

Secondo workshop di formazione

ELEMENTI DI EPIDEMIOLOGIA CLINICA
12-13 dicembre 2022

- **Validità e riproducibilità delle misure diagnostiche**
- **Screening**

Prof.ssa Maria Sofia Cattaruzza

Dipartimento di Sanità Pubblica e Malattie Infettive

La Sapienza Università

mariasofia.cattaruzza@uniroma1.it

Test ideale

Per la diagnosi (e lo screening) di una malattia ci si avvale di test clinici, strumentali, di laboratorio.

Un test ideale dovrebbe essere:

- **VALIDO o ACCURATO**

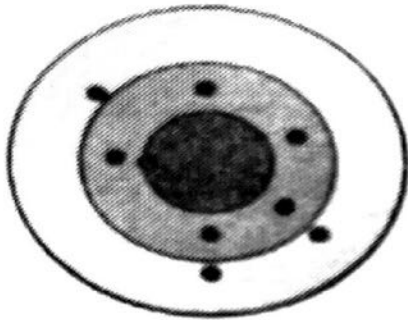
cioè capace di fornire un valore (medio) corrispondente al valore reale e distinguere le persone sane da quelle malate

- **PRECISO o ATTENDIBILE o AFFIDABILE**

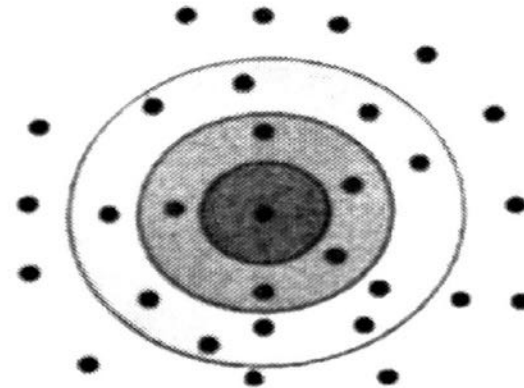
cioè capace di dare sempre lo stesso risultato in misurazioni ripetute (RIPRODUCIBILE); dipende dalla bontà dello strumento e/o operatore.

