Evidence Based Laboratory Medicine - *a cura di EBLM SIMeL*

A simple and effective bayesian calculator for daily practice of Laboratory Medicine

D. Giavarina[ab*], G. Pistolato[c], R.M. Dorizzi[ad]

[a]*On behalf of the Evidence Based Laboratory Medicine group of the Italian Society of Laboratory Medicine*
[b]*Clinical Pathology Dept., San Bortolo Hospital, Vicenza*
[c]*Free lance professional informatician*
[d]*Clinical Chemistry and Hematology Laboratory, Hospital of Verona*
**To whom can be requested the free Bayesian calculator*

Summary

The extension of Evidence Based Medicine (EBM) principles to diagnostics has been widely advocated and the assessment of accuracy is essential for an "evidence based" evaluation of diagnostic procedures. The practice of Evidence Based Laboratory Medicine (EBLM) requires not only the calculation of likelihood ratios, predictive values and number needed to diagnose but also the estimation of their imprecision. However, the calculation of these parameters is not very easy nor common and the inadequacy of the available tools limits EBLM adoption in general practice and specialized environment. Information technology is making more and more available accessible easy-to-use, accurate and fast tools for these needs. Here we summarize the Bayesian theory for the probabilistic evaluation of a diagnostic test and briefly review several Web available Bayesian calculators. A simple calculator written in JAVA is presented that allows entering the data in a single page in Italian and English, prepared on behalf of the Evidence Based Laboratory Medicine group of the Italian Society of Laboratory Medicine. The diagnostic performance of BNP assay at different cut-off values reported in a classical paper are used in order to clarify the possible use of the proposed calculator. The calculator instantaneously shows the very poor performance of the cut-off of 17.9 ng/L (number needed to diagnose = 4.72, i.e. almost five assays are needed for a single diagnosis) and the much better performance of the cut-off of 76 ng/L (number needed to diagnose = 1.91).

Riassunto

Un calcolatore Bayesiano semplice ed efficace per la pratica quotidiana della medicina di laboratorio

L'estensione dei principi dell'Evidence Based Medicine (EBM) alla diagnostica è stata auspicata da più parti e la stima dell'accuratezza è essenziale per una valutazione "basata sull'evidenza" delle procedure diagnostiche. La pratica dell'Evidence Based Laboratory Medicine (EBLM) richiede non solo il calcolo dei quozienti di probabilità (LR), dei valori predittivi e del numero necessario per la diagnosi (NND), ma anche la stima della loro imprecisione. Tuttavia, il calcolo di questi parametri non è molto facile né comune, e l'inadeguatezza degli strumenti disponibili limita l'adozione dell'EBLM nella medicina sia di base che specialistica. L'Information Technology sta divenendo sempre più accessibile, disponibile e facile da usare, fornendo mezzi rapidi ed accurati per queste necessità. Questo lavoro riassume la teoria Bayesiana per una valutazione dei test diagnostici in termini probabilistici e descrive brevemente alcuni dei più interessanti calcolatori Bayesiani disponibili su Web. Viene inoltre presentato un semplice calcolatore, preparato per conto del Gruppo di studio sulla EBLM della Società Italiana di Medicina di Laboratorio, scritto in JAVA, che permette l'inserimento dei dati, sia in italiano che in inglese, in una singola pagina. Sono state utilizzate le prestazioni diagnostiche del dosaggio del BNP a differenti cut-off, ripresi da un articolo classico al fine di esemplificare il possibile uso del calcolatore proposto. Il calcolatore calcola immediatamente le prestazioni molto cattive del cut-off di 17.9 ng/L (NND=4.72, cioè, almeno cinque dosaggi positivi per avere una singola diagnosi vera) e le prestazioni decisamente migliori del cut-off di 76 ng/L (NND=1.91).

