

L'Appropriatezza e l'EBM

D. Giavarina

Laboratorio di Chimica Clinica ed Ematologia, Ospedale "San Bortolo", Vicenza

Abstract

Since both appropriateness and Evidence Based Medicine (EBM) depend on the efficacy, they are interconnected. It is more difficult to define Evidence Based Laboratory Medicine (EBLM) than to define EBM regarding the efficacy of care. A first step to define EBLM is the accuracy. A test accurate in detecting a disease, that can be effectively treated, is appropriate.

The accuracy allows to measure efficiency, but efficacy requires further features (e.g. is it possible to assess if the result helps diagnostic and/or treatment decision? are there resources for this diagnostic procedure?...).

There are several tools to determine the accuracy of a test while it is more difficult to measure the effects of a test in the outcome, because the outcome depends also by the treatment prescribed after the test. EBM is helpful in evaluating diagnostic procedures; e.g. in estimating the accuracy of a test, in establishing guidelines and in auditing performances.

Laboratory audit is the review against the benchmark of current best practice i.e. the review of the use of a test. Eventually, EBM can play a role in the identifying the value of a diagnostic service for a better allocation of resources.

Esistono diversi livelli di definizione dell'appropriatezza. Alcune definizioni tendono a descrivere un prerequisito, in senso positivistico (appropriato è tutto ciò che può essere benefico per il paziente, o in cui i benefici possibili superano le eventuali conseguenze negative, di un margine sufficientemente ampio che la procedura valga la pena di essere praticata, compreso il costo)¹, o in senso negativo (è inappropriato quanto può essere eliminato senza particolare nocimento per il paziente)². In questo secondo approccio, ma in parte anche nel primo, è percepibile una preoccupazione per le risorse "finite" del sistema sanitario, e misurare l'appropriatezza può essere quindi utile alle ragioni economiche. La necessità di misurare l'appropriatezza riconosce però an-

che ragioni più ampie: in una medicina sempre più complessa, occorre evitare che alcuni pazienti non ottengano l'intervento di cui hanno bisogno ed altri abbiano l'intervento di cui non necessitano³.

Definizioni più moderne di appropriatezza spostano l'attenzione da caratteristiche intrinseche alla procedura all'effetto finale che la procedura determina. Per gli esami di laboratorio, appropriato può essere il test in cui il risultato fornisce una risposta al quesito clinico e mette in grado di prendere una decisione o intraprendere un'azione⁴.

Esiste di fatto un legame tra questo tipo di definizione e l'Evidence Based Medicine (EBM). infatti L'avvento dell'EBM ha introdotto una metodologia di valutazione degli interventi sanitari incentrata non su caratteristiche intrinseche delle procedure stesse, ma sull'efficacia. Evidence based Medicine è infatti tradotto in italiano non solo come *Medicina basata sull'evidenza*, ma anche *Medicina basata sulle prove di efficacia*.

Quando si considerano i trattamenti medici o chirurgici, le prove di efficacia si ottengono essenzialmente attraverso due strumenti:

1. le *revisioni sistematiche*, una metodologia che cerca di garantire che tutte le informazioni disponibili siano prese in considerazione;
2. le *metanalisi*, processi di sintesi statistica dei risultati ottenuti da diversi studi, pesati per tipologia dello studio (studio controllato randomizzato, studio di coorte, caso-controllo, ecc.) e numerosità del campione.

I test diagnostici rappresentano un problema più complesso, per molti motivi: spesso i test diagnostici sono valutati negli studi sperimentali come argomento secondario o minore; non esiste una letteratura composta di studi randomizzati controllati sulla comparazione di test diversi per la stessa patologia; gli obiettivi dei test diagnostici sono molto più numerosi degli obiettivi terapeutici, comprendendo la diagnosi, il monitoraggio, il follow-up, la stratificazione di rischio, la prevenzione attraverso screening, la prognosi, ...