Accuratezza e Precisione

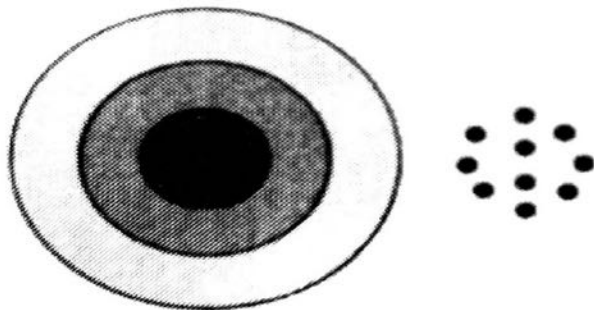
A Sia accuratezza che precisione



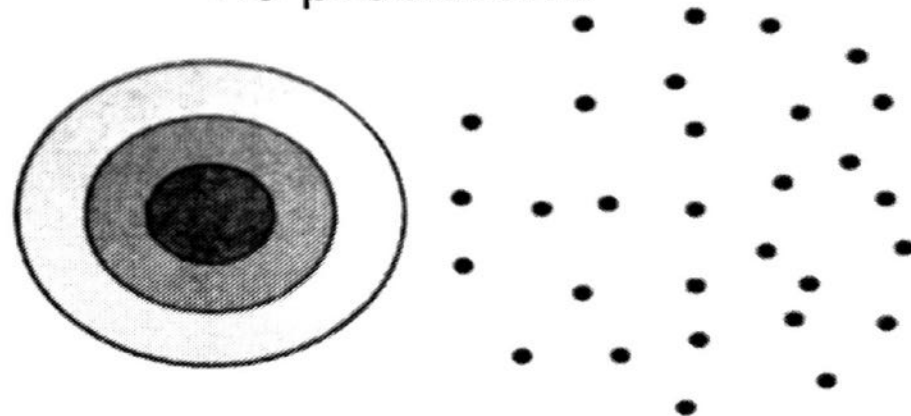
B Solo accuratezza



C Solo precisione

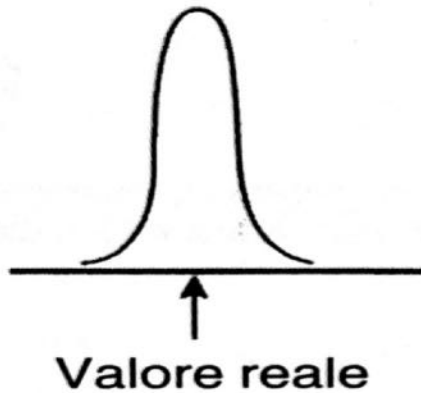


D Né accuratezza né precisione

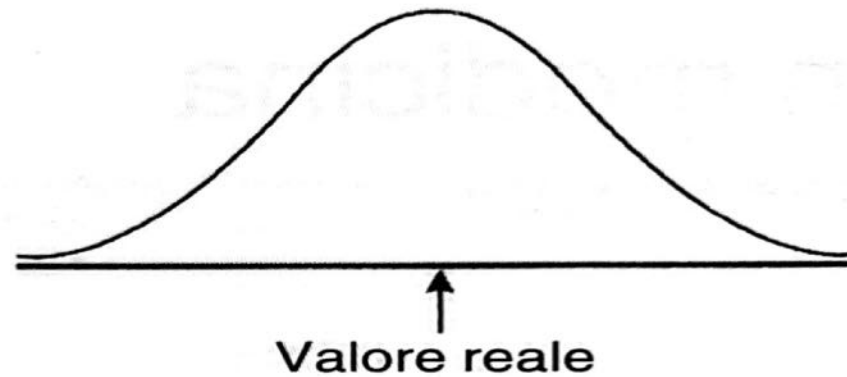


Accuratezza e Precisione

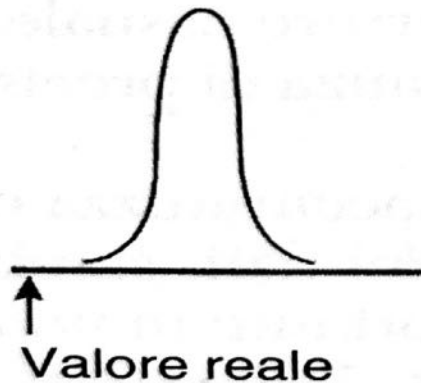
A Sia accuratezza
che precisione



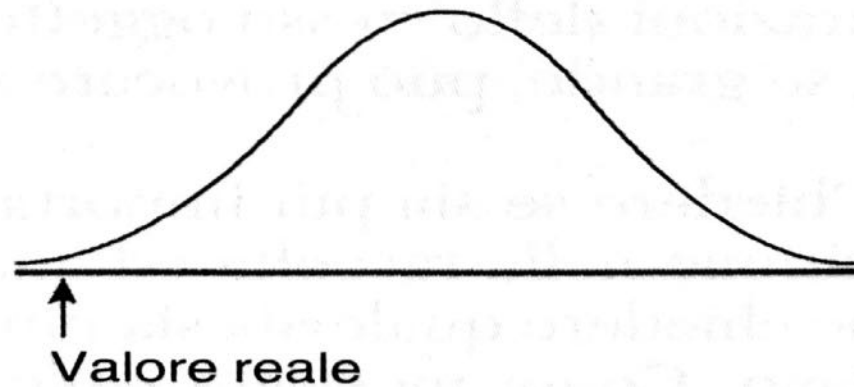
B Solo accuratezza



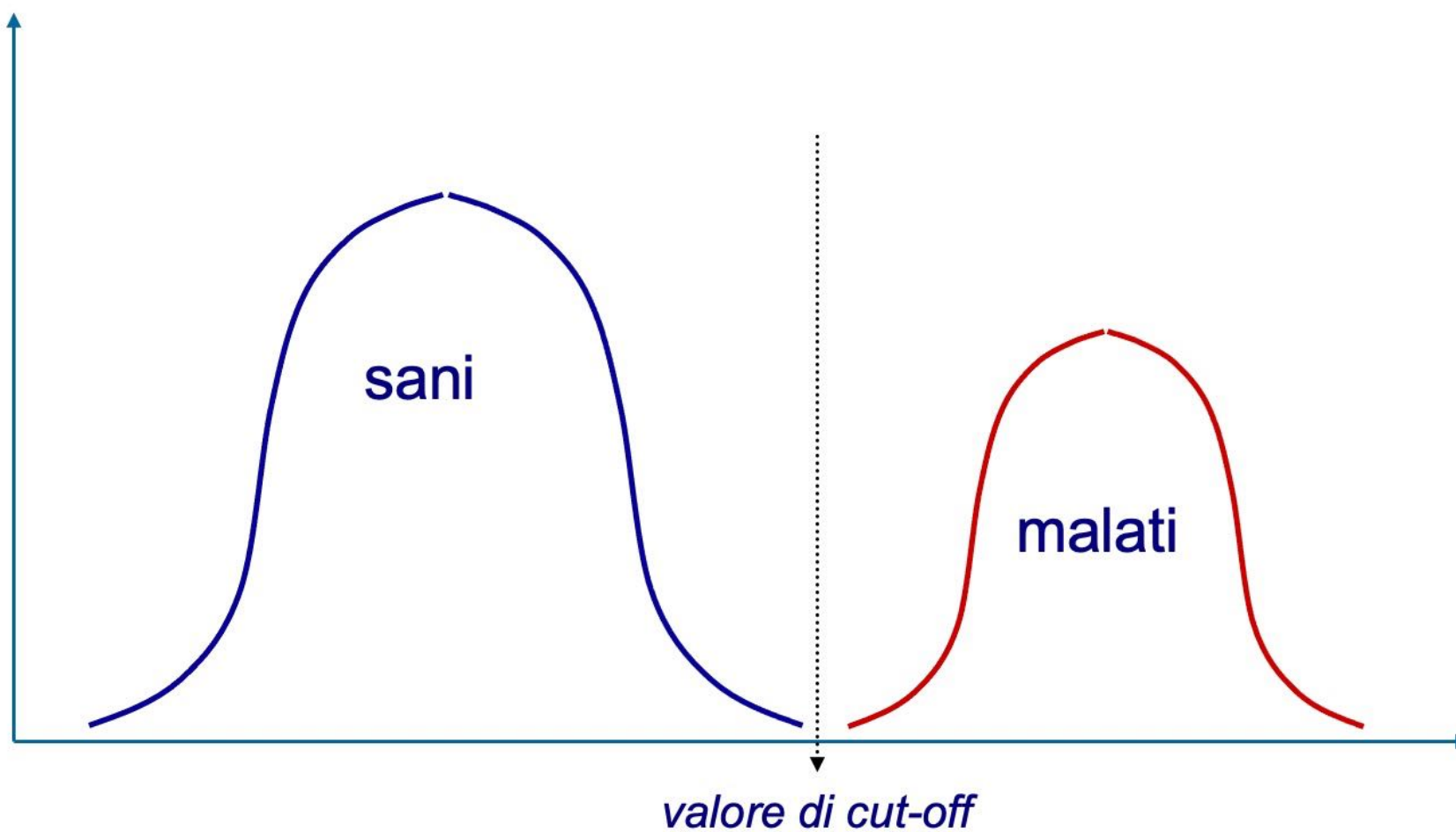
C Solo precisione



D Né accuratezza
né precisione



Il test ideale



Lo screening (definizione OMS)

Lo screening è definito come l'identificazione **presuntiva** di una malattia **non riconosciuta** in una popolazione **apparentemente sana e asintomatica** mediante test, esami o altre procedure che possono essere applicati **rapidamente e facilmente** alla popolazione target.

Lo **screening è un processo:** inizia con l'**invito** a partecipare e termina con il **trattamento efficace** per le persone con diagnosi di malattia.

Lo **screening è lo strumento della prevenzione secondaria**

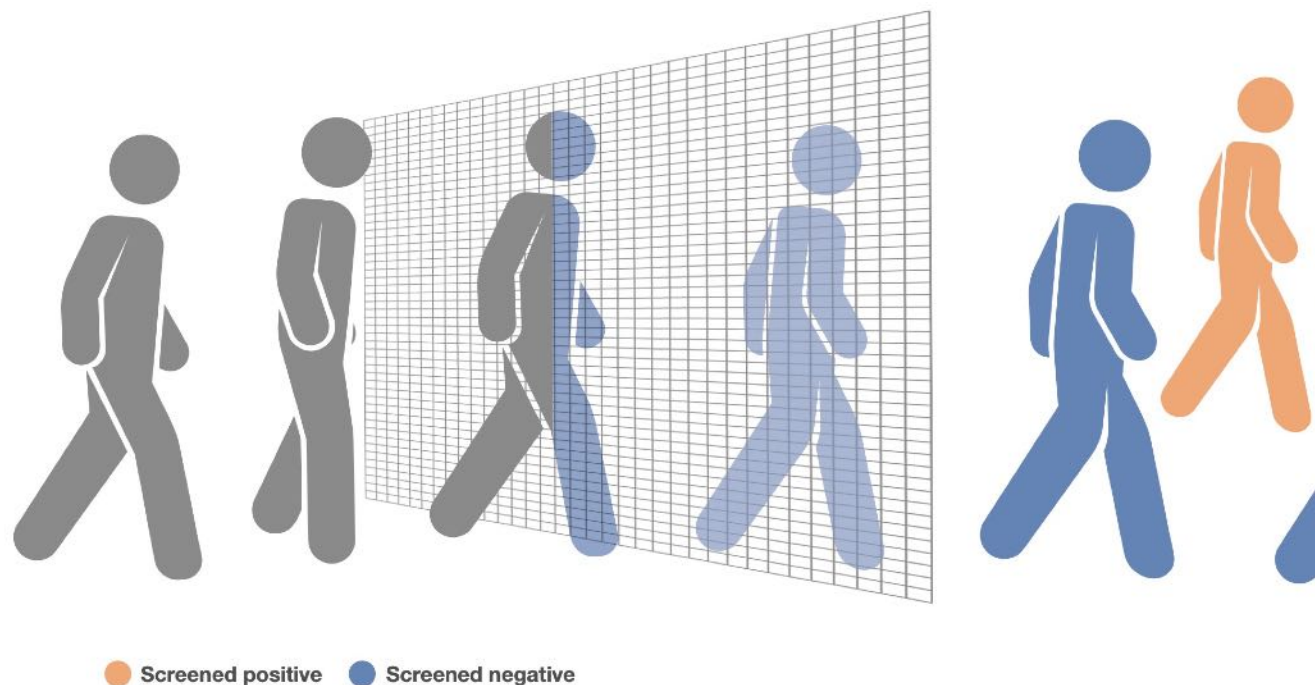
Screening (definizione OMS)

Lo screening è un processo di **selezione approssimativo**.

Funziona come un setaccio, separando le persone che probabilmente hanno la condizione da quelle che probabilmente non ce l'hanno.

