

Presentazione del repertorio sull'efficienza diagnostica degli esami di laboratorio

R.M. Dorizzi^a, D. Giavarina^b, M. Venturini^c

^aLaboratorio Analisi Chimico-Cliniche ed Ematologiche, Azienda Ospedaliera, Verona

^bLaboratorio di Chimica Clinica ed Ematologia, Ospedale "San Bortolo", Vicenza

^cGruppo di Studio Evidence Based Laboratory Medicine della SIMeL

Introduzione

L'Evidence Based Medicine (EBM) è stata definita come "l'uso coscienzioso, esplicito e giudizioso della migliore evidenza corrente nel prendere decisioni relativamente alla cura del paziente". Questa definizione si può applicare anche al laboratorio in quanto parte integrante nel processo decisionale clinico (1). Più recentemente lo stesso gruppo ha precisato che l'EBM è l'integrazione della migliore evidenza clinica, dell'esperienza clinica e delle preferenze del paziente (2). *Migliore evidenza clinica:* è ricavata dalla ricerca di base, ma soprattutto dalla ricerca clinica, è concentrata sul paziente e si basa sull'accuratezza e sulla precisione degli esami diagnostici, sulla potenza degli indicatori prognostici e sull'efficacia e sulla sicurezza degli interventi terapeutici e riabilitativi. La ricerca clinica dimostra continuamente che nuovi esami diagnostici ed interventi terapeutici superano in validità quelli in uso e devono sostituirli (non affiancarsi ad essi come spesso è accaduto ed accade).

Esperienza clinica: rappresenta l'essenza dell'attività del clinico. È infatti l'abilità di usare le conoscenze e le precedenti esperienze per identificare il particolare stato di salute (o di non salute) di quel paziente, ed il beneficio (o il danno) che egli potrà avere da un intervento terapeutico.

Preferenze del paziente: l'EBM non promuove l'aspirazione della tecnologia in medicina; anzi ha accentuato l'importanza di preferenze, preoccupazioni ed attese del paziente. Nel corso dell'incontro clinico il paziente deve essere adeguatamente "informato" e le sue preferenze debbono essere integrate nelle decisioni cliniche.

Il movimento dell'EBM, fondato nel 1982 da Gordon Guyatt della McMaster University origina in qualche modo da Pierre Louis che, nella Parigi post-rivoluzionaria, respinse le affermazioni delle "autorità scientifiche" del tempo e promosse la ricerca della verità attraverso l'osservazione sistematica dei pazienti. L'EBM ha trovato rapidamente molti proseliti e praticanti; è stata oggetto di un singolo articolo nel 1992 e 1000 nel 1998. Sono stati dedicati al tema un con-

vegno della Società Italiana di Medicina di Laboratorio, un terzo delle relazioni del Congresso Europeo di Chimica Clinica nel 2001 e trenta tra relazioni e workshop al Congresso della American Association for Clinical Chemistry nel 2001.

Vi sono numerose spiegazioni di questo rapidissimo aumento dell'interesse:

- Ogni laboratorista, come ogni medico, deve quotidianamente ricorrere ad una fonte valida di informazioni relative a diagnosi (soprattutto), prognosi, terapia e prevenzione.
- Le tradizionali fonti di informazione risultano spesso inadeguate perchè obsolete (trattati), frequentemente non corrette (i cosiddetti esperti), inefficaci (non strutturate in forma prontamente utilizzabili) o troppo voluminose, disperse e di valore variabile (giornali).
- Il medico si accorge di una divaricazione sempre maggiore tra la sua abilità diagnostica e clinica che aumenta con gli anni ed il suo aggiornamento che diminuisce con il tempo.
- È impossibile dedicare più di qualche minuto ad ogni paziente e più di mezz'ora alla settimana per lo studio e la lettura di giornali scientifici e professionali.

Fino ad oggi l'EBM ha avuto un impatto modesto nella pratica quotidiana della medicina di laboratorio ed è stato dimostrato che il rispetto di criteri di solida evidenza per l'uso degli esami diagnostici è scarsa (3), come è dimostrato anche dalle grandi variazioni nelle strategie diagnostiche usate dai diversi laboratori per la stessa patologia e dall'ampia letteratura dedicata all'uso non appropriato degli esami di laboratorio.

Il laboratorio rischia così di non essere percepito come erogatore di prestazioni importanti per l'*outcome* del paziente; questo rischio, anche se basato su una scarsa conoscenza da parte del clinico, ma in alcuni casi anche del laboratorista, del reale significato degli esami di laboratorio potrebbe rivelarsi esiziale per il futuro della disciplina (4).

È stato ampiamente documentato che l'aumento degli esami diagnostici induce un aumento delle procedure terapeutiche anche se la loro utilità non è stata dimostrata (5).

