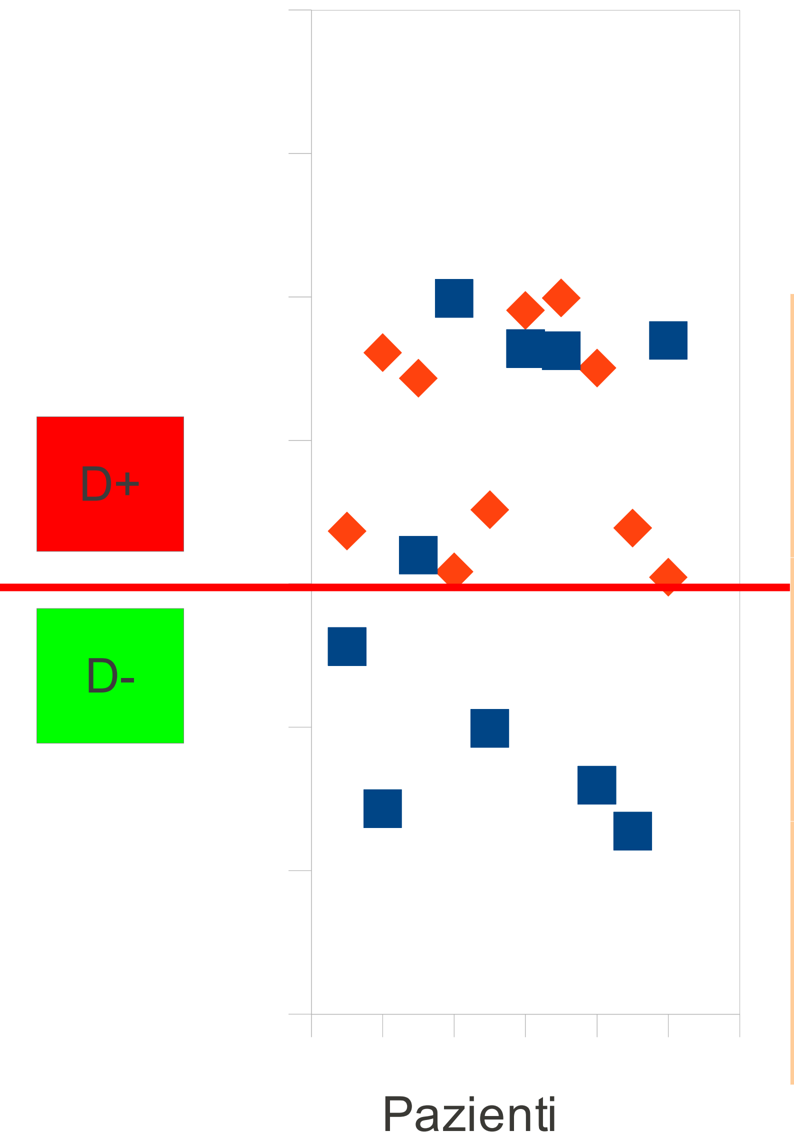
Sensibilità



In questo caso, i due gruppi non sono chiaramente distinti

Non è possibile evitare di individuare falsi positivi

Sensibilità

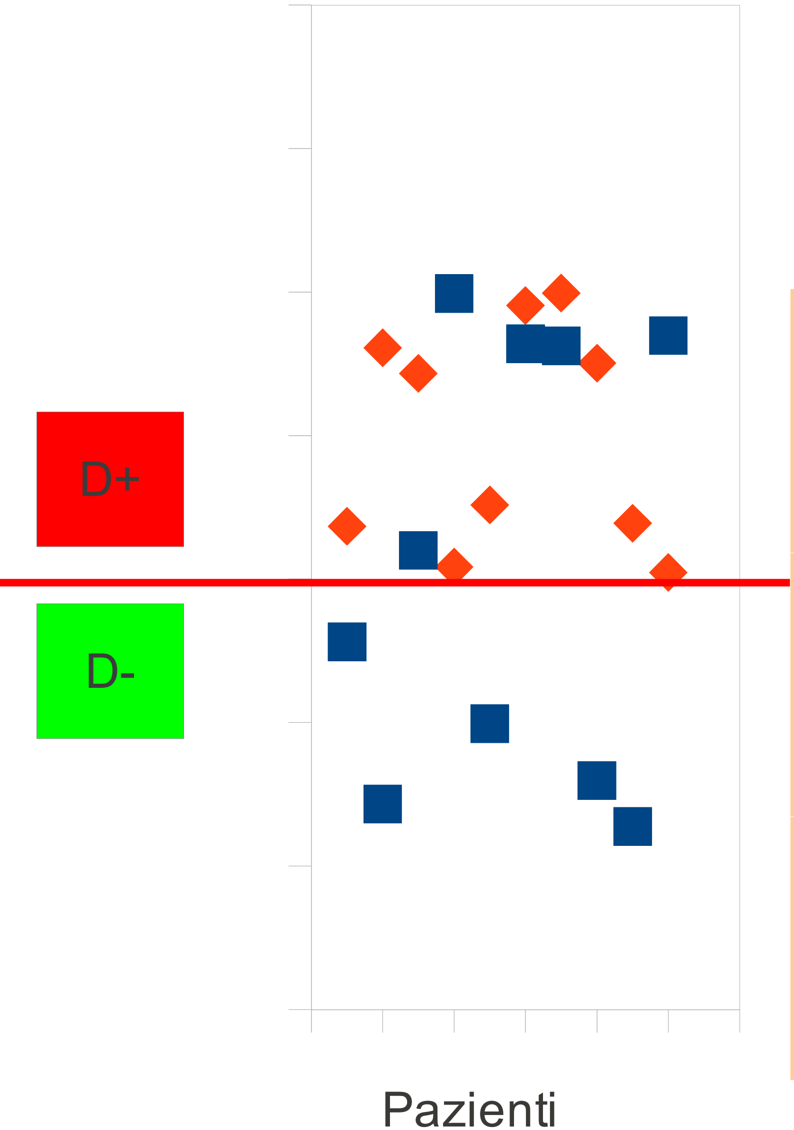


In questo caso, i due gruppi non sono chiaramente distinti

Non è possibile evitare di individuare falsi positivi

	Malattia presente	Malattia assente
Malattia diagnosticata	10	5
Malattia non diagnosticata	0	5

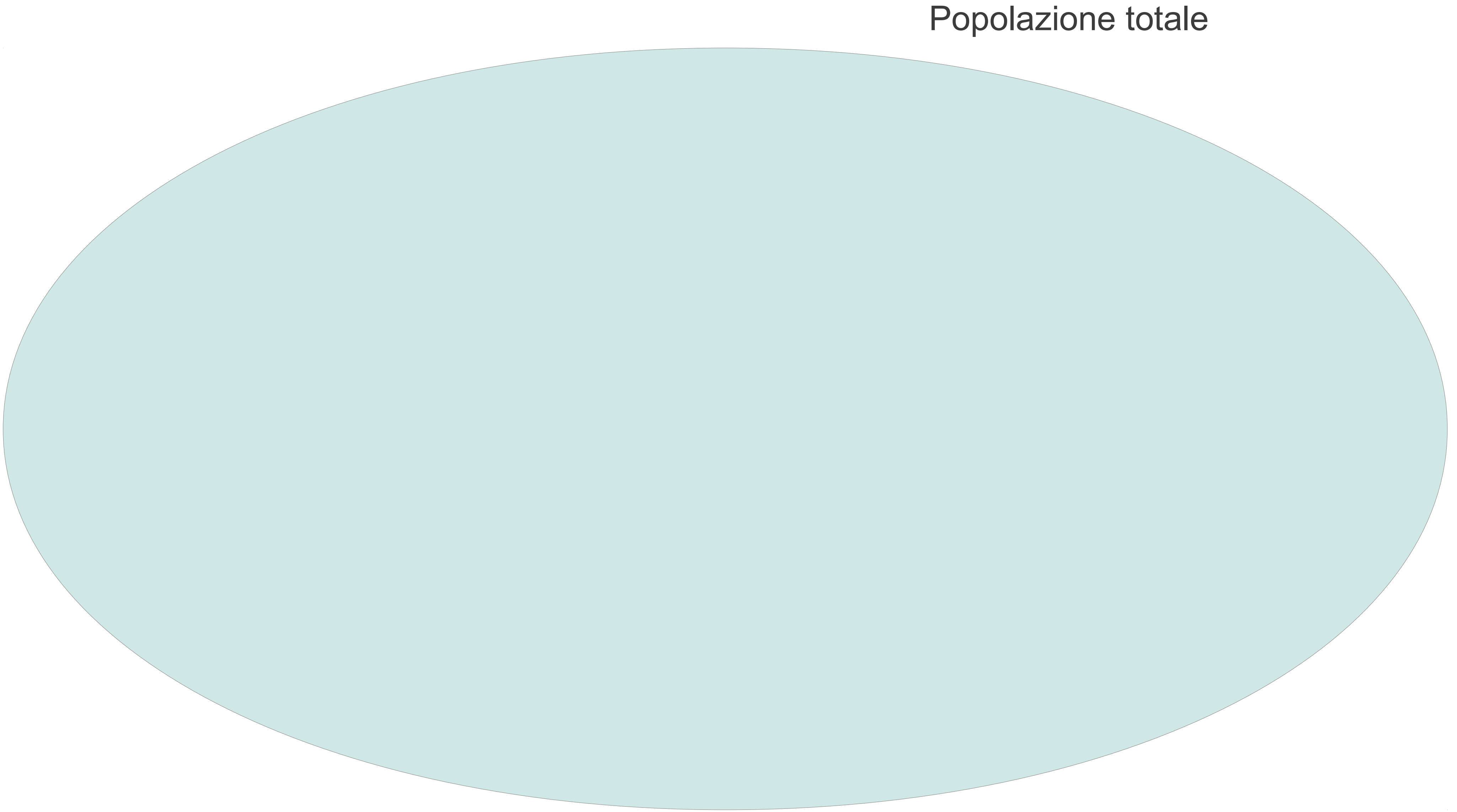
Sensibilità



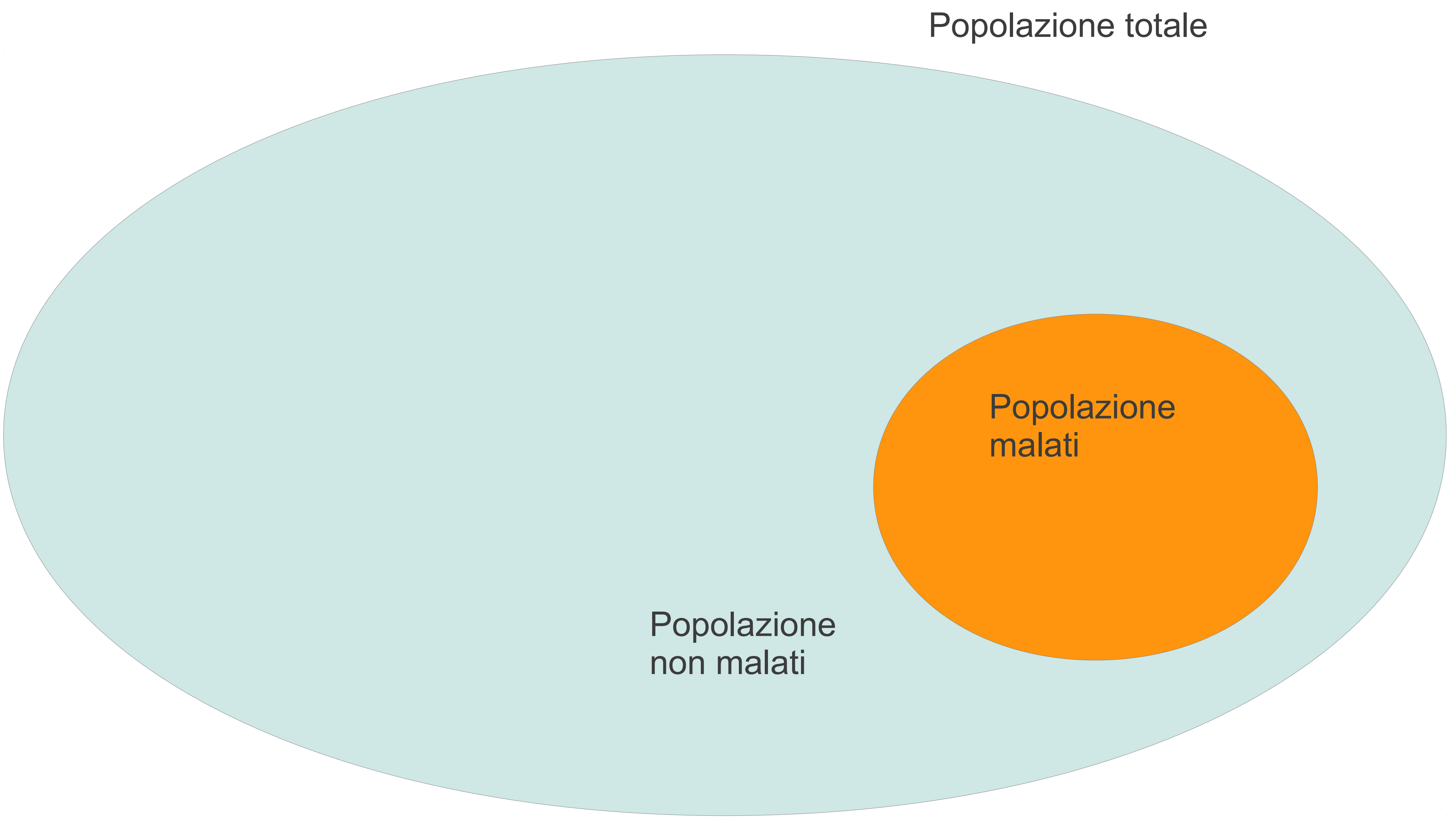
	Malattia presente	Malattia assente
Malattia diagnosticata	10	5
Malattia non diagnosticata	0	5

Tutti gli individui con patologia sono positivi al test
Alcuni individui senza patologia sono positivi al test (FALSI POSITIVI)

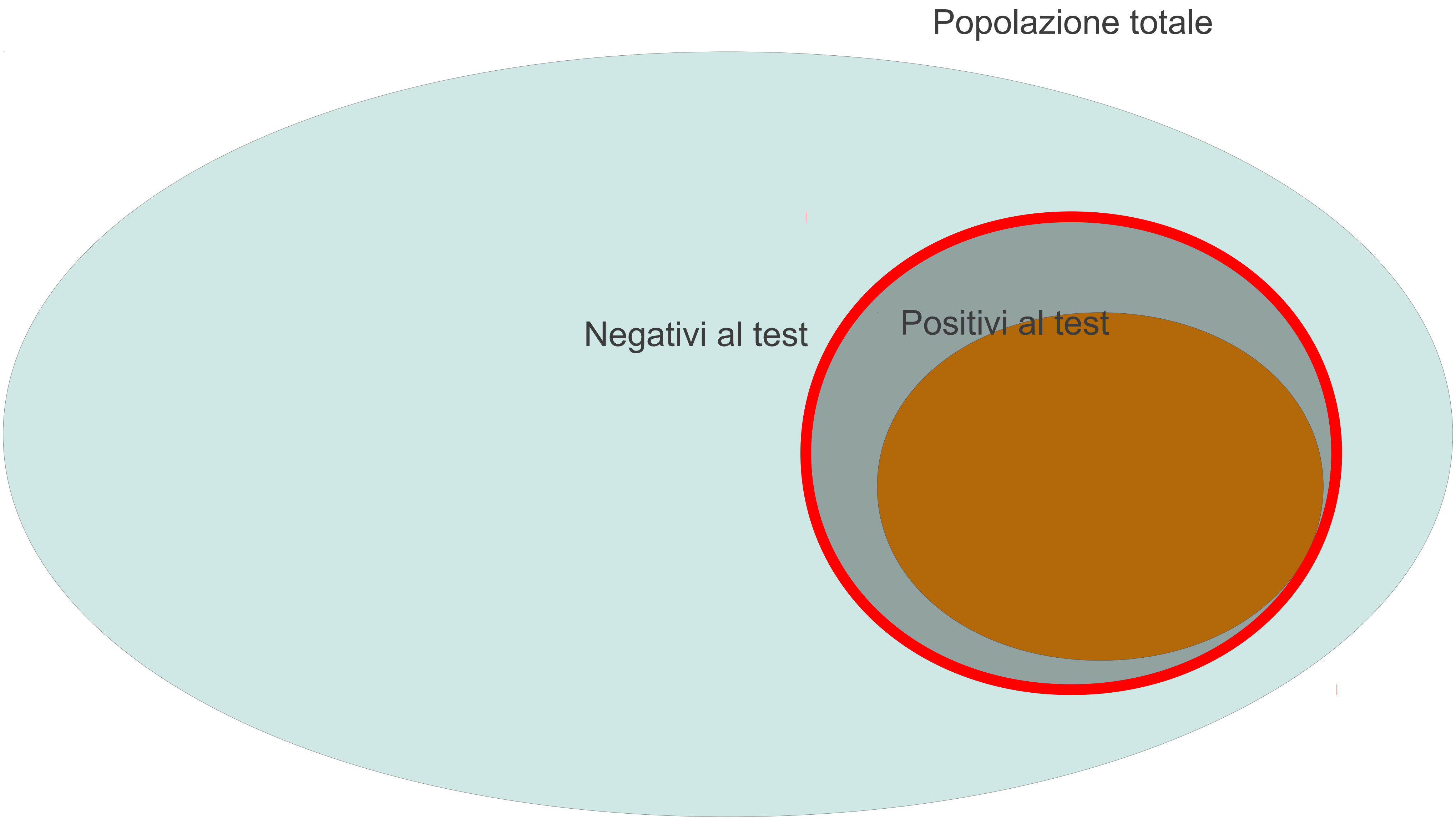
Il rapporto tra individui sani, malati, e l'intera popolazione può essere rappresentato nel modo seguente:



Data la popolazione totale degli individui



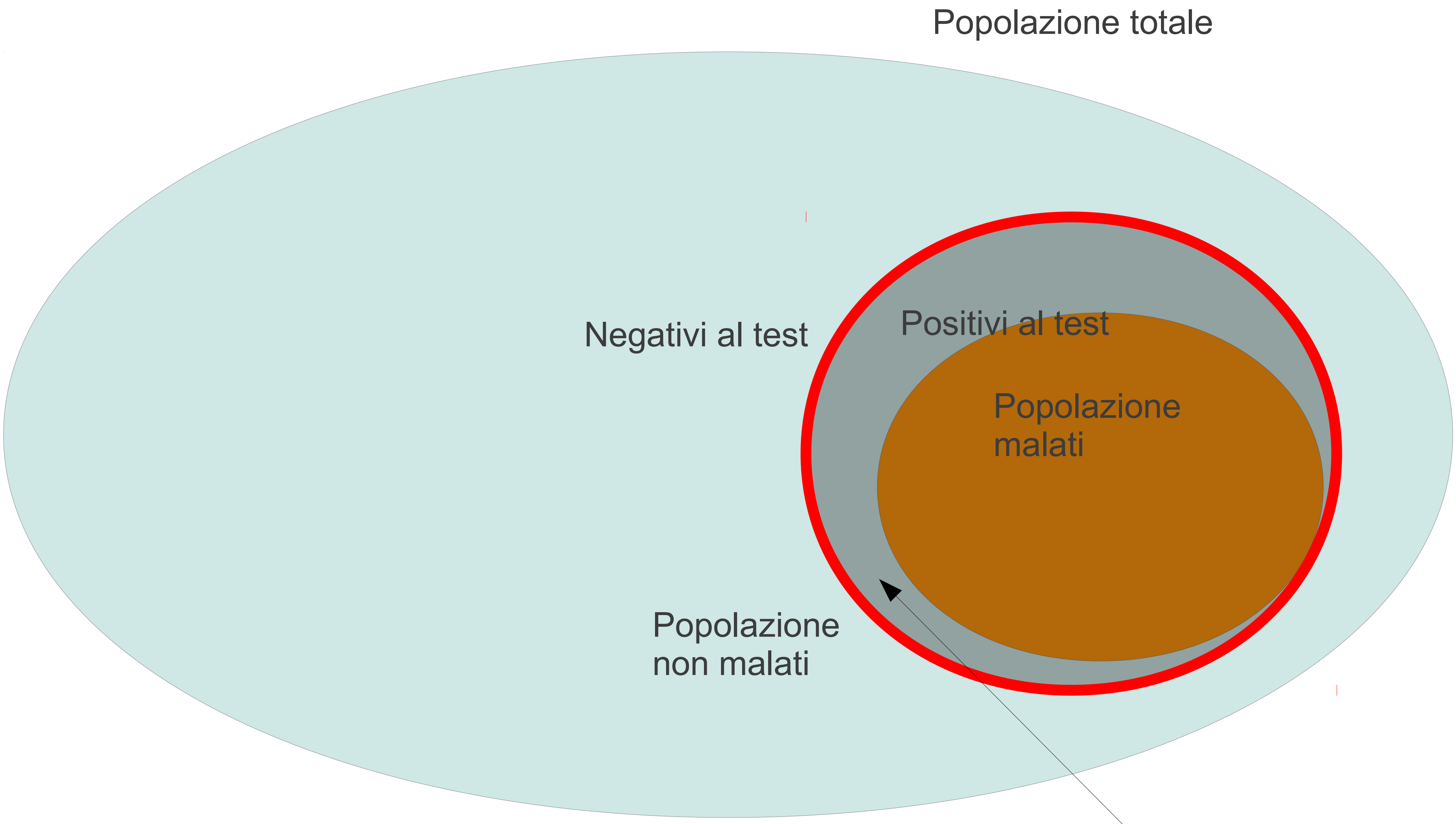
Data la popolazione totale degli individui, che comprende sia individui sani che malati



Data la popolazione totale degli individui, che comprende sia individui sani che malati

Il cerchio rosso rappresenta gli individui positivi al test

Sensibilità



Tutti gli individui con patologia sono positivi al test
Alcuni individui senza patologia sono positivi al test (FALSI POSITIVI)

Un test si definisce SENSIBILE quando rileva una alta percentuale di individui con patologia (VERI POSITIVI)

SENSIBILITÀ = proporzione di individui con malattia che ha un risultato positivo del test

In questo caso, 100% (tutti gli individui con malattia sono individuati dal risultato del test)

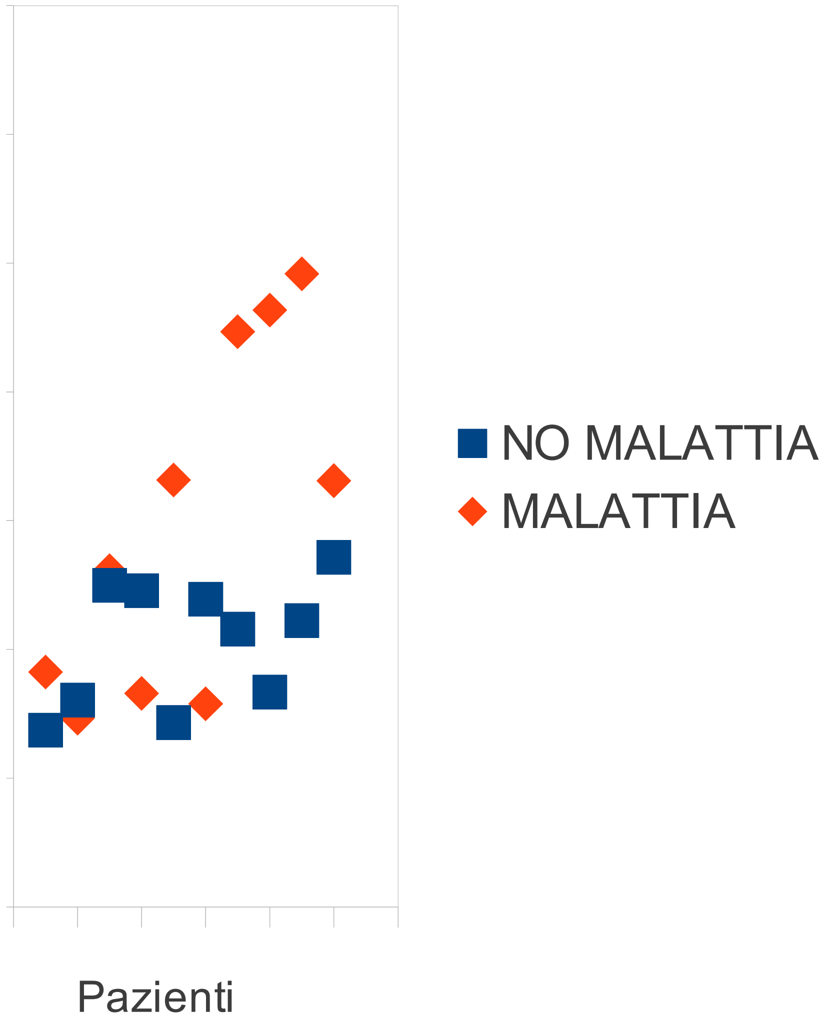
$$\text{Sensibilità} = \frac{VP}{VP + FN}$$

SPECIFICITÀ

Specificità

In questo caso, i due gruppi non sono chiaramente distinti

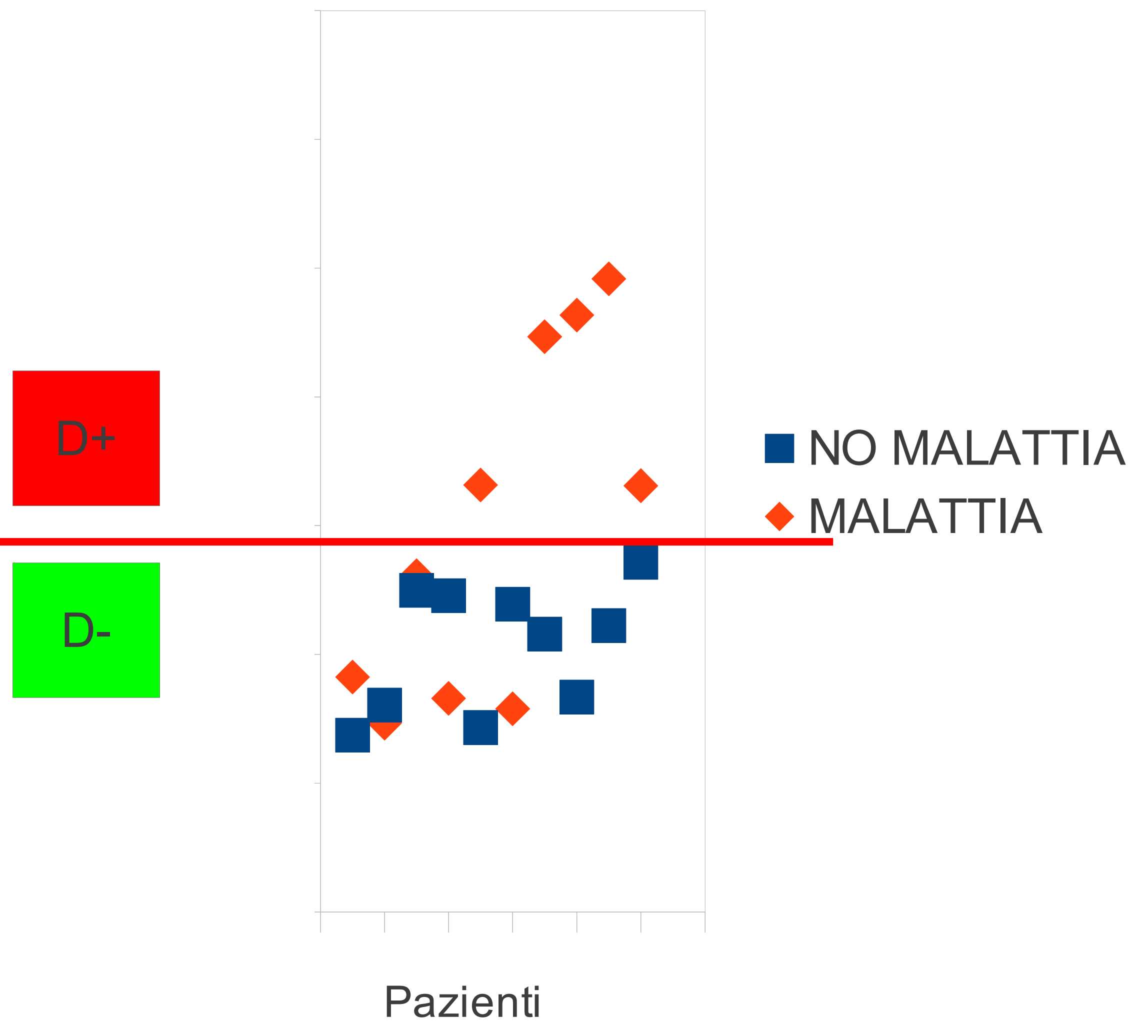
Non è possibile evitare di individuare falsi negativi



Specificità

In questo caso, i due gruppi non sono chiaramente distinti

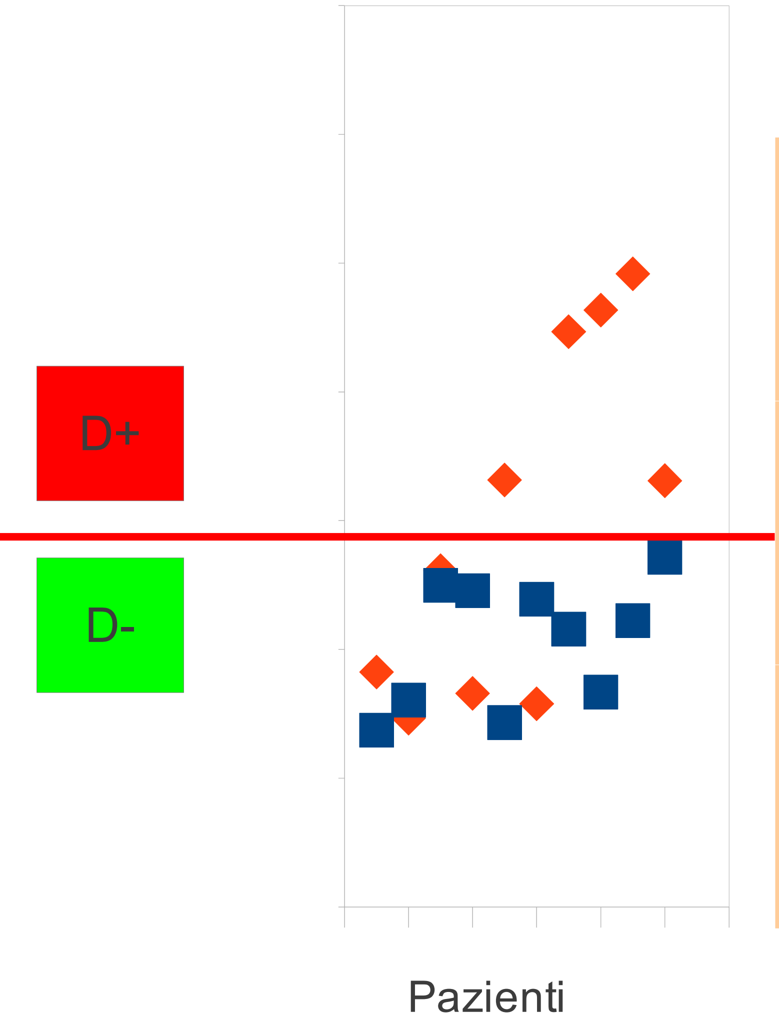
Non è possibile evitare di individuare falsi negativi



Specificità

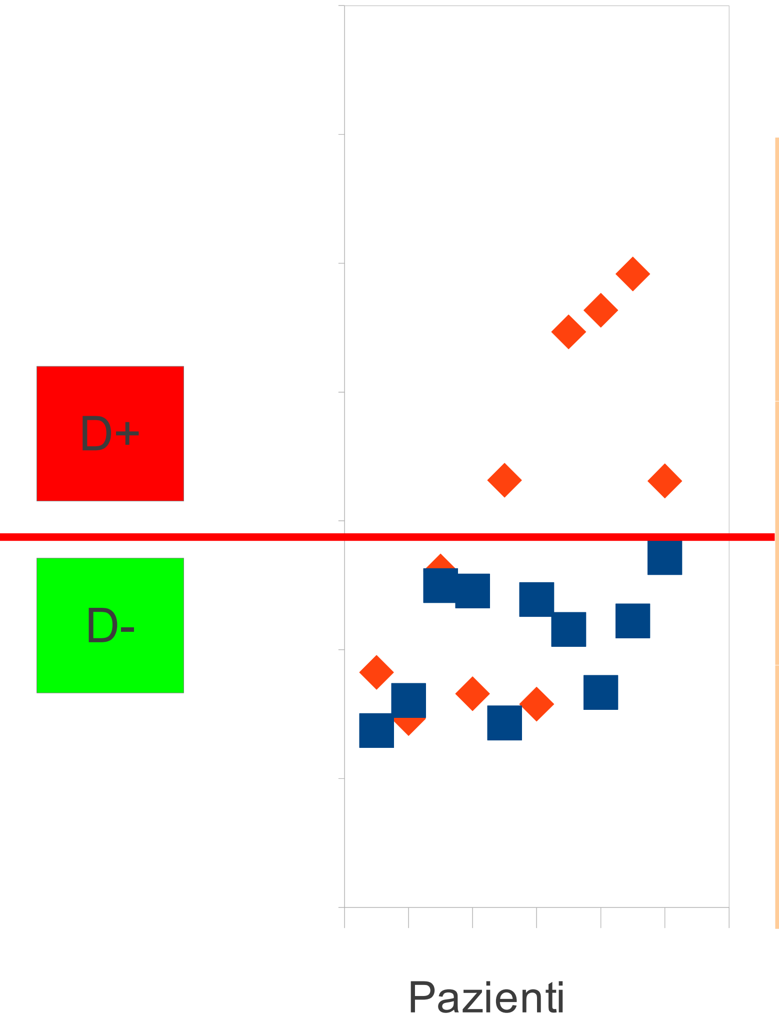
In questo caso, i due gruppi non sono chiaramente distinti

