

Dall'accuratezza all'appropriatezza: il documento STARD

D. Giavarina

Laboratorio di Chimica clinica ed Ematologia, Ospedale San Bortolo, Vicenza

Riassunto: L'accuratezza è alla base di ogni processo Evidence Based nella diagnostica e quindi anche della appropriatezza. Una migliore definizione EBM dell'accuratezza comporta una migliore appropriatezza e una più efficiente attività diagnostica.

La qualità dei lavori scientifici sulla accuratezza è determinante in questo processo. Lo Standard per la produzione di articoli relativi all'accuratezza dia-

gnostica (STARD) è una lista di controllo che ha l'obiettivo di rendere gli studi sull'accuratezza diagnostica completi e comparabili tra loro, garantendo al lettore tutte le informazioni per una corretta e completa valutazione.

Una migliore qualità nella pubblicazione degli studi sull'accuratezza potrà migliorare l'efficienza e la qualità del servizio sanitario.

Summary: Accuracy is the first step in the Evidence Based of diagnostic test and also of the "appropriateness". A better EBM definition of the accuracy get a better appropriateness and a better diagnostic efficiency.

The quality of the scientific publications regarding the accuracy is a priority in this field.

The Standard for the Reporting of Diagnostic

Accuracy studies (STARD) is a checklist with the aim to improve the quality of reporting, allowing the reader to detect the potential for bias in the study and to assess the generalisability and applicability of the results.

Complete and informative reporting can only lead to better decisions in healthcare.

Accuratezza e appropriatezza sono temi ricorrenti quando si vuole applicare la Medicina Basata sull'Evidenza (EBM) alle problematiche della diagnostica.

La metodologia EBM, nata agli inizi degli anni '70 dal pensiero dell'epidemiologo inglese Archibald Cochrane,¹ ha iniziato ad avere rilevanza per la disciplina medica solamente nell'ultimo decennio. Le citazioni su Pubmed per "evidence base medicine" sono oggi oltre 13.000, mentre erano meno di 10 nel 1993 (Figura 1).²

Secondo la classica definizione, l'EBM è l'uso consapevole, esplicito e giudizioso della migliore evidenza disponibile circa la cura dei singoli pazienti. Essa è quindi la sintesi della conoscenza scientifica verificata per le possibilità di applicazione, nelle condizioni in cui si opera.

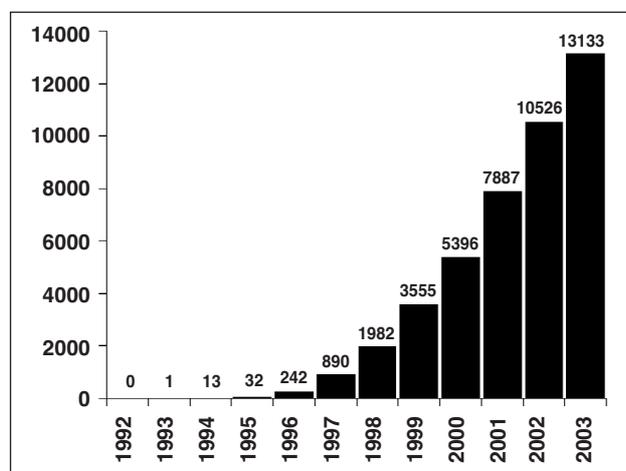
Nella valutazione dell'efficacia dei trattamenti, l'EBM ha trovato ampia applicazione, con metodologie specifiche. Le revisioni sistematiche sono lo strumento principale. Mediante la definizione esplicita

degli obiettivi, la chiara descrizione della strategia di ricerca, la descrizione dei criteri d'inclusione e di esclusione degli studi primari, la valutazione della qualità metodologica degli stessi e una particolare presentazione dei risultati, esse si propongono come documenti sintetici e completi per l'aggiornamento e la verifica delle nozioni e delle procedure in sanità.

La situazione nell'ambito delle problematiche della diagnostica invece non è altrettanto definita e consolidata. Scrive J.A. Knottnerus nel suo "The evidence based of clinical diagnosis": «Sebbene lo sviluppo delle tecnologie diagnostiche sia grandemente accelerato, la metodologia della ricerca diagnostica è ancora molto lontana rispetto alla valutazione dei trattamenti».³

Il problema principale risiede nella mancanza o scarsità della materia prima per revisioni sistematiche, ossia di studi scientifici robusti, possibilmente controllati e randomizzati.