Wenneberg et al., hanno studiato la relazione tra numero di coronarografie eseguite e numero di procedure di rivascularizzazione, trovando una correlazione indipendentemente dalla prevalenza della patologia e dalla disponibilità di servizi chirurgici nella regione (6). In un altro studio è stato dimostrato un aumento del 50-300% in 7 anni delle indagini (caterismo cardiaco, imaging della colonna vertebrale, mammografia, biopsia prostatica) a cui corrispondeva un aumento delle procedure terapeutiche come rivascularizzazione, interventi chirurgici alla colonna, gastrostomia percutanea, biopsia della mammella, prostatectomia (7). Recentemente in Canada è stata valutata l'appropriatezza dell'impiego degli esami di laboratorio al momento del ricovero ospedaliero, sia di elezione che in urgenza, in terapia intensiva, nella pratica ambulatoriale (8). La percentuale degli esami richiesti non appropriatamente andava dal 5 al 95% e gli esami richiesti più frequentemente in modo non appropriato erano esami comuni come tempo di protrombina, calcemia, VDRL nel liquido cerebrospinale, determinazione della concentrazione dei farmaci antiepilettici. Lo stesso gruppo ha studiato come interventi diversi (produzione di linee guida, modifica dei moduli di richiesta degli esami, cambiamento delle modalità di rimborso delle prestazioni) hanno modificato tra il 1991 ed il 1997 la richiesta di esami come velocità di eritrosedimentazione, esame microscopico delle urine, esami di funzionalità renale, esami di funzionalità tiroidea (9). Le richieste di VES ed urea sono diminuite rispettivamente del 58% e del 57%, la richiesta della determinazione del ferro è diminuita dell'80% e quella di ferritina del 34%, la richiesta di analisi delle urine non accompagnate da esame microscopico è aumentato del 1700%, quella di analisi microscopica delle urine è diminuito del 14%, la determinazione della tiroxina totale è diminuita del 96% e quella del TSH del 12%. George Lundberg ha commentato questi articoli sostenendo che l'uso degli esami può essere modificato se il direttore del laboratorio ha ben chiaro come intervenire, se ha il coraggio di farlo ed è appoggiato dalla istituzione in cui opera (10). Per il clinico invece vale ancora l'esortazione di A. Cochrane; deve avere chiaro, prima di richiedere una qualunque indagine come procederà nel caso che il risultato sia "normale" e come procederà nel caso sia "patologico" (11).

La medicina di laboratorio, usata male, non incide sulla salute dei cittadini e risulta addirittura controproducente; l'Evidence Based Laboratory Medicine (EBLM) consente: 1) il miglioramento della qualità degli esami di laboratorio; 2) il "raffreddamento" dell'aumento del numero delle procedure diagnostiche di efficacia non sufficientemente dimostrata; 3) la documentazione, come richiesto dai Servizi Sanitari Nazionali e dalle Società di Assicurazione, che il costo degli esami è proporzionato alla loro utilità.

L'EBLM presenta problematiche particolari come la frequente mancanza di un "gold standard" assoluto con cui confrontare i nuovi esami e la grande diver-

sità tra i valori numerici prodotti da metodi diversi per la misurazione di uno stesso analita (basti pensare a troponina I e PSA).

Reid et al. (3) hanno riportato che solo il 24% di 112 articoli relativi ad esami radiologici ed immunometrici pubblicati tra il 1978 ed il 1993 in giornali autorevoli come Lancet, British Medical Journal, New England Journal of Medicine e Journal of American Medical Association, rispettava 4 standard dei 7 considerati importanti ed addirittura solo il 6% ne rispettava 6. Il dato positivo è che la percentuale degli standard soddisfatti negli articoli pubblicati nel 1993 è risultata maggiore rispetto a quelli pubblicati nel 1978. Il capitolo dedicato a diagnosi e screening del classico volume di Sackett et al (2) ha un approccio simile ed esordisce con l'affermazione che l'EBLM deve aiutare a rispondere a 3 domande:

- le prove di accuratezza di un esame diagnostico sono valide?
- questo esame è in grado di distinguere in modo accurato i pazienti con una determinata malattia da quelli che non la presentano?
- questo esame può essere usato in un mio paziente particolare?

A seconda dei casi può essere preferibile rispondere prima alla seconda domanda (se l'impatto di un esame è modesto, a chi interessa se è valido?) o alla prima (se un esame non è accurato, cosa importa se è importante?). È invece ineludibile l'assunto che l'applicabilità di un esame ad uno specifico paziente non può essere valutata prima che siano stati considerati i primi due punti.