Questo spiega perchè l'approccio EBM ai test clinici è arrivato dopo; i centri Cochrane, forse la più autorevole organizzazione di revisioni sistematiche, affrontano il problema per la prima volta nel 1996, limitando l'attenzione al solo, ma primario, aspetto diagnostico dei test clinici.

Nel documento del gruppo Cochrane sulle *Revisioni Sistematiche dei Test di Screening e Diagnostici*⁵ sono presentati alcuni punti in cui si possono riconoscere le definizioni dell'appropriatezza: nella valutazione di un test diagnostico occorre maggiore attenzione sull'esito finale; l'esito finale centrato sullo specifico paziente (sopravvivenza, qualità della vita) è il fine ultimo dell'utilizzo di un test (definito come ogni misura diretta ad identificare gli individui che possono beneficiare del trattamento).

Sul versante della metodologia EBM, il gruppo di lavoro del Cochrane riconosce che, pur potendo valutare gli effetti di un test sugli esiti mediante trial randomizzati, non è fattibile rivalutare mediante questo tipo di studio ogni nuovo test. Piuttosto, sembra più semplice e pratico valutare l'accuratezza di test indirizzati alla diagnosi di condizioni per le quali esistono già evidenze dell'efficacia delle cure.

Nella ricerca dell'appropriatezza, quindi, l'accuratezza diventa elemento fondante.

Gli strumenti per la misura dell'accuratezza sono lo studio delle qualità intrinseche dei test (sensibilità e specificità) e delle loro condizioni operative caratteristiche al variare delle probabilità pre-test di condizione patologica o meno; il corpus delle teorie e degli strumenti per la determinazione dell'accuratezza ha raggiunto ormai una robustezza teorica ed applicativa consolidata. La metodologia EBM ha incontrato e incontra invece notevoli difficoltà nella valutazione "evidente" dell'accuratezza dei singoli test, per la difformità sostanziale e formale dei diversi studi. La necessità di incrementare la qualità degli studi di accuratezza ed anche di standardizzare i modelli di studio, al fine di permettere la comparazione tra diverse esperienze, ha indotto la ricerca di uno standard. Dopo alcuni tentativi isolati dal parte di importanti riviste scientifiche e seguendo le orme del CONSORT, standard per il miglioramento dei report dei trial randomizzati, nel 2003 si è concretizzato lo STARD^{6,7}, standard per la produzione di studi di accuratezza diagnostica. Il fatto che numerose riviste abbiano aderito rapidamente al progetto⁸, dimostra l'interesse ed il bisogno della comunità scientifica nell'incrementare l'evidenza anche per i processi diagnostici e non solo per l'efficacia dei trattamenti, e fa sperare in un rapido futuro miglioramento delle evidenze.

La sola accuratezza però non può essere comprensiva di tutta la medicina di laboratorio basata sull'evidenza, e quindi di appropriatezza. Se il centro ed il fine della medicina è il paziente, occorre che non solo il test riconosca accuratamente una condizione,

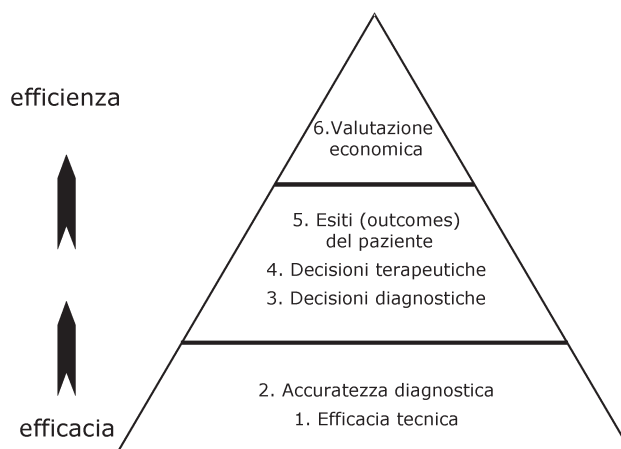


Figura 1. Modello gerarchico per la valutazione dei test diagnostici

ma anche che tale condizione possa essere curata o che il riconoscimento determini un vantaggio per il paziente.