Ricevuto: 18-04-2007 Pubblicato on-line: 21-05-2007

Corrispondenza a: Dott. Davide Giavarina, Laboratorio di Chimica clinica ed Ematologia, ULSS 6 Vicenza, Viale Rodolfi n. 37, 36100 Vicenza. Tel. 0444-753372, e-mail: davide.giavarina@ulssvicenza.it

Introduction

Diagnostics comprise only a small fraction of total hospital costs (less than 5%), but may influence as much as 60-70% of health care decision-making. An evidence-based use of diagnostics is integral to high quality health care, including informing earlier, better targeting health care interventions and averting adverse health outcomes and unnecessary costs[1]. Several clinical decision support system applications have been proposed to assist with clinical decisions: preventive and monitoring tasks, prescribing of drugs, and diagnosis. Applications in the third category are by far the most challenging and require sophisticated probabilistic mathematical models and in their more mature versions require an expansive knowledge base covering the full range of diseases and conditions, detailed patient specific clinical information (e.g., history, physical examination, laboratory data), and a powerful computational engine that employs some form of probabilistic decision analysis[2]. However, since the adoption of Evidence-Based Medicine (EBM) tenets and tools is becoming more and more relevant and urgent also in clinical laboratories, some simpler devices have been developed for helping in this context. The need to measure the efficacy of treatments is extending to the diagnostic tests since the clinical decision making is influenced by the results of more and more innovative tests. The development and the marketing of genomics and proteomics technologies are opening unexplored diagnostic pathways; however, "the need for a well-reasoned and evidence based approach to the choice of efficient diagnostic strategies has never been greater"[3]. The interest in improving the assessment of diagnostic accuracy is growing and the STARD initiative for the reporting of diagnostic studies has been discussed and accepted by several clinical and laboratory journals[2-8].The importance of assessing the diagnostic tests in clinical relevant and homogeneous population is well known and reinforces the need that the single laboratorian and clinician learn to estimate the "evidence" of his/her practice in the correct environment[9]. The aim of this study was to develop a simple tool which can help, in our opinion, in this relevant task.

EBM tools for diagnostics

The Bayesian approach, which defines the diagnosis in probability terms, is a valid EBM tool for choosing and assessing a test. The heart of the diagnosis is to convert pre-test probability of disease (p) through sensitivity (the probability of positive test results in disease) (P[T+|D+]) and specificity (the probability of negative test results without disease) (P[T-|D-]) into post-test probability of disease (P[D+|T+]).

$$P[D+|T+] = \frac{p \cdot (P[T+|D+])}{p \cdot (P[T+|D+]) + (1-p) \cdot (1-(P[T-|D-])} \quad (1)$$

The formula (1) can be simplified by the likelihood ratio (LR), the ratio between sensitivity and the reciprocal of specificity:

$$LR = \frac{(P[T+|D+])}{1-(P[T-|D-])} \quad (2)$$

$$P[D+|T+] = \frac{p}{p + (1-p)/LR} \quad (3)$$

However, since estimation of the probability of disease after a positive result is not immediate nor easy, nomograms have been proposed for assisting in this task. Fagan proposed in 1975 one of the most popular of these devices[10].

Bayesian calculators in the Web

Several bayesian calculators of different power and simplicity can be found in the Web. A sample of some of the most interesting, according to us, follows:

- *http://araw.mede.uic.edu/cgi-bin/testcalc.pl* (last accessed April 10th 2007): a good Fagan's nomogram presentation coupled to calculations and confidence intervals;
- *http://statpages.org/bayes.html* (last accessed April 10th 2007): calculator with a mathematical approach; it can be used only by researchers with a strong statistical background;
- *http://psych.fullerton.edu/mbirnbaum/bayes/BayesCalc.htm* (last accessed April 10th 2007): essential calculator, presented in a very technical form, provided with a complete explanation;
- *http://www.healthcare.ubc.ca/calc/bayes.html* (last accessed April 10th 2007) and
- *http://www.intmed.mcw.edu/clincalc/bayes.html* (last accessed April 10th 2007): produced respectively by the University of British Columbia and by the Medical College of Wisconsin, allow the entering of raw data in a 2x2 table format or Pre-test likelihood ratio, sensitivity and specificity or Pre-test likelihood and the positive (LR+) and negative (LR-) likelihood ratios. Their features are similar and no information about the confidence of the results is provided;
- *http://www.cebm.net/dxtable.asp* (last accessed April 10th 2007) and *http://www.cebm.net/nomogram.asp* (last accessed April 10th 2007): the graphical nomogram contained in the very effective teaching site for EBM cured by the Centre of Evidence Medicine of Oxford requires Shockwave platform is one of the most intuitive and attractive example of Fagan's nomogram.
- *http://www.cebm.utoronto.ca/practise/ca/statscal/* (last accessed April 10th 2007): produced by Mount Sinai Hospital, University of Toronto, is probably the best EBM Calculator designed to calculate relevant stati-