Secondo Sackett un esame diagnostico è valido se:

- È stato confrontato in cieco con il metodo di riferimento (*gold standard*). Si deve comunque essere cauti nel considerare il metodo di riferimento e tenere presente che perfino nella interpretazione di biopsie di mammella, cute e fegato la concordanza tra esperti non raggiunge il 50% (12). La normalità di un risultato secondo l'EBLM va inteso in senso diagnostico: non indica cioè se un soggetto è sano o malato, ma indica quell'ambito di risultati oltre il quale la patologia in esame diventa altamente probabile. La definizione gaussiana (media \pm 2 Deviazioni Standard) o in percentili (che possono andare dal percentile 2.5 al percentile 97.5 o dal percentile 5 al percentile 95) della normalità si concentra esclusivamente sul risultato dell'esame diagnostico. Queste definizioni non solo implicano che le anomalie si verificano con la stessa frequenza ma comportano che se eseguiamo un numero elevato di esami in un paziente aumentiamo progressivamente la probabilità di trovare un risultato anomalo. Il paziente viene quindi classificato "malato" e viene avviato ad una ulteriore serie di accertamenti, visite, indagini (non appropriati e poco efficaci dal punto di vista economico). Va sempre ricordato che il 64% dei soggetti in cui sono eseguiti 20 esami di laboratorio indipendenti ne presenta almeno uno al di fuori dell'intervallo di riferimento calcolato nel modo tradizionale [$100 \times (1 - 0.9520) = 1 - 0.358 = 100 \times 0.642 = 64.2\%$] (13).

- L'esame di riferimento è stato applicato a tutti i pazienti (sia in quelli in cui l'esame in valutazione è risultato positivo che in quelli in cui è risultato negativo).
- L'esame è stato valutato in uno spettro appropriato di pazienti.
- L'esame è stato validato in un secondo (ed indipendente) gruppo di pazienti.

Sensibilità, specificità e quoziente di probabilità

Il risultato dell'esame può essere *positivo* o *negativo*. È difficile stabilire la correttezza, la *verità*, del risultato anche se si ricorre, quando è disponibile, ad un metodo di riferimento. La tabella a doppia entrata in Figura 1 indica i rapporti tra risultati dell'esame e presenza o assenza della malattia e consente di ricavare: *Sensibilità* (percentuale dei veri positivi o percentuale di risultati positivi nella malattia) cioè la percentuale dei pazienti affetti dalla malattia in cui l'esame è positivo che si può calcolare dal rapporto tra *a* (veri positivi) e *a+c* (veri positivi + falsi negativi).

Specificità (percentuale dei veri negativi o percentuale di risultati negativi nei sani) cioè la percentuale dei soggetti non affetti dalla malattia in cui l'esame è negativo che si può calcolare dal rapporto tra *d* (veri negativi) e *b+d* (veri negativi + falsi positivi).

La classificazione dei risultati dell'esame come veri e falsi positivi e negativi dipende dall'esistenza di un "gold standard" in grado di classificare correttamente il soggetto come affetto o meno dalla malattia. Abbiamo già segnalato come anche i più classici esempi di *gold standard* come l'esame biptico di un nodulo polmonare che è stato classificato come benigno o maligno dalla Tomografia Assiale Computerizzata è ben lontano da assicurare la certezza ed è piagato da una notevole variabilità interosservatore. Altre situazioni complesse sono, per esempio:

- l'impiego di esami diagnostici in una fase molto precoce della malattia;
- la valutazione di un esame diagnostico innovativo con prestazioni migliori a quelle dell'esame considerato fino a quel momento il *gold standard*. Se il presunto *gold standard* non classifica correttamente i pazienti, *sensibilità* e *specificità* del nuovo esame possono risultare sottostimate.

Conoscere sensibilità e specificità di un esame aiuta il clinico a stabilire se un esame è più utile per confermare (*ruling in*) o escludere (*ruling out*) una malattia. Per esempio, un esame molto sensibile mancherà molto raramente di diagnosticare pazienti con una malattia; un risultato negativo di un esame con elevata sensibilità è quindi utile per escludere una diagnosi. Di converso, un esame molto specifico classificherà molto raramente soggetti senza malattia come malati; un risultato positivo di un esame con elevata specificità è quindi utile per confermare una diagnosi (anche se in questa situazione un risultato negativo non esclude necessariamente la malattia). Uno degli aspetti

più rilevanti della EBLM è l'approccio pratico che introduce nell'impiego dei risultati di laboratorio nella gestione del singolo paziente. Due esempi di questo approccio sono gli acronimi SnNout e SpPin (14).

SnNout deriva dall'aforisma *When a test has a very high Sensitivity, a Negative result rules out the diagnosis* (Quando un esame ha una sensibilità molto elevata un risultato negativo esclude virtualmente la diagnosi); **SpPin** deriva invece dall'aforisma *When a test has a very high Specificity, a Positive result rules in the diagnosis* (Quando un esame ha una specificità molto elevata un risultato positivo conferma virtualmente la diagnosi).

