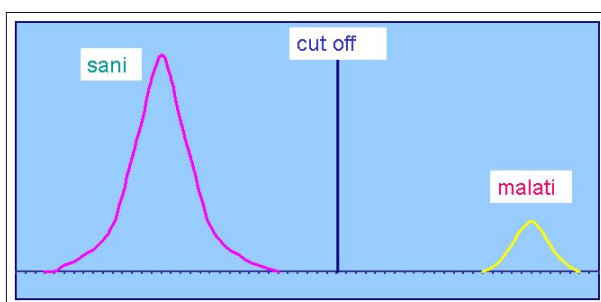


## Sensibilità, specificità, falsi positivi e falsi negativi

### Veri Positivi, Veri Negativi, Falsi Positivi e Falsi Negativi

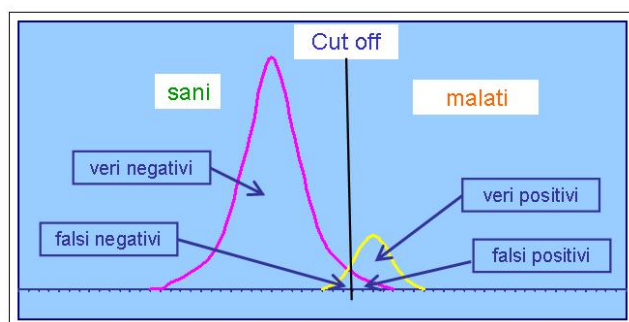
In una situazione ideale ci si aspetta che un test sia in grado di discriminare perfettamente due popolazioni (sani e malati) non sovrapponibili (mutuamente esclusive), come rappresentato nella figura sotto, dove il 'cut off' rappresenta il valore soglia del test.

#### LA SITUAZIONE IDEALE



In realtà quello che solitamente avviene è che le due popolazioni si sovrappongono in parte, ed il test necessariamente identificherà come positivi alcuni soggetti non malati (Falsi Positivi) e come negativi alcuni soggetti invece malati (Falsi Negativi).

#### LA SITUAZIONE REALE



Introduciamo quindi i concetti di **Veri Positivi**, **Veri Negativi**, **Falsi Positivi** e **Falsi Negativi**.

Così ad esempio se si considera un campione di 981 donne in cui viene eseguito, per la diagnosi di lesioni della cervice uterina, il Pap test e si comparano i risultati del Pap test con la diagnosi istologica dopo biopsia (il gold standard) [1], tralasciando i numeri si avrà:

	Biopsia Anormale (gold standard)	Biopsia Normale (gold standard)	
Pap test Anormale	a	b	a + b test positivi
Pap test Normale	c	d	c + d test negativi
	a + c cervice alterata	b + d cervice normale	a + b + c + d totale cervici valutate

Le pazienti rappresentate nella cella "a" sono dette **Vere Positive (VP)**, cioè sono quelle pazienti che risultano positive al test e sono veramente malate.

Le pazienti rappresentate nella cella "d" sono dette **Vere Negative (VN)**, cioè sono quelle pazienti che risultano negative al test e sono veramente sane.

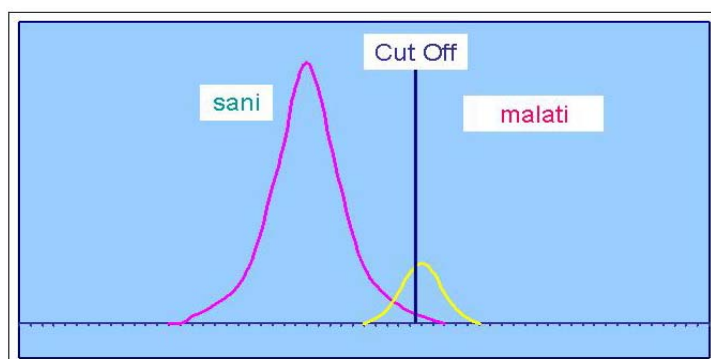
Le pazienti rappresentate nella cella "b" sono dette **False Positive (FP)**, cioè sono quelle pazienti che risultano positive al test ma che nella realtà sono sane.

Le pazienti rappresentate nella cella "c" sono dette **False Negative (FN)**, cioè sono quelle pazienti che risultano negative al test ma che nella realtà sono malate.

### Come si modifica la sensibilità e la specificità al modificarsi del cut off del test?

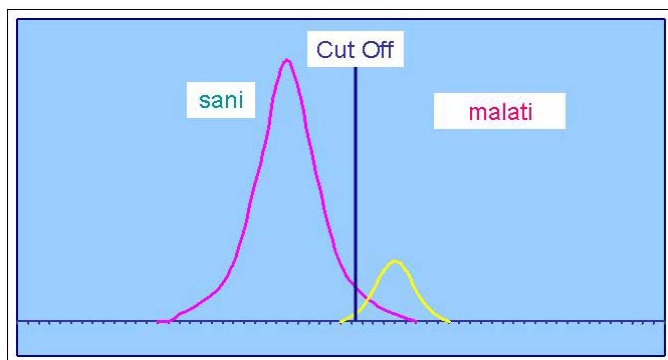
Come abbiamo già detto non esiste un test che distingua nettamente le due popolazioni dei sani e malati. E' chiaro che se la soglia per definire positivo o negativo un soggetto viene spostata a destra o sinistra (cioè innalzata o abbassata) avrà un rischio maggiore o minore di FP e di FN. Questo influenzerà la sensibilità e la specificità del test stesso.

### CUT OFF ELEVATO



Un **cut off elevato** permetterà di identificare correttamente la maggior parte dei sani, conferendo al test un'**elevata specificità** (quindi pochi FP), ma sottostimerà la proporzione dei malati, conferendo al test una **bassa sensibilità** (quindi molti FN).

