Performance di un test

Data 14 marzo 2008
Categoria scienze_varie

Come si può calcolare la performance di un test?

E' possibile determinare la performance di un determinato test diagnostico usato per diagnosticare una certa malattia. Per poterlo fare è necessario conoscere tre parametri: la sensibilità del test, la specificità e la prevalenza della malattia inquestione.

La [b]sensibilità[/b] di un test si applica ai soggetti affetti dalla malattia e determina quanti di questi soggetti il test riesce a individuare: per esempio se un esame ha una sensibilità del 60% vuol dire che su 100 malati 60 avranno il test positivo (veri positivi o VP) e 40 l'avranno negativo (falsi negativi o FN).

La [b]specificità[/b] di un test si applica invece ai soggetti sani: se un esame ha una specificità del 60% vuol dire che su 100 sani 60 avranno il test negativo (veri negativi o VN) e 40 l'avranno positivo (falsi positivi o FP).

Ovviamente per calcolare il numero di VP, FP, VN e FN bisogna conoscere la **prevalenza** della malattia nella popolazione in esame. Il metodo viene schematizzato nella figura 1 in cui si ipotizza un test con sensibilità del 70%, specificità del 90% e prevalenza della malattia del 10% (su 100 soggetti di quella popolazione sottoposta al test 10 hanno la malattia e 90 non ce l'hanno quindi sono catalogati come sani).

Figura1 Calcolo di VP, V, FP e FN in caso di sensibilità del 70%, specificità del 90% e prevalenza della malattia del 10%.

A questo punto è possibile calcolare i tre parametri che più interessano al medico: il valore predittivo positivo del test (VPP), il valore predittivo negativo (VPP) e l'accuratezza globale del test (overall accuracy).

Il **VPP** risponde alla domanda: su 100 soggetti con test positivo quanti sono quelli affetti dalla malattia (cioè i VP)? Basta dividere il numero dei VP trovato per il numero dei positivi totale (VP +FP) e moltiplicare per 100. Nel caso in esame 7/16 = 0,43 x 100 = 43%. In altre parole ogni 100 test positivi ci si deve aspettare che 43 siano veri e 57 siano dei falsi positivi.

Il **VPN** risponde invece alla domanda: su 100 soggetti con test negativo quanti sono quelli che effettivamente non hanno la malattia (VN)? Basta dividere il numero dei VN trovato per il numero dei negativi totali (VN + FN) e moltiplicare il tutto per 100. Nel caso in esame 81/84 = 0,96 x100 = 96%. Su 100 test negativi ci si deve aspettare che 96 siano corretti mentre 4 sono dei malati (FN).

L'**accuratezza** totale del test (overall accuracy) risponde alla domanda: su 100 test eseguiti quanti sono quelli corretti? E' evidente che la risposta si ottiene sommando i VP e i VN. Nel caso in esame rispettivamente 7 + 81 = 88. Questo vuol dire che il test restituisce una risposta corretta nell'88% dei casi e in 12 casi sbaglia. Si può esprimere l'accuratezza o come percentuale (88%) oppure come corrispondente numero decimale (0,88).
Si noti come il valore ottimale di accuratezza di un test sia 100% (= 1) in quanto in questo caso il numero dei veri positivi e dei veri negativi corrisponde al totale dei soggetti sottoposti al test e non vi sono né falsi positivi né falsi negativi. Questa si verifica, ovviamante, quando sensibilità e specificità sono entrambe del 100%.

Primoscenario

Si applichi ora quanto appena appreso ad un caso reale: il test "fattore reumatoide" e il test "anticorpi anticitrullina" per la diagnosi di artrite reumatoide.
Il fattore reumatoide ha una sensibilità media del 60% e una specificità media del 79%.
Gli anticorpi anticitrullina hanno una sensibilità media del 65% e una specificità media del 95%.
La prevalenza dell'artrite reumatoide stimata nella popolazione italiana è dello 0,5%.
Si otterranno i seguenti risultati

Fattore reumatoide:
VPP = 1,41%
VPN = 99%
Accuracy = 78,9%

Anticorpi anticitrullina
VPP = 6,1%
VPN = 99,8%
Accuracy = 94,8%

Come si può vedere entrambi i test sono molto utili se negativi in quanto escludono quasi con certezza la presenza di artrite reumatoide. Invece se i test risultano positivi la probabilità di avere una artrite reumatoide è bassa, soprattutto con il test "fattore reumatoide".
Tra i due test quello che performa meglio è il dosaggio degli anticorpi anticitrullina perché restituisce una risposta esatta 95 volte su 100 mentre il fattore reumatoide sbaglia 21 volte su 100.

Secondoscenario

Si supponga adesso di somministrare i due test non ad una generica popolazione ma a soggetti che presentano determinati sintomi che lasciano sospettare un' artrite reumatoide. In questo caso ovviamente la prevalenza della malattia non sarà più dello 0,5%, ma molto più elevata perché si sta esaminando una popolazione selezionata. Si ipotizzi che la prevalenza in questa popolazione sia del 50%.
Si otterranno i seguenti risultati:

Fattore reumatoide
VPP = 74%
VPN = 66%
Overall accuracy = 69,5%

Anticorpi anticitrullina
VPP = 92,8%
VPN = 73%
Overall accuracy = 80%

Come si può vedere in questo caso aumenta il VPP dei due test ma diminuisce il VPN. Inoltre in entrambi i casi si riduce l'accuratezza diagnostica. Come si spiega questo fenomeno? A prima vista si potrebbe pensare che se il test viene somministrato ad una popolazione selezionata in cui la prevalenza dell'artrite reumatoide è più elevata il test dovrebbe avere una performance migliore. In realtà, apparentemente, non è così: siccome i due test hanno una sensibilità non elevata si ha, rispetto, allo scenario in cui il test viene somministrato ad una popolazione indifferenziata, un aumento cospicuo dei falsi negativi. Ovviamente questo va ad incidere sulla performance del test. Se per ipotesi la sensibilità dei test fosse stata del 95% l'overall accuracy sarebbe stata simile nei due scenari.
L'apparente riduzione della overall accuracy nel secondo scenario non deve trarre in inganno. Per esempio riferendoci al test "anticorpi anticitrullina" è vero che quando viene somministrato ad una popolazione indifferenziata fornisce un risultato esatto nel 94% dei casi mentre se viene somministrato ad una popolazione selezionata i risultati esatti sono "solo" 80 su 100. Però ipotizziamo un medico che decida di screenare con il test tutti i suoi assistiti: dovrà richiedere moltissimi test e ne avrà quindi, in valore assoluto, molti di sbagliati; al contrario un medico che richiede il test solo a chi presenta sintomi sospetti di artrite reumatoide chiederà il test pochissime volte e di conseguenza il numero di test sbagliati sarà, in valore assoluto, molto minore.
Dal che consegue che gli esami, al di fuori di ben specifici programmi di screening, andrebbero richiesti solo in presenza di un motivato sospetto clinico.

RenatoRossi