

Ragionamento diagnostico, Evidence-based Medicine e linee guida

A. Cartabellotta

Direttore Scientifico, GIMBE® - Gruppo Italiano per la Medicina Basata sulle Evidenze, Palermo

Premessa

Nel corso degli ultimi vent'anni la progressiva diffusione di tecnologie sempre più avanzate ha contribuito ad identificare la diagnosi con la frenetica ed immotivata richiesta di test di laboratorio e strumentali, stimolando un accanimento diagnostico che, oltre ad essere poco professionale e scientifico, riveste aspetti etici ed economici di grande rilievo sociale.

Infatti, l'uso indiscriminato di test diagnostici scatena un meccanismo a cascata con gravi ripercussioni sull'efficacia ed efficienza dell'assistenza sanitaria: sia perché il numero elevato d'indagini richieste determina una congestione dei servizi con allungamento sia dei tempi di degenza che di quelli di attesa; sia perché i risultati falsamente positivi, la cui frequenza cresce esponenzialmente con il numero dei test richiesti, scatenano una cascata di indagini inutili e "peripezie diagnostiche", configurando la cosiddetta sindrome di Ulisse¹, causa di incremento dei costi e fonte di apprensione per il paziente (*labeling effect*). In realtà, ad eccezione del beneficio psicologico nel rassicurare medico e paziente, l'uso indiscriminato di test diagnostici è spesso fonte di informazioni ridondanti, l'utilità clinica non riflette l'entusiasmo e la celerità con cui si diffondono ed, oltre al lievitamento dei costi, aumenta il rischio di gravi errori specie se il test è poco accurato o scarsamente riproducibile. Purtroppo il numero di indagini utilizzate nella pratica clinica è in costante aumento perché all'estrema disponibilità ad accettare le novità - anche in assenza di evidenze - si contrappone un'ostinata riluttanza ad abbandonare i test obsoleti, nonostante rigorose valutazioni metodologiche li abbiano relegati alla storia della medicina.

Questo delicato aspetto dell'assistenza sanitaria consegue alla variabile interazione di vari fattori²:

- gli stimoli allo sviluppo tecnologico, oltre a derivare dal progresso scientifico, sono fortemente influenzati dall'interesse dell'industria a conquistare consistenti spazi di mercato;
- la formazione universitaria e l'aggiornamento professionale continuano a dedicare poco spazio sia

agli strumenti per valutare la performance dei test diagnostici, sia alle tecniche di analisi decisionale;

- l'insicurezza mista all'eccesso di prudenza - per una maggiore tutela medico-legale - generano la convinzione che all'aumento delle indagini richieste corrispondono maggiori probabilità di giungere ad una diagnosi corretta;
- la mania di grandezza culturale che mira al completamento totale delle indagini, un atteggiamento figlio della "strategia esaustiva" che inizia il ragionamento diagnostico dopo aver raccolto in maniera afinalistica una grande quantità di dati;
- l'insistenza del paziente e la disponibilità del medico ad assecondare le sue richieste, quasi sempre conseguenti all'informazione troppo poco scientifica dei media;
- la mancata rigorosa valutazione dei test diagnostici prima di una loro diffusione in larga scala (d'altra parte, indagini di utilità clinica ampiamente dimostrate sono poco utilizzate per lo scarso interesse editoriale e/o commerciale);
- l'estensione del concetto di screening al "*case-finding*", ossia il tentativo di identificare dati patologici grazie all'esecuzione di un certo numero di test "di routine". In realtà, la speranza di identificare "troppe malattie in tutti i pazienti" ha dato risultati molto deludenti: infatti, per diversi test (ECG, rx torace, emocromo, formula leucocitaria, profili biochimici, esami delle urine) esistono buone evidenze che la loro richiesta senza sospetto clinico raramente consente di formulare nuove diagnosi e quasi mai di istituire nuovi trattamenti.

Ragionamento diagnostico ed utilizzo dei test

La strategia esaustiva concepisce il processo diagnostico in due fasi: nella prima il medico costruisce, in maniera acritica, il *database* del paziente; quindi, sulla scorta di una quantità rilevante di dati raccolti, inizia il ragionamento diagnostico. Tale strategia, tipica del medico principiante ed invalidata da numerosi studi, dovrebbe essere definitivamente abbandonata.