Frequentemente, i risultati di un esame sono refertati come variabili continue o ordinali piuttosto che positivi o negativi. In questo caso variando il valore di cut-off tra risultati "normali" e "patologici" si avranno valori diversi di sensibilità e specificità. In generale all'aumento della sensibilità diminuisce la specificità e viceversa

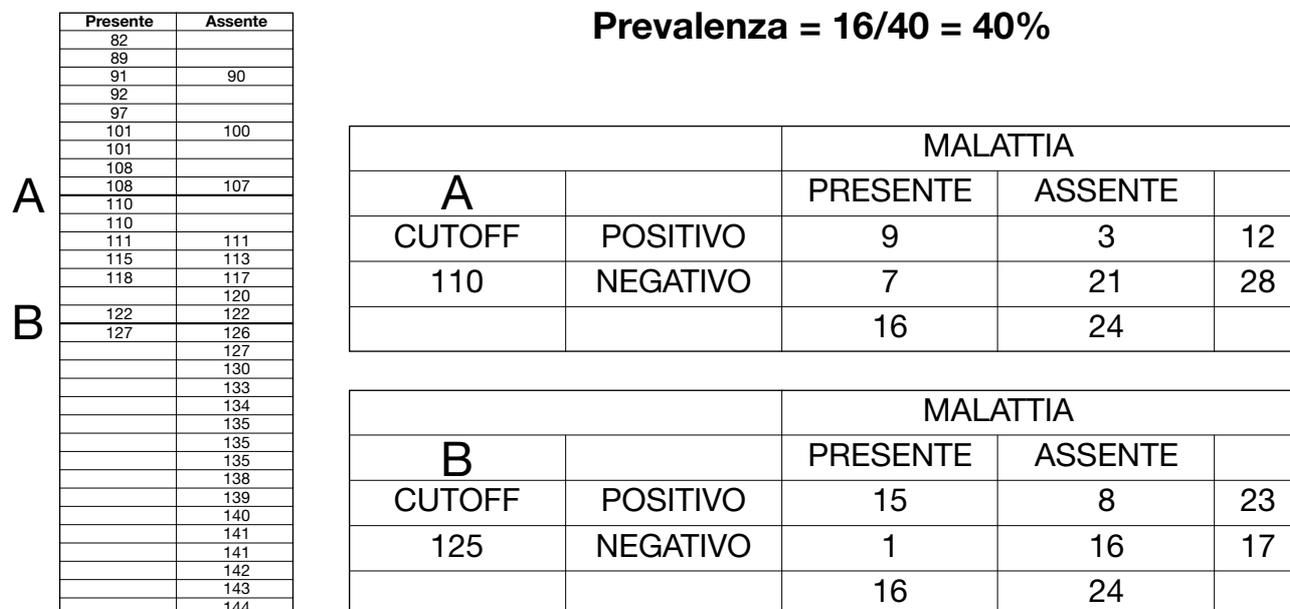
Nella Figura 2 è riportato l'esempio della determinazione dell'emoglobina correlata alla presenza di depositi di ferro nella biopsia del midollo osseo (il *gold standard*). Se il valore di cut-off per l'anemia sideropenica è fissato ad un livello più alto (come in B) più pazienti con anemia sideropenica sono classificati come malati e la sensibilità risulta relativamente elevata ($15/16 = 0.94$); aumenta anche il numero di soggetti sani che sono stati classificati erroneamente come malati come evidenziato dalla specificità relativamente bassa ($16/24 = 0.67$). Di converso se viene adottato un valore di cut-off per l'anemia sideropenica più basso (come in A) più pazienti con anemia sideropenica sono misclassificati come non anemici come indicato da una sensibilità più bassa ($7/16 = 0.56$); minore è invece il numero di soggetti sani che sono stati classificati erroneamente come malati come evidenziato dalla specificità aumentata ($21/24 = 0.88$). Questo approccio assiste nel definire il valore di cut-off ottimale per un esame; per esempio quando viene messo a punto un esame di screening si desidera di solito ottenere una sensibilità elevata, in grado cioè di identificare tutti i soggetti con elevata probabilità di malattia e che richiedono esami di conferma. I soggetti invece in cui l'esame è negativo hanno scarsa probabilità di avere la malattia e non richiedono ulteriori indagini.

Figura 1. Tabella a doppia entrata per il calcolo di sensibilità e specificità.

		MALATTIA (VERITÀ)	
		PRESENTE	ASSENTE
RISULTATO	POSITIVO	VERO POSITIVO	FALSO POSITIVO
ESAME	NEGATIVO	FALSO NEGATIVO	VERO NEGATIVO

		MALATTIA (VERITÀ)	
		PRESENTE	ASSENTE
RISULTATO	POSITIVO	a	b
ESAME	NEGATIVO	c	d

Figura 2. Effetto della variazione del cut-off su sensibilità e specificità.



