

Introduzione alla valutazione di un *test* diagnostico: sensibilità, specificità, valore predittivo

Francesco Franco¹, Anteo Di Napoli²

¹Regione Lazio, Roma

²Comitato Tecnico-Scientifico RIDT, Roma

An introduction to the evaluation of a diagnostic test: sensitivity, specificity, predictive value

Starting from a 2×2 cross tabulation of “test response” vs “true disease status” it is possible to calculate the sensitivity, the specificity, the positive predictive value and the negative predictive value of that test. Sensitivity is the probability that a person with the disease will be correctly classified by the test. Specificity is the probability that a person without the disease will be correctly classified by the test. Positive predictive value is the probability that a person with a positive test result is affected by the disease. Negative predictive value is the probability that a person with a negative test result is not affected by the disease. Sensitivity and specificity are characteristic of the test, while predictive values are influenced by the prevalence of the disease in the tested population.

Keywords: Diagnostic test, Predictive values, Sensibility, Specificity



Francesco Franco

Il dibattito in corso sulla cosiddetta “appropriatezza prescrittiva”, avviatosi in coincidenza della presentazione, da parte del Ministro della Salute, del decreto che individua le condizioni di erogabilità e le indicazioni per oltre duecento prestazioni radiologiche, di laboratorio e non solo, ribadisce l’importanza per il clinico di porsi in maniera critica rispetto alla valutazione di un *test* diagnostico. Il presente articolo si propone di fornire ai clinici, e non solo, qualche informazione

introduttiva a un argomento così cruciale.

Come è ben noto, la valutazione dello stato di salute di un paziente è il risultato di un insieme di informazioni ottenute attraverso l’anamnesi, l’esame obiettivo e gli eventuali esami di laboratorio e radiologici prescritti sulla base di uno specifico quesito diagnostico.

È proprio la possibilità che uno o più esami di laboratorio supportino un medico di medicina generale o uno specialista nel definire un’ipotesi diagnostica a giustificarne la prescri-

zione. Il *test* diagnostico, per essere utile, deve accrescere le conoscenze sulle condizioni di salute del paziente. Pertanto, l’utilità del *test* è massima se esso è applicato nei casi di maggiore incertezza diagnostica; viceversa, l’utilità del *test* si riduce se esso è utilizzato nei casi in cui la probabilità di malattia è già molto alta o molto bassa, indipendentemente dalla sua effettuazione. In sostanza, un *test* diagnostico è utile quando è in grado di fornire al medico informazioni tali da modificare la probabilità che un paziente sia affetto da una determinata patologia, a seconda che il *test* risulti positivo o negativo.

Il ruolo del *test* diagnostico nel supportare il medico nella decisione clinica si fonda sull’assunto che individui malati e sani possano essere discriminati in maniera valida e riproducibile dall’esame prescritto.

I requisiti che un *test* diagnostico deve necessariamente possedere, per la sua applicabilità con successo nella pratica clinica, sono il chiaro significato in riferimento a un appropriato *standard* e l’attendibilità dei risultati ottenuti.

I *test* diagnostici possono essere classificati come qualitativi o quantitativi. Si dicono qualitativi quando classificano gli individui come malati o sani in accordo con la presenza o con l’assenza di sintomi o segni clinici (per esempio, un esame radiologico). Si definiscono, invece, quantitativi i *test* che forniscono risultati misurabili su una scala numerica continua (per esempio, la misurazione della pressione arteriosa); la classificazione dicotomica degli individui come malati o sani a partire da un *test* quantitativo si può realizzare scegliendo un opportuno valore soglia (*cut-off*) che li collochi al di sopra o al di sotto del *cut-off* scelto (la dicotomia di un *test* quantitativo porta a un *test* qualitativo derivato).

Accepted: February 2, 2016

Published online: February 23, 2016

Indirizzo per la corrispondenza:

Dr. Francesco Franco
franco_francesco@fastwebnet.it
Dr. Anteo Di Napoli
anteo.dinapoli@libero.it

La validità di un *test* altro non è che la sua capacità di misurare ciò che corrisponde alla realtà; in termini più formali dal punto di vista statistico, la validità è la capacità della misura di essere corretta in media. È sinonimo di accuratezza e può dirci in che misura quel *test* riproduce i risultati di un altro *test* assunto come riferimento, detto “*gold standard*”, al quale un *test* dovrebbe essere comparato.

Il *gold standard* è un *test* di riferimento definito come migliore approssimazione possibile della reale presenza o assenza della malattia. Da questo punto di vista, la validità di un *test* si esprime come confronto dei risultati ottenuti del *test* che stiamo valutando rispetto ai risultati che si otterrebbero utilizzando il *test* assunto come “*gold standard*”. In altre circostanze, il riferimento può essere costituito da un’“evidenza biologica” non misurata attraverso un *test*.

Nella valutazione di un nuovo *test* diagnostico ci sono anche altre dimensioni che vanno considerate: se, a parità di validità, il nuovo *test* ha un costo minore rispetto a quello correntemente utilizzato (ottimizzazione della spesa in termini di costo-beneficio), se utilizza una metodica più semplice per l’operatore, se fornisce una diagnosi più rapida o più precoce, se comporta un minor disagio per il paziente e se presenta un minor numero di complicanze.