Il metodo che meglio si accorda con il concetto di efficienza diagnostica, è quello ipotetico-deduttivo: i primi dati del paziente (sintomi/segni principali, età, sesso, fattori di rischio) suggeriscono precocemente una o più ipotesi, che indirizzano l'ulteriore ricerca dei dati, sia clinici, sia di laboratorio e strumentali. La loro progressiva acquisizione rimodella l'ipotesi iniziale ed indirizza la ricerca di nuovi elementi sino alla formulazione della diagnosi corretta. L'ipotesi iniziale è comunque molto flessibile, soggetta a continui aggiustamenti ed, eventualmente, ad essere sostituita da altre ipotesi se emergono nuovi dati (secondo Kassirer³ la generazione delle ipotesi diagnostiche è più simile alle sequenze di un film che ad un singolo fotogramma). Secondo la strategia ipotetico-deduttiva, i test sono utili quando il grado di certezza diagnostica che deriva dalla storia e dall'esame obiettivo è insufficiente per prendere decisioni cliniche. Tuttavia, la capacità intrinseca di un test di ridurre i margini di incertezza è ampiamente variabile, per cui la sua richiesta a scopi diagnostici deve risultare dalla valutazione di numerosi fattori:

- *Probabilità pre-test di malattia.* Se la probabilità di malattia prima di eseguire un test diagnostico è troppo bassa o troppo alta, il suo risultato difficilmente modificherà la probabilità *post-test* e quindi la decisione clinica
- *Sensibilità e specificità del test.* Sono caratteristiche proprie di ciascun test, che riflettono le informazioni attese in pazienti con o senza la malattia presa in considerazione. Un risultato positivo non fornisce la certezza assoluta di malattia, considerato che può esserlo sia in pazienti malati, proporzionalmente alla sua sensibilità, sia in soggetti sani, in relazione inversa alla sua specificità; analogamente la negatività di un test, proporzionale alla sua specificità in un soggetto sano ed in relazione inversa alla sensibilità nei soggetti malati, non è sempre associata all'assenza di malattia. Pertanto i tests molto sensibili sono utili per escludere, se negativi, una malattia, mentre quelli ad elevata specificità, quando positivi, servono per confermare l'ipotesi diagnostica. Considerato che per numerosi test di laboratorio i risultati si dispongono lungo una scala di valori numerici (ad es. glicemia, creatininemia), è opportuno distinguerli in tre gruppi: quelli che rientrano ampiamente in un range di normalità, che indicano verosimilmente l'assenza di malattia, quelli nettamente patologici, ed una fascia intermedia da interpretare adeguatamente in relazione al contesto clinico del paziente. In questo gruppo spesso rientrano risultati falsamente-positivi, che per un fenomeno puramente statistico (regressione verso la media) risultano normali ad una successiva determinazione.
- *Accettabilità del test in relazione al vantaggio diagnostico.* Un'attenta valutazione del rischio, dell'invasività e dei costi del test diagnostico può suggerire l'ipotesi di prendere una decisione terapeutica senza eseguirlo: ad esempio, alla luce delle nuove evidenze sull'efficacia delle eparine a basso peso molecolare nei pazienti con sospetta embolia

polmonare, è opportuno iniziare il trattamento anche a un livello intermedio di probabilità diagnostica.

- *Effetti del test sulla decisione clinica.* Se si prevede che questa non venga modificata dal risultato del test è inutile eseguirlo, specie se costoso e/o invasivo. Ad esempio, in pazienti con metastasi da neoplasia di origine sconosciuta, le evidenze indicano discrete probabilità di identificare la sede primitiva del tumore, ma prospettive terapeutiche molto scarse. Infatti, in una valutazione cumulativa su 2283 pazienti⁴, la sede primitiva del tumore è stata identificata nel 21% dei casi, ma solo 25 pazienti (1.1%) erano affetti da neoplasie potenzialmente curabili (linfomi, tumori del testicolo).

Evidence-based Medicine: quale contributo al processo diagnostico?

I promotori della Evidence-based Medicine⁵ (EBM) hanno trascurato che il procedimento clinico è una "creatura" estremamente complessa ed il potenziale contributo offerto dall'EBM alle decisioni del medico non è omogeneo. La recente revisione sistematica sui limiti della EBM⁶, conferma che questo aspetto è stato approfondito in letteratura solo da Autori italiani⁷, che si sono soffermati sul contributo che l'EBM può offrire nelle varie fasi del processo decisionale.