Non è possibile evitare di individuare falsi negativi



	Malattia presente	Malattia assente
Malattia diagnosticata	5	0
Malattia non diagnosticata	5	10

Specificità

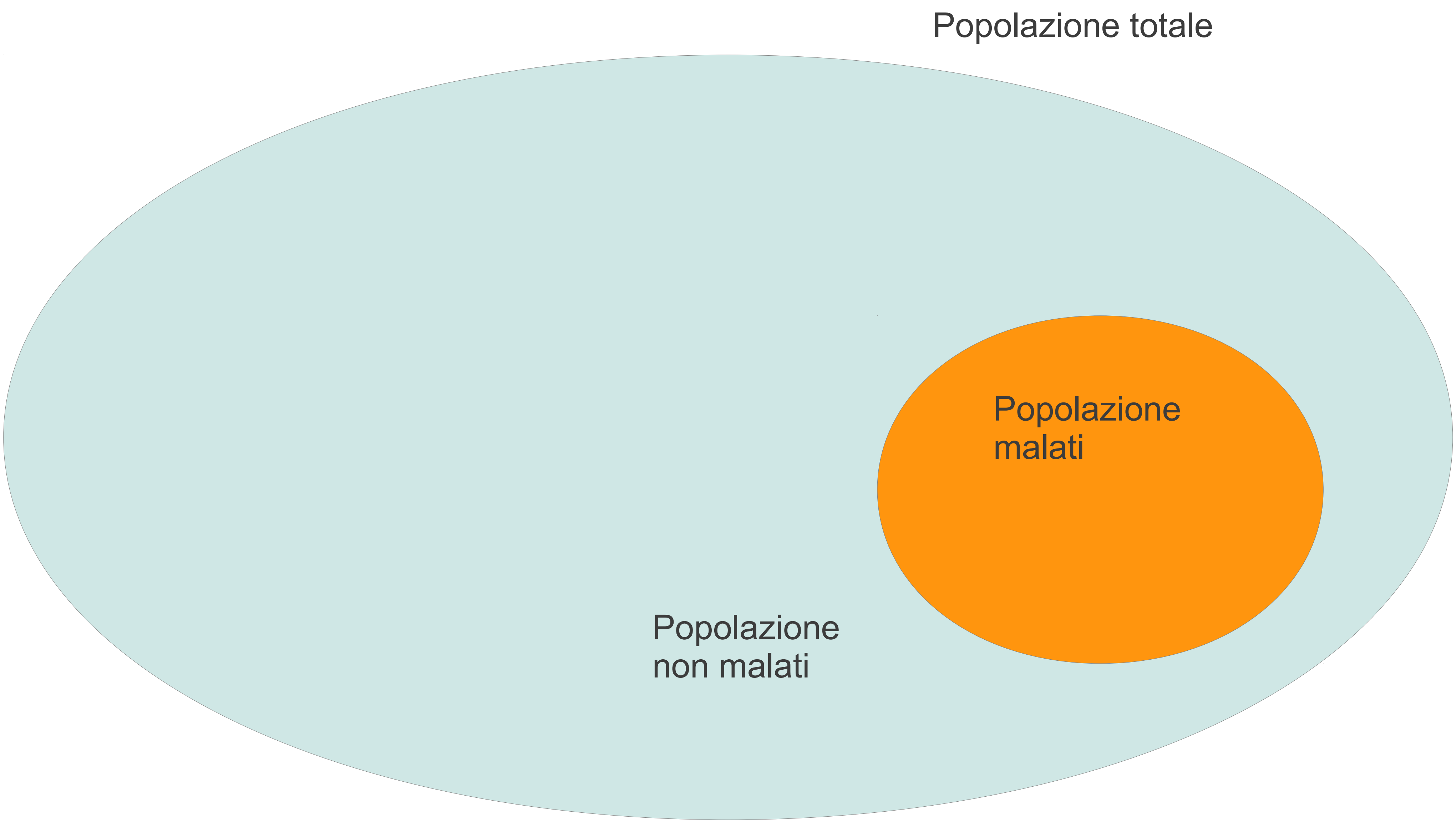


	Malattia presente	Malattia assente
Malattia diagnosticata	5	0
Malattia non diagnosticata	5	10

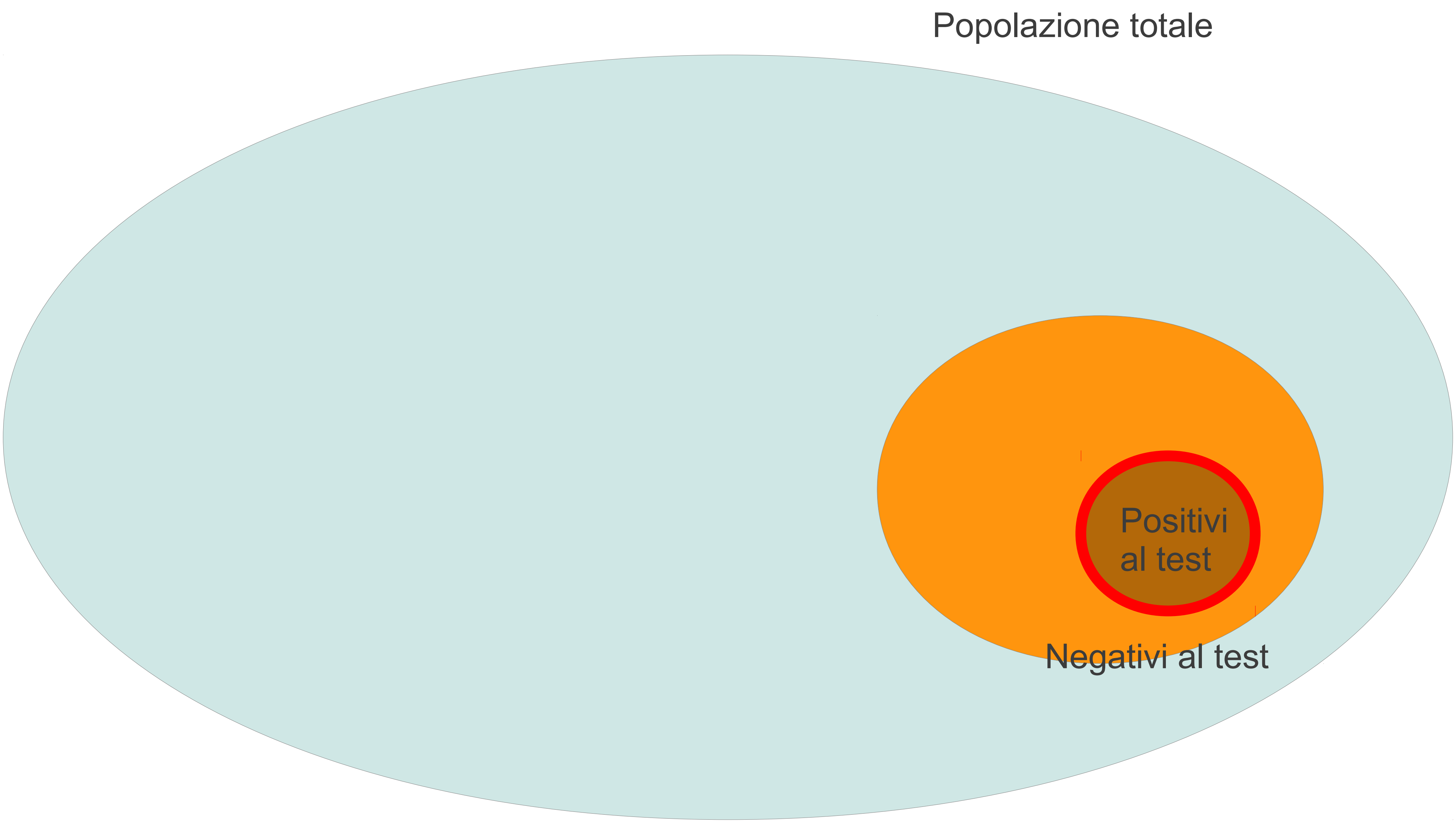
Tutti gli individui positivi al test presentano la patologia
Alcuni individui con patologia sono negativi al test (FALSI NEGATIVI)

Popolazione totale

Data la popolazione totale



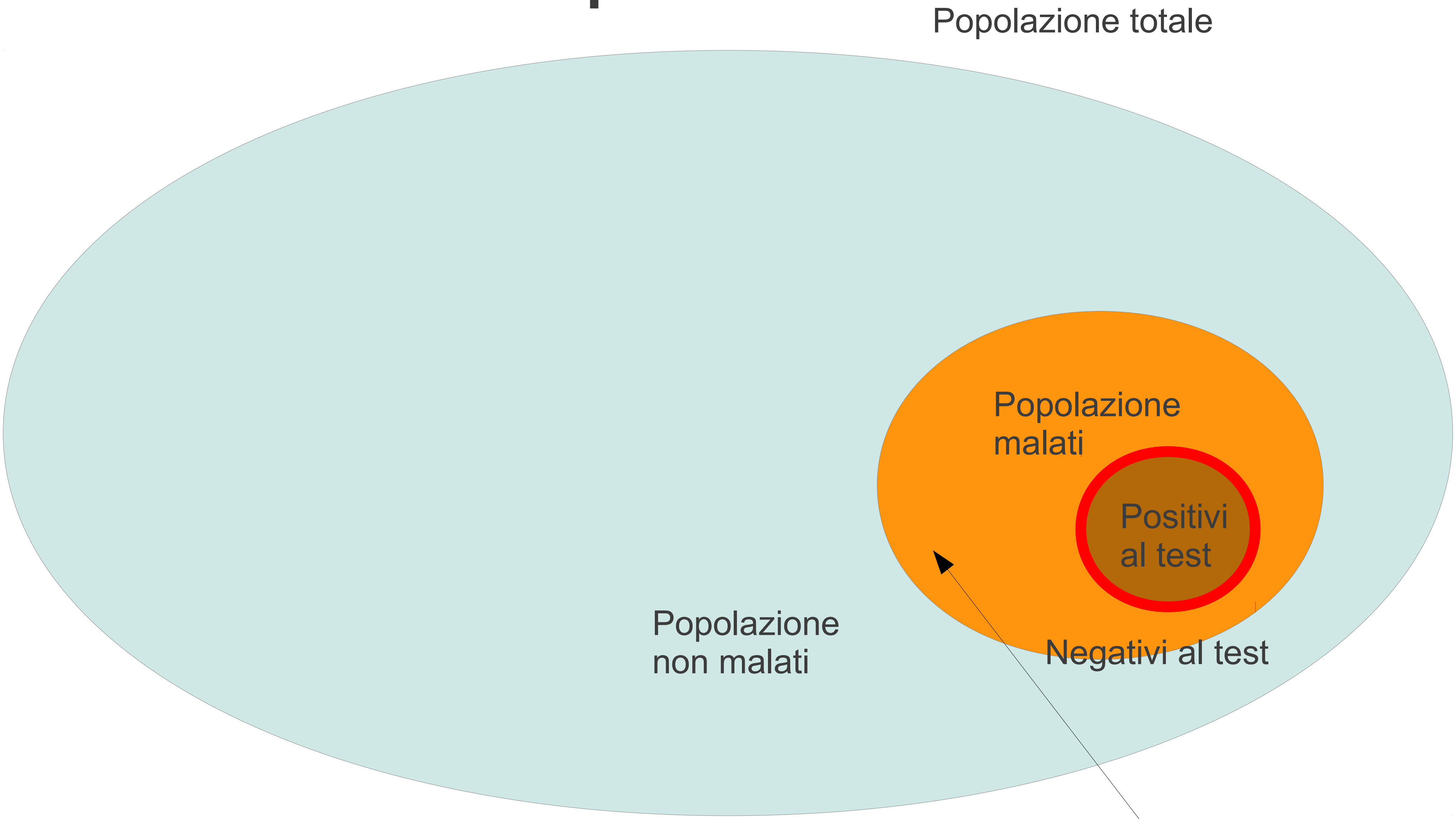
Data la popolazione totale, che comprende sia individui sani che individui con malattia,



Data la popolazione totale, che comprende sia individui sani che individui con malattia,

Il cerchio rosso rappresenta gli individui positivi al test

Specificità



Solo gli individui con patologia sono positivi al test
Alcuni individui con patologia sono negativi al test (FALSI NEGATIVI)

Un test si definisce SPECIFICO quando indica come sani una alta percentuale di individui senza patologia (VERI NEGATIVI)

SPECIFICITÀ = proporzione di individui senza malattia che ha un risultato negativo del test

In questo caso, 100% (tutti gli individui indicati come sani non sono malati)

$$\text{Specificità} = \frac{VN}{VN + FP}$$

È NECESSARIO CONOSCERE I VALORI DI SENSIBILITÀ E SPECIFICITÀ DI UN TEST

SENSIBILITÀ = proporzione di individui con malattia che ha un risultato positivo del test

	Malattia presente	Malattia assente
Malattia diagnosticata	10	5
Malattia non diagnosticata	0	5

$$\frac{10 \text{ (positivi al test)}}{10 \text{ (malattia presente)}} = 100\%$$

SENSIBILITÀ = proporzione di individui con malattia che ha un risultato positivo del test

	Malattia presente	Malattia assente
Malattia diagnosticata	10	5
Malattia non diagnosticata	0	5

$$\frac{10 \text{ (positivi al test)}}{10 \text{ (malattia presente)}} = 100\%$$

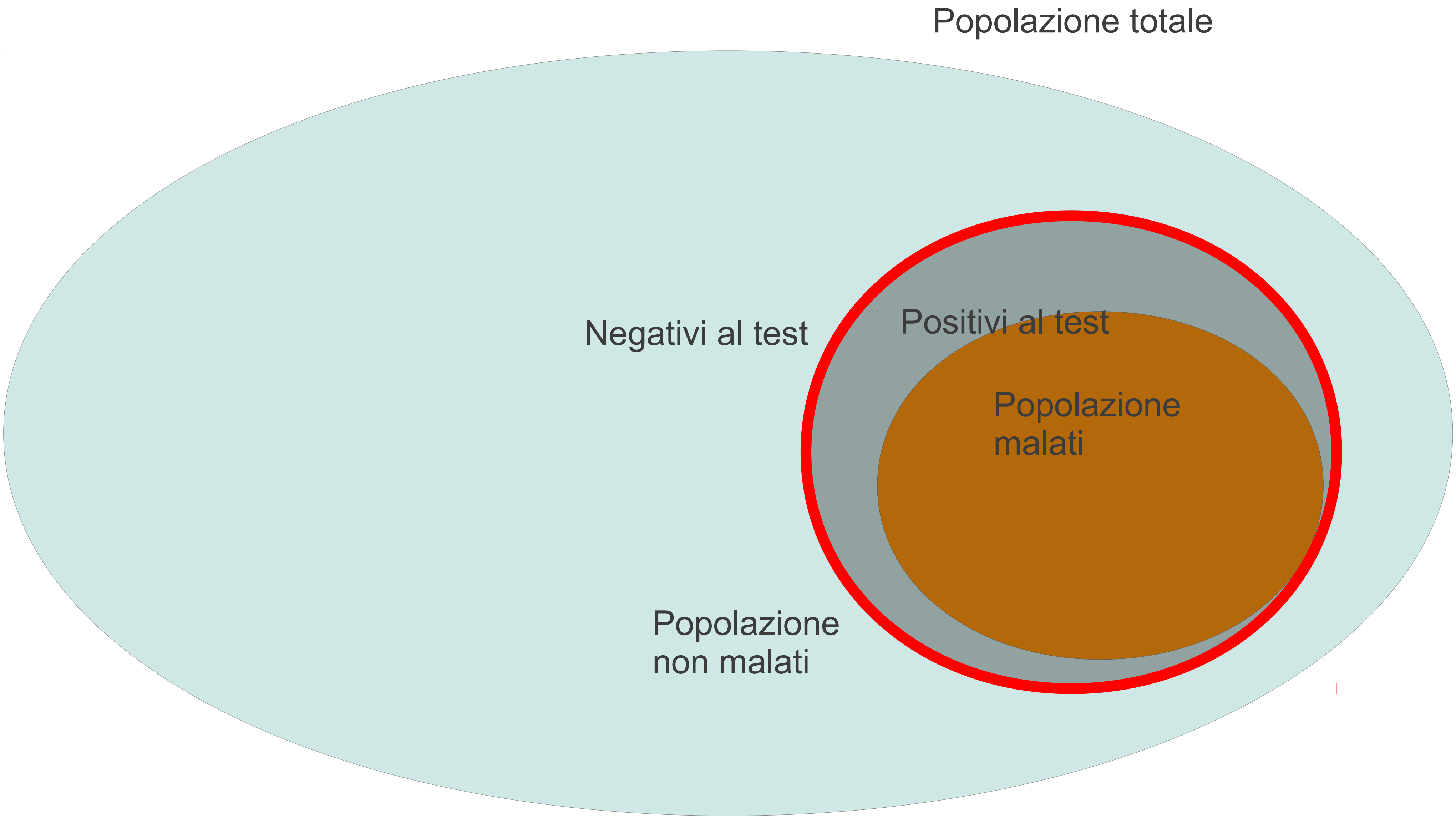
Interpretazione:

Se il test è NEGATIVO, allora la malattia NON È presente

ma:

Se il test è POSITIVO, allora la malattia PUÒ ESSERE presente

Sensibilità

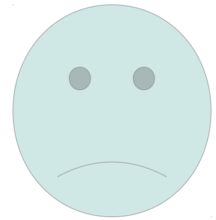


Interpretazione:

Se il test è NEGATIVO, allora la malattia NON È presente

ma:

Se il test è POSITIVO, allora la malattia PUÒ ESSERE presente



TEST <u>sensibile</u>

POSITIVO → ULTERIORI ACCERTAMENTI...