La validità di un *test* diagnostico, in definitiva, si misura in base al risultato positivo o negativo rispetto al quesito diagnostico e alla presenza o all’assenza della malattia.

Nell’articolo relativo alle applicazioni della curva “normale” al *test* diagnostico era stata presentata la cosiddetta Tabella 2 × 2, in un certo senso punto di partenza e di arrivo dell’epidemiologia, nella quale si incrociano l’informazione relativa al risultato positivo/negativo del *test* e quella sullo stato di presenza/assenza della malattia (Tab. I).

La Tabella di contingenza 2 × 2 ha una grande utilità operativa, in quanto, a partire dalla relazione tra risultato del *test* e stato di salute, è possibile calcolare la sensibilità, la specificità e il valore predittivo (positivo e negativo) di un *test* diagnostico. Sensibilità e specificità sono indicatori di validità relativa di un *test* diagnostico, rappresentando la proporzione di pazienti classificati correttamente rispetto a un *gold standard* (1).

Convenzionalmente, nella cella **a** sono classificati i veri positivi, vale a dire i soggetti con la malattia che risultano positivi al *test*, e, nella cella **b**, sono classificati i falsi positivi, vale a dire i soggetti senza la malattia che risultano positivi al *test*; nella cella **c** sono classificati i falsi negativi, vale a dire i soggetti con la malattia che risultano negativi al *test* e, nella cella **d**, sono classificati i veri negativi, vale a dire i soggetti senza la malattia che risultano negativi al *test*.

Nell’ambito dei *test* di verifica delle ipotesi statistiche, i falsi positivi rappresentano l’errore di primo tipo (probabilità dell’errore α) mentre i falsi negativi rappresentano l’errore di secondo tipo (probabilità dell’errore β)(2-4).

La sensibilità misura la probabilità condizionale di essere positivi al *test* essendo malati (potenza diagnostica del *test*).

TABELLA I - Test diagnostico e condizione del paziente

	Malati	Non malati	Totale	
TEST	Positivi	a	b	(a + b)
	Negativi	c	d	(c + b)
	Totale	(a + c)	(b + d)	(a + b + c + d)
	Malati	Non malati	Totale	
TEST	Positivi	VP	FP	(VP + FP)
	Negativi	FN	VN	(FN + VN)
	Totale	(VP + FN)	(FP + VN)	(VP + FP + FN + VN)

VP = veri positivi; FP = falsi positivi; VN = veri negativi; FN = falsi negativi.

Essa definisce la capacità del *test* di identificare i soggetti affetti da una determinata malattia e si calcola come proporzione di persone malate che risultano positive al *test* diagnostico. La sensibilità è data dal rapporto tra veri positivi e il totale degli ammalati: **sensibilità = $a/(a + c)$** . Il complemento della sensibilità è la proporzione di falsi negativi, definiti come la frazione di malati negativi al *test*: **$c/(a + c)$** .

Un *test* con elevato valore di sensibilità, poiché determina un numero minimo di falsi negativi, è necessario quando è indispensabile escludere un forte sospetto diagnostico; è fondamentale per confermare l’assenza di malattia quando il clinico deve essere certo, come, per esempio, per la diagnosi precoce di una malattia grave ma per la quale si ha un trattamento efficace (*test* di *screening*). Viceversa, un *test* con basso valore di sensibilità determina un elevato numero di falsi negativi e, poiché non identifica come malati soggetti che hanno la malattia, può determinare il rischio di non trattare persone ammalate.

La specificità misura la probabilità condizionale di essere negativi al *test* essendo sani. Essa definisce la capacità del *test* di identificare i soggetti non affetti da una determinata malattia. La specificità si calcola come proporzione di persone non malate che risultano negative al *test* diagnostico. Tale misura è data dal rapporto tra veri negativi e il totale dei non ammalati: **specificità = $d/(b + d)$** . Il complemento della specificità è la proporzione di falsi positivi, definiti come la frazione di non malati positivi al *test*: **$b/(b + d)$** .

Un *test* con elevato valore di specificità, poiché determina un numero minimo di falsi positivi, è necessario quando è indispensabile avere la massima certezza diagnostica (*test* di conferma) ed è fondamentale per verificare la presenza di malattia con una bassa probabilità a priori. Viceversa, un *test* con basso valore di specificità determina un elevato numero di falsi positivi e, poiché identifica come malati soggetti che non hanno la malattia, può determinare il rischio di trattare persone sane.