Decisioni diagnostiche

Riconoscimento di presentazioni cliniche e generazione d'ipotesi diagnostiche. Quando il medico è esperto e/o il caso clinico gli è familiare, il riconoscimento di malattia è immediato grazie ad un matching con le caratteristiche memorizzate di pazienti visti in precedenza⁸. Questa capacità rappresenta la caratteristica essenziale del processo diagnostico: infatti, la memoria si espande progressivamente accogliendo nuove esperienze e conoscenze, che vengono organizzate in una rete sempre più ampia e ricca di connessioni: il *semantic network*. Se il medico è meno esperto e/o nei casi insoliti o più complessi, il ragionamento diagnostico si basa sul processo ipotetico-deduttivo, attraverso la generazione di una o più ipotesi e la ricerca di ulteriori dati che dovrebbero essere presenti o assenti. Ad esempio, se in una paziente con dolore addominale acuto e febbre il medico sospetta una colecistite, presume che siano presenti storia di coliche biliari e/o familiarità, ed assenti i disturbi dell'alvo.

In questa fase al medico viene richiesta la capacità di: a) percepire i dati rilevanti per la diagnosi, tra i numerosi segni, sintomi e risultati di eventuali indagini; b) aggregare i dati ottenuti in modelli nosografici di malattia ben definiti. Considerato che questa fase è basata sia sulle capacità percettive e cognitive del medico, sia sull'ampiezza e l'organizzazione del *semantic network*, l'EBM ha un ruolo molto limitato.

Verifica delle ipotesi diagnostiche: la selezione-interpretazione dei test. Le ipotesi formulate devono essere confermate o smentite in seguito all'acquisizione di ulteriori dati, che generalmente prevedono l'utilizzo di test diagnostici, in grado di valutare anche la severità della malattia e le condizioni individuali che possono influenzare le scelte terapeutiche. In questa fase, l'errore più frequente è interpretare i dati acquisiti come conferma delle ipotesi già formulate, anche se insignificanti o con esse incompatibili. In questa fase l'EBM è di discreta utilità, sia per meglio conoscere una malattia ipotizzata o diagnosticata, sia - soprattutto - per la scelta e l'interpretazione dei test diagnostici.

Decisioni terapeutiche: scelta e valutazione dei trattamenti. Raggiunta una probabilità diagnostica sufficientemente certa, la decisione terapeutica può essere favorevolmente influenzata dalle prove di efficacia dei trattamenti disponibili, dalle evidenze di tollerabilità/rischio, dalla storia naturale della malattia non trattata. In questa fase, l'EBM è di massima utilità perché informa sull'entità/precisione delle prove di efficacia dei trattamenti; rimane l'incertezza legata all'applicabilità clinica delle sperimentazioni controllate e randomizzate (RCTs) e la frequente difficoltà a valutare il rischio basale dei pazienti a causa del numero limitato di studi di storia naturale.

Oltre alla limitata utilità della EBM nel processo diagnostico, l'interesse delle numerose pubblicazioni si concentra prevalentemente sulle prove di efficacia dei trattamenti, nonostante il bisogno d'informazione dei clinici riguardi con la stessa frequenza problemi diagnostici e terapeutici^{9,10}.

- *Best Evidence 5* è un CD-ROM che contiene circa 1500 abstract strutturati selezionati dalle pubblicazioni secondarie *ACP Journal Club* ed *Evidence-based Medicine*. Oltre 80% di tali abstract riguardano l'efficacia dei trattamenti.
- *Clinical Evidence*, il nuovo "evidence-compendium" organizzato per malattie, costituisce di fatto una mappa delle prove di efficacia dei trattamenti disponibili per varie condizioni cliniche, e la *Cochrane Library* contiene, virtualmente, solo revisioni sistematiche sull'efficacia degli interventi terapeutici.
- Nei vari studi che hanno verificato in vari setting specialistici l'aderenza della pratica clinica ai principi della EBM, la diagnosi è sempre un dato acquisito. In altre parole questi studi hanno solo verificato se le scelte terapeutiche erano in accordo con RCTs, con evidenze non sperimentali, oppure non supportate da alcuna prova.
- L'ampia letteratura dedicata agli aspetti tecnici della EBM^{11,12,13,14}, trascurando l'obiettivo di fornire strumenti per l'esercizio del ragionamento clinico, prende in considerazione solo la componente del processo diagnostico relativa alla conferma delle ipotesi formulate, attraverso la valutazione delle caratteristiche operative (sensibilità, specificità,

valori predittivi, rapporti di verosimiglianza) dei dati clinici, di laboratorio e strumentali. Infatti, l'esiguo spazio dedicato alla diagnosi in *Best Evidence* e *Clinical Evidence* e la serie *The Rational of Clinical Examination*¹⁵ (oltre 30 revisioni sistematiche sui dati della storia e dell'esame fisico), si concentrano esclusivamente sulle caratteristiche operative di dati clinici e di test diagnostici.