Nel 1996 la Cochrane Collaboration tentò una raccomandazione per le revisioni sistematiche riguar-

Figura 1. Citazioni su Medline per "Evidence Based Medicine"

danti gli screening e i test diagnostici. Il *Methods Working Group* propose una definizione operativa di un test come ogni misura diretta ad identificare quegli individui che possono beneficiare dal trattamento.⁴

Il documento raccomanda di aumentare l'attenzione sull'esito finale, centrato sullo specifico paziente (sopravvivenza, qualità della vita); suggerisce che gli effetti di un test sugli esiti possano essere valutati mediante trial randomizzati, ma tuttavia ritiene non fattibile rivalutare mediante trial randomizzati ogni nuovo test, sia per problemi economici e di tempo che per problemi etici.

Può essere sufficiente valutare semplicemente l'accuratezza del test, a patto che esista già evidenza dell'efficacia delle cure per la patologia riconosciuta dal test in valutazione.

Ecco dunque la prima fonte autorevole che riconosce l'accuratezza come base per l'evidenza nella diagnostica. Ma EBM non è solo "evidenza" ma anche "fattibilità". Nel suo libro "Come praticare e come insegnare l'EBM" D.L. Sackett⁵ ha dedicato interi capitoli alle problematiche dei test. In particolare, propone una "query-list" per verificare la propria pratica EBM, basata su tre fondamentali quesiti:

1. Esiste evidenza che l'accuratezza del test è idonea?
 - a. C'era un confronto indipendente e cieco con un sistema di riferimento ("gold") diagnostico?
 - b. Il test è stato valutato in un ampio spettro di pazienti (simili a quelli che io avrò)?
 - c. E' stato applicato uno standard di riferimento indipendente dal risultato del test?
 - d. Il test è stato validato in un secondo gruppo indipendente di pazienti?
2. È dimostrato che questo (evidente) valido test discrimina tra chi ha e chi non ha una determinata patologia?
 - a. Sensibilità, specificità
 - b. Valore predittivo
 - c. Likelihood ratio
 - d. Curve ROC

3. Posso applicare questo valido, importante test diagnostico ad uno specifico paziente?
 - a. Il test è disponibile, "economico", accurato e preciso nella mia organizzazione?
 - b. Possiamo stimare con sensibilità la probabilità pre-test dei nostri pazienti?
 - c. La risultante probabilità post-test cambia il trattamento e la cura del paziente?

Nello svolgimento di questi tre quesiti si può vedere come il processo parta dalla valutazione dell'accuratezza per arrivare alla scelta di ciò che è opportuno, giusto, appropriato.

Anche la parte economica rientra nella valutazione della scelta e quindi indirettamente nell'appropriatezza.

C. Price, nel suo "Evidence Based Laboratory Medicine", classifica la valutazione dei test diagnostici con una piramide gerarchica che muove dall'efficacia verso l'efficienza:⁶

1. efficacia tecnica
2. accuratezza diagnostica
3. decisione diagnostica
4. decisione terapeutica
5. esito per paziente
6. valutazione economica

L'accuratezza è quindi alla base di ogni processo EBM ed è anche alla base dell'appropriatezza stessa. Non si può parlare di appropriatezza per test non sufficientemente accurati.

La definizione dell'accuratezza di un test abbisogna di studi sperimentali che abbiano valutato e comparato il nuovo test.

Come tutti i lavori scientifici, le problematiche della qualità della comunicazione in termini di efficienza sono rilevanti. Sono inoltre importanti metodi di standardizzazione tra gli studi, al fine di rendere comparabili i diversi lavori.

Il primo importante tentativo di standardizzazione e miglioramento della qualità dei report scientifici è il CONSORT (CONSolidated Standards Of Reporting Trials).

Si tratta di uno strumento che prende un approccio evidence-based per migliorare la qualità dei trial randomizzati. In una checklist di 22 punti sono definite ed esemplificate le varie parti che debbono essere descritte in un lavoro, dal titolo alla discussione.