Un test di screening non è mai accurato al 100%

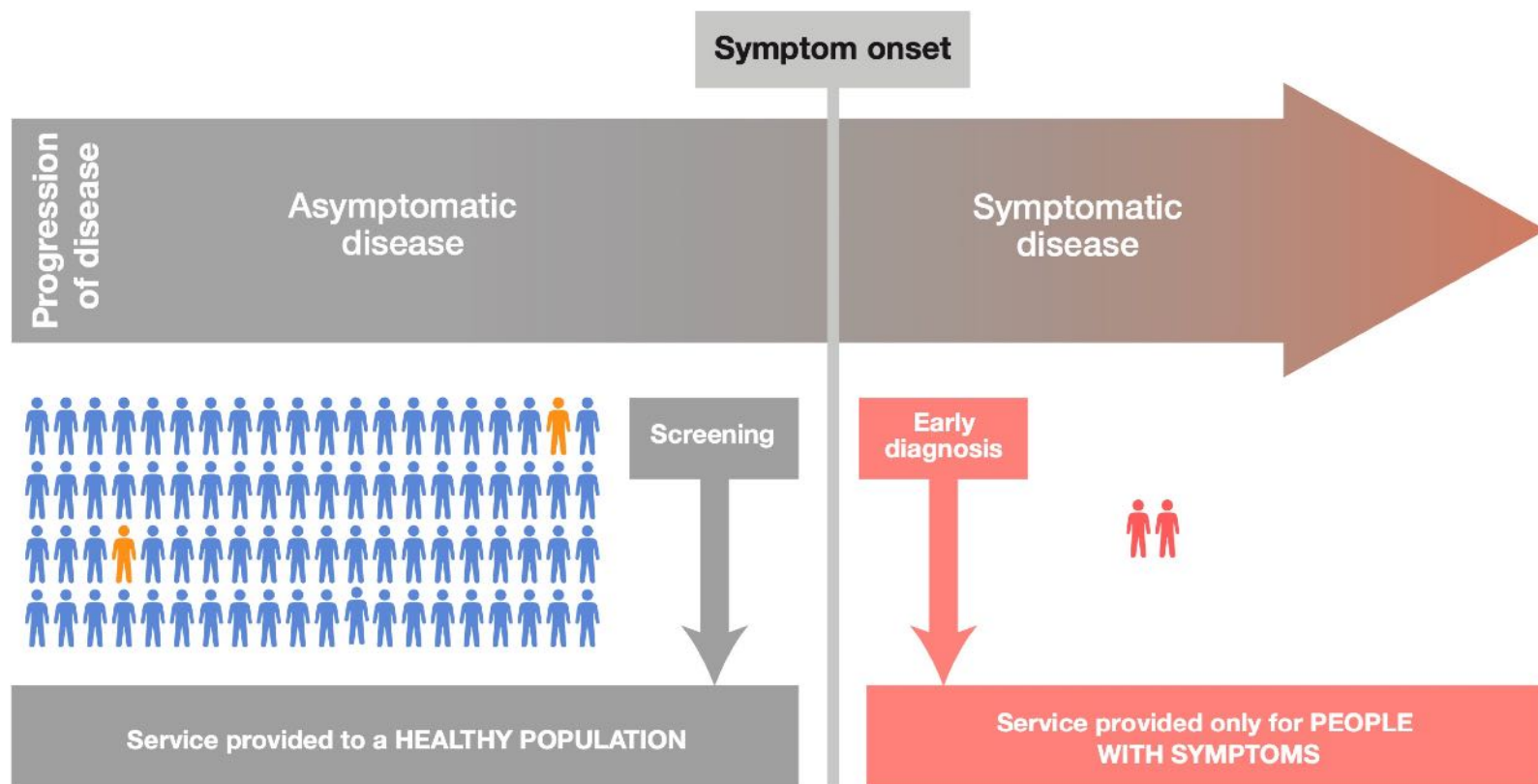
non fornisce la certezza ma solo una probabilità che una persona sia a rischio (o senza rischio) per la condizione di interesse.



Screening è diverso dalla diagnosi precoce

Screening invita le persone che **non hanno sintomi** a sottoporsi al test

Diagnosi precoce ha lo scopo, **tra le persone con sintomi**, di identificare la malattia il più presto possibile



Screening e diagnosi precoce

I programmi di **screening** testano **un gran numero di persone**.

Ciò richiede **notevoli investimenti** in attrezzature, personale e tecnologie informatiche, che possono mettere a dura prova il sistema sanitario.

La **diagnosi precoce**, invece, è una strategia che si concentra **solo sulle persone con sintomi**, che sono **molto meno numerose** e quindi utilizzano **meno risorse**.

Prerequisiti e caratteristiche che rendono uno screening applicabile

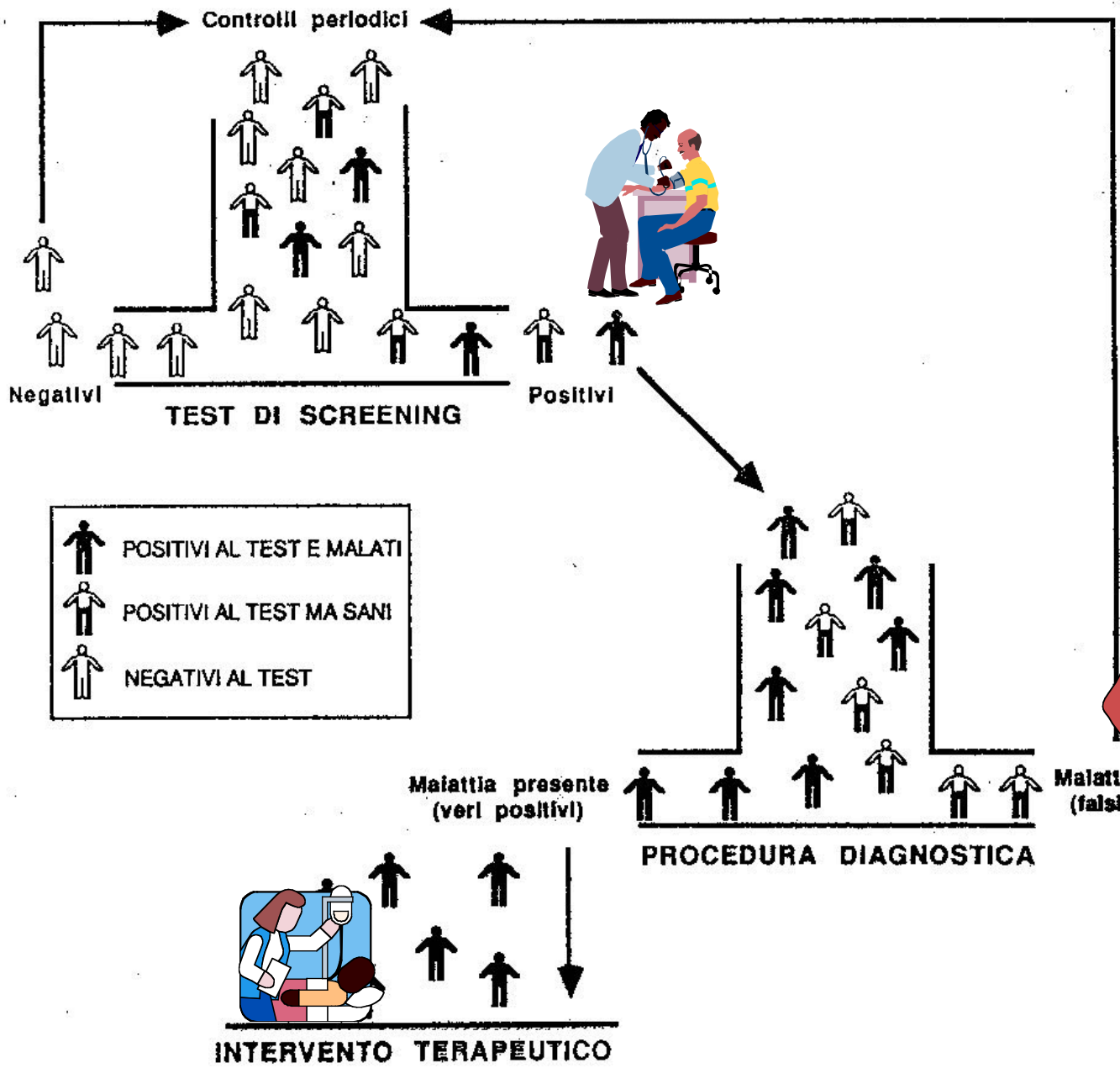


1. La condizione patologica da sottoporre a screening deve essere frequente, grave e con elevati costi individuali e sociali
2. deve esistere una possibilità di intervento per es. deve passare un tempo sufficientemente lungo tra il riconoscimento diagnostico della patologia e la comparsa dei sintomi
3. l'applicazione dell'intervento in epoca precoce (prima della comparsa dei sintomi clinici) deve fornire garanzia di miglior risultato, deve cioè esserci evidenza scientifica che il trattamento o l'intervento precoce possano portare a buoni risultati di salute

Prerequisiti e caratteristiche che rendono uno screening applicabile



4. deve essere disponibile un test:
 - accurato e preciso
 - di facile esecuzione
 - non o minimamente invasivo
 - accettabile dall'utente
 - poco costoso
5. devono esserci risorse economiche, strutturali e di personale



To screen =
passare al
setaccio



Fig. 5.1 – Schema delle varie fasi di una campagna di screening.