- Solo cinque anni dopo dalla pubblicazione delle due *Users' Guide* sulla diagnosi^{16,17} - centrate sulla valutazione dei test diagnostici - sono stati approfonditi gli aspetti relativi alla diagnosi differenziale¹⁸ ed alla generazione delle ipotesi¹⁹.
- L'unico testo che correla i principi del ragionamento diagnostico con la valutazione delle caratteristiche operative dei test è *Diagnostic Strategies for Common Medical Problems*²⁰, la cui prima edizione risale all'epoca pre-EBM.

Anche se pochi riconoscono la scarsa utilità della EBM nel processo diagnostico^{21,22}, tale "ipotrofia" può essere ricondotta a varie motivazioni.

- La rigorosa valutazione di nuovi test diagnostici è ancora poco diffusa, soprattutto in relazione al livello metodologico raggiunto dai RCTs: infatti, diverse valutazioni metodologiche della ricerca clinica nell'area dei test diagnostici dimostrano che siamo ancora ben lontani da standard accettabili^{23,24}.
- I criteri per la conduzione delle meta-analisi nell'area dei test diagnostici sono stati definiti solo recentemente^{25,26}.
- La maggior parte dei promotori della EBM - di area epidemiologico-statistica²⁷ - sono estranei al processo diagnostico ed alla relazione cognitiva che collega i dati di una presentazione clinica con una malattia nosograficamente definita, uno degli aspetti fondamentali per la generazione delle ipotesi diagnostiche. Infatti, a fare scuola in quest'area, sono stati, in epoca pre-EBM, Elstein e coll, Pauker e Kasirer, Sox, e più recentemente Regehr e Norman²⁸.

Linee guida: necessarie, ma non sufficienti

Dopo anni di totale indifferenza assistiamo anche nel nostro Paese ad un crescente interesse per le linee guida (LG), sia da parte degli amministratori sanitari - che verosimilmente intravedono un mezzo per il contenimento della spesa sanitaria - sia tra la classe medica che inizia a prenderle in considerazione quale strumento per migliorare la pratica clinica. Ritengo opportuno, pertanto, soffermarmi su tre aspetti principali che possono condizionare il successo delle LG.

Innanzitutto, il processo di elaborazione di linee guida (LG) dovrebbe seguire metodologie oggettive, condivise e facilmente riproducibili. In realtà, numerosi studi^{29,30,31,32} dimostrano che la qualità media delle LG internazionali - in particolare quelle prodotte da società scientifiche - è ancora ben lontana da standard metodologici accettabili. Infatti, questi studi hanno documentato come nella

Tabella I. Livelli di evidenza e forza delle raccomandazioni.

Livelli di Evidenza	
Ia	Informazioni derivate da revisioni sistematiche o meta-analisi di studi clinici randomizzati
Ib	Informazioni derivate da almeno uno studio clinico randomizzato
Ila	Informazioni derivate da almeno uno studio clinico controllato non randomizzato di buona qualità
Ilb	Informazioni derivate da altri tipi di studi controllati di buona qualità
III	Informazioni derivate da studi non controllati di buona qualità (serie di casi, studi di correlazione, studi descrittivi)
IV	Informazioni basate unicamente sull'opinione di esperti di altri organismi o su quella dei componenti del gruppo di lavoro multidisciplinare
Forza delle Raccomandazioni	
Grado A	Raccomandazioni basate su informazioni scientifiche di livello Ia o Ib
Grado B	Raccomandazioni basate su informazioni scientifiche di livello Ila, Ilb o III
Grado C	Raccomandazioni basate unicamente su opinioni di esperti (livello IV)

maggior parte delle LG valutate non vi sia alcuna garanzia che le raccomandazioni siano formulate sulla base di una revisione sistematica delle evidenze disponibili e che la loro interpretazione sia stata condotta in modo trasparente e con criteri di valutazione espliciti.