Il documento è stato pubblicato per la prima volta su JAMA nel 1996,⁷ poi rivisto nel 2001⁸ ed in ultimo nel 2004,⁹ limitatamente per la parte degli studi randomizzati a gruppi di individui.

Sempre nel 1996, un gruppo di laboratoristi ha iniziato a sviluppare una serie di argomenti da includere in una checklist per gli studi dell'accuratezza diagnostica. Dopo una fase di revisione, questa lista di controllo è stata pubblicata per la prima volta nel 2000 su Clinical Chemistry ed incorporata nelle istruzioni per gli autori della stessa rivista.¹⁰

Nel gennaio del 2003 una lista di controllo più generale, lo STARD, Standard for the Reporting of Diagnostic Accuracy studies, è pubblicata contemporaneamente da diverse importanti riviste mediche (Tabella I). E' un documento, composto da una checklist di 25 punti e da un albero decisionale per la conduzione di uno studio sull'accuratezza, scaturito dal consenso tra esperti della materia e ottenuto partendo da una revisione delle varie liste di controllo più o meno attinenti, reperibili su Medline, Embase, BIOSIS e banca dati metodologica Cochrane.

Da 33 liste si sono ottenuti 75 punti, poi ridotti per consenso a 25 nella stesura finale. Lo STARD mantiene la necessità di adesione al CONSORT per gli aspetti generali, mentre descrive in maniera specifica i punti per gli studi di accuratezza diagnostica.

Due sono gli obiettivi di questa lista di controllo: migliorare la qualità degli studi sull'accuratezza e renderli confrontabili. Numerosi punti della lista di controllo hanno lo scopo di assicurare che un articolo contenga una descrizione chiara dei criteri di inclusione dei pazienti, i protocolli di analisi, i criteri di positività adottati ed una relazione adeguata sui soggetti inseriti nello studio corredata dai risultati.

Il documento inoltre presenta un diagramma di flusso che fornisce informazioni sulla modalità di reclutamento dei pazienti; sull'ordine di esecuzione degli esami; sul numero di pazienti sottoposti all'esame in valutazione e/o allo standard di riferimento. Una versione in italiano è disponibile su questa rivista,^{11,12}

anche on-line, attraverso il portale della società.

Lo STARD ha avuto accoglienza ben maggiore del CONSORT, presso riviste di medicina di laboratorio, di radiologia, ma anche di medicina generale.

Può una lista di controllo migliorare la qualità degli articoli dedicati all'accuratezza diagnostica a vantaggio di clinici, ricercatori, revisori, giornali e pazienti?

In uno studio comparativo prima-dopo, gli articoli pubblicati dalle riviste che avevano aderito al CONSORT (BMJ, JAMA e Lancet), hanno dimostrato un notevole miglioramento nella qualità dei reports, rispetto a riviste altrettanto prestigiose che però non avevano aderito all'uso della lista di controllo, quali il New England Journal of Medicine.¹³

Anche il gruppo STARD, già al momento della pubblicazione del documento, aveva in progetto di valutarne l'impatto sulla qualità degli articoli pubblicati dedicati all'accuratezza diagnostica con uno studio prima-dopo, sul modello dello studio del CONSORT. Alcuni dati preliminari sono stati recentemente presentati da Lunbreras-Lacarra e coll.¹⁴ Il miglioramento è dimostrabile per Clinical Chemistry, che aveva già introdotto una metodologia simile allo STARD a partire dal 2001, mentre un simile incremento della qualità non è ancora evidente per Clinical Chemistry and Laboratory Medicine, che non usava una lista di controllo prima dello STARD. Questi dati sono incoraggianti e dimostrano la necessità di un ulteriore lavoro per introdurre maggiormente le raccomandazioni dello STARD.