Quando si mette a punto l'esame diagnostico definitivo (da usare, per esempio, per confermare i risultati positivi ottenuti nell'esame di screening) interessa che l'esame abbia una elevata specificità, vale a dire che classifichi raramente i soggetti senza malattia come malati.

L'impiego di sensibilità e specificità come unico strumento per interpretare gli esami diagnostici presenta dei limiti; descrivono, infatti, attributi dell'esame quando la "verità" è nota (per esempio la percentuale dei malati in cui l'esame è positivo o la percentuale dei sani in cui l'esame è negativo).

Tuttavia il medico dispone raramente di questa informazione; richiede l'esame appunto per sapere se un paziente è o meno affetto da una malattia. Il clinico riceve un risultato positivo e vuole conoscere la probabilità che il paziente sia malato ovvero, di converso, riceve una risposta negativa e vuole conoscere la probabilità che il paziente sia sano. In questo contesto sono quindi importanti la predittività di un risultato positivo (PPV) e di un risultato negativo (NPV).

La predittività di un risultato positivo è data dal rapporto tra i pazienti con un risultato positivo, i veri positivi, e tutti i soggetti con risultato positivo; $a/(a+b)$, secondo la terminologia usata nella Figura 1. La predittività di un risultato negativo è data dal rapporto tra i soggetti sani con un risultato negativo, i veri negativi, e tutti i soggetti con risultato negativo; $d/(c+d)$, secondo la terminologia usata nella Figura 1.

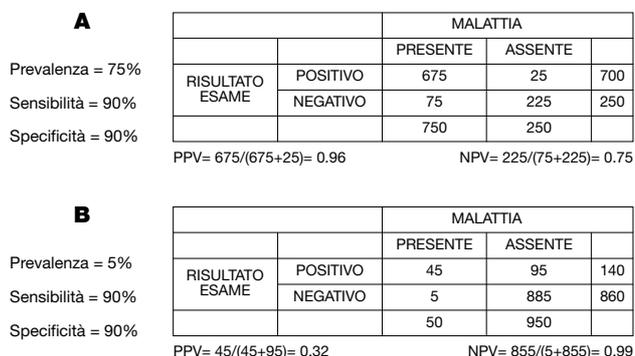
I valori di predittività sono calcolati orizzontalmente lungo le righe della tabella in Figura 1 e variano con il variare della prevalenza della malattia (la percentuale dei soggetti in cui viene eseguito l'esame che soffrono della malattia). Sensibilità e specificità, che descrivono caratteristiche dell'esame, rimangono

invece più stabili indipendentemente dalla prevalenza della malattia. Questo spiega perché un esame diagnostico è molto informativo per lo specialista che lo richiede in una popolazione di soggetti ad elevata prevalenza di malattia ma risulta molto meno informativo per il medico di medicina generale che lo richiede in una popolazione di soggetti a prevalenza molto più bassa (Fig. 3).

In questo contesto, infatti, i risultati falsi positivi superano di molto i veri positivi anche quando l'esame ha una elevata specificità (quindi è più adatto a confermare la presenza di malattia).

La EBLM si serve dei classici concetti di sensibilità e specificità soprattutto come mezzo per ottenere il Quoziente di Probabilità (Likelihood Ratio, LR), che consente di calcolare come la probabilità di una diagnosi sia modificata dal risultato di un esame; ovvero come si ricava la probabilità di una diagnosi dopo che è stato eseguito un esame (probabilità post-test) a partire dalla probabilità di una diagnosi prima della esecuzione (probabilità pre-test o prevalenza).

Figura 3. Effetto della prevalenza sulla predittività di un esame.



Questo è importante dal punto di vista pratico sia perché questi strumenti consentono di valutare se un esame diagnostico può essere utile ovvero quale esame è da preferire quando ve ne siano numerosi disponibili. Nella Figura 4 è sintetizzato il classico esempio della concentrazione di ferritina nel siero nella diagnosi di anemia sideropenica (15). Se la sensibilità è del 90% e la specificità dell'85% equivale a dire che la concentrazione della ferritina è inferiore a 30 µg/L 6 volte più frequentemente (90%/15%) nei pazienti con anemia sideropenica rispetto ai soggetti che non presentano questa patologia; questo concetto è definito Quoziente di probabilità Positivo (LR+) e può essere calcolato dividendo la sensibilità per il reciproco della specificità (2).

Figura 4. Come calcolare quoziente di probabilità positivo e negativo e probabilità pre-odds e post-odds (15).