Risultati possibili del test

	malattia +	malattia -	
Test +	Veri Pos	Falsi Pos	Totale Pos
Test -	Falsi Neg	Veri Neg	Totale Neg
	Totale M+	Totale M-	N

Risultati possibili del test

	malattia +	malattia -	
Test +	a	b	a+b
Test -	c	d	c+d
Totale	a+c	b+d	N

Sensibilità

	M+	M-	
Test +	a VP	b FP	a+b
Test -	c FN	d VN	c+d
Totale	a+c	b+d	N

Capacità del test di individuare come positivi i soggetti malati in una popolazione.

E' la proporzione di soggetti realmente malati e positivi al test (VP) sul totale dei malati

$$\text{SENSIBILITA}' = a/(a+c)$$

Specificità

	M+	M-	
Test +	a VP	b FP	a+b
Test -	c FN	d VN	c+d
Totale	a+c	b+d	N

Capacità del test di individuare come negativi i soggetti sani in una popolazione.

E' la proporzione di soggetti realmente sani e negative al test (VN) sul totale dei sani

$$\text{SPECIFICITA}' = d/(b+d)$$

Paragone con allarme antifurto dell'auto 1

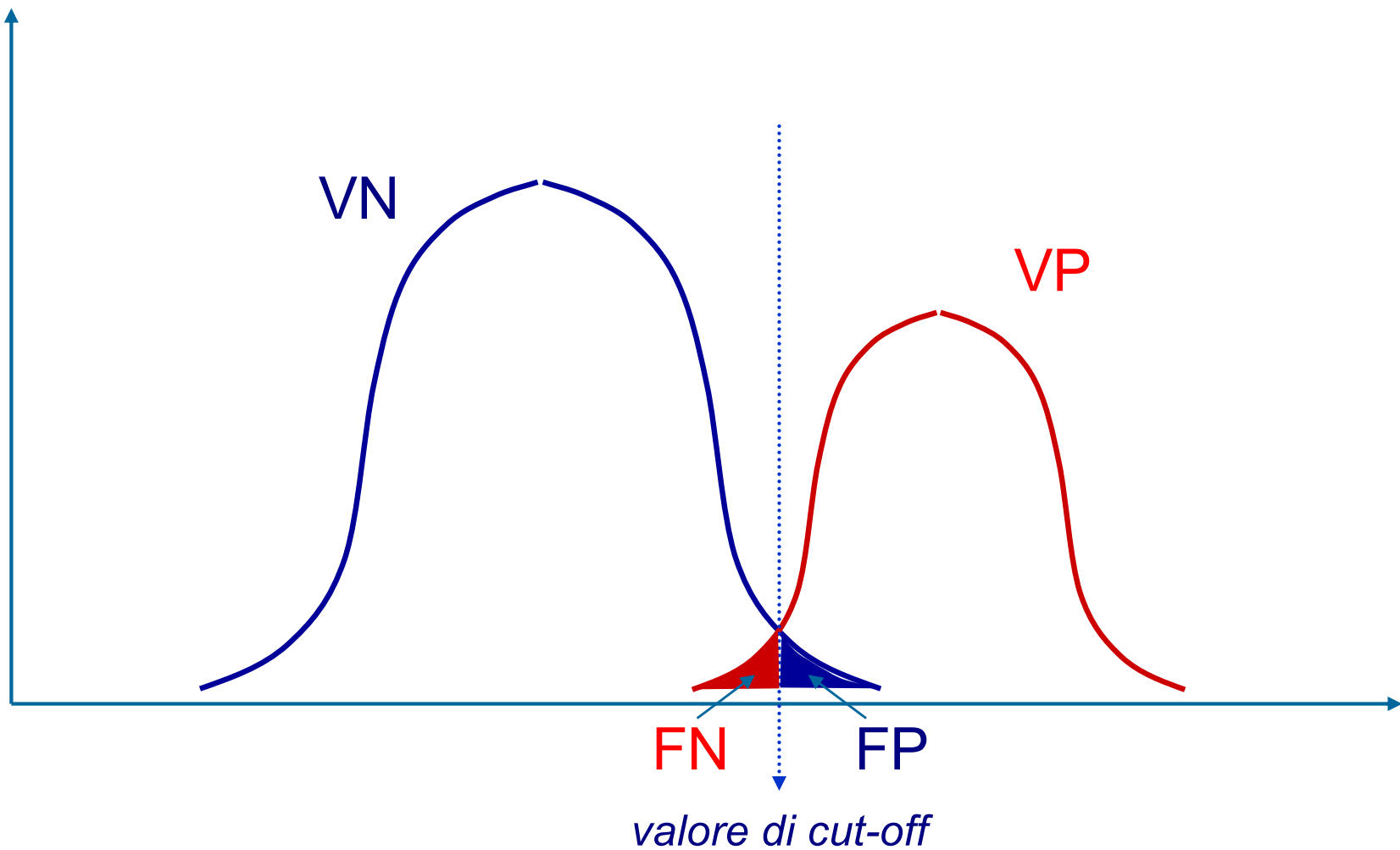
tratto da «Epidemiologia Facile» Lopalco-Tozzi

Supponiamo che il proprietario possa tarare sensibilità e specificità dell'allarme antifurto della propria auto.

Valori molto alti di Sensibilità farebbero scattare l'allarme molto spesso (volo di mosca, spostamento d'aria per passaggio di mezzo pesante...), quindi si potrebbero verificare **molti falsi allarmi = falsi positivi** con conseguente scarsa specificità

Valori molto bassi di Sensibilità riducono i falsi positivi, aumentano la specificità, ma si corre il rischio di avere **falsi negativi** cioè che l'allarme non suoni quando il ladro ruba veramente la macchina

Il test reale...