NEGATIVO → LA MALATTIA NON È PRESENTE

SPECIFICITÀ = proporzione di individui senza malattia che ha un risultato negativo del test

	Malattia presente	Malattia assente
Malattia diagnosticata	5	0
Malattia non diagnosticata	5	10

$$\frac{10 \text{ (negativi al test)}}{10 \text{ (malattia assente)}} = 100\%$$

SPECIFICITÀ = proporzione di individui senza malattia che ha un risultato negativo del test

	Malattia presente	Malattia assente
Malattia diagnosticata	5	0
Malattia non diagnosticata	5	10

$$\frac{10 \text{ (negativi al test)}}{10 \text{ (malattia assente)}} = 100\%$$

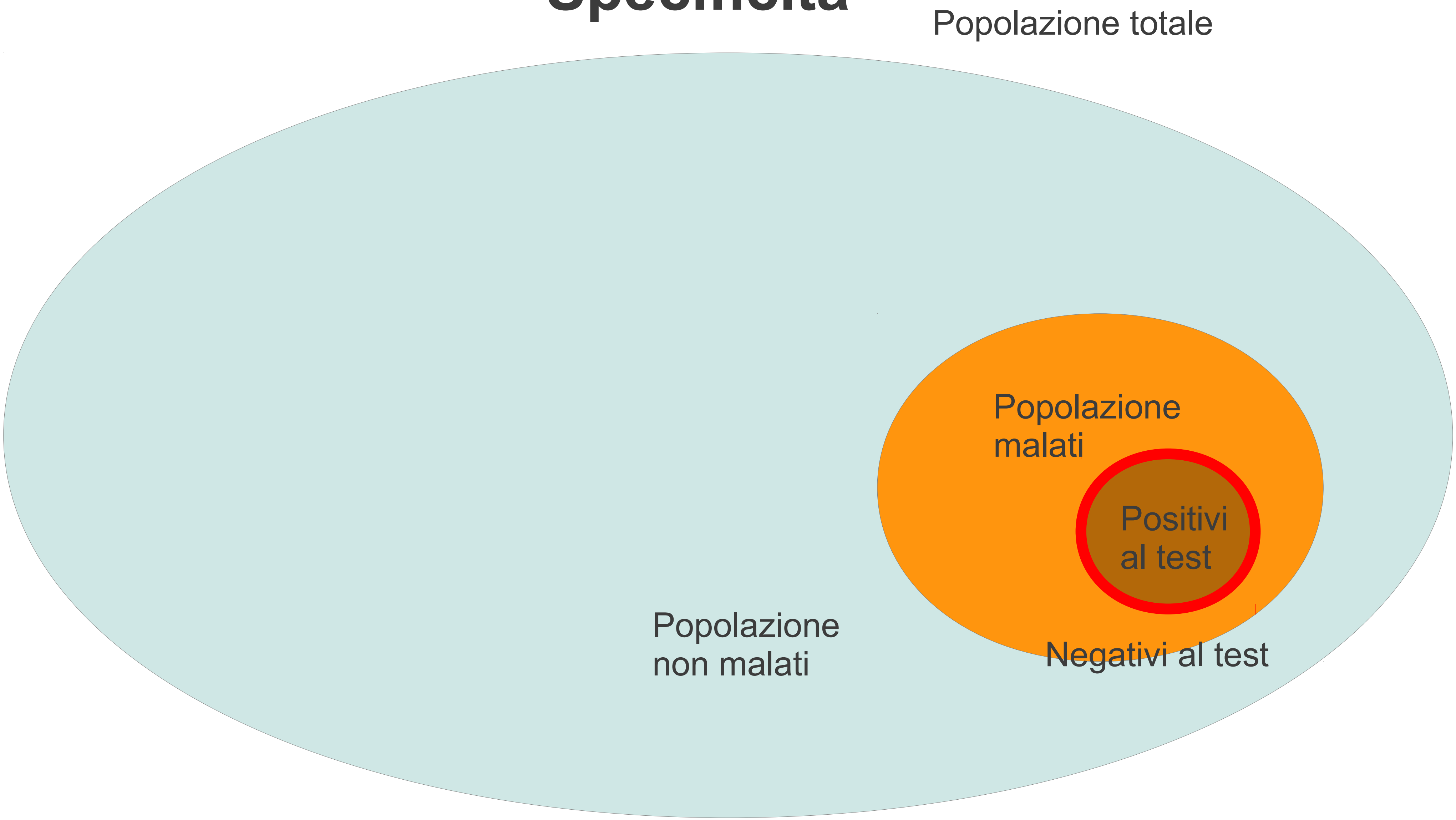
Interpretazione:

Se il test è POSITIVO, allora la malattia È presente

ma:

Se il test è NEGATIVO, allora la malattia PUÒ ESSERE assente

Specificità

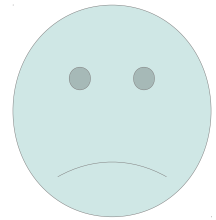


Interpretazione:

Se il test è POSITIVO, allora la malattia È presente

ma:

Se il test è NEGATIVO, allora la malattia PUÒ ESSERE assente



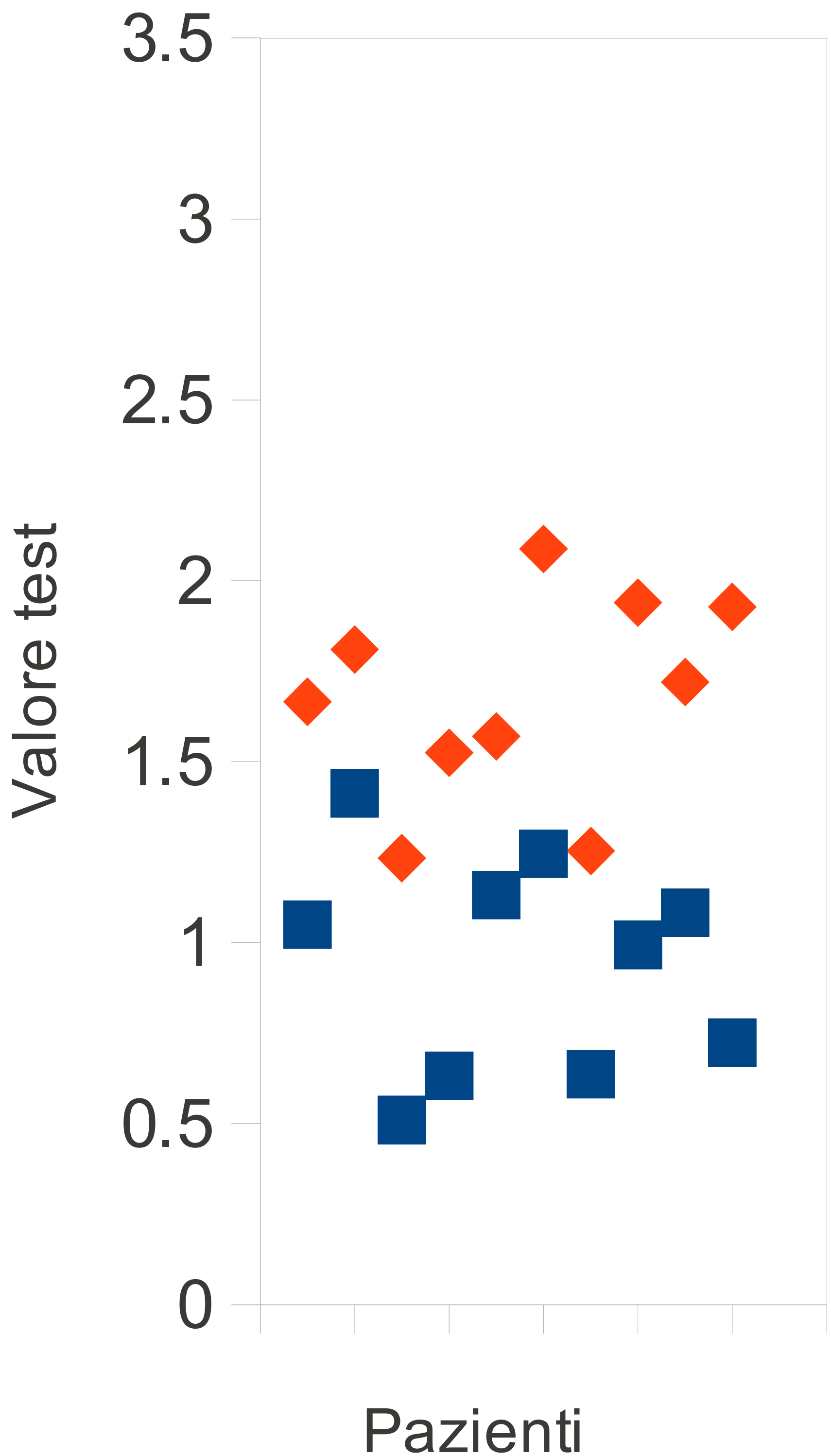
TEST specifico

POSITIVO

LA MALATTIA È PRESENTE

NEGATIVO

ULTERIORI ACCERTAMENTI...



Dove fissereste la soglia se...

Malattia poco grave e senza conseguenze nella popolazione di cui il paziente fa parte

Malattia ad alto rischio e potenzialmente contagiosa e mortale

Per ogni scenario, specificare il numero di VP, VN, FP, FN

Per i due scenari che avete indicato, trovate i valori di sensibilità e specificità

Sommario

Falsi positivi e Falsi negativi

Prevalenza

Effetti della prevalenza sulla decisione medica

Sensibilità

Specificità

Uso delle informazioni sulla sensibilità e specificità nel processo diagnostico

Un problema tipico

Spiegazione grafica

Il teorema di Bayes

Errori di ragionamento

Generalizzazione

Sommario

Falsi positivi e Falsi negativi

Prevalenza

Effetti della prevalenza sulla decisione medica

Sensibilità

Specificità

Uso delle informazioni sulla sensibilità e specificità nel processo diagnostico

Un problema tipico

Spiegazione grafica

Il teorema di Bayes

Errori di ragionamento

Generalizzazione

UN PROBLEMA TIPICO

Un problema tipico

Assumiamo che, per le donne di 40 anni, la probabilità di avere un tumore al seno sia del 1%. Se una donna ha questo tipo di tumore e si sottopone a screening mammografico di routine, la probabilità che l'esito della mammografia sia positivo è dell'80%. Se non lo ha, la probabilità che la mammografia dia un esito comunque positivo è del 10%.

Se doveste comunicare al medico curante di una donna di 40 anni, la cui mammografia ha avuto un esito positivo, la probabilità che la paziente abbia effettivamente un tumore al seno, quale valore indichereste?

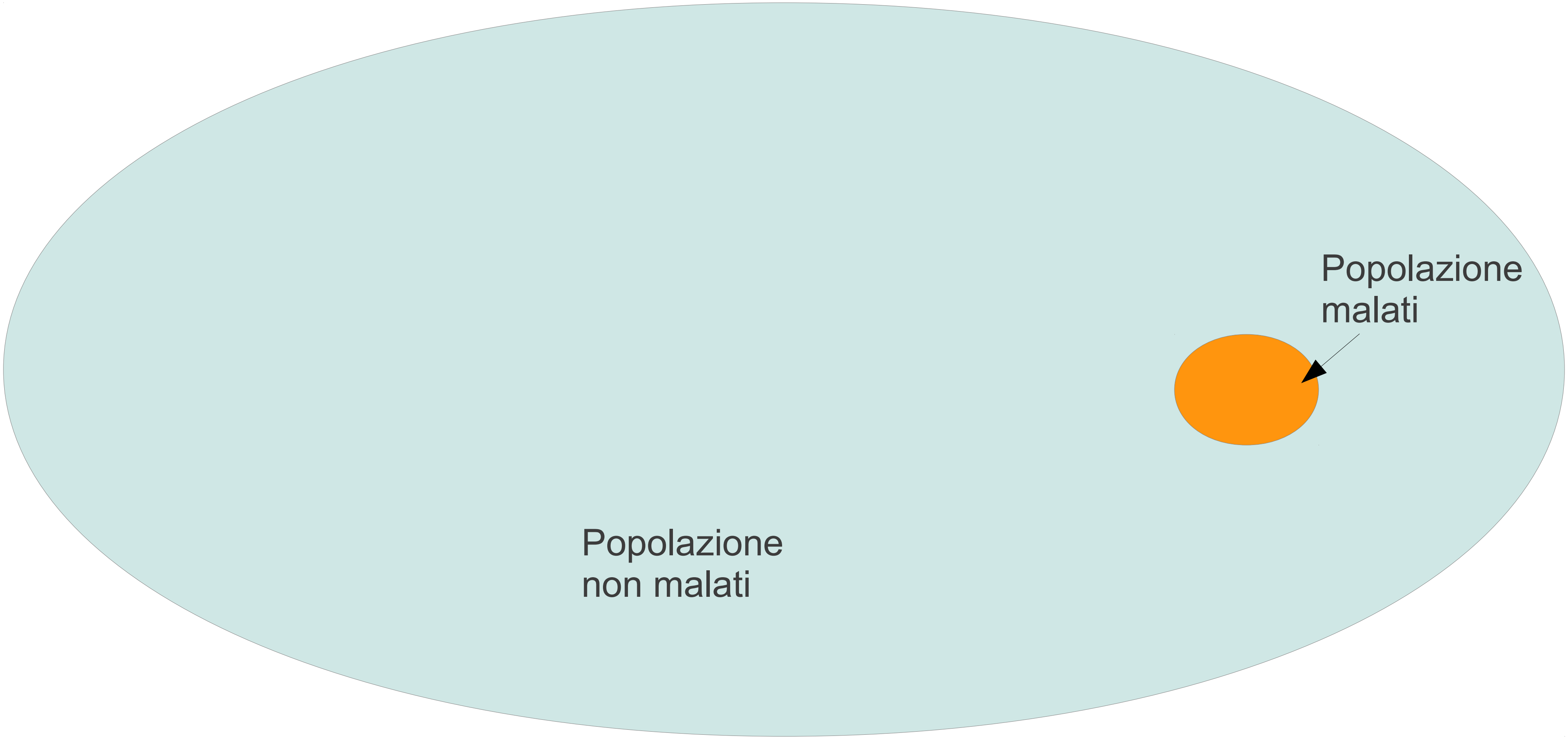
Un problema tipico

Assumiamo che, per le donne di 40 anni, *la probabilità di avere un tumore al seno sia del 1%.* prevalenza

Se una donna ha questo tipo di tumore e si sottopone a screening mammografico di routine, *la probabilità che l'esito della mammografia sia positivo è dell'80%.* sensibilità