CONCENTRAZ. FERRITINA	ANEMIA SIDEROOPENICA		
	PRESENTE	ASSENTE	
< 30 mg/L	731 ^a	270 ^b	1001 ^{a+b}
≥ 30 mg/L	78 ^c	1500 ^d	1578 ^{c+d}
	809 ^{a+c}	1770 ^{b+d}	2579 ^{a+b+c+d}

$a/(a+c) = \text{Sen} = 731/809 = 90\%$
 $d/(b+d) = \text{Spe} = 1500/1770 = 85\%$
 $\text{LR}+ = \text{Sens}/(1-\text{Spec}) = 90\%/15\% = 6$
 $\text{RR} = (1-\text{Sens})/\text{Spec} = 10\%/85\% = 0.12$
 $\text{PPV} = a/(a+b) = 731/1001 = 73\%$
 $\text{NPV} = d/(c+d) = 1500/1578 = 95\%$
 $\text{PREVALENZA} = a+c/a+b+c+d = 809/2579 = 31\%$
 $\text{PREODDS} = \text{PRE} / (1-\text{PRE}) = 31\%/69\% = 0.45$
 $\text{POSTODDS} = \text{PREODDS} \times \text{LR} = 0.45 \times 6 = 2.7$
 $\text{PROB. POSTODDS} = 2.7/(2.7+1) = 73\%$

Si può esprimere anche come il rapporto tra risultati Veri Positivi (sensibilità) e Falsi Positivi, i due assi del grafico delle curve ROC (Receiving Operating Characteristics).

Il clinico è interessato però alla probabilità che un paziente con una concentrazione, per esempio, di 26 µg/L di ferritina sia affetto da anemia sideropenica. La prevalenza della malattia consente di calcolare gli Odds pretest che sono dati dal rapporto tra prevalenza e reciproco della prevalenza (nel caso dell'esempio $31/69\% = 0.45$). Moltiplicando gli Odds pre-test per il Quoziente di Probabilità (0.45 x 6) si ottengono gli Odds post-test (2.7); gli Odds (la probabilità) post-test favoriscono l'anemia sideropenica nel rapporto di 2.7:1. Gli Odds possono essere facilmente convertiti in percentuale come segue $2.7/(2.7+1) = 2.7/3.7 = 73\%$.

Il Quoziente di Probabilità consente anche di tenere conto di come la concentrazione più o meno alta del risultato influenza la probabilità di un risultato. Mentre impiegando i concetti di sensibilità e specificità (rispettivamente 90% e 85% nell'esempio considerato) il risultato ha due soli livelli (positivo e negativo), l'LR consente di attribuire un peso diverso al risultato numerico ottenuto. Quando la prevalenza della malattia è dell'ordine di grandezza dell'esempio sopra riportato un esame con un LR positivo superiore a 10 raggiunge una probabilità post-test superiore all'80% e quindi utile

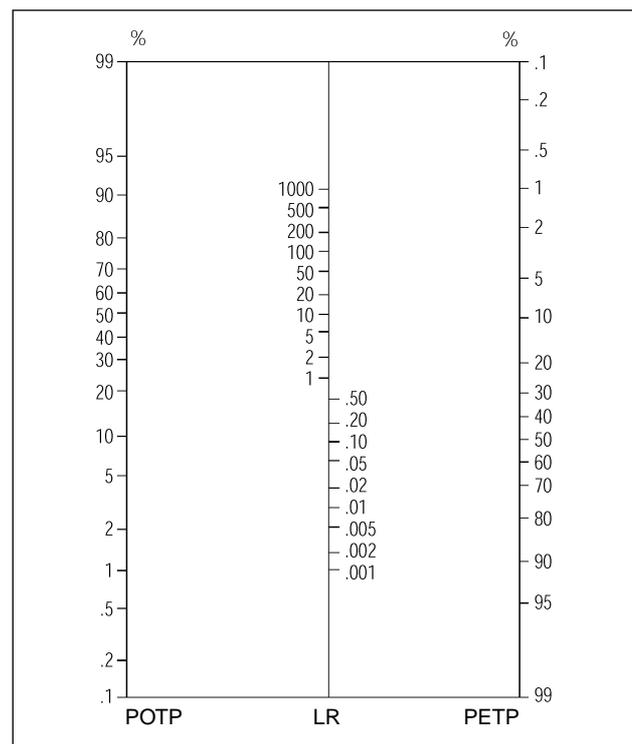
dal punto di vista diagnostico. LR negativi inferiori a 0.1 quando la probabilità pre-test è quella dell'esempio, portano ad una probabilità post-test molto bassa, inferiore al 5% (Odds post-test = $0.45 \times 0.1 = 0.045$; probabilità post-test $0.045/1.045 = 4.5\%$). Jaeschke et al. (16) hanno classificato in modo empirico, ma utile nella pratica, l'impatto dell'LR nel modo seguente:

- LR+ > 10 ed LR- < 0.1 modificano in modo spesso conclusivo la probabilità di malattia
- LR+ > 5 ed < 10 o LR- > 0.1 e < 0.2 modificano in modo discreto la probabilità di malattia
- LR+ > 2 ed < 5 o LR- > 0.2 e < 0.5 modificano modestamente la probabilità di malattia
- LR+ > 1 ed < 2 o LR- > 0.5 e < 1 modificano scarsamente la probabilità di malattia e solo di rado sono importanti.