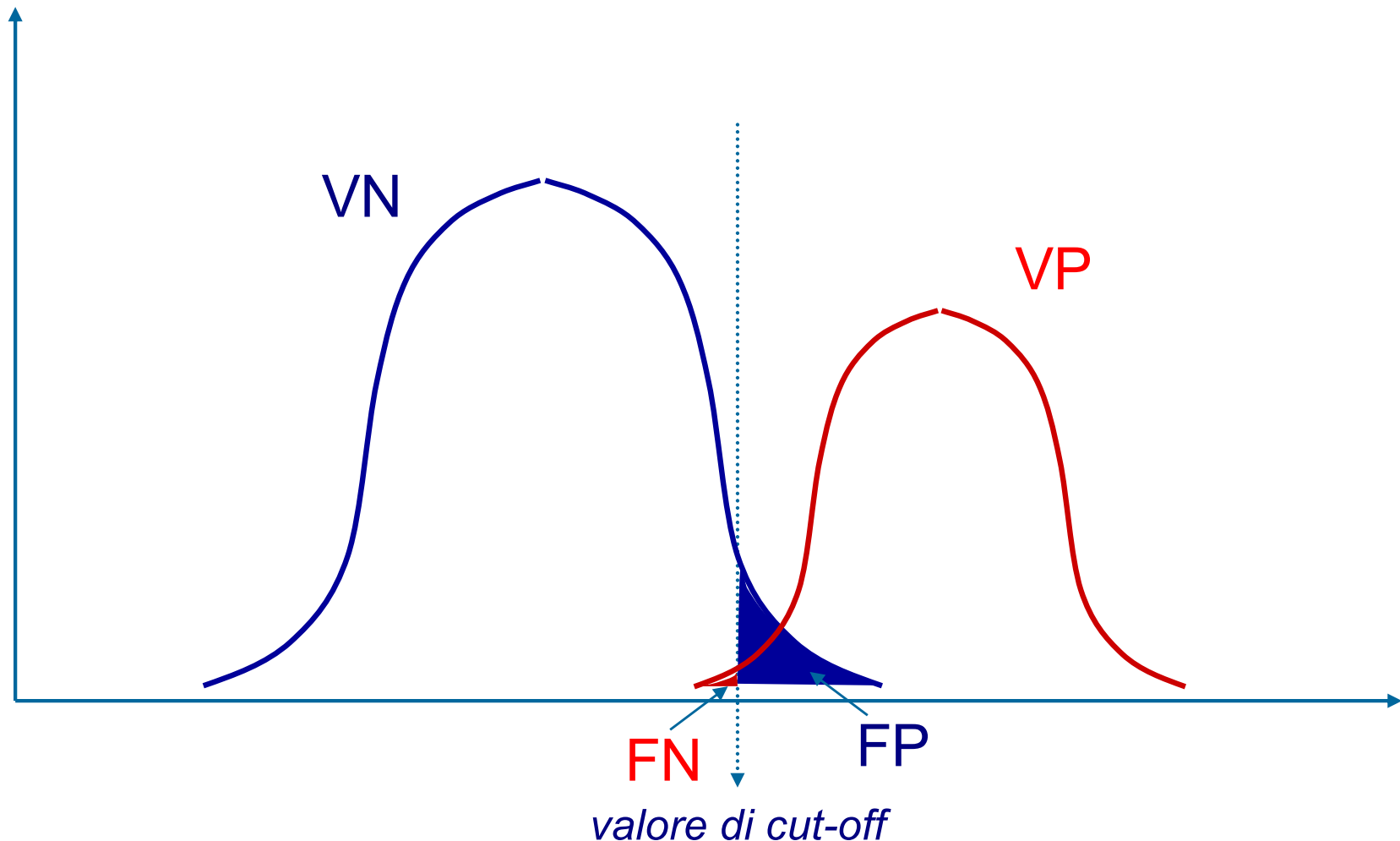


Se si aumenta la sensibilità

Obiettivo: riconoscere il numero più alto possibile di malati (\downarrow FN)

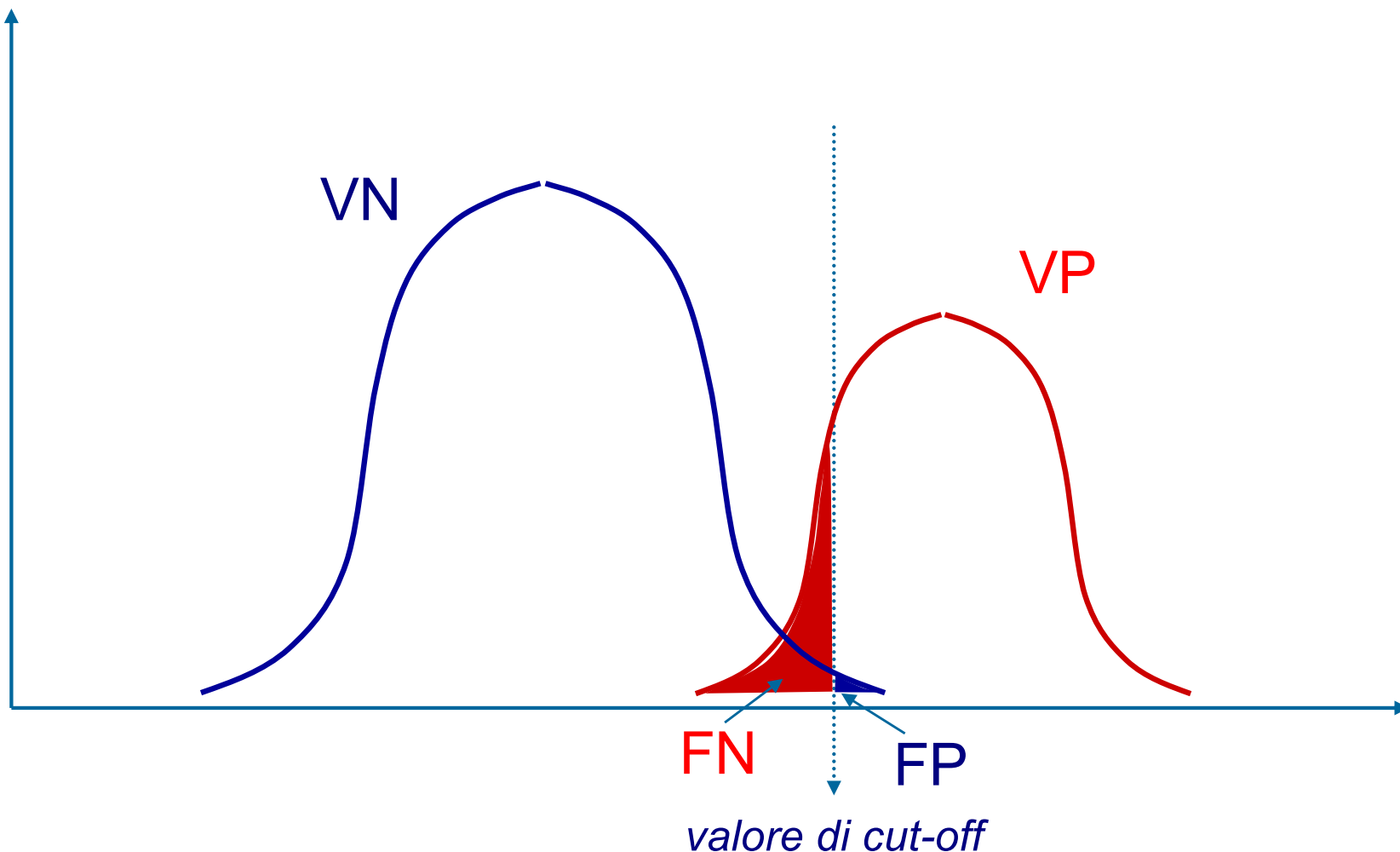
Effetti: abbassa il livello di cut-off e si includono più sani nel gruppo dei positivi (\uparrow FP)

→ DIMINUISCE SPECIFICITA'



Se si aumenta la specificita'

Obiettivo: riconoscere il numero più alto possibile di sani (\downarrow FP)
Effetti: aumenta il livello di cut-off e si “perdono più malati” che vengono classificati come negativi (\uparrow FN),
→ DIMINUISCE SENSIBILITA'



Come scegliere il migliore valore di cut-off? Le curve ROC

Sensibilità e Specificità sono tra loro dipendenti

Modifiche nei valori di cut-off determinano le loro variazioni

Per scegliere il valore di cut-off che massimizzi la sensibilità e la specificità, si usano le curve ROC (Receiver Operating Characteristics).

Le Curve ROC riportano su un grafico, sull'asse y, i diversi valori di sensibilità e sull'asse x, di percentuale di errore falso positivo (cioè $1 - \text{specificità}$) misurati a diversi valori di cut-off.