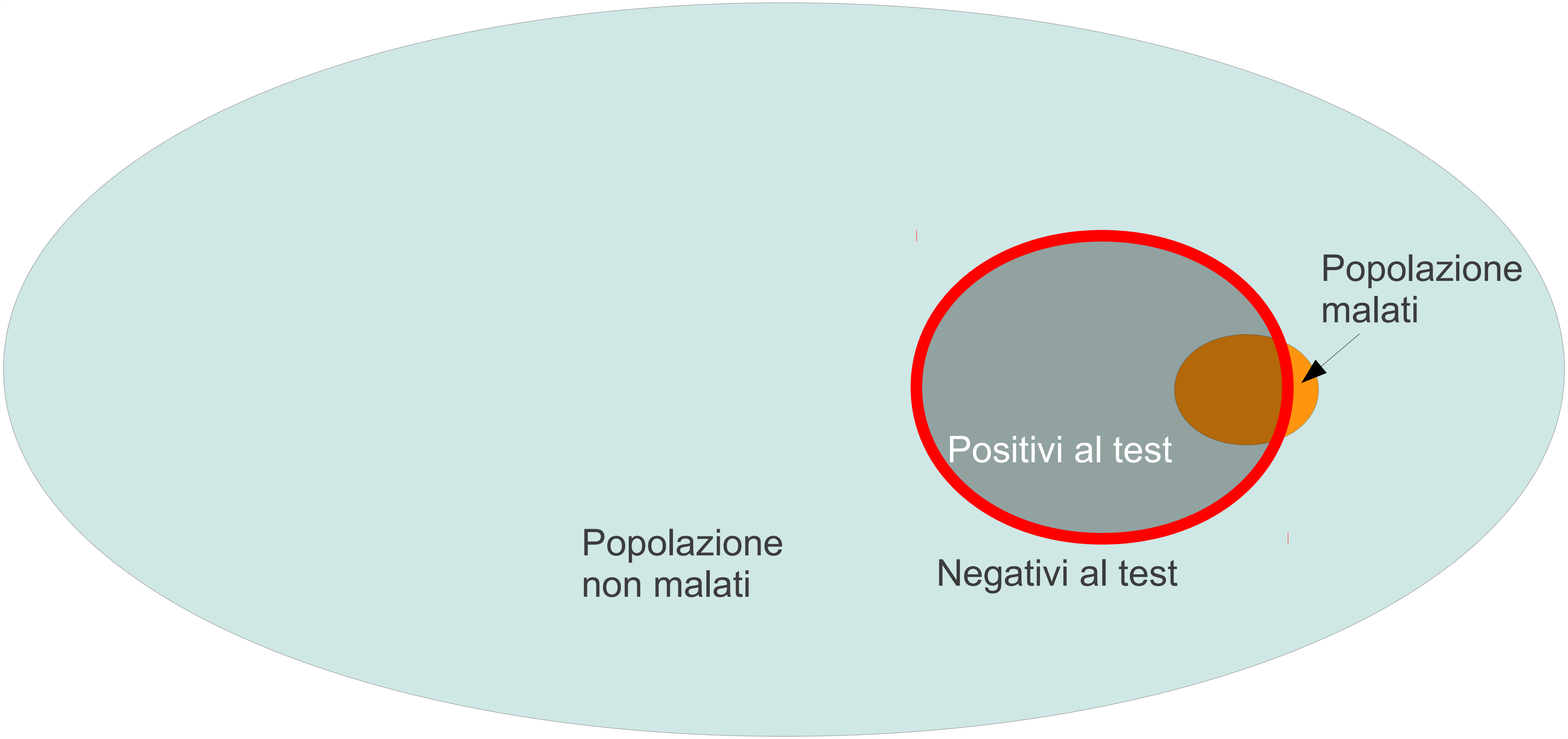
Se non lo ha, *la probabilità che la mammografia dia un esito comunque positivo è del 10%.* Falsi positivi: 100% - specificità

Se doveste comunicare al medico curante di una donna di 40 anni, la cui mammografia ha avuto un esito positivo, la probabilità che la paziente abbia effettivamente un tumore al seno, quale valore indichereste?

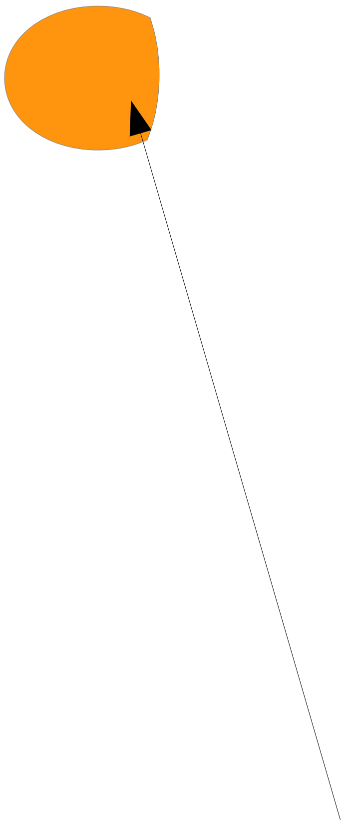


Prevalenza: 1% della popolazione totale

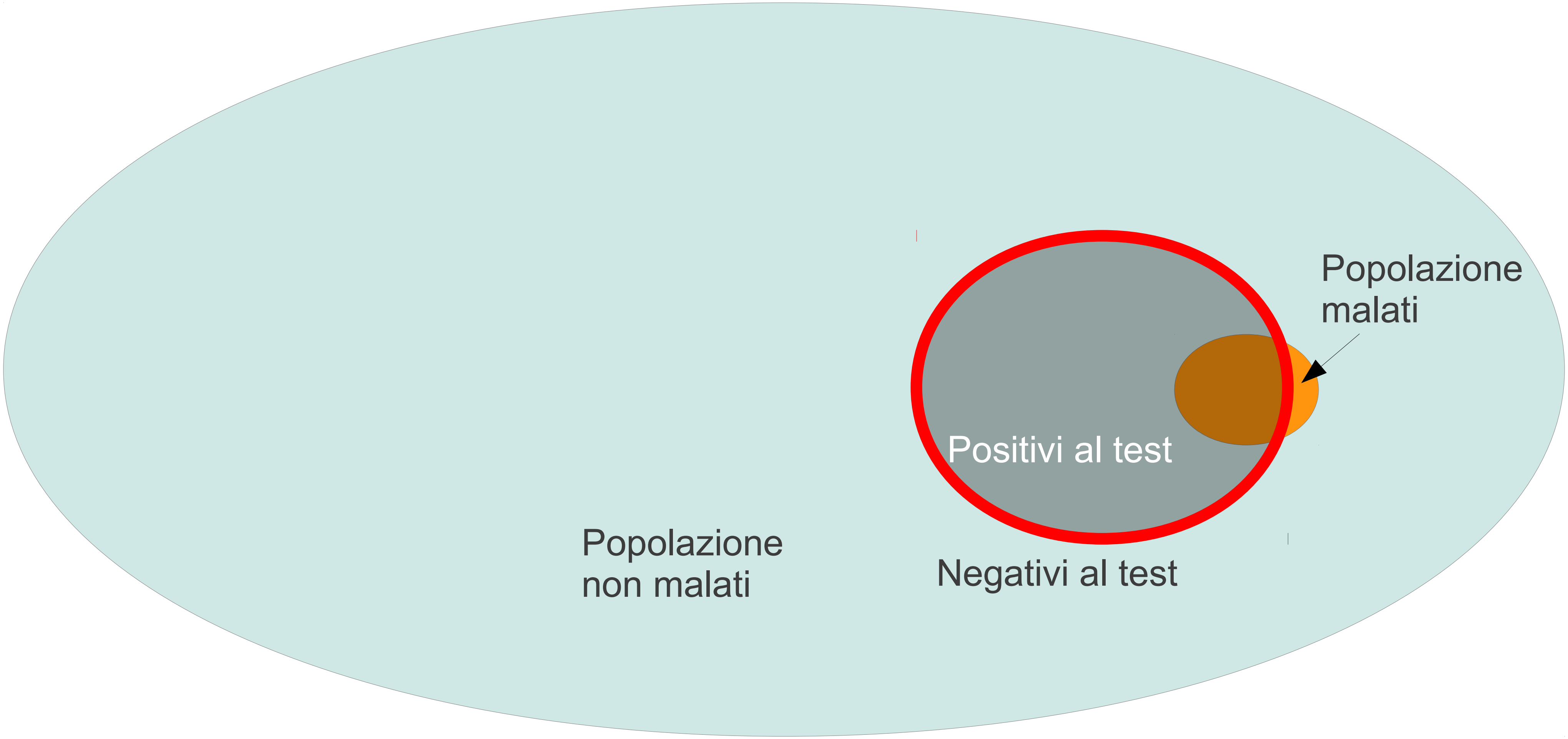
Assumiamo che, per le donne di 40 anni, la probabilità di avere un tumore al seno sia del 1%.



<u>Se una donna ha questo tipo di tumore</u> e si sottopone a screening mammografico di routine, *la probabilità che l'esito della mammografia sia positivo è dell'80%*

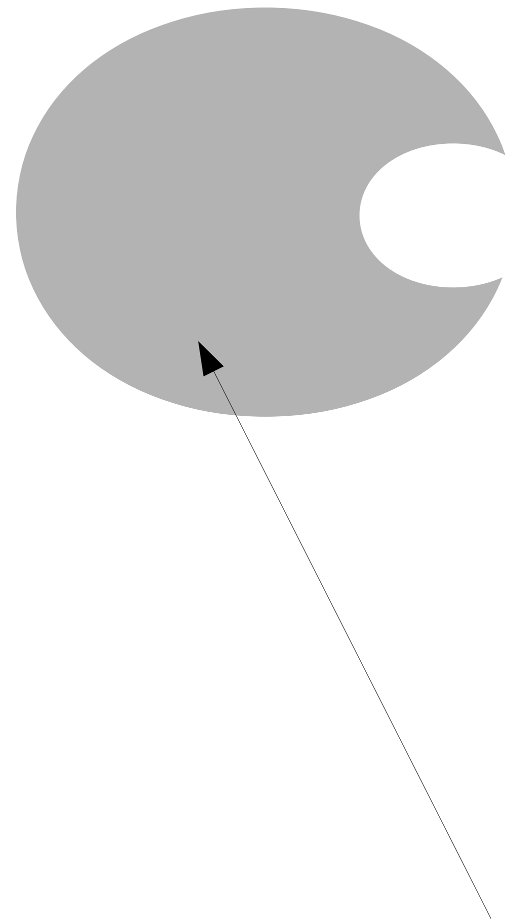


<u>Se una donna ha questo tipo di tumore</u> e si sottopone a screening mammografico di routine, *la probabilità che l'esito della mammografia sia positivo è dell'80%*

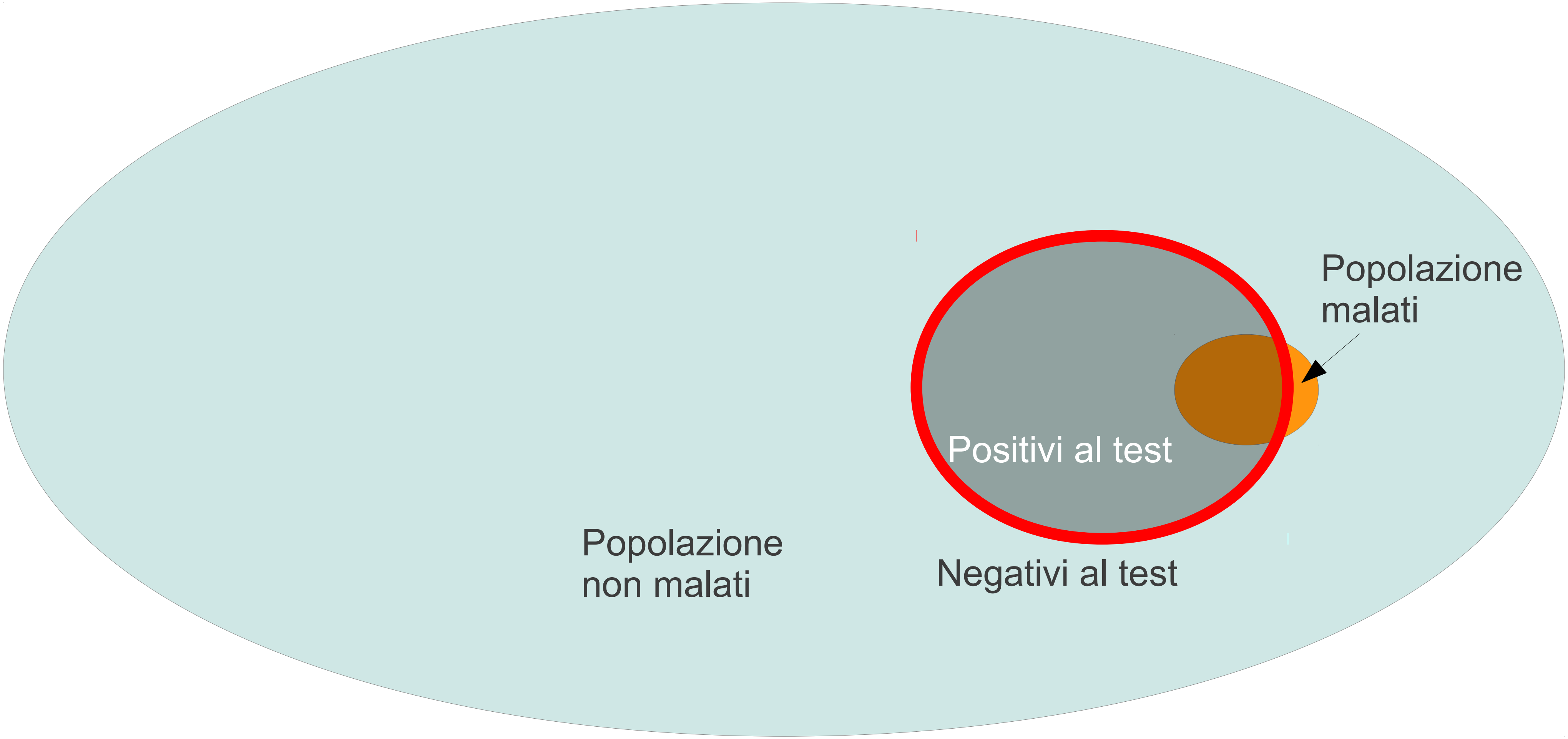


<u>Se non lo ha</u>, *la probabilità che la mammografia dia un esito comunque positivo è del 10%.*

<u>Se non lo ha</u>, *la probabilità che la mammografia dia un esito comunque positivo è del 10%.*

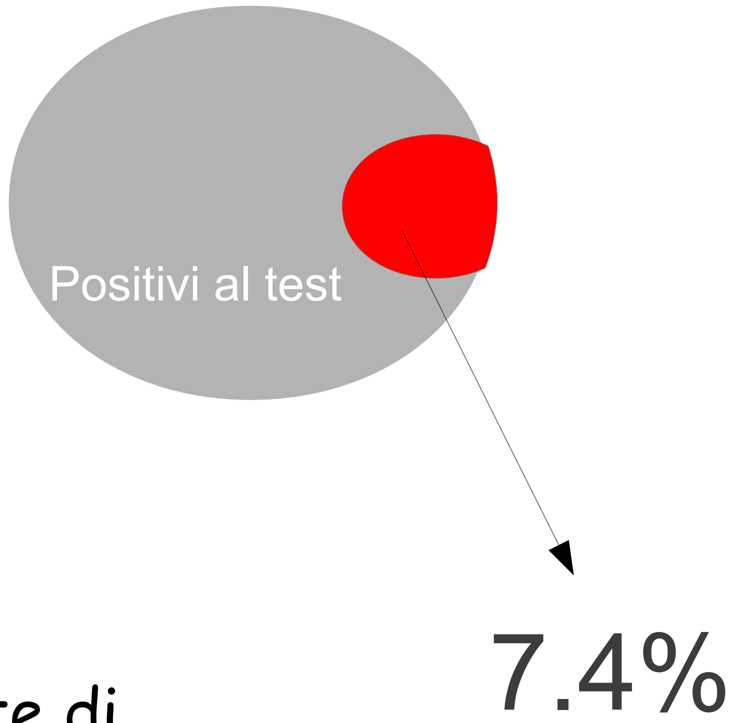


FALSI POSITIVI:
10% sani positivi al test
(specificità = 100%-10% = 90%)



Se doveste comunicare al medico curante di una donna di 40 anni, <u>la cui mammografia ha avuto un esito positivo,</u> *la probabilità che la paziente abbia effettivamente un tumore al seno,* quale valore indichereste?