Tabella I. Elenco di riviste che hanno aderito allo STARD

Rivista	International Standard Serial Number	Abbreviazione
Archives of biochemistry and biophysics	pISSN: 0003-9861	Arch Biochem Biophys
Clinical chemistry	pISSN: 0009-9147	Clin Chem
Clinical biochemistry	pISSN: 0009-9120	Clin Biochem
American journal of clinical pathology	pISSN: 0002-9173	Am J Clin Pathol
Clinical chemistry and laboratory medicine: CCLM / FESCC	pISSN: 1434-6621	Clin Chem Lab Med
Clinica chimica acta; international journal of clinical chemistry	pISSN: 0009-8981	Clin Chim Acta
Annals of clinical biochemistry	pISSN: 0004-5632	Ann Clin Biochem
Lancet	pISSN: 0140-6736	Lancet
JAMA: the journal of the American Medical Association	pISSN: 0098-7484	JAMA
Annals of internal medicine	pISSN: 0003-4819	Ann Intern Med
BMJ (Clinical research ed.)	pISSN: 0959-8138	BMJ
Journal of clinical epidemiology	pISSN: 0895-4356	J Clin Epidemiol
Croatian medical journal	pISSN: 0353-9504	Croat Med J
Radiology	pISSN: 0033-8419	Radiology
Clinical radiology	pISSN: 0009-9260	Clin Radiol
Academic radiology	pISSN: 1076-6332	Acad Radiol
European journal of nuclear medicine	pISSN: 0340-6997	Eur J Nucl Med
Neurology	pISSN: 0028-3878	Neurology
Journal of personality assessment	pISSN: 0022-3891	Pers Assess
Rivista di medicina di laboratorio – Journal of laboratory medicine	pISSN: 1592-3479	Riv Med Lab – JLM

Il miglioramento delle pubblicazioni scientifiche sull'accuratezza costituisce la base per una migliore evidence based dei test diagnostici. Ciò determina migliori definizioni dell'appropriatezza dei test impiegati, con la promessa di poter migliorare l'efficienza e la qualità del Servizio Sanitario.¹⁵

Ringraziamenti

Si ringrazia il Dott. Giuliano dall'Olio per l'attenta revisione del testo.

Bibliografia

1. Cochrane A. Effectiveness and efficiency. Random reflections on health service. London: Nuffield Provincial Hospital Trust, 1972.
2. Straus SE. What's the E for EBM? *BMJ* 2004; 328:535-6.
3. Knottnerus JA. The evidence base of clinical diagnosis. London: BMJ Books; 2002.
4. Cochrane Methods Group on Systematic Review of Screening and Diagnostic Tests: Recommended Methods. Cochrane Collaboration, 1996. <http://www.cochrane.org/cochrane/sadtdoc1.htm>. (data ultima consultazione 20.04.2004).
5. Sackett DL, Straus SE, Richardson Ws, Rosenberg W, Haynes RB. Evidence Based Medicine. How to practice and teach EBM. China: Churchill Livingstone; 2000.
6. Price CP, Christenson Rh. Evidence-Based Laboratory Medicine: from principle to outcomes. Washington D.C.: AACC Press; 2003.
7. Begg C, Cho M, Eastwood S, Horton R, Mother D, Olkin I. *et al.* Improving the quality of reporting of randomised controlled trial. The CONSORT statement, *JAMA* 1996; 276:637-9.
8. Moher D, Shulz KF, Altman DG, for the CONSORT Group. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet* 2001; 357:1191-4.
9. Campbell MK, Elbourne DR, Altman DG, for the CONSORT Group. CONSORT statement: extension to cluster randomised trials. *BMJ* 2004; 328:702-8.
10. Bruns DE, Edward JH, Magid E, Young DS. Toward a checklist for reporting of studies of diagnostic accuracy of medical tests. *Clin Chem* 2000; 46:893-5.
11. 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.
12. 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.
13. Moher D, Jones A, Lepade F, for the CONSORT group. Use of CONSORT statement and quality of reports of randomised trials. A comparative before-and-after evaluation. *JAMA* 2001; 285:1992-5.
14. Lumberras-Lacraa B., Ramos-Rincón JM, Hernández-Aguado I. Methodology in Diagnostic Laboratory Test Research in *Clinical Chemistry and Clinical Chemistry and Laboratory Medicine*. *Clin Chem* 2004; 50:530-6.
15. Bossuyt PMM. The quality of reporting in diagnostic test research: getting better, still not optimal. *Clin Chem* 2004; 50:465-6.