Occorre che esistano anche le condizioni per l'utilizzo del test. Fryback e Thornbury⁹ hanno sintetizzato in uno schema gerarchico il processo di valutazione di un test. (Figura 1). Muovendo dalla base della piramide al vertice si procede dall'efficacia all'efficienza. L'efficacia tecnica e l'accuratezza diagnostica definiscono il livello di efficacia (la probabilità che il test lavori bene in condizioni ideali d'uso); la possibilità conseguente di prendere decisioni cliniche, terapeutiche o di modificare il senso positivo l'outcome del paziente definiscono un livello intermedio, di "effectiveness" (la probabilità che il test "lavori" bene in condizioni ordinarie d'uso); la valutazione economica permette, al livello più alto, il giudizio di efficienza (l'uso ottimale delle limitate risorse sanitarie). Ogni livello considerato è più importante del livello sottostante, ma non può esistere, in termini di evidenza, senza il sottostante.

L'approccio metodologico EBM ai test diagnostici è in gran parte centrato sulla definizione e misura dell'accuratezza, perché è elemento fondamentale, ma non sufficiente. La valutazione delle decisioni diagnostiche e terapeutiche e specialmente degli outcome risulta essere una parte cruciale dello sviluppo e del mantenimento di una base evidente per l'utilizzo dei test di laboratorio. Purtroppo, la letteratura estremamente limitata sugli outcome diagnostici contrasta con quella degli outcome terapeutici, dove entrano a far parte integrante dello sviluppo, dell'approvazione e del marketing di nuovi farmaci.

Pur essendo anche i test diagnostici interventi medici, gli esiti conseguenti sono spesso più lontani nel tempo, a volte meno evidenti, in ogni caso dipendenti sempre dalle azioni conseguenti al risultato del test (Figura 2)¹⁰. Non sempre gli interventi sono congruenti. Un medico può iniziare un trattamento nonostante un risultato negativo o, al contrario, può non fare nulla nonostante un risultato positivo. La

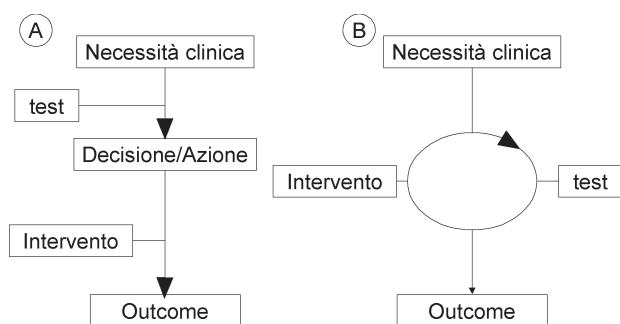


Figura 2. Test e interventi terapeutici intervengono nel determinismo dell'outcome, sia nella situazione critica, dove il test determina la decisione (A), sia nei processi complessi, in cui si instaura un ciclo di interventi modulati e controllati dai risultati dei test, per determinare l'outcome (B). Da Price CP, modificato.

grande variabilità tra i clinici nella cura dei pazienti e la grande mole di informazioni disponibili è una delle cause della produzione di sempre più numerose linee guida.

Le linee guida possono essere definite come raccomandazioni, sviluppate con metodo sistematico definito, che hanno l'obiettivo di assistere le decisioni il medico e/o il paziente per appropriate cure mediche di specifiche condizioni cliniche. L'EBM entra in modo importante in questo ambito, fornendo gli strumenti per le misure delle evidenze che sostengono le raccomandazioni. Sono oggi comuni e correntemente utilizzati sistemi di gradazione delle evidenze e delle raccomandazioni date.