### CUT OFF BASSO



Un **cut off basso** al contrario permetterà di identificare correttamente la maggior parte dei malati, conferendo al test un'**elevata sensibilità** (pochi FN), ma sottostimerà la proporzione dei sani, conferendo al test una **bassa specificità** (molti FP).

Riprendiamo l'esempio del Pap test e della biopsia per la diagnosi di lesione della cervice uterina [1]. Sappiamo che in realtà sia il risultato del Pap test che quello della biopsia viene classificato in più di due classi [link sistema Bethesda 2001 [http://www.saperidoc.it/file/303\\_1.pdf](http://www.saperidoc.it/file/303_1.pdf)].

Esempio cut off basso: se viene considerato anormale e quindi indicativo di rischio di cancro un Pap test ogni qualvolta sia non normale, definendo quindi positive tutte le lesioni di tipo ASCUS/AGC, LSIL, HSIL oltre che al cancro vero e proprio, ottengo:

	<b>Biopsia Anormale (gold standard)</b>	<b>Biopsia Normale (gold standard)</b>	
<b>Pap test Anormale</b>	<b>125</b> a	86 b	<b>211</b> a + b test positivi
<b>Pap test Normale</b>	126 c	<b>644</b> d	<b>770</b> c + d test negativi
	<b>251</b> a + c cervice alterata	<b>730</b> b + d cervice normale	<b>981</b> a + b + c + d totale cervici valutate

**Sensibilità del test**  $125/251 = 49,8\%$

**Specificità del test**  $644/730 = 88,2\%$

Esempio di cut off elevato: se si desidera essere più specifici ed identificare come positive solo quelle lesioni altamente sospette per malignità, si possono classificare i risultati del Pap test e considerare normale sia una biopsia normale che una biopsia positiva per ASCUS/AGC, considerando anormale solo le lesioni LSIL, HSIL e il cancro. In pratica si sposta il cut off verso l'alto. La tabella si modificherà in questo modo:

	<b>Biopsia Anormale (gold standard)</b>	<b>Biopsia Normale (gold standard)</b>	
<b>Pap test Anormale</b>	<b>99</b> a	44 b	<b>143</b> a + b test positivi
<b>Pap test Normale</b>	152 c	<b>686</b> d	<b>838</b> c + d test negativi
	<b>251</b> a + c cervice alterata	<b>730</b> b + d cervice normale	<b>981</b> a + b + c + d totale cervici valutate

**Sensibilità del test**  $99/251 = 39,4\%$

**Specificità del test**  $686/730 = 93,9\%$

In questo caso innalzando il cut off si riduce la sensibilità mentre la specificità aumenta.

La decisione di fissare il cut off di un test in modo da essere più sensibili o più specifici dipende dall'obiettivo del test stesso. Se si vuole fare uno screening per identificare una situazione frequente, letale e curabile si preferirà essere meno specifici ma più sensibili (non rischiare di perdere nessun caso anche a costo di avere più falsi positivi). Invece nel caso di una malattia rara e poco invalidante, o in presenza di una scarsità di risorse, o nel caso la terapia da instaurare sia gravata da elevati effetti collaterali, o la diagnosi sia particolarmente stigmatizzante, sarà importante garantire un'elevata specificità nella diagnosi, anche a costo di sacrificare la sensibilità del test stesso, quindi avere più falsi negativi ma pochissimi falsi positivi.

Nella realtà dei paesi con elevate risorse spesso si procede con test in serie, con una prima batteria di test molto sensibili e poco specifici per identificare un primo gruppo di persone a rischio da sottoporre ad ulteriori accertamenti, ed una successiva serie di test più specifici (e solitamente più costosi) che siano in grado di discriminare i malati veri dai falsi positivi. E' quello che avviene, ad esempio nel caso di vogliono identificare i feti con sindrome di Down. Si inizia selezionando la popolazione di donne ritenute a rischio in base al criterio dell'età (primo test), si procede poi con il test ecografico della traslucenza nucale (TN): questo è un test sensibile e poco specifica. Nel caso in cui la TN risulti positiva si procede con l'amniocentesi e con lo studio del DNA (specifico per sindrome di Down). [2] Un altro esempio è rappresentato

dalla diagnosi di infezione da HIV: si inizia con un test immunoenzimatico per la ricerca degli anticorpi (ELISA, meno costoso e molto sensibile) e poi sui pazienti risultati positivi si esegue un altro test di conferma come il Western blot o l'immunofluorescenza (più costosi ma più specifici) oppure si eseguono test ELISA che utilizzano antigeni differenti in serie. [3]

## Referenze

[1] Longatto Filho A, Pereira SMM, Di Loreto C, et al. DCS liquid-based system is more effective than conventional smears to diagnosis of cervical lesions: study in high-risk population with biopsy-based confirmation. *Gynecol Oncol* 2005;97:497-500

[2] Malone FD, Canick JA, Ball RH, Nyberg DA, Comstock CH, Bukowski R, Berkowitz RL, Gross SJ, Dugoff L, Craigo SD, Timor-Tritsch IE, Carr SR, Wolfe HM, Dukes K, Bianchi DW, Rudnicka AR, Hackshaw AK, Lambert-Messerlian G, Wald NJ, D'Alton ME; First- and Second-Trimester Evaluation of Risk (FASTER) Research Consortium. First-trimester or second-trimester screening, or both, for Down's syndrome. *N Engl J Med*. 2005;353:2001-11

[3] Babona DV, Slama G, Puiahi E. Laboratory diagnosis of HIV infection in Papua New Guinea. *P N G Med J*. 1996;39:200-4