Computer peggiora accuratezza mammografia nello screening del k al seno

Data 12 marzo 2008
Categoria oncologia

L'uso della lettura assistita dal computer è associata a una riduzione dell'accuratezza nell'interpretazione delle mammografie di screening.

La lettura assistita da computer (computer-aided detection CAD) può aiutare i radiologi ad identificare reperti mammografici sospetti. Da quando la Food and Drug Administration ha approvato questa tecnologia nel 1998 essa si è diffusa nella pratica, ma al momento non sono noti i suoi effetti sull'accuratezza dell'interpretazione.
Per valutare l'associazione fra l'uso della CAD e la performance degli screening mammografici sono stati revisionati i dati di 429.345 mammografie effettuate in 43 centri diagnostici su 222.135 donne e che avevano rilevato 2352 casi di neoplasia mammaria.
Sono state calcolate la specificità, la sensibilità e il valore predittivo positivo della mammografia di screening con e senza lettura aiutata da computer, così come i tassi delle biopsie e della rilevazione di tumori alla mammella e l'accuratezza complessiva, misurati come l'area sotto la curva ROC (receiver-operating-characteristic).
Sette strutture (16%) hanno implementato la lettura aiutata da computer durante il periodo oggetto di studio costituendo il gruppo sperimentale rispetto alle rimanenti 36 che hanno rappresentato il gruppo di controllo. In tali strutture la specificità diagnostica è scesa dal 90,2% prima dell'implementazione all'87,2% dopo l'implementazione ($p<0,001$), il valore predittivo positivo è sceso dal 4,1% al 3,2% ($p=0,01$) e il tasso di biopsie è cresciuto del 19,7% ($p<0,001$). L'incremento della sensibilità, passato dall'80,4% prima dell'implementazione della lettura aiutata da computer all'84,0% dopo l'implementazione non è stato significativo ($p=0,32$). La variazione nel tasso di rilevazione dei tumori (compresi i tumori alla mammella invasivi e i carcinomi duttali in situ) non è stata significativa (4,15 casi per 1000 mammografie di screening prima dell'implementazione e 4,20 casi dopo l'implementazione, $p=0,90$). Le analisi dei dati provenienti da tutte le 43 strutture hanno mostrato che l'uso della lettura aiutata da computer era associata a un'accuratezza complessiva significativamente inferiore rispetto al suo non utilizzo (area sotto la curva ROC, 0,871 contro 0,919; $p=0,005$).
L'uso della lettura assistita dal computer è associata a una riduzione dell'accuratezza nell'interpretazione delle mammografie di screening. L'incremento nel tasso delle biopsie con l'uso della lettura assistita da computer non è associato in modo chiaro a un aumento nella rilevazione di tumori alla mammella invasivi.

Fonte: New Eng J Med 2007;356:1399-1409

Commento di Marco Grassi

Negli USA lo screening mammografico è una delle pratiche diagnostiche a più alto rischio di causa legale per malpratica in radiologia. Ciò non deve sorprendere, visto che in una percentuale variabile dal 25 al 50% dei casi di neoplasia mammaria, è possibile identificare retrospettivamente la neoplasia in una mammografia precedente. La lettura di mammografie di screening è pertanto una attivita molto stressante e associata ad alto tasso di burn-out fra i radiologi statunitensi, tanto che i giovani radiologi scansano accuratamente questa sub-specialità, preferendo attività piu remunerative e meno pericolose.
La CAD è una tecnologia computerizzata che analizza le immagini digitalizzate di una mammografia e identifica aree sospette da sottoporre alla valutazione del radiologo che stila il referto finale. In base a studi di valutazione di efficacia, nel 1998 la nuova tecnologia è stata approvata dall'FDA per l'uso clinico. Per le considerazioni fatte in precedenza la CAD è stata salutata molto favorevolmente dai radiologi e nel giro dei primi 3 anni di diffusione, il 10% delle strutture diagnostiche statunitensi si è dotata di questo ausilio. Inoltre, l'adozione sempre più estesa della mammografia digitale ha contribuito notevolmente anche alla adozione di questa tecnologia.
I dati preliminari di accuratezza diagnostica avevano mostrato un aumento del 10-15% nel numero di diagnosi di neoplasia mammaria rispetto alla lettura di un singolo radiologo. Questi risultati sono tuttavia sovrapponibili a quelli ottenibili con la doppia lettura di due radiologi, pratica diffusa nel Regno Unito e in molti paesi europei. Pratica tuttavia difficilmente implementabile negli USA per i motivi già accennati.
Questo studio sarà verosimilmente accolto con delusione dai radiologi che si occupano di mammografia perchè mostra non solo che l'accuratezza diagnostica complessiva della lettura umana è ancora superiore a quella della macchina ma che l'uso della lettura guidata, oltre a non aumentare significativamente il tasso di diagnosi di neoplasia, può perfino essere pericolosa per l'aumentato numero di falsi positivi che si traduce in più biopsie e procedimenti chirurgici. Tutto ciò ha poi dei costi umani oltre che economici, tutt'altro che indifferenti.
Durante i 4 anni di implementazione della tecnica, infatti, le sette strutture che hanno adottato la lettura guidata delle mammografie hanno visto accrescere la sensibilità dell'atto diagnostico fino all'84% rispetto all'80,4% delle strutture che hanno continuato la normale lettura delle mammografie. Ciò significa che su 100 casi di neoplasia mammaria che si presentavano, la lettura guidata era in grado di rilevarne 3,6 in più rispetto al metodo tradizionale. Risultato tutt'altro che disprezzabile anche in considerazione del fatto che una strategia di screening, che si propone come compito fondamentale di scoprire casi asintomatici, più casi riesce a rilevare più è efficace. Cioè più alta è la sensibilità del test meno casi sfuggono. Per inciso, e rifacendosi alla situazione americana, significa verosimilmente anche meno cause medico-legali. Ogni test diagnostico ha tuttavia un altro lato da considerare oltre alla capacità di rilevare le vere neoplasie (sensibilità). Deve essere anche in grado di diagnosticare correttamente i seni sani (specificità). E' infatti su