Se doveste comunicare al medico curante di una donna di 40 anni, la cui mammografia ha avuto un esito positivo, *la probabilità che la paziente abbia effettivamente un tumore al seno,* quale valore indichereste?



Se doveste comunicare al medico curante di una donna di 40 anni, la cui mammografia ha avuto un esito positivo, *la probabilità che la paziente abbia effettivamente un tumore al seno*, quale valore indichereste?

TEOREMA DI BAYES

per stimare questa probabilità bisogna usare il **TEOREMA DI BAYES**:

$$p(T/O) = \frac{p(O/T) \times p(T)}{p(O/T) \times p(T) + p(O/-T) \times p(-T)}$$

per stimare questa probabilità bisogna usare il **TEOREMA DI BAYES**:

$$p(T/O) = \frac{p(O/T) \times p(T)}{p(O/T) \times p(T) + p(O/-T) \times p(-T)}$$

T rappresenta un'ipotesi teorica (esistenza del tumore);

O è una qualsiasi osservazione o dato empirico (mammografia positiva);

p(T/O) la probabilità a posteriori che sia vera T se è vera O (probabilità di avere il tumore al seno se la mammografia è positiva);

p(O/T) la probabilità a posteriori che si verifichi l'evento O se è vera T (probabilità che la mammografia sia positiva se si ha il tumore al seno);

p(T) la probabilità di base o a priori che sia vera T;

p(O/-T) la probabilità che si verifichi l'evento O se T NON è vera (probabilità che la mammografia sia positiva se NON si ha il tumore al seno);

p(-T) la probabilità di base o a priori che NON sia vera T .

per stimare questa probabilità bisogna usare il **TEOREMA DI BAYES**:

$$p(T/O) = \frac{p(O/T) \times p(T)}{p(O/T) \times p(T) + p(O/-T) \times p(-T)}$$

T rappresenta un'ipotesi teorica (esistenza del tumore);

O è una qualsiasi osservazione o dato empirico (mammografia positiva);

p(T/O) la probabilità a posteriori che sia vera T se è vera O (probabilità di avere il tumore al seno se la mammografia è positiva);

p(O/T) la probabilità a posteriori che si verifichi l'evento O se è vera T (probabilità che la mammografia sia positiva se si ha il tumore al seno); ***SENSIBILITÀ***

p(T) la probabilità di base o a priori che sia vera T; ***PREVALENZA***

p(O/-T) la probabilità che si verifichi l'evento O se T NON è vera (probabilità che la mammografia sia positiva se NON si ha il tumore al seno); ***1-SPECIFICITÀ***

p(-T) la probabilità di base o a priori che NON sia vera T . ***1-PREVALENZA***

per stimare questa probabilità bisogna usare il **TEOREMA DI BAYES**:

$$p(T/O) = \frac{p(O/T) \times p(T)}{p(O/T) \times p(T) + p(O/-T) \times p(-T)} = \frac{0{,}8 \times 0{,}01}{0{,}8 \times 0{,}01 + 0{,}1 \times 0{,}99} \boxed{= 0{,}074}$$

T rappresenta un'ipotesi teorica (esistenza del tumore);

O è una qualsiasi osservazione o dato empirico (mammografia positiva);

p(T/O) la probabilità a posteriori che sia vera T se è vera O (probabilità di avere il tumore al seno se la mammografia è positiva);

p(O/T) la probabilità a posteriori che si verifichi l'evento O se è vera T (probabilità che la mammografia sia positiva se si ha il tumore al seno);

p(T) la probabilità di base o a priori che sia vera T;

p(O/-T) la probabilità che si verifichi l'evento O se T NON è vera (probabilità che la mammografia sia positiva se NON si ha il tumore al seno);

p(-T) la probabilità di base o a priori che NON sia vera T .

Applicando il teorema di Bayes, si ottiene che la probabilità che una donna abbia il tumore
al seno dato un test diagnostico positivo è del 7,4%.

Questa risposta non viene quasi mai data. Ad esempio, in una ricerca
95 clinici su 100 ritenevano che la probabilità fosse del 70-80 %.

Sommario

Falsi positivi e Falsi negativi

Prevalenza

Effetti della prevalenza sulla decisione medica

Sensibilità

Specificità

Uso delle informazioni sulla sensibilità e specificità nel processo diagnostico

Un problema tipico

Spiegazione grafica

Il teorema di Bayes

Errori di ragionamento

Generalizzazione

Sommario

Falsi positivi e Falsi negativi

Prevalenza

Effetti della prevalenza sulla decisione medica

Sensibilità

Specificità

Uso delle informazioni sulla sensibilità e specificità nel processo diagnostico

Un problema tipico

Spiegazione grafica

Il teorema di Bayes

Errori di ragionamento

Generalizzazione

ERRORI DI RAGIONAMENTO

ERRORI DI RAGIONAMENTO

Il ragionamento umano è basato su una serie di meccanismi (euristiche)

Sono DIVERSI dai meccanismi di calcolo delle probabilità e di logica!

Permettono di ragionare in maniera rapida, flessibile

Espongono ad una serie di errori nel momento in cui ci si confronta con problemi logici o probabilistici

Quando pensiamo che un risultato positivo ad uno screening ci dia un'altissima probabilità che quel paziente ABBIA EFFETTIVAMENTE SVILUPPATO/CONTRATTO quella patologia, compiamo un ERRORE DI RAGIONAMENTO.

Abbiamo visto che l'esito positivo all'esame sopra descritto si traduce in una probabilità effettiva (ovvero, a posteriori) di avere un tumore pari al 7,4%.

Tale probabilità (che è comunque molto più alta della probabilità dell'1% stimata a priori sulla popolazione generale di riferimento) è tuttavia molto inferiore a quell'80% che la maggior parte dei clinici (e di tutti noi) tende ad indicare come probabilità effettiva, e che viene riportata "sulla base dell'esperienza".

Infatti, i medici tendono a riportare che:

"in situazioni analoghe, circa l'80% dei pazienti che erano positivi al test avevano effettivamente la malattia".

L'ERRORE che si compie è quello di una GENERALIZZAZIONE ARBITRARIA a tutta la popolazione che si sottopone al test, a partire da CASI PARTICOLARI (i pazienti realmente seguiti), che sono i pazienti effettivamente malati E che si sono sottoposti al test.

Viene compiuto un ragionamento a posteriori su base empirica, dimenticando che, per definizione, il medico non seguirà quei pazienti che sono sì risultati POSITIVI al test di screening, ma che NON avevano effettivamente il tumore.

Invece, SE si considerasse l'intera casistica (che dovrebbe comportare lo screening di tutta la popolazione o di un suo campione rappresentativo), ALLORA si avrebbe la riprova empirica di quale sia la p(T/O), che dovrebbe coincidere con la probabilità del 7,4% individuata dalla formula.

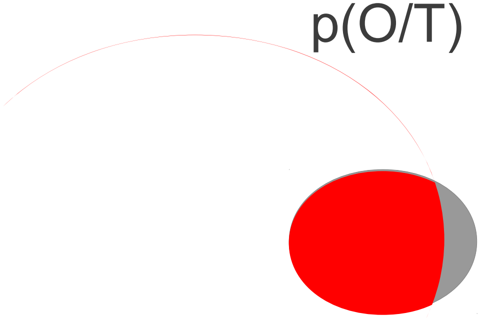
Qualora la probabilità non coincidesse (essendo per esempio del 9%), allora andrebbe modificata la p(T).

Dunque, l’affermazione abituale dei medici sopra riportata è errata, e dovrebbe essere riformulata in:

“<u>l’80% dei pazienti che avevano la malattia (e che poi sono stati seguiti dall’equipe medica) erano risultati positivi al test”.</u>

Tale probabilità è appunto la p(O/T).

p(O/T)



P(T/O)

ESEMPI

Test di gravidanza

Un medico che lavora in un consultorio in cui vengono effettuati test di gravidanza vuole sapere in che misura un risultato positivo sia predittivo di una effettiva gravidanza..

Il medico ha a disposizione queste informazioni:

La probabilità che una donna che si rivolge al consultorio per fare il test sia incinta è, sulla totalità delle occorrenze, del 2%.

Se una donna che fa il test di gravidanza è effettivamente incinta, la probabilità che il suo test sia positivo è del 90%.

Se una donna che fa il test non è incinta, la probabilità che il suo test risulti positivo è comunque dello 0,5%.

Qual è la probabilità che una donna sia effettivamente incinta, se risulta positiva al test di gravidanza?

P(O/T): PROBABILITA' DI RISULTARE POSITIVA AL TEST di una donna che è effettivamente incinta:

90%

P(T): PROBABILITA' CHE UNA DONNA CHE FA IL TEST SIA INCINTA:

2%

P(O/-T): PROBABILITA' DI RISULTARE POSITIVA AL TEST di una donna che NON è effettivamente incinta:

0,5%

P(T/O): Probabilità di essere incinta di una donna positiva al test:

??**%**

per stimare questa probabilità bisogna usare il **TEOREMA DI BAYES**:

$$p(T/O) = \frac{p(O/T) \times p(T)}{p(O/T) \times p(T) + p(O/-T) \times p(-T)} = \frac{0,9 \times 0,02}{0,9 \times 0,02 + 0,005 \times 0,98} \boxed{= 0,786}$$

T rappresenta un'ipotesi teorica (GRAVIDANZA);

O è una qualsiasi osservazione o dato empirico (test positivo);

p(T/O) la probabilità a posteriori che sia vera T se è vera O (probabilità di gravidanza se il test è positivo);

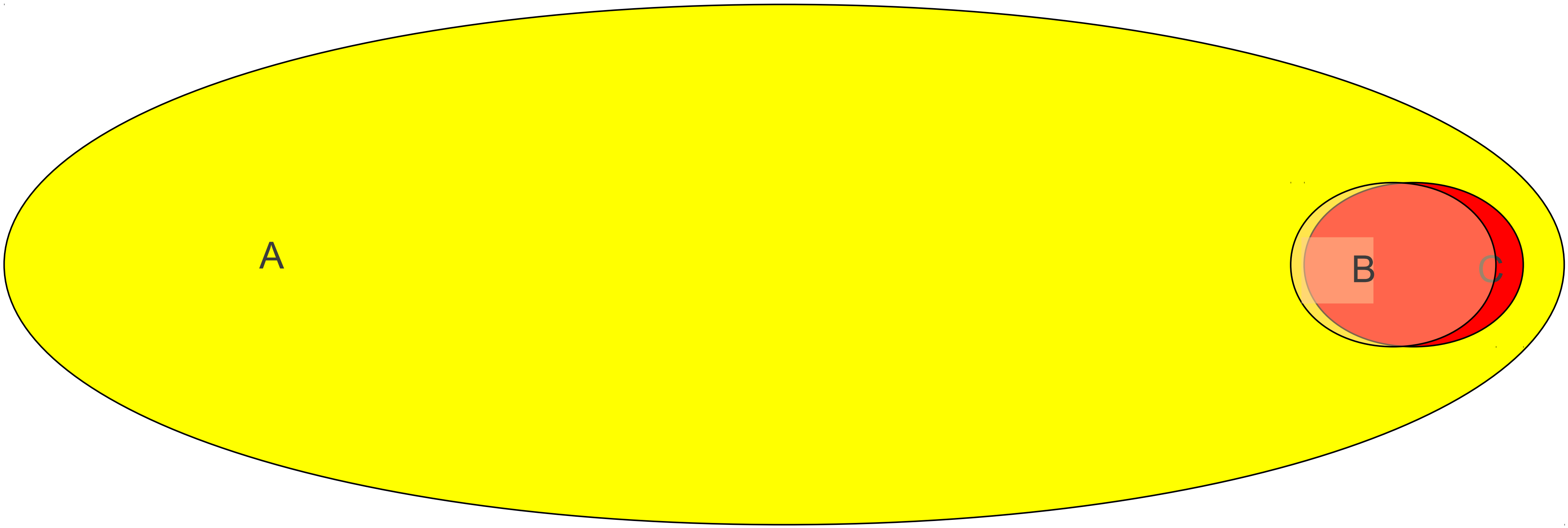
p(O/T) la probabilità a posteriori che si verifichi l'evento O se è vera T (probabilità che il test sia positivo se sussiste gravidanza);

p(T) la probabilità di base o a priori che sia vera T;

p(O/-T) la probabilità che si verifichi l'evento O se T NON è vera (probabilità che il test sia positivo se NON sussiste gravidanza);

p(-T) la probabilità di base o a priori che NON sia vera T.

Dati A (donne che fanno il test) e C (donne in gravidanza), chiamiamo B l'insieme delle donne che risultano positive al test.



Data l'elevata sensibilità (pochi FN) e specificità del test (pochi FP), quasi tutti i B (donne che risultano positive al test) appartengono anche a C.

Il test di gravidanza risulta dunque essere un test molto attendibile, con alto valore predittivo positivo (VPP). Infatti, la probabilità che un B (donna con test positivo) sia anche un C (donna in gravidanza) risulta essere del 78,6%.

Setticemia

In una struttura ospedaliera il 10% dei pazienti dell'ultimo anno ha sviluppato setticemia.

Un paziente manifesta febbre alta e lesioni alla pelle.

Il medico sa che:

Se il paziente ha la setticemia, nell'80% dei casi manifesta questi sintomi.

Se il paziente NON ha la setticemia, nel 10% dei casi manifesta questi sintomi.

Qual è la probabilità che il paziente abbia effettivamente la setticemia, dal momento che manifesta questi sintomi?

P(O/T): PROBABILITA' DI AVERE FEBBRE E LESIONI DI UN PAZIENTE CON SETTICEMIA:

80%

P(T): PERCENTUALE DI PAZIENTI CHE HANNO SVILUPPATO SETTICEMIA NELL'ULTIMO ANNO:

10%

P(O/-T): PROBABILITA' DI AVERE FEBBRE E LESIONI DI UN PAZIENTE SENZA SETTICEMIA: **10%**

P(T/O): Che probabilità ha di avere effettivamente la setticemia un paziente con questi sintomi: **??%**

per stimare questa probabilità bisogna usare il **TEOREMA DI BAYES**:

$$p(T/O) = \frac{p(O/T) \times p(T)}{p(O/T) \times p(T) + p(O/-T) \times p(-T)} = \frac{0{,}8 \times 0{,}1}{0{,}8 \times 0{,}1 + 0{,}1 \times 0{,}9} \boxed{= 0{,}470}$$

T rappresenta un'ipotesi teorica (setticemia);

O è una qualsiasi osservazione o dato empirico (sintomi);

p(T/O) la probabilità a posteriori che sia vera T se è vera O (probabilità di avere la setticemia, dati i sintomi);

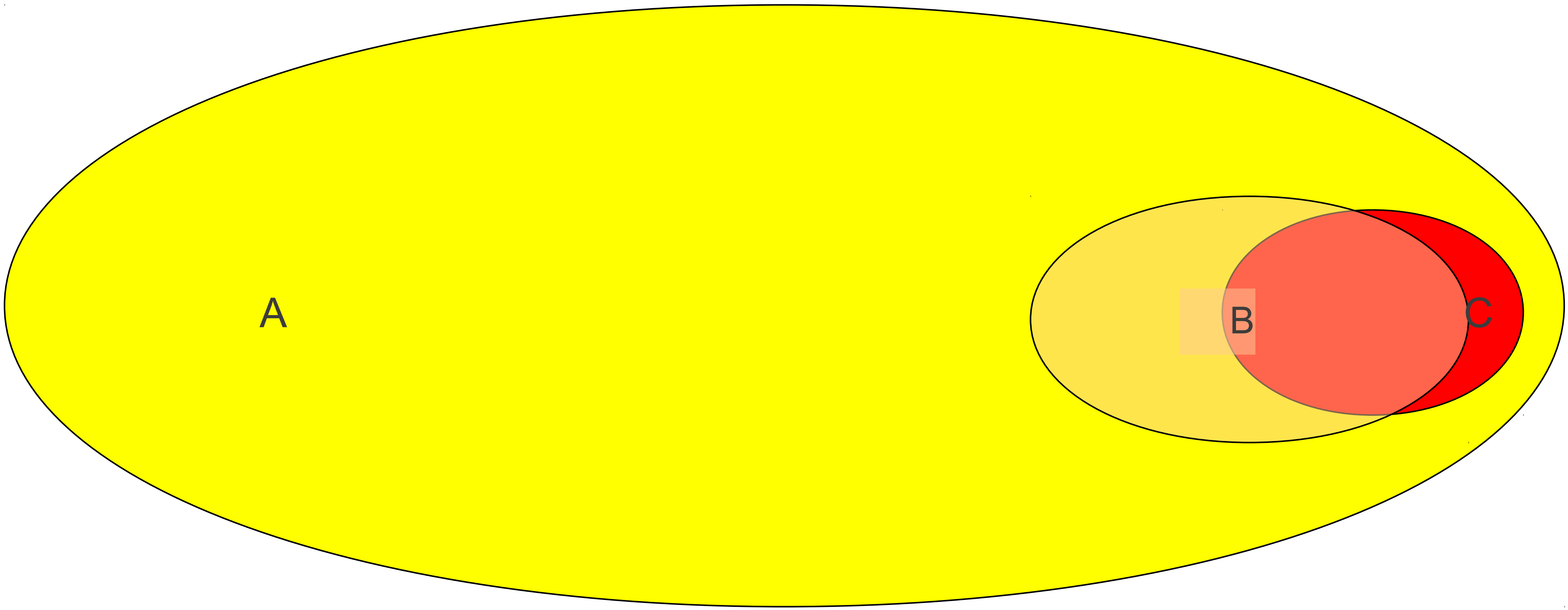
p(O/T) la probabilità a posteriori che si verifichi l'evento O se è vera T (probabilità di manifestare questi sintomi se si ha la setticemia);

p(T) la probabilità di base o a priori che sia vera T;

p(O/-T) la probabilità che si verifichi l'evento O se T NON è vera (probabilità di manifestare questi sintomi se NON si ha la setticemia);

p(-T) la probabilità di base o a priori che NON sia vera T .