Secondo Grilli et al. esistono tre criteri base che permettono di definire la qualità metodologica delle LG:

- **Multidisciplinarietà del gruppo di lavoro.** Necessaria per garantire un "sano equilibrio" nell'interpretazione delle evidenze e nella formulazione delle raccomandazioni; infatti è documentato che le LG prodotte da gruppi monodisciplinari di specialisti tendono a sovrastimare l'efficacia degli interventi terapeutici e ad estendere le indicazioni dei test diagnostici. In un gruppo di lavoro multidisciplinare dovrebbero essere rappresentati: tutte le specialità cliniche coinvolte nella gestione del problema assistenziale, medici di medicina generale, infermieri, medici di direzione sanitaria, epidemiologi, statistici, esperti nella ricerca di informazioni scientifiche, economisti, rappresentanti dei pazienti.
- **Processo esplicito di ricerca bibliografica.** Considerato che le LG sono un prodotto di ricerca secondaria, dovrebbe essere descritto nei dettagli il processo di identificazione delle evidenze: banche dati e strategie di ricerca utilizzate, range temporale della ricerca bibliografica, eventuali altre fonti.

- **Grading delle evidenze.** Le regole formali dell'epidemiologia clinica assegnano a ciascun tipo di studio una specifica posizione nella "gerarchia delle evidenze scientifiche". Attraverso il *grading* viene "mappata" la qualità della ricerca disponibile che, attraverso la "forza delle raccomandazioni", viene esplicitamente legata alle singole raccomandazioni fornite dalla LG (Tab. I).

In secondo luogo, poiché è ormai evidenza consolidata che la semplice diffusione delle LG non garantisce l'aderenza dei clinici alle sue raccomandazioni³³, le LG devono essere implementate con adeguate strategie³⁴. Ad esempio, le richieste inappropriate relative alla richiesta di test diagnostici possono essere modificate grazie alle LG^{35,36}, a condizione che vengano "rinforzate" da adeguate strategie d'implementazione: modifiche ai moduli di richiesta (reminder) e/o disposizioni di politica sanitaria locale (Tab. II).

È corretto, infine, ricordare che esistono numerosi problemi legati alla trasferibilità clinica delle LG: infatti, poiché dirette a guidare gli interventi medici su classi di malattia e non su pazienti singoli, l'utilità pratica delle LG dipende dalla capacità del medico di interpretarle con la flessibilità richiesta dalle caratteristiche del paziente individuale.

Tabella II. Effetti sulla richiesta di test diagnostici ottenuti dalle linee guida (LG).

Area	Tipo intervento	Test	Effetti dell'intervento
Ematologia	LG, MM	VES	Riduzione del 58% (da 2000 a 500 per 100.000 persone)
Funzionalità renale	LG, MM, PS	Analisi urine chimico-fisica (AU-CF) ± microscopica (AU-M)	Incremento di AU-CF e riduzione di AU-M
		Azotemia	Notevole riduzione (da 1800 a 400 per 100.000 persone)
		Creatinina	Nessun incremento significativo
Depositi marziali	LG, PS	Sideremia, TIBC	Notevole riduzione (80%)
		Ferritina	Nessun incremento significativo
Funzione tiroidea	LG, MM, PS	T4	In 3 anni il T4 diminuisce da 1300 per 100.000 persone a zero
		TSH	Nessun incremento significativo

LG = linee guida; MM = modifiche ai moduli di richiesta; PS = disposizioni di politica sanitaria.