Anche in laboratorio le linee guida sono da tempo strumento ampiamente diffuso. Per esempio, tra il 1997 ed il 2003 ben 542 articoli di *Clinical Chemistry* contengono la parola "linea guida". La recente attenzione all'appropriatezza e all'EBM, essendo legata strettamente ai trattamenti conseguenti ai test per la determinazione degli outcome, sta spostando sempre più le linee guida di laboratorio da problematiche analitiche a indicazioni e raccomandazioni medico/cliniche, indirizzate ad assistere il medico (ed il paziente) a prendere decisioni appropriate per la cura della specifica condizione clinica. Delle ultime 20 linee guida pubblicate da *Clinical Chemistry*, 19 sono linee guida cliniche.

Il problema del conflitto di interessi da parte di chi definisce le linee guida è evidente ed importante. Esso riguarda i clinici e ricercatori, che possono avere interessi privati, i finanziatori e sponsor, i responsabili dei sistemi sanitari. Il legame con metodologie EBM può aiutare ad una valutazione indipendente di quanto è appropriato, ma il limite delle risorse economiche "finite" costituisce argomento sostanziale dell'efficienza.

In una fase di grande fermento diagnostico, sostenuto dalle nuove frontiere della genomica, della proteomica e della biologia molecolare, la sostenibilità economica dei processi diagnostici è sempre più rilevante. Nella piramide di Fryback e Thornbury la valutazione economica è al vertice e non è esterna

alle logiche dell'EBM. Una valutazione economica completa comprende l'identificazione, la misura e la valutazione sia dei benefici clinici di una nuova procedura, sia i costi associati. Esistono oggi sistemi riconosciuti ed accettati per la valutazione economiche (Tabella I).

La connessione tra end-point intermedi (accuratezza, possibilità di decisione clinica/terapeutica) e finali (outcome) risulta necessaria anche in questa valutazione.

Tabella I. Criteri per la valutazione della qualità metodologica di una valutazione economica di test diagnostici¹¹

1. È specificata una ben definita questione economica, che esamini sia i costi che gli outcome di un intervento in un contesto di possibili alternative, con una chiara e definita popolazione bersaglio, un chiaro punto di vista e un chiaro orizzonte temporale?
2. È fornita una descrizione completa delle possibili alternative?
3. È definita l'efficacia dell'intervento?
4. Sono identificati tutti i più importanti e rilevanti costi e i relativi benefici per ogni alternativa?
5. I costi e i benefici sono misurati accuratamente in appropriate unità di misura fisiche?
6. I costi e i benefici sono valutati credibili e appropriati per il tipo di analisi considerato?
7. I costi e i benefici sono stati corretti per il differenziale di tempo?
8. È stata eseguita un'analisi incrementale dei costi e dei benefici degli interventi alternativi?
9. È possibile valutare l'incertezza della stima dei costi mediante l'analisi della sensibilità?
10. La presentazione e la discussione dello studio tratta tutti i temi e tutte le possibili preoccupazioni degli utilizzatori?

Conclusioni

L'EBM permette di definire l'appropriatezza non solo in termini di efficacia, ma anche, e soprattutto, in termini di efficienza. In un contesto in cui i carichi di lavoro aumentano, compaiono costantemente nuove tecnologie, la domanda sembra essere incontrollabile, è continua la richiesta di troppi test non necessari ed è difficile percepire immediatamente la relazione tra test ed outcome, una metodologia come l'EBM può aiutare nel mantenere e nel far progredire la medicina di laboratorio. La definizione corretta dell'appropriatezza collabora a definire e a consolidare il ruolo ed il valore dei servizi diagnostici, anche garantendo il miglior utilizzo delle risorse, per il paziente e per la società.

Non basta però la sola definizione di quanto è appropriato e di quanto non è appropriato per realizzare processi diagnostici basati sull'evidenza, ma sono

essenziali la comunicazione e la condivisione con gli utilizzatori del laboratorio delle evidenze disponibili e delle strategie diagnostiche implementate.