Sensibilità e specificità sono due caratteristiche interdipendenti del *test* diagnostico, per cui non è possibile aumen-

TABELLA II - Valore predittivo positivo e negativo in due regioni con diversa prevalenza della patologia nella popolazione, a parità di sensibilità e specificità del *test* diagnostico

Regione 1: Prevalenza della malattia nella popolazione: 10%					
		Gold standard			
		Positivo	Negativo	Sensibilità	0.8
Test Diagnostico	Positivo	800	900	Specificità	0.9
	Negativo	200	8100	VPP	0.471
				VPN	0.976
Regione 2: Prevalenza dalla malattia nella popolazione: 1%					
		Gold standard			
		Positivo	Negativo	Sensibilità	0.8
Test Diagnostico	Positivo	80	990	Specificità	0.9
	Negativo	20	8910	VPP	0.075
				VPN	0.998

VP = Valore Predittivo Positivo; VN = Valore Predittivo Negativo

tare la sensibilità se non diminuendo la specificità e viceversa.

Sensibilità e specificità, come pure la proporzione dei falsi positivi e dei falsi negativi, non sono influenzate dalla prevalenza della malattia nella popolazione (non variano in funzione della proporzione di sani e ammalati nel gruppo in studio), ma dipendono dalle caratteristiche del *test* diagnostico.

Il medico, di solito, più che alla sensibilità e alla specificità di un *test* è interessato al suo valore predittivo. In sostanza, data la positività o negatività del *test*, è interessato a sapere qual è la probabilità che il paziente sia davvero malato o sano.

Il valore predittivo positivo misura la probabilità condizionale di essere malati essendo risultati positivi al *test* diagnostico; rappresenta la proporzione di soggetti positivi al *test* che hanno la malattia. Il valore predittivo positivo di un *test* è dato dal rapporto tra i veri positivi e il totale dei soggetti positivi al *test*: **valore predittivo positivo = $a/(a + b)$** .

Il valore predittivo negativo misura la probabilità condizionale di essere sani essendo risultati negativi al *test* diagnostico; rappresenta la proporzione di soggetti negativi al *test* che non hanno la malattia. Il valore predittivo negativo di un *test* è dato dal rapporto tra i veri negativi e il totale dei soggetti negativi al *test*: **valore predittivo negativo = $d/(c + d)$** (5).

La sensibilità e la specificità di un *test* diagnostico dipendono esclusivamente dalle sue caratteristiche. Invece, il valore predittivo di un *test* dipende, oltre che da sensibilità e specificità del *test*, quindi dalle sue caratteristiche, anche dalla prevalenza nella popolazione della malattia oggetto del quesito diagnostico.

Il valore predittivo (positivo e negativo) può essere anche calcolato come (6):

$$VPP = (sensibilità \times prevalenza) / (sensibilità \times prevalenza + (1 - specificità) \times (1 - prevalenza))$$

$$VPN = specificità \times (1 - prevalenza) / ((1 - sensibilità) \times prevalenza + specificità \times (1 - prevalenza))$$

Quando aumenta la prevalenza, si hanno un aumento del valore predittivo positivo (relazione direttamente proporzionale) e un decremento del valore predittivo negativo (relazione inversamente proporzionale). Viceversa, quando diminuisce la prevalenza si hanno un aumento del valore predittivo negativo e un decremento del valore predittivo positivo.

Nella Tabella II è presentato un esempio che dimostra come, a parità di sensibilità e specificità di un *test* diagnostico, il suo valore predittivo positivo si modifichi notevolmente al variare della prevalenza della malattia di interesse nella popolazione.

In sostanza, la prevalenza di una malattia nella popolazione rappresenta la probabilità di un individuo di quella popolazione di avere quella malattia prima di effettuare un eventuale *test* diagnostico. La prevalenza di una malattia nella popolazione altro non è che la probabilità a priori per un individuo di esserne affetto. Successivamente all'esecuzione di un *test* diagnostico, finalizzato a verificare la presenza di una data patologia, il risultato del *test* rappresenta la verosimiglianza che modifica la probabilità a priori in probabilità a posteriori per un individuo di essere affetto dalla malattia.

Pertanto, un *test* diagnostico è tanto più utile quanto maggiore è la sua capacità di modificare la probabilità a priori di malattia in una a posteriori significativamente diversa dalla prevalenza della patologia nella popolazione di riferimento. La differenza tra la probabilità di malattia a posteriori e a priori rappresenta il guadagno di probabilità determinato dal *test* diagnostico considerato.

Disclosures

Financial support: No financial support was received for this submission. Conflict of interest: The authors have no conflict of interest.

Bibliografia

1. Morabia A. L'Epidemiologia Clinica. Il Pensiero Scientifico Editore, Roma 1999;2:11-4.
2. Jekel JF, Katz DL, Elmore JG. Epidemiologia, Biostatistica e Medicina preventiva. II Edizione. EDISes, Napoli 2005;7:122-8.
3. Armitage P, Berry G. Statistica Medica, metodi statistici per la ricerca in Medicina. III Edizione. Mc Graw-Hill, Milano 1996; 16:527-32.
4. Pagano M, Gauvreau K. Biostatistica. Guido Gnocchi Editore, Napoli 1994;6:96-7.
5. Dunn G, Everitt B. Biostatistica Clinica. Il Pensiero Scientifico Editore, Roma 1999;2:9-14.
6. Altman DG, Bland JM. Diagnostic tests 2: predictive values. BMJ 1994;309:102.