stics for diagnostic studies. Entering raw data or LRs in a 2x2 table, Receiver Operating Characteristic (ROC) curve, sensitivity and specificity, Positive Predictive Values (PPV), Negative Predictive Values (NPV), LR+ e LR-, and the relative confidence intervals can be obtained. Post Test Probability is not directly provided but can be indirectly obtained by ROC curve. The software, in JAVA environment, can be freely downloaded both for palm and pocket PC–device and requires Plug-in:Java™ Version 1.3.1_06 (JRE 1.3.1_06 Java HotSpot™ Client VM). Problems of compatibility between some browsers and JAVA script may be encountered.

– *http://faculty.vassar.edu/lowry/clin1.html* (last accessed April 10th 2007): a somewhat confused calculator, produced by Vassar University, NY; the structure of the 2x2 table is unusual (i.e. the positions of "condition present" cells are at the right and those of "condition absent" are at the left). On the contrary, this calculator should be recommended for its weighted for prevalence Likelihood ratio.
– *http://www.fammed.ouhsc.edu/robhamm/cdmcalc.htm#Disease* (last accessed April 2nd 2007): a list of calculators and spreadsheets (Excel™ format; many links to the frameworks are by FTP protocol) produced by the Department of Family and Preventive Disease at the University of Oklahoma Health Science Center. The user can transpose calculations into his/her own spreadsheets.

EBM and confidence interval

LRs, ROC curves and Predictive values are based on Sensitivity, Specificity and the Prevalence of disease in the particular population of interest. Statistical procedures cannot avoid the misleading of the transference of predictive values obtained in populations with substantially different prevalences. Furthermore, the estimation of the confidence intervals allows to assess the imprecision of the sample measurement[11]. Notwithstanding this obvious limitation, the diagnostic accuracy assessment does not take in account the sampling problem and the need to quantify the uncertainty of the accuracy measurement. Really, most of the calculators available in the web do not provide the confidence intervals of their results and this represent a relevant limitation.

Handheld computers and clinical utilities

Palm devices and EBM tools for diagnosis are beginning to make a dent in clinical practice. Healthcare Information and Management Systems Society found that 64% of the surveyed US physicians was using personal digital assistants (PDAs) and their facilities, and 9% was using them to download clinical laboratory data[12]. Ebell was a pioneer in this field and, almost 10 years ago, reported the use of a palmar PC providing practical information for the general practitioner (e.g. data about drugs and clinical cases, tools and algorithms for diagnosis including a Bayesian calculator and 420 systemic reviews)[13]. Straus e Sackett, two gurus of EBLM, commented this article reporting the use of handheld computers by a General Medicine clinical team[14]. They concluded that clinicians could use evidence when it was readily available in the place of their practice, "but we still have a way to go in getting it to them quickly enough".

It must be pointed out, in this regard, that the technology evolution makes quickly obsolescent many electronic devices and the comparability issues between operating systems and softwares are particularly relevant. For example, the software discussed in the paper by Ebell et al. is no more available[13].

The explosive diffusion of the web helps in facing the technological innovation and the simplest languages can be often graphically interpreted by different browsers. However, even a simple calculator can more easily handle a text than to run a program; this explains the success encountered by the programming languages. The simplest applications of JAVA scripts do not require plug-ins; the programs can run with any browsers and any device.