L'LR- può essere facilmente calcolato con il rapporto tra il reciproco della sensibilità e la specificità ovvero, analogamente all'LR +, come il rapporto tra risultati falsi negativi e veri negativi (specificità) (nell'esempio relativo alla ferritina, $10/85 = 0.12$).

Uno dei vantaggi degli LR è che possono essere impiegati anche se sono stati originariamente calcolate in una popolazione diversa. Questo può essere ottenuto attraverso il calcolo degli odds come visto precedentemente o, più semplicemente, con il nomogramma di Fagan, uno degli strumenti più utili tra quelli raccomandati dall'EBLM (17). Il nomogramma si usa ancorando una linea retta in corrispondenza della probabilità pre-test indicata sull'asse di destra e ruotandola finché raggiunge il valore l'LR relativo all'esame che si sta considerando sull'asse al centro (Fig. 5).

Figura 5. Nomogramma di Fagan (PETP = Probabilità pre-test; POTP = Probabilità post-test).



Proseguendo la retta si incrocia sull'asse di sinistra la probabilità post-test (18). LR e nomogramma di Fagan presentano numerosi vantaggi; per esempio:

- sono più comprensibili e più facili da usare rispetto a sensibilità e specificità;
- possono essere calcolati a livelli diversi di risultato;
- possono essere usati in modo sequenziale (la probabilità post-test che si ottiene da un esame può diventare la probabilità pre-test per il successivo);
- consentono di embricare medicina di laboratorio e clinica in modo molto efficace;
- forniscono gli strumenti per introdurre forme di audit clinici nella pratica valutando l'efficacia diagnostica degli esami richiesti dal clinico ed eseguiti dal laboratorio.

Il clinico usa infatti metodi impliciti per interpretare i risultati degli esami. Se un esame è chiaramente positivo e negativo, può usarlo per orientare diagnosi e terapia; se il risultato è indeterminato può eseguire ulteriori esami per chiarire il quadro. La EBLM può consentire da una parte di dare maggiore evidenza a queste decisioni e dall'altra di aiutare il clinico meno esperto ad interpretare i risultati di laboratorio in modo efficiente.

Il laboratorio di qualunque tipologia e dimensione deve attrezzarsi per partecipare a questi nuovi processi decisionali. È auspicabile che l'EBLM entri nella attività quotidiana e che alla conoscenza ed all'uso oramai consolidati di concetti come sensibilità e specificità si aggiungano concetti come probabilità pre- e post-test e quoziente di probabilità positivo e negativo. Introdurre gli strumenti dell'EBLM nell'attività quotidiana del laboratorio consente infatti di valutare gli esami diagnostici (sia tra quelli già in uso che quelli in via di introduzione) dotati della maggiore efficienza diagnostica ed affidabilità, meglio tollerati dal paziente e più vantaggiosi dal punto di vista economico. D'altra parte non è pensabile che oggi il laboratorio produca o impieghi strumenti complessi come, per esempio, i Prevalence-Value-Accuracy-Plots, grafici che considerano nella valutazione degli esami accuratezza, effetto della prevalenza e costo della misclassificazione (19).

In uno studio recente sono state poste a medici di 6 specialità diverse domande del tipo "Quando richiedete ed interpretate esami di laboratorio considerate i valori di sensibilità e specificità?": i metodi bayesiani erano usati dal 3% dei medici e curve ROC ed LR dall'1% (da nessuno dei ginecologi e dei medici di famiglia che hanno partecipato allo studio). Gli Autori concludono che è necessario presentare in modo sempre più chiaro ed intuitivo i risultati degli esami di laboratorio (20).

Il repertorio

L'EBLM ha anche lo scopo di diffondere gli approcci e gli strumenti di lavoro descritti precedentemente sia tra i laboratoristi che tra i clinici, in modo che le informazioni che il laboratorio fornisce siano comprese e, soprattutto, impiegate correttamente.

Questa considerazione ha originato la preparazione di un repertorio degli esami di laboratorio che fornisce le caratteristiche (sensibilità, specificità, quoziente di probabilità positivo e negativo, valore di cut-off nei casi in cui è disponibile, e riferimento bibliografico) di oltre 1000 esami diagnostici diversi.