Bibliografia

1. Rang M. The Ulysses syndrome. *Can Med Assoc J* 1972;106:122-3.
2. Lundemberg GD. Perseveration of laboratory test ordering: a syndrome affecting clinicians. *JAMA* 1983;249:639-44.
3. Kassirer JP. Diagnostic reasoning. *Ann Intern Med* 1989;110:893-900.
4. Shapira DV, Jarret AR. The need to consider survival, outcome, and expense when evaluating and treating patients with unknown primary carcinoma. *Arch Intern Med* 1995;155:2050-4.
5. Evidence-Based Medicine Working Group. Evidence-based medicine: a new approach to teaching the practice of medicine. *JAMA* 1992;268:2420-5.
6. Straus SE, McAlister FA. Evidence-based medicine: a commentary on common criticisms. *Can Med Ass J* 2000;163:837-41.
7. Pagliaro L, Cappello M, Malizia G, et al. La medicina delle prove di efficacia nell'assistenza al paziente individuale. In: Liberati A, ed. *La medicina delle prove di efficacia: utilità e limiti della evidence-based medicine*. Roma: Il Pensiero Scientifico Editore, 1997:33-56.
8. Richardson WS. Where do pretest probabilities come from? *Evidence-Based Medicine* 1999;4:68-9.
9. Ely JW, Osheroff JA, Gorman PN, et al. A taxonomy of generic clinical questions: classification study. *BMJ* 2000;321:429-32.
10. Leclerc H, Beaulieu MD, Bordage G, et al. Why are clinical problems difficult? General practitioners' opinions concerning 24 clinical problems. *Can Med Ass J* 1990;143:1305-15.
11. Sackett DL, Straus S. On some clinically useful measures of the accuracy of diagnostic tests. *ACP J Club* 1998;129(Suppl 2):A17-A19.
12. McKibbin KA, Walker-Dilks CJ. Beyond ACP Journal Club: how to harness MEDLINE for diagnostic problems. *ACP J Club* 1994;121(Suppl 2):A10-A12.
13. Haynes RB. Change in criteria for studies of diagnostic tests. *ACP Journal Club* 1991;115(Suppl. 3):A-13.
14. Moore RA. On the need for evidence-based clinical biochemistry. *Evidence-Based Medicine* 1998;3:7-8.
15. Sackett DL, Rennie D. The science of the art of the clinical examination. *JAMA* 1992;267:2650-2.
16. Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? *JAMA* 1994;271:703-7.
17. Jaeschke R, Guyatt G, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. A. Are the results of the study valid? *JAMA* 1994;271:389-91.
18. Richardson WS, Wilson MC, Guyatt GH, et al. Users' guides to the medical literature: XV. How to use an article about disease probability for differential diagnosis. *JAMA* 1999;281:1214-9.
19. Richardson WS, Wilson MC, Williams JW Jr, et al. Users' guides to the medical literature: XXIV. How to use an article on the clinical manifestations of disease. *JAMA* 2000 ;284:869-75.
20. Black ER, Bordley DR, Tape TG, Panzer RJ. *Diagnostic strategies for common medical problem*. 2nd edition. Philadelphia: American College of Physicians, 1999.
21. Richardson WS. Evidence-based diagnosis: more is needed. *Evidence-Based Medicine* 1997;2:70-1.
22. Fowler PB. Evidence-based diagnosis. *J Eval Clin Pract* 1997;3:153-9.
23. Carrington Reid M, Lachs MS, Feinstein AR. Use of methodological standards in diagnostic test research. Getting better but still not good. *JAMA* 1995;274:645-51.
24. Lijmer JG, Mol BW, Heisterkamp S, Bossel GJ, et al. Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA* 1999;282:1061-6.
25. Irwig L, Tosteson ANA, Gatsonis C, et al. Guidelines for meta-analyses evaluating diagnostic test. *Ann Intern Med* 1994;120:666-76.
26. Cochrane Methods Working Group on Systematic Review of Screening and Diagnostic Tests. *Recommended Methods*, updated 6 June 1996. Disponibile a: <http://som.flinders.edu.au/FUSA/COCHRANE/cochrane/sadtdoc1.htm>
27. Marwick C. Proponents gather to discuss practicing evidence-based medicine. *JAMA* 1997;278:531-2.
28. Regehr G, Norman GR. Issues in cognitive psychology: implications for professional education. *Acad Med* 1996;71:988-1001.
29. Shaneyfelt TM, Mayo-Smith MF, Rothwangl J. Are guidelines following guidelines? The methodological quality of clinical practice guidelines in the peer-reviewed medical literature. *JAMA* 1999;281:1900-5.
30. Grilli R, Magrini N, Penna A, Mura G, Liberati A. Quality of practice guidelines developed by specialty societies. The need for a critical appraisal. *Lancet* 2000; 355:103-6.
31. Lacasse Y, Ferreira I, Brooks D, Newman T, Goldstein RS. Critical appraisal of clinical practice guidelines targeting chronic obstructive pulmonary disease. *Arch Intern Med* 2001;161:69-74.
32. Graham ID, Beardall S, Carter AO, et al. What is the quality of drug therapy clinical practice guidelines in Canada? *Can Med Ass J* 2001;165:157-63.
33. Cabana MD, Rand CS, Powe NR, Wu AW, Wilson MH, Abboud PAC, et al. Why don't physicians follow clinical practice guidelines? A framework for improvement. *JAMA* 1999; 282: 1458-65.
34. Grimshaw J, Shirran E, Thomas R, Mowatt G, Fraser C, Bero L, et al. Getting evidence into practice. *NHS Centre for Review & Dissemination. Effective Health Care* vol 1, 1999.
35. Solomon DH, et al. Techniques to improve physicians' use of diagnostic test. *JAMA* 1998;280:2020-7.
36. van Walraven C, et al. Effect of population-based interventions on laboratory utilization. *JAMA* 1998; 280: 2028-33.