Software description

The Evidence Based Laboratory Medicine group of the Italian Society of Laboratory Medicine prepared a Bayesan calculator. The main goals of the software were:

1. to provide a Bayesian calculator that could run in most, including palmar, devices with current browsers, not requiring dedicated softwares;
2. to provide an Italian and English version of the calculator;
3. to be of limited size in order to facilitate the web and local use even with small devices
4. to provide clear and simple post-test probability information;
5. to provide the results with the confidence intervals even for the parameters such as LR not commonly provided by the others calculators available in the web;
6. to make it available in the web through the portal of the Italian Society of Laboratory Medicine (http://www.simel.it /calcolatoreBayesiano.asp)

The program has been written in JAVA by one of the authors (GP); the user can enter the data in a single page and can change the language from Italian to English and viceversa without scrolling or changing the page.

The raw data are entered into a classical 2x2 table and 90, 95 (the default one) and 99% confidence intervals can be chosen. The software calculates: Sensitivity, Specificity, LR+, LR-, PPV, NPV, Efficiency, Pre-test probability (prevalence), Pre-Test odds, Post-test odds, Post-test probability, Number Needed to Diagnose.

EBM Tools for Diagnostic Tests, Ver 1.2		TARGET DISORDER		
		Present	Absent	TOT.
TEST Result	Positive	35	57	92.00
	Negative	5	29	34.00
○90% ⦿95% ○99%	TOT.	40.00	86.00	126.00

Calc | Clear | ○Italiano | ⦿English

		Confidence interval 95%	
Sensitivity (%)	87.500	73.888	94.541
Specificity (%)	33.721	24.616	44.218
Likelihood Ratio for a positive test	1.320	1.091	1.598
Likelihood Ratio for a negative test	0.371	0.155	0.886
Positive Predictive Value (%)	38.043	28.791	48.254
Negative Predictive Value (%)	85.294	69.872	93.551
Efficiency (%)	50.794	42.171	59.370
Pre test Probability (prevalence) (%)	31.746	24.261	40.311
Pre Test odds	0.465	0.320	0.675
Post Test odds	0.614	0.349	1.079
Post test Probability (%)	38.043	25.893	51.903
Number Needed to Diagnose (NND)	4.712	-66.844	2.580

Figure 1. The information obtained with the reported calculator using the BNP concentration cut-off of 17.9 ng/L[16].

EBM Tools for Diagnostic Tests, Ver 1.2		TARGET DISORDER		
		Present	Absent	TOT.
TEST Result	Positive	26	11	37.00
	Negative	14	75	89.00
○90% ⦿95% ○99%	TOT.	40.00	86.00	126.00

Calc | Clear | ○Italiano | ⦿English

		Confidence interval 95%	
Sensitivity (%)	65.000	49.506	77.866
Specificity (%)	87.209	78.531	92.705
Likelihood Ratio for a positive test	5.082	2.798	9.231
Likelihood Ratio for a negative test	0.401	0.261	0.617
Positive Predictive Value (%)	70.270	54.217	82.511
Negative Predictive Value (%)	84.270	75.311	90.392
Efficiency (%)	80.159	72.349	86.184
Pre test Probability (prevalence) (%)	31.746	24.261	40.311
Pre Test odds	0.465	0.320	0.675
Post Test odds	2.364	0.896	6.234
Post test Probability (%)	70.270	47.262	86.177
Number Needed to Diagnose (NND)	1.915	3.567	1.417

Figure 2. The information obtained with the reported calculator using the BNP concentration cut-off of 76.0 ng/L[16].

The confidence interval for Sensitivity, Specificity, PPV, NPV, Efficiency, Pre-test probability (prevalence), has been calculated using the method proposed by Wilson[15], which is more reliable compared to the traditional one based on the standard approach of taking a multiple of the standard error either side of the estimated quantity. The confidence interval for the LR has been calculated after logarithmic transformation.

The "confidence" for Pre-Test odds, Post-test odds, Post-test probability, Number needed to diagnose, have been indirectly calculated using the confidence