Dati A (popolazione ospedaliera generale) e C (popolazione con setticemia), **definiamo B gli individui con sintomi che POSSONO essere indice di setticemia.**



Dei B, poco meno della metà è anche C. I SINTOMI in esame sono infatti dotati di alta sensibilità (pochi FN), ma relativamente bassa specificità (molti FP).

La probabilità che un appartenente a B appartenga anche a C risulta dunque del 47%. Tale probabilità è comunque decisamente più elevata del rischio di setticemia nella popolazione ospedaliera generale (10%).

Attacco ischemico

Un gruppo di neurologi del Nebraska ha riscontrato che, nella popolazione sopra i 60 anni, la probabilità di un attacco ischemico è maggiore nelle donne che negli uomini. Tuttavia, in quella fascia d'età la popolazione femminile è maggiore di quella maschile, e la diversa incidenza di ischemia nei due generi potrebbe dipendere esclusivamente da questa sproporzione campionaria.

Per controllare il possibile effetto confondente dell'età, è stata esaminata l'incidenza di ischemia in un gruppo di persone residenti a Omaha, con età compresa tra 30 e 60 anni. In questa fascia d'età, infatti, la numerosità degli uomini e delle donne è circa la stessa.

Secondo questo studio epidemiologico:

Il 10% della popolazione aveva avuto almeno un attacco ischemico.

Nella popolazione con attacco ischemico, la percentuale di donne era del 20%.

Nella popolazione SENZA attacco ischemico, la percentuale di donne era del 55%.

Qual è dunque il rischio di avere un attacco ischemico se si ha un'età compresa tra 30 e 60 anni e si è di sesso femminile?

P(O/T): PERCENTUALE DI DONNE ALL'INTERNO DELLA POPOLAZIONE CON ATTACCO ISCHEMICO :

20%

P(T): PERCENTUALE DI POPOLAZIONE CON ATTACCO ISCHEMICO NEL CAMPIONE GENERALE:

10%

P(O/-T): PERCENTUALE DI DONNE ALL'INTERNO DELLA POPOLAZIONE SENZA ATTACCO ISCHEMICO :

55%

P(T/O): Probabilità di avere un attacco ischemico per una donna appartenente al campione: **??%**

per stimare questa probabilità bisogna usare il **TEOREMA DI BAYES**:

$$p(T/O) = \frac{p(O/T) \times p(T)}{p(O/T) \times p(T) + p(O/-T) \times p(-T)} = \frac{0{,}2 \times 0{,}1}{0{,}2 \times 0{,}1 + 0{,}55 \times 0{,}9} \boxed{= 0{,}039}$$

T rappresenta un'ipotesi teorica (verificarsi dell'attacco ischemico);

O è una qualsiasi osservazione o dato empirico (essere di sesso femminile);

p(T/O) la probabilità a posteriori che sia vera T se è vera O (probabilità di avere un attacco ischemico se si è una donna);

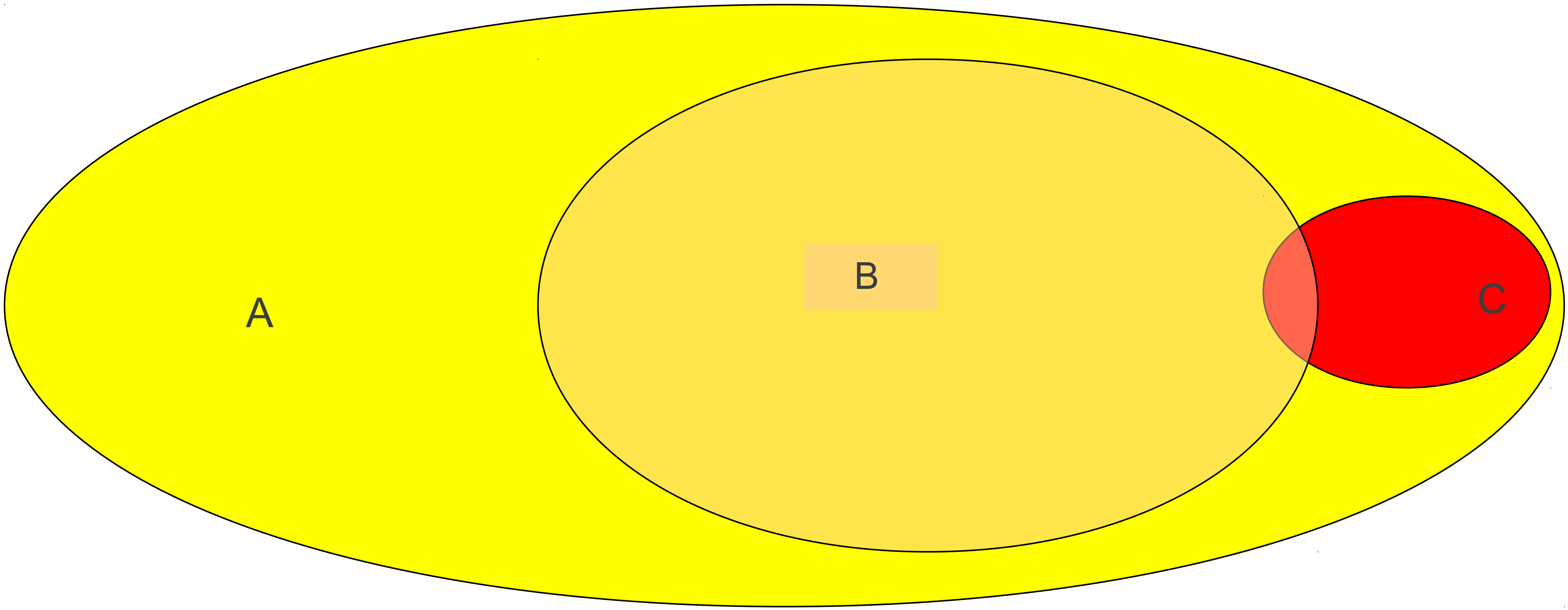
p(O/T) la probabilità a posteriori che si verifichi l'evento O se è vera T (probabilità di essere una donna se si ha avuto un attacco ischemico);

p(T) la probabilità di base o a priori che sia vera T;

p(O/-T) la probabilità che si verifichi l'evento O se T NON è vera (probabilità di essere una donna se NON si ha avuto un attacco ischemico);

p(-T) la probabilità di base o a priori che NON sia vera T .

Dati A (popolazione generale) e C (popolazione con attacco ischemico), chiamiamo B l'insieme che rappresenta la popolazione femminile



Poiché nella popolazione generale si ha il 55% di donne, mentre nella popolazione con ischemia il solo 20%, risulta evidente come il sottoinsieme "popolazione femminile" sia scarsamente sovrapponibile a quello "popolazione con ischemia".

Infatti, la probabilità che una donna abbia un attacco ischemico risulta essere del 3,9%, una percentuale di rischio più bassa di quella riferita alla popolazione generale (10%).

Attacco cardiaco

Un medico vuole valutare il rischio di disturbi cardiaci per un paziente maschio, di 55 anni, il cui padre è deceduto per infarto.

Il medico ha a disposizione le seguenti informazioni:

Il 20% dei maschi compresi tra 50 e 60 anni soffrono di disturbi cardiaci.

Il 25% dei maschi compresi tra 50 e 60 anni che soffrono di disturbi cardiaci hanno avuto parenti maschi morti per infarto.

L'8% dei maschi compresi tra 50 e 60 anni che non soffrono di disturbi cardiaci hanno avuto parenti maschi morti per infarto.

Basandosi su tali informazioni, qual è la probabilità che un paziente soffra effettivamente di disturbi cardiaci, se ha avuto almeno un congiunto maschio morto per infarto?

P(O/T): PERCENTUALE DI MASCHI TRA 50 E 60 ANNI CHE HANNO AVUTO UN CONGIUNTO MORTO PER INFARTO E CHE SOFFRONO DI DISTURBI CARDIACI:

25%

P(T): PERCENTUALE DI MASCHI TRA 50 E 60 ANNI CON DISTURBI CARDIACI:

20%

P(O/-T): PERCENTUALE DI MASCHI TRA 50 E 60 ANNI CHE HANNO AVUTO UN CONGIUNTO MORTO PER INFARTO E CHE NON SOFFRONO DI DISTURBI CARDIACI:

8%

P(T/O): Probabilità di avere un attacco cardiaco per un uomo che ha avuto un congiunto maschio morto per infarto:

??%

per stimare questa probabilità bisogna usare il **TEOREMA DI BAYES**:

$$p(T/O) = \frac{p(O/T) \times p(T)}{p(O/T) \times p(T) + p(O/-T) \times p(-T)} = \frac{0{,}25 \times 0{,}20}{0{,}25 \times 0{,}20 + 0{,}08 \times 0{,}80} \boxed{= 0{,}439}$$

T rappresenta un'ipotesi teorica (attacco cardiaco);

O è una qualsiasi osservazione o dato empirico (congiunto maschio deceduto per infarto);

p(T/O) la probabilità a posteriori che sia vera T se è vera O (probabilità di avere un attacco cardiaco se si ha avuto un congiunto morto per infarto);

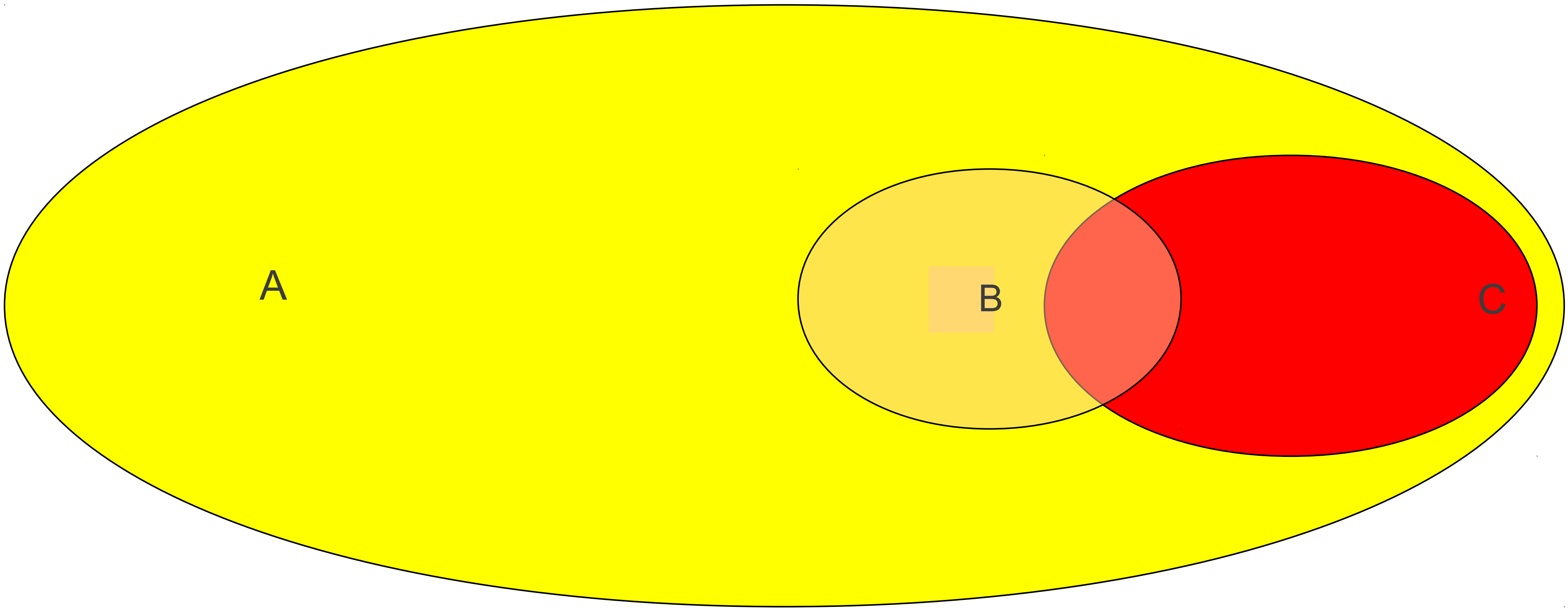
p(O/T) la probabilità a posteriori che si verifichi l'evento O se è vera T (probabilità che si abbia un congiunto morto per infarto, se si hanno disturbi cardiaci);

p(T) la probabilità di base o a priori che sia vera T;

p(O/-T) la probabilità che si verifichi l'evento O se T NON è vera (probabilità che si abbia un congiunto morto per infarto, se NON si hanno disturbi cardiaci);

p(-T) la probabilità di base o a priori che NON sia vera T.

Dati A (popolazione maschile tra i 50 e i 60 anni) e C (sottopopolazione con disturbi cardiaci), chiamiamo B la sottopopolazione maschile che ha avuto un congiunto maschio morto per infarto.



In questo caso, siamo di fronte a una patologia con alta prevalenza (20%) e a un indice predittivo (famigliarità per la malattia) dotato di sensibilità molto scarsa (molti FN), e di media specificità (buon numero di FP).

Tuttavia, è anche vero che la probabilità di avere disturbi cardiaci è più elevata per chi ha avuto un congiunto morto per infarto (43,8%) che per la popolazione maschile generale (20%).

Tumore alla pelle

Un medico in Kansas visita un bambino con ustioni solari così gravi da presentare vesciche sulla pelle ustionata. I genitori del bambino chiedono al medico che probabilità abbia il figlio di sviluppare un tumore alla pelle in seguito all'ustione.
Il medico dispone delle seguenti informazioni:

Il 17% degli abitanti del Kansas sviluppa una qualche forma di tumore alla pelle in età adulta.
Il 40% degli abitanti del Kansas che sviluppano una qualche forma di tumore alla pelle in età adulta, hanno avuto da bambini gravi ustioni solari.
Il 20% degli abitanti del Kansas che NON sviluppano tumore in età adulta, hanno avuto da bambini gravi ustioni solari.

Qual è la probabilità che un abitante del Kansas sviluppi una qualche forma di tumore alla pelle in età adulta, se da bambino ha subito gravi ustioni solari?