Il repertorio ha lo scopo di servire da "primer" per l'EBLM per laboratoristi, clinici, medici di famiglia, studenti e tecnici di laboratorio che possono trovare una sorta di "innesco" (primer) per il loro primo approccio alla metodologia. Sarà sufficiente stimare anche approssimativamente la prevalenza di una malattia nella specifica popolazione di interesse e, ricavato il Quoziente di Probabilità dal Repertorio e usando il nomogramma di Fagan (Fig.5) si potrà avere la probabilità post-test.

Il repertorio non ha la presunzione di comprendere tutti gli esami ma nasce dagli interessi principali di quanti hanno lavorato al progetto che sono sostanzialmente i componenti del gruppo di studio sull'Evidence Based Laboratory Medicine della Società Italiana di Medicina di Laboratorio. Si tratta, comunque, del lavoro più completo di questo tipo a nostra conoscenza compiuto fino ad oggi. Una delle finalità del neonato gruppo di studio è quello di compiere una attenta ed accurata manutenzione in modo da renderlo in breve un documento di uso quotidiano per laboratoristi e clinici.

Peabody raccomandava 80 anni fa a Boston *Ogni ospedale ha il dovere di impedire che un medico riceva il suo diploma a meno che non abbia dimostrato di conoscere come usare i risultati degli esami nella cura dei suoi pazienti. La buona medicina non consiste nell'indiscriminata esecuzione degli esami di laboratorio, ma piuttosto in una chiara comprensione delle probabilità di una patologia e di quale esame può essere utile.* Il repertorio ha lo scopo di dare qualche elemento per favorire il raggiungimento di questo risultato.

Bibliografia

1. Sackett DL, Rosenberg WMC, Gray JAM, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. Br Med J 1996;312: 71-2.
2. Sackett DL, Straus SE, Richardson WS, Rosenberg W, Haynes RB. Evidence based medicine. 2nd ed. Edinburgh: Churchill Livingstone, 2000.
3. Reid MC, Lachs MS, Feinstein AR. Use of methodological standards in diagnostic test research: getting better but still not good. JAMA 1995; 274: 645-51.
4. Price C. Evidence-based laboratory medicine: supporting decision-making. Clin Chem 2000; 46: 1151-60.
5. Muir Gray JA, ed. Evidence-based healthcare. How to make health policy and management decisions. London: Churchill Livingstone, 1997.
6. Wenneberg DE, Kellet, Dickens JD, Malenka DJ, Keilson LM, Keller RB. The association between local diagnostic testing and invasive cardiac procedures. JAMA 1996; 275: 1161-4.
7. Verrilli D, Welch G. The impact of diagnostic testing on therapeutic interventions. JAMA 1996; 275: 1189-91.

8. Van Walraven C, Naylor D. Do we know what inappropriate laboratory utilization is? *JAMA* 1998; 280: 550-8.
9. Van Walraven C, Goel V, Chan B. Effect of population-based interventions of laboratory utilization. *JAMA* 1998; 280: 2028-33.
10. Lundberg G. Changing physician behavior in ordering diagnostic tests. *JAMA* 1998; 280: 2036.
11. Lundberg G. The need for an outcomes research agenda for clinical laboratory testing. *JAMA* 1998; 280: 565-6.
12. Anonimo. Melanoma diagnosis. *Bandolier* 1998; 5 (1): 8.
13. Solberg HE. Establishment and use of reference values. In: Burtis CA, Ashwood ER, eds. *Tietz Textbook of clinical chemistry*. Philadelphia: Saunders, 1999.
14. Fleming K. Evidence based pathology. *Evidence-Based Medicine* 1997; 2: 132.
15. Guyatt GH, Oxman AD, Ali M, William A, Mallory W, Patterson C. Laboratory diagnosis of iron-deficiency anemia: an overview. *J Gen Intern Med* 1992; 7: 145-53.
16. Jaeschke R, Guyatt GR, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. *JAMA* 1994; 271: 703-7.
17. Fagan TJ. Nomogram for Bayes's theorem. *N Engl J Med* 1975; 293:257.
18. Sackett DL, Straus S. On some clinically useful measures of the accuracy of diagnostic tests. *Evidence-Based Medicine* 1998; 3: 68-70.
19. Remaley AT, Sampson ML, De Leo JM, Remaley NA, Farsi BD, Zweig MH. Presence-value-accuracy plots: a new method for comparing diagnostic tests based on misclassification costs. *Clin Chem* 1999; 45: 934-41.
20. Reid MC, Lane DA, Feinstein AR. Academic calculations versus clinical judgements: practising physicians' use of quantitative measures of test accuracy. *Am J Med* 1998; 104: 374-80.