questo punto che casca la nuova tecnologia diagnostica: la lettura guidata delle mammografia perde 3 punti secchi di specificità rispetto alla lettura convenzionale. Significa che su 100 mammografie eseguite su seni normali, 3 mammografie in più vengono decretate come anomale quando la lettura è assistita dal computer.
Perchè 3,6 punti di guadagno di sensibilità (che è quello che ci dovrebbe interessare quando è importante individuare il maggior numero possibile di neoplasie) sono ininfluenti e la perdita di 3 punti di specificità sono in grado di diminuire significativamente l'accuratezza diagnostica complessiva di questa nuova tecnica?
Per capire l'arcano occorre definire il concetto di accuratezza diagnostica complessiva (overall accuracy). Si tratta di una frazione che prende in considerazione al numeratore i casi correttamente diagnosticati (veri positivi, cioè in questo caso i tumori, più i veri negativi, vale a dire il numero di seni normali correttamente valutati come tali) e al denominatore la somma di veri positivi, veri negativi, falsi positivi (seni normali valutati come affetti da neoplasia) e falsi negativi (presenza di neoplasia non rilevata).
La prevalenza della malattia, secondo i dati riferiti dal NEJM, dovrebbe essere attorno allo 0,517%. Perciò se si esegue una mammografia con lettura convenzionale su 100.000 donne si rilevano 415 tumori su 517 con 102 falsi negativi e 9.750 falsi positivi.
Se invece la mammografia è letta dal computer si rilevano 434 tumori su 517 con una riduzione dei falsi negativi che saranno 83 ma contemporaneamente, siccome cala la specificità di 3 punti, ben 12.734 mammografie verranno dichiarate come anomale senza in realtà esserlo (falsi positivi), con un incremento di ben 3.000 casi.
Insomma tradotto tutto su 100 donne il computer permette di scoprire 0,019 tumori in più a scapito di 3 falsi positivi. Ne consegue una accuratezza complessiva (overall accuracy) della metodica inferiore rispetto alla lettura esclusivamente umana (87,2 vs 90,1).
Dai conti fatti e riassunti in tabella 1 risulta chiaramente che nella lettura guidata da computer il numero di veri negativi prodotti in meno, poichè vengono conteggiati al numeratore della frazione e non sono
controbilanciati da analogo aumento di veri positivi, sono in grado di diminuire fortemente il valore della frazione stessa, cioè della accuratezza complessiva.

Un modo metodologicamente più elegante e di immediato impatto visivo è la costruzione della curva ROC (figura 1) delle due metodiche diagnostiche.
Per convenzione il grafico della curva ROC mostra il tasso di falsi positivi (o anche 1-specificità) sull'asse delle ascisse (asse X) e la sensibilità sull'asse delle ordinate (asse Y) per ogni valore di risultato del test.
La curva ROC indica la qualità globale di un test. La curva ROC di un test che abbia una sensibilità e specificità eccellenti ha una forma che tende a far avvicinare la sua convessità verso l'angolo superiore sinistro del grafico. Al contrario la curva ROC di un test di media o debole validità tende ad avvicinarsi alla diagonale che va dall'angolo inferiore sinistro all'angolo superiore destro. La curva che ha l'area sotto la curva (cosiddetta AUROC) più ampia identifica quindi il test più accurato.

Referenze

1) Fenton J. et al. Influence of Computer-Aided Detection on Performance of Screening Mammography. New Eng J Med 2007;356:1399-1409