P(T): PERCENTUALE DI POPOLAZIONE DEL KANSAS CON SVILUPPO DI TUMORE ALLA PELLE IN ETA' ADULTA:

17%

P(O/T): PERCENTUALE DI POPOLAZIONE DEL KANSAS CHE SVILUPPANO TUMORE ALLA PELLE IN ETA' ADULTA E CHE HANNO AVUTO GRAVI USTIONI SOLARI NELL'INFANZIA :

40%

P(O/-T): PERCENTUALE DI POPOLAZIONE DEL KANSAS CHE SVILUPPANO TUMORE ALLA PELLE IN ETA' ADULTA E CHE NON HANNO AVUTO GRAVI USTIONI SOLARI NELL'INFANZIA :

20%

P(T/O): Probabilità di sviluppare un tumore alla pelle in età adulta di un bambino che ha ricevuto gravi ustioni solari:

??%

per stimare questa probabilità bisogna usare il **TEOREMA DI BAYES**:

$$p(T/O) = \frac{p(O/T) \times p(T)}{p(O/T) \times p(T) + p(O/-T) \times p(-T)} = \frac{0{,}40 \times 0{,}17}{0{,}40 \times 0{,}17 + 0{,}20 \times 0{,}83} \boxed{= 0{,}291}$$

T rappresenta un'ipotesi teorica (tumore alla pelle);

O è una qualsiasi osservazione o dato empirico (gravi ustioni solari nell'infanzia);

p(T/O) la probabilità a posteriori che sia vera T se è vera O (probabilità di sviluppare tumore alla pelle **in seguito** a gravi ustioni solari nell'infanzia);

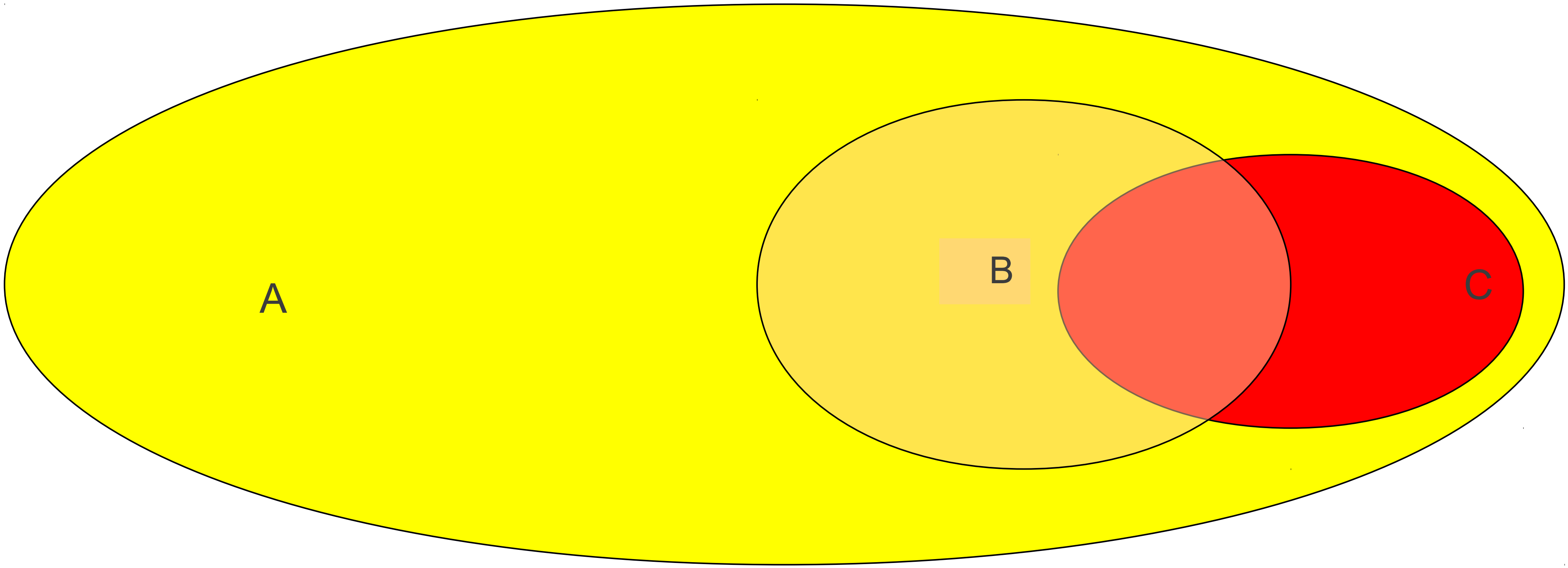
p(O/T) la probabilità a posteriori che si verifichi l'evento O se è vera T (probabilità di aver avuto in infanzia gravi ustioni solari se si ha, in età adulta, un tumore alla pelle);

p(T) la probabilità di base o a priori che sia vera T (NB: **in questa specifica popolazione**);

p(O/-T) la probabilità che si verifichi l'evento O se T NON è vera T (probabilità di aver avuto in infanzia gravi ustioni solari se NON si ha, in età adulta, un tumore alla pelle);

p(-T) la probabilità di base o a priori che NON sia vera T .

Dati A (popolazione generale del Kansas) e C (popolazione con tumore alla pelle in età adulta), chiamiamo B la popolazione che ha avuto gravi ustioni solari nell'infanzia.



In questo caso, abbiamo una patologia ad alta prevalenza (18%) e un indice predittivo con sensibilità non molto elevata (buon numero di FN) e bassa specificità (alto numero di FP).

Infatti, la probabilità che un bambino che riceve gravi ustioni solari sviluppi un tumore alla pelle in età adulta risulta essere del 29%, a fronte di una prevalenza generale del 17%. Aumenta quindi il rischio, ma non in misura considerevole.

TEST HIV

Un centro per la ricerca sull'AIDS somministra il test anti-HIV ai pazienti di un centro di disintossicazione, popolazione in cui l'incidenza di AIDS è notoriamente alta.

Gli psicologi del centro, prima di informare i pazienti risultati positivi, vogliono sapere qual è il rischio effettivo che l'infezione da HIV venga confermata.

Essi hanno a disposizione questi dati:

La probabilità che una persona seguita dal centro abbia effettivamente contratto l'infezione è del 18%.

Se una persona seguita dal centro non ha contratto l'infezione , la probabilità che il test risulti comunque positivo è del 5%.

Se una persona seguita dal centro ha realmente contratto l'infezione, la probabilità che il test sia positivo è del 100%.

Qual è la probabilità che una persona il cui test risulta positivo abbia effettivamente contratto l'infezione da HIV?

P(O/T): PROBABILITA' DI RISULTATO POSITIVO AL TEST DI UNA PERSONA CHE HA REALMENTE L'INFEZIONE:

100%

P(T): PROBABILITA' DI AVERE L'INFEZIONE DA HIV NELLE PERSONE DEL CENTRO:

18%

P(O/-T): PROBABILITA' DI RISULTATO POSITIVO AL TEST DI UNA PERSONA CHE NON HA REALMENTE L'INFEZIONE:

5%

P(T/O): Probabilità di avere l'infezione da HIV per una persona che risulta positiva al test:

??%

per stimare questa probabilità bisogna usare il **TEOREMA DI BAYES**:

$$p(T/O) = \frac{p(O/T) \times p(T)}{p(O/T) \times p(T) + p(O/-T) \times p(-T)} = \frac{1{,}00 \times 0{,}18}{1{,}00 \times 0{,}18 + 0{,}05 \times 0{,}82} \boxed{= 0{,}814}$$

T rappresenta un'ipotesi teorica (presenza dell'infezione da HIV);

O è una qualsiasi osservazione o dato empirico (test positivo);

p(T/O) la probabilità a posteriori che sia vera T se è vera O (probabilità di avere l'infezione da HIV se il test è positivo);

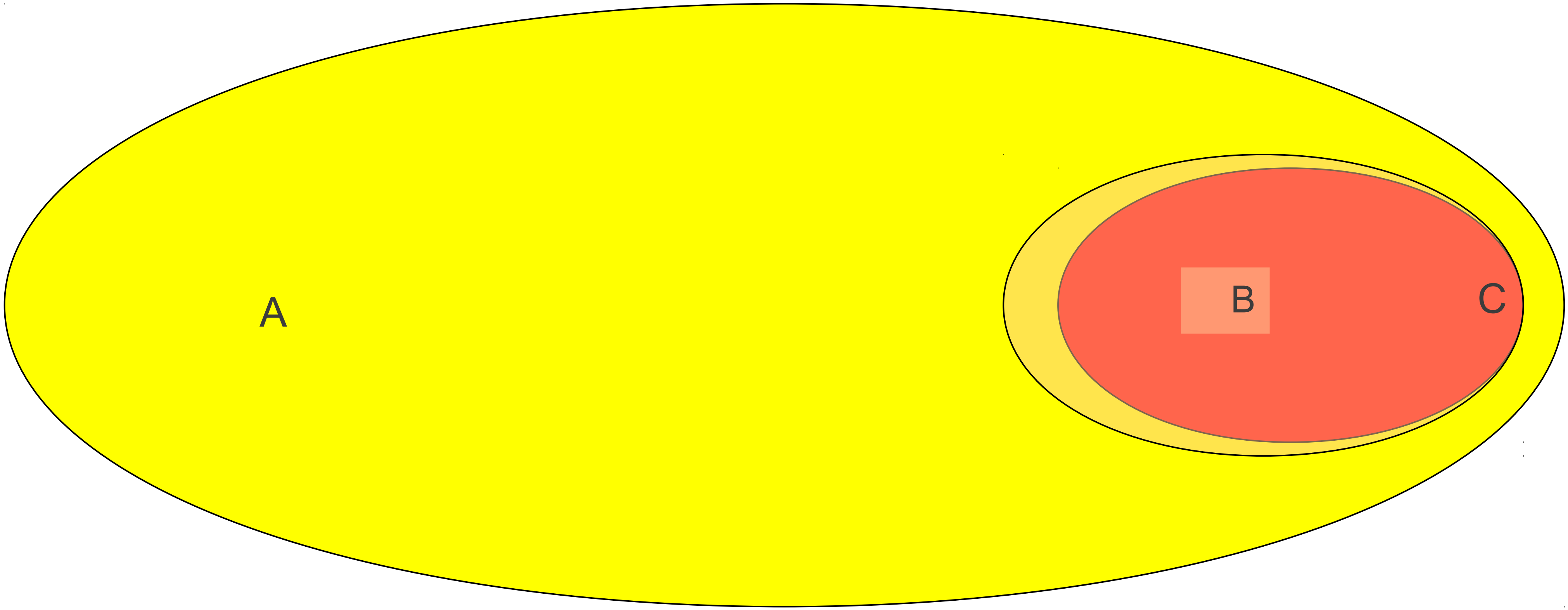
p(O/T) la probabilità a posteriori che si verifichi l'evento O se è vera T (probabilità che il test sia positivo se si ha effettivamente l'infezione);

p(T) la probabilità di base o a priori che sia vera T (NB: **in questa specifica popolazione**);

p(O/-T) la probabilità che si verifichi l'evento O se T NON è vera (probabilità che il test sia positivo se NON si ha effettivamente l'infezione);

p(-T) la probabilità di base o a priori che NON sia vera T .

Dati A (pazienti di un centro di disintossicazione) e C (popolazione con infezione da HIV), chiamiamo B la popolazione che risulta positiva al test anti-HIV.



In questo caso, abbiamo due informazioni molto rilevanti: la sensibilità del test è del 100% (non esistono FN). La patologia in esame, IN QUESTA SPECIFICA POPOLAZIONE, ha un'alta prevalenza.

Risulta dunque estremamente alta (81,4%) la probabilità che una persona positiva al test abbia effettivamente contratto l'infezione

TEST HIV NELLA POPOLAZIONE GENERALE: L'importanza della prevalenza

Come cambia la stima di probabilità, se ci spostiamo all'interno della popolazione generale?

Sappiamo che la prevalenza dell'infezione da HIV nella popolazione generale italiana è dello 0,1%.

Se una persona non ha contratto l'infezione, la probabilità che il test risulti comunque positivo rimane del 5%.

Se una persona ha realmente contratto l'infezione, la probabilità che il test sia positivo rimane del 100%.

Qual è, in questo caso, la probabilità che una persona il cui test risulta positivo abbia effettivamente contratto l'infezione da HIV?

P(O/T): PROBABILITA' DI RISULTATO POSITIVO AL TEST DI UNA PERSONA CHE HA REALMENTE L'INFEZIONE:

100%

P(T): PROBABILITA' DI AVERE L'INFEZIONE DA HIV NELLA POPOLAZIONE GENERALE:

0,1%

P(O/-T): PROBABILITA' DI RISULTATO POSITIVO AL TEST DI UNA PERSONA CHE NON HA REALMENTE L'INFEZIONE:

5%

P(T/O): Probabilità di avere l'infezione da HIV per una persona che risulta positiva al test:

??%

per stimare questa probabilità bisogna usare il **TEOREMA DI BAYES**:

$$p(T/O) = \frac{p(O/T) \times p(T)}{p(O/T) \times p(T) + p(O/-T) \times p(-T)} = \frac{1{,}00 \times 0{,}001}{1{,}00 \times 0{,}001 + 0{,}05 \times 0{,}999} = 0{,}019$$

T rappresenta un'ipotesi teorica (presenza dell'infezione da HIV);

O è una qualsiasi osservazione o dato empirico (test positivo);

p(T/O) la probabilità a posteriori che sia vera T se è vera O (probabilità di avere l'infezione da HIV se il test è positivo);

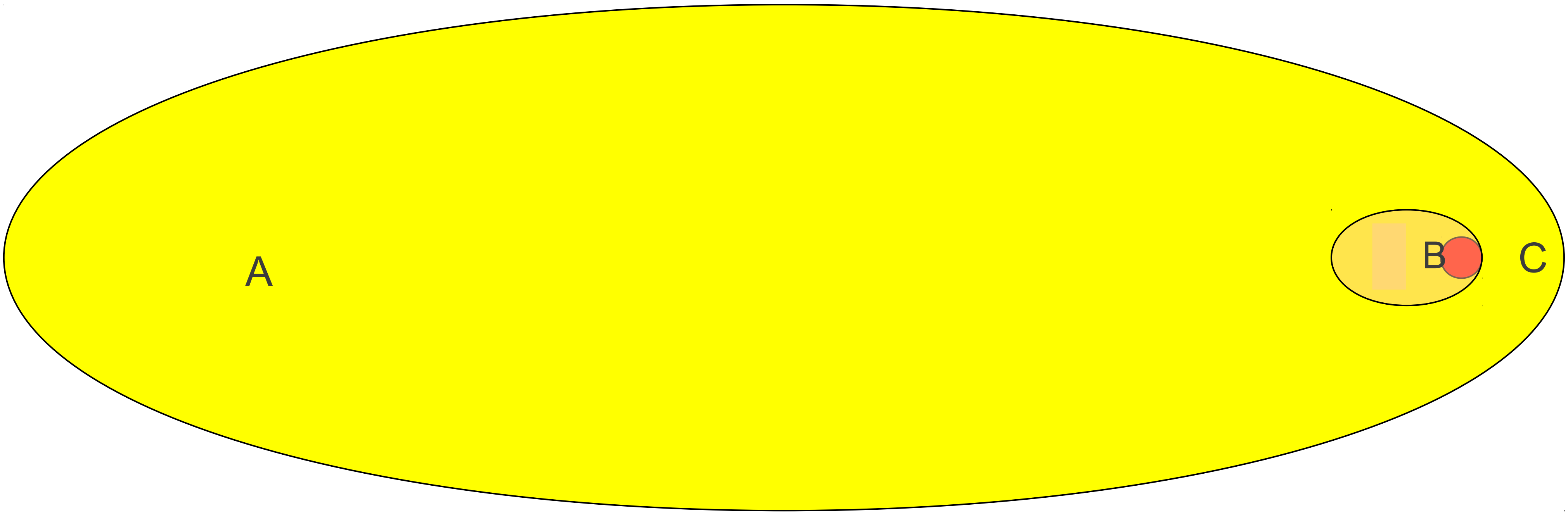
p(O/T) la probabilità a posteriori che si verifichi l'evento O se è vera T (probabilità che il test sia positivo se si ha effettivamente l'infezione);

p(T) la probabilità di base o a priori che sia vera T (NB: **nella popolazione generale**);

p(O/-T) la probabilità che si verifichi l'evento O se T NON è vera (probabilità che il test sia positivo se NON si ha effettivamente l'infezione);

p(-T) la probabilità di base o a priori che NON sia vera T .

Dati A (POPOLAZIONE GENERALE) e C (popolazione con infezione da HIV), chiamiamo B la popolazione che risulta positiva al test anti-HIV.



la sensibilità del test è del 100% (non esistono FN). Tuttavia, la patologia in esame, NELLA POPOLAZIONE GENERALE, ha prevalenza MOLTO BASSA, e la specificità del test, in rapporto alla prevalenza, risulta scarsa (molti FP in rapporto ai VP).

In questa particolare condizione (se cioè sottoponessimo al test l'intera popolazione), la probabilità di aver effettivamente contratto l'infezione di una persona con risultato positivo al test sarebbe dell'1,9% (comunque decisamente maggiore dello 0,1% relativo alla popolazione generale).

Sommario

Quando prendiamo una decisione medica, lo facciamo sulla base della sua probabilità

La probabilità è basata su:

- Informazioni pregresse (es. caratteristiche della popolazione e prevalenza)
- Caratteristiche di sensibilità e specificità del test / sintomatologia su cui ci basiamo

Il ragionamento umano non applica le regole della probabilità e della logica

È necessario

- Essere consapevoli delle informazioni su cui ci basiamo (es. prevalenza, sensibilità, specificità) e del contesto (costi e benefici di VP, VN, FP, FN)
- Utilizzare strumenti che consentano di stimare correttamente la probabilità di una ipotesi diagnostica: ad es., il teorema di Bayes