Stringere alleanze con gruppi di clinici riconosciuti, trovare l'accordo sulle azioni, sulla base dell'evidenza e del consenso, implementare i cambiamenti per via amministrativa, aggiungere aspetti educazionali ampi e aggiustare su questa base le azioni di cambiamento, essere aperti alla comunicazione in tutti i modi essa avvenga, sono alcuni dei punti che Lundberg, 21 anni dopo il suo *brain to brain loop*, indica per il miglioramento dell'appropriatezza¹².

I processi di comunicazione e condivisione possono essere condotti con un certo rigore metodologico solo attraverso la continua revisione dei modi in cui i test di laboratorio vengono utilizzati. Molto spesso questo può essere provocato o stimolato dal rapido aumento del numero di richieste di test o dal fallimento nel raggiungere definite misure di outcome. Questa revisione continua prende il nome di *Audit di Laboratorio*. L'obiettivo è quello di ottenere una serie di informazioni legate alla richiesta e all'uso del test, del tipo: qual'è il quesito clinico cui questo test deve rispondere? Quali sono le indicazioni cliniche? Perché il test è richiesto in urgenza? Quali decisioni e/o azioni alternative verranno prese sulla base del risultato? Qual è il rischio relativo a non avere questo risultato? Qual è l'esito che ci si aspetta dal risultato? Qual è il livello di evidenza per l'utilità di questo test?

L'audit è, di fatto, uno strumento fondamentale per lo spostamento dell'attenzione dalla produttività alle performance cliniche dei laboratori. Gli strumenti di comparazione tra i servizi di laboratorio debbono focalizzarsi sugli outcome ma la loro valutazione è spesso complessa, poiché da una parte sono non facilmente misurabili, dall'altra essi non possono essere separati dagli interventi (Figura 2).

L'attività di Audit è uno strumento EBM che permette il continuo controllo dei propri processi diagnostici, la discussione e la condivisione con gli uti-

lizzatori delle basi teoriche e delle applicazioni nella pratica diagnostica, la costruzione di dati per un corretto *benchmarking*, ma è anche, e forse soprattutto, uno strumento che può agire come volano per un migliore e più appropriato utilizzo del laboratorio da parte dei clinici.

Bibliografia

1. <http://www.rand.org> (data di consultazione: 09.10.2004).
2. Smellie WSA. Appropriateness of test use in pathology: a new era or reinventing the wheel? *Ann Clin Biochem* 2003; 40:585-92.
3. Brook RH. Appropriateness: the next frontier. *BMJ* 1994; 308:218-9.
4. Price CP. Evidence-based Laboratory Medicine: supporting decision-making. *Clin Chem* 2000; 46:1041-50.
5. <http://www.cochrane.org/docs/sadtdoc1.htm> (data di consultazione: 13.10.2004).
6. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. L'iniziativa STARD per la produzione di studi completi ed accurati di accuratezza diagnostica. *Riv Med Lab - JLM* 2003; 4:74-9.
7. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. L'iniziativa STARD per la produzione di studi di accuratezza diagnostica: spiegazione e commenti. *Riv Med Lab - JLM* 2003; 4:80-93.
8. Giavarina D. Dall'accuratezza all'appropriatezza: il documento STARD. *Riv Med Lab - JLM*, Vol, 2004; 5:143-6.
9. Fryback DG, Thornbury JR. The efficacy of diagnostic imaging. *Med Decis Making* 1991; 11:88-94.
10. Price CP. Medical and economic outcomes of point of care testing. *Clin Chem Lab Med* 2002; 40:246-51.
11. Drummond MF, O'Brien B, Stoddart GL, Torrance GW. *Methods for the economic evaluation of health care programs*, 2nd ed. Toronto: Oxford University Press;1997.
12. Lundberg GD. Changing physician behavior in ordering diagnostic tests. *JAMA* 1998; 280: 2036.