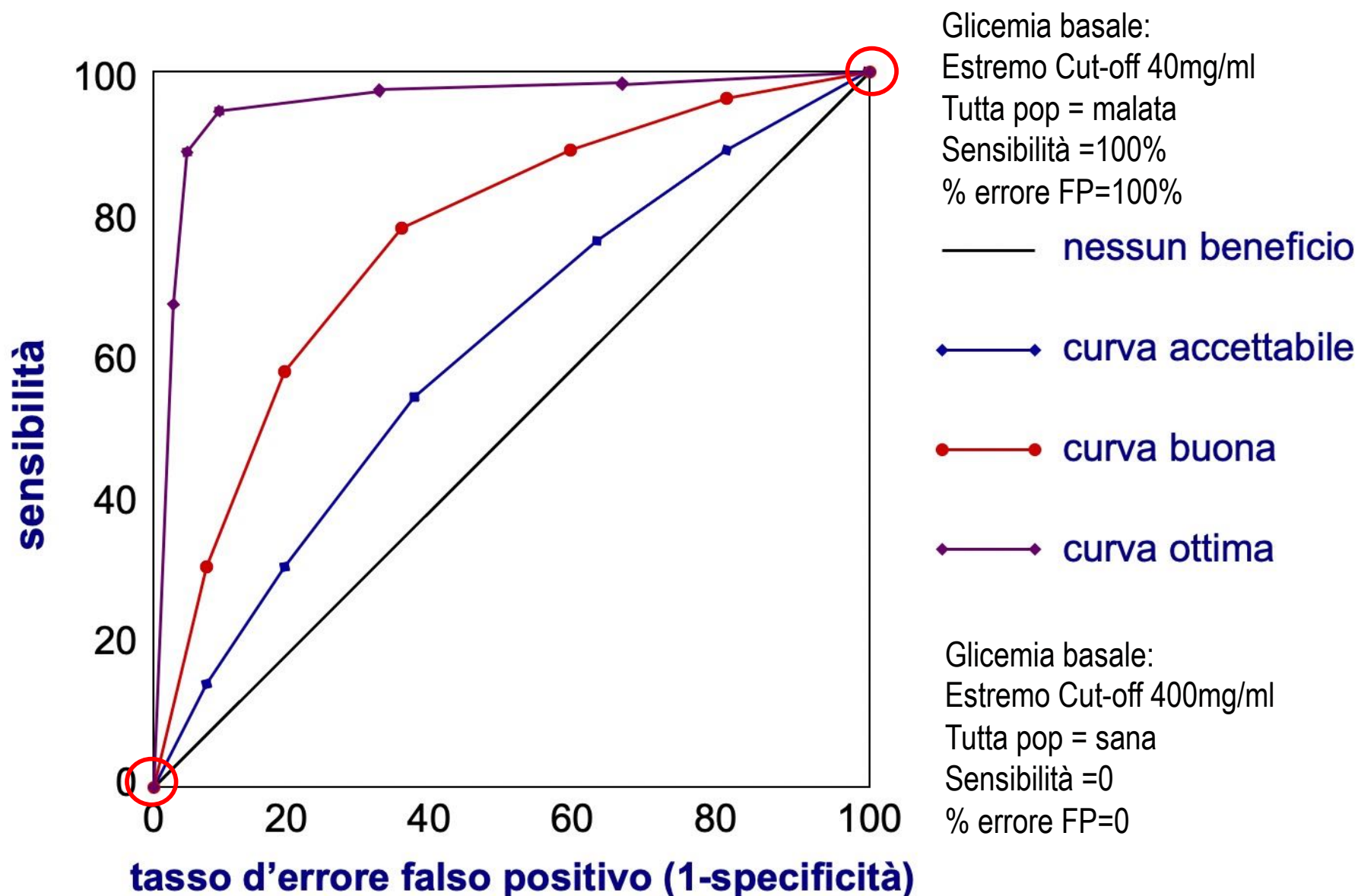
Come scegliere il migliore valore di cut-off? Le curve ROC

Le Curve ROC riportano su un grafico, sull'asse y, i diversi valori di sensibilità e sull'asse x, le percentuali di errore falso positivo (cioè $1 - \text{specificità}$) misurati a diversi valori di cut-off.

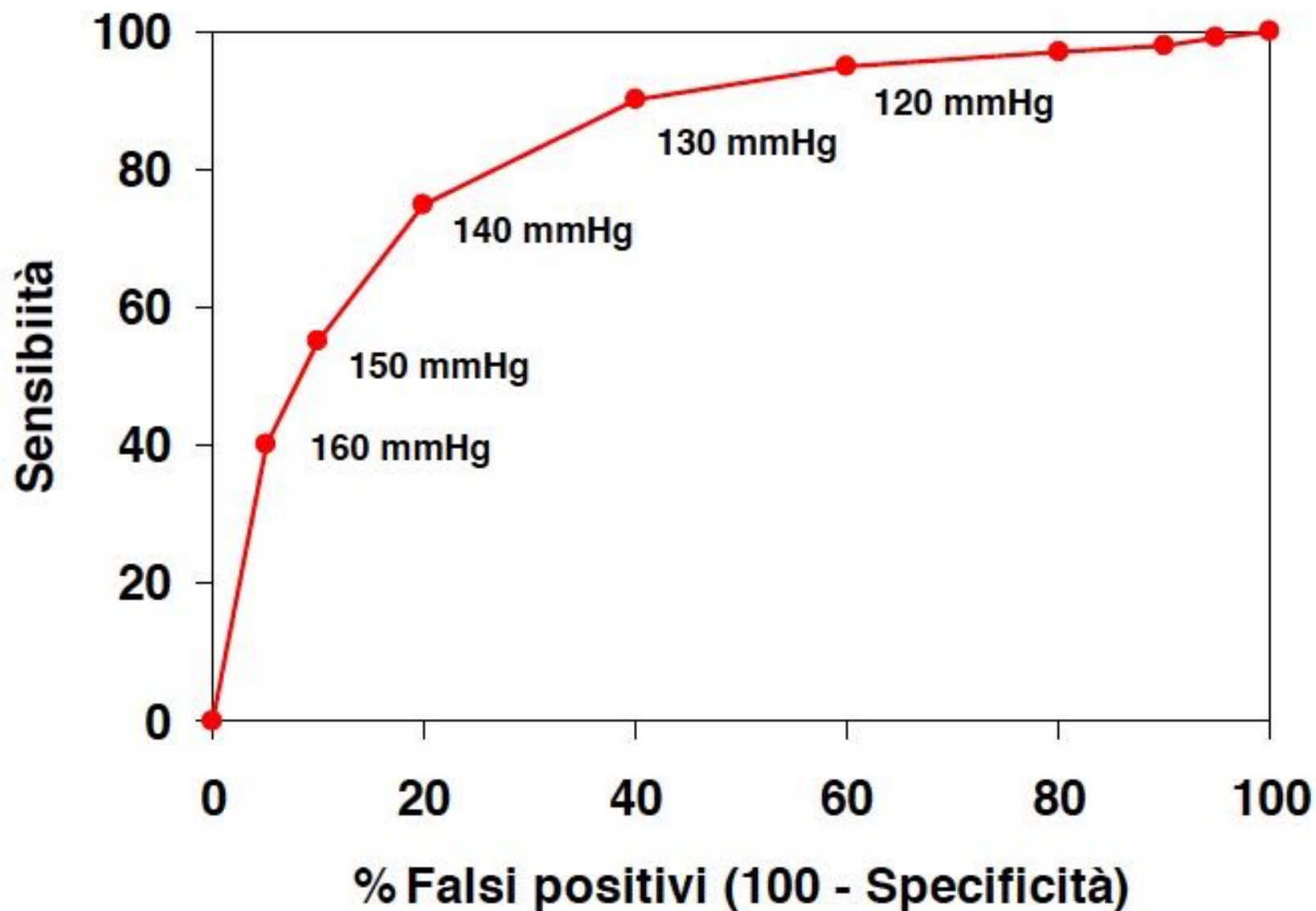
La diagonale è la linea di «nessun beneficio» in cui la sensibilità ha valori sempre uguali ai valori della percentuale di falsi positivi.

Più le curve ROC si allontanano dalla diagonale e si avvicinano all'angolo superiore sinistro, più sono buone perché danno il massimo della sensibilità (100%) e il minimo di errore falso positivo (0%).

Come scegliere il migliore valore di cut-off? Le curve ROC



Curva ROC (Receiver Operating Characteristic) di uno studio per stabilire il miglior valore soglia (cut-off) per un programma di screening per l'ipertensione (dati fittizi)



Analisi delle curve ROC per confrontare test differenti

Le Curve ROC vengono usate anche per confrontare test differenti.

Si calcola l'area sottesa dalla curva ROC per ciascun test, e poi si usa un test statistico di significatività per decidere se l'area sottesa da una curva differisce significativamente dall'area sottesa dall'altra curva.

Maggiore è la percentuale di area sottesa alla curva, migliore è il test.

Valore Predittivo Positivo (VPP)

	M+	M-	
Test +	a VP	b FP	a+b
Test -	c FN	d VN	c+d
Totale	a+c	b+d	N

E' la probabilità che una persona risultata positiva al test, sia realmente malata.

E' la proporzione di soggetti veri positivi al test (VP) sul totale dei positivi (VP+FP)

$$\text{VPP} = a/(a+b)$$

Valore Predittivo Negativo (VPN)

	M+	M-	
Test +	a VP	b FP	a+b
Test -	c FN	d VN	c+d
Totale	a+c	b+d	N

E' la probabilità che una persona risultata negativa al test, sia realmente sana.

E' la proporzione di soggetti veri negativi al test (VN) sul totale dei negativi (VN+FN)

$$\text{VPN} = d / (c + d)$$

ESEMPIO

 tratto da «Epidemiologia Facile» Lopalco-Tozzi

	M+	M-	
Test +	25	2	27
Test -	4	55	59
Totale	29	57	86

Prevalenza=

Sensibilità=

Specificità=

VPP=

VPN=

ESEMPIO

 tratto da «Epidemiologia Facile» Lopalco-Tozzi

	M+	M-	
Test +	25	2	27
Test -	4	55	59
Totale	29	57	86

Prevalenza = $29/86 = 33,7\%$

Sensibilità = $25/29 = 86,2\%$

Specificità = $55/57 = 96,5\%$

VPP = $25/27 = 92,6\%$

VPN = $55/59 = 93,2\%$

Predittività del test e Prevalenza della malattia

I valori predittivi sono parametri che riguardano la pratica clinica.

Nel corso della propria attività, il clinico sceglie i test da utilizzare per giungere alla diagnosi, in funzione di sensibilità e specificità, ma poi deve gestire risultati positivi o negativi del test riferiti ai pazienti.

La predittività di un test non è una caratteristica intrinseca del test, ma dipende dalla prevalenza (frequenza) della malattia.

Il VPP sarà maggiore quando la prevalenza è più elevata e viceversa.

ESEMPIO

 tratto da «Epidemiologia Facile» Lopalco-Tozzi

	M+	M-	
Test +	18	38	56
Test -	2	342	344
Totale	20	380	400

Prevalenza= 5% (0,05)

Sensibilità= 90%

Specificità= 90%

VPP= $18/56 = 32,1\%$

VPN= $342/344 = 99,4\%$

ESEMPIO

 tratto da «Epidemiologia Facile» Lopalco-Tozzi

	M+	M-	
Test +	72	32	104
Test -	8	288	296
Totale	80	320	400

Prevalenza= 20% (0,2)

Sensibilità= 90%

Specificità= 90%

VPP= $72/104 = 69,2\%$

VPN= $288/296 = 97,3\%$

Paragone con allarme antifurto dell'auto 2

tratto da «Epidemiologia Facile» Lopalco-Tozzi

A parità di sensibilità, **la predittività positiva dipende dal contesto = prevalenza!**

La probabilità che ad un allarme corrisponda effettivamente un tentativo di furto (e non un falso allarme) dipende dalla città in cui mi trovo:

- se i furti sono rari, è più probabile che si tratti di un falso allarme
- se i furti sono molto frequenti, è più probabile che si tratti di un vero furto, quindi preoccuparsi e **CONTROLLARE!**

Test confirmatori

Se si risulta positivi ad un primo test, si deve effettuare un secondo test di «conferma».

Generalmente il secondo test è dotato di **maggiore specificità** perché ha lo **scopo di identificare i falsi positivi** ed escluderli.

Questo secondo test si applica in un **contesto** con alta prevalenza perché è una popolazione selezionata dal primo test, quindi il VPP sarà elevato.

Test confirmatori

Lo stesso test diagnostico funzionerà in modo diverso secondo la popolazione che viene ad esso sottoposto, cioè secondo il **CONTESTO** in cui si trova la persona.

- Se si applica alla popolazione generale, la probabilità di identificare un caso di malattia sarà uguale alla prevalenza.
- Se si applica ai pazienti che afferiscono ad un ambulatorio specialistico, la probabilità di identificare un caso di malattia sarà maggiore rispetto a quella osservata nella popolazione generale.

La probabilità di identificare un caso di malattia si chiama **PROBABILITA' PRE-TEST** che dipende da: prevalenza della malattia nella popolazione, età, sesso, presenza di sintomi tipici, contesto in cui si trova la persona.

Positive Likelihood ratios o Rapporto di verosimiglianza positivo

	M+	M-	
T+	a (VP)	b (FP)	TP
T-	c (FN)	d (VN)	TN
	TM+	TM-	N

LR+ proporzione di veri positivi rispetto alla proporzione di falsi positivi $\frac{\%VP \text{ cioè } (a/a+c)}{\%FP \text{ cioè } (b/b+d)} = \frac{\text{sensibilità}}{1-\text{specificità}}$

Poiché LR+ è dato dal rapporto tra ciò che i clinici desiderano (sensibilità) e ciò che vorrebbero evitare (proporzione di errori di falsi positivi), più elevato è il valore di questo rapporto, migliore è il test (test buono se LR+ >10)

NB non è influenzato dalla prevalenza della malattia (Jekel, Katz, Elmore, Wild)

Negative Likelihood ratios o Rapporto di verosimiglianza negativo

	M+	M-	
T+	a (VP)	b (FP)	TP
T-	c (FN)	d (VN)	TN
	TM+	TM-	N

LR- = $\frac{\text{proporzione di falsi negativi rispetto alla proporzione di veri negativi}}{\text{proporzione di falsi negativi rispetto alla proporzione di veri negativi}} = \frac{1 - \text{sensibilità}}{\text{specificità}}$

Poiché LR- è dato dal rapporto tra una quantità che i clinici non desiderano (proporzione di errori di falsi negativi) e una quantità che invece desiderano (specificità), più è piccolo LR- (cioè più vicino a 0) meglio (test buono se LR- < 0,1)

NB non è influenzato dalla prevalenza della malattia

(Jekel, Katz, Elmore, Wild)

Likelihood ratios o Rapporti di verosimiglianza

	M+	M-	
T+	a (VP)	b (FP)	TP
T-	c (FN)	d (VN)	TN
	TM+	TM-	N

LR+ proporzione di veri positivi rispetto alla proporzione di falsi positivi $\frac{\%VP \text{ cioè } (a/a+c)}{\%FP \text{ cioè } (b/b+d)} = \frac{\text{sensibilità}}{1-\text{specificità}}$

LR- proporzione di falsi negativi rispetto alla proporzione di veri negativi $\frac{\%FN \text{ cioè } (c/a+c)}{\%VN \text{ cioè } (d/b+d)} = \frac{1-\text{sensibilità}}{\text{specificità}}$

Test Multipli

➤ IN PARALLELO:

Si tratta di una serie di test che il soggetto effettua tutti insieme.

Il paziente sarà positivo, nel caso in cui risulti positivo ad almeno 1 dei test.

➤ IN SERIE:

Si tratta di una serie di test da effettuarsi in sequenza, ciascuno dei quali riservato ai soggetti risultati positivi al test precedente.

Test Multipli

➤ IN PARALLELO:

Per diagnosi corrette in tempi brevi o per uso di test con scarsa sensibilità (es. diagnosi di appendicite in Pronto Soccorso)

➤ IN SERIE:

Per uso di test con scarsa specificità (programmi strutturati di screening es. oncologici)

TEST IN PARALLELO: SENSIBILITA'

➤ Diagnosi di appendicite in P.S.

	Sensibilità	Specificità
Manovre Semeiotiche	60,0%	75,2%
Formula Leucocitaria	66,9%	79,4%
Combinazione Dei test	86,7%	60,0%

TEST IN SERIE: SPECIFICITA'

➤ Preferenzione secondaria del K. Mammario

	Sensibilità	Specificità
Mammografia	75,0%	92,3%
Citologia	87,7%	94,7%
Entrambi	66,7%	99,6%
i test in sequenza		

Valutazione efficacia screening (1)

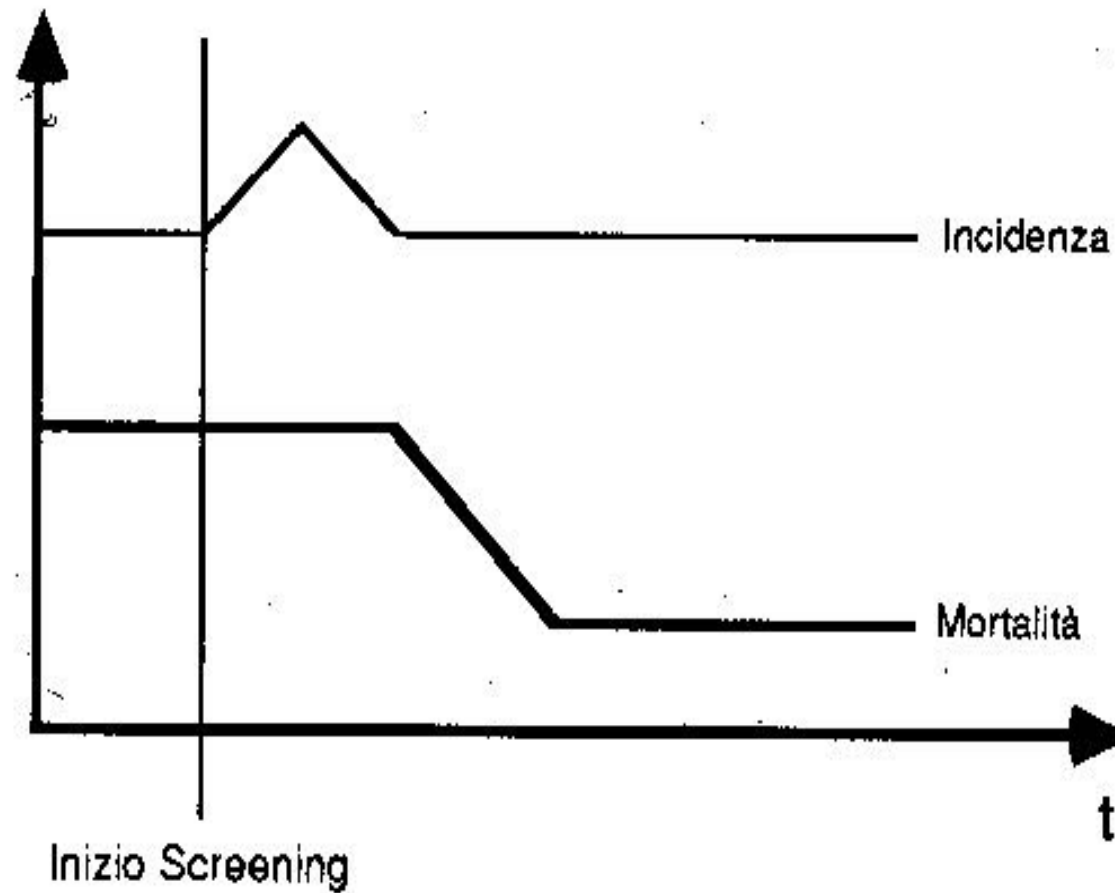


Fig. 5.4 - Andamento dell'incidenza e della mortalità dopo aver intrapreso una campagna di screening basata sulla identificazione precoce

Valutazione efficacia screening (2)

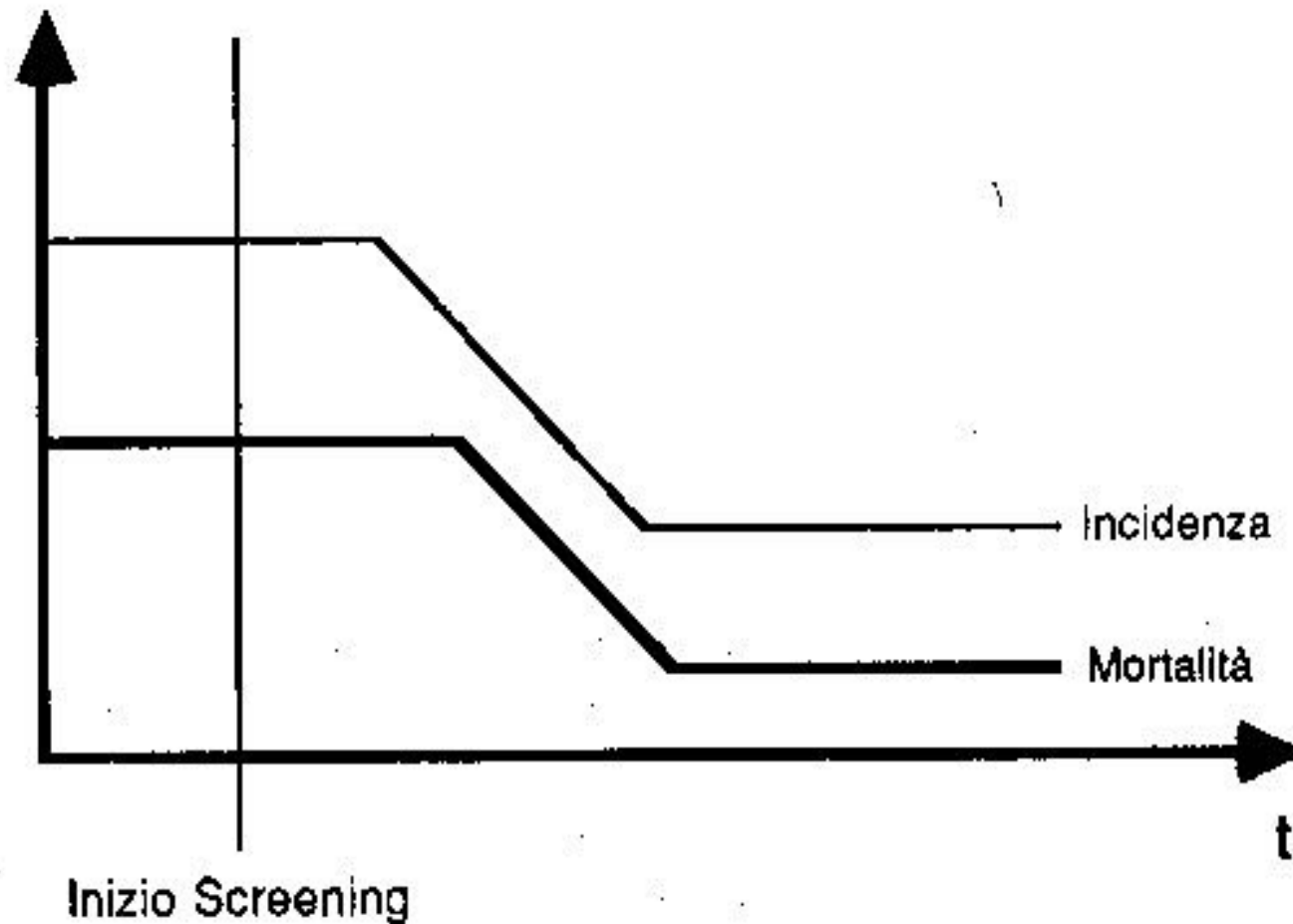


Fig. 5.5 - Andamento dell'incidenza e della mortalità dopo aver intrapreso una campagna di screening basata sull'identificazione di condizioni a rischio.

Esercitazione in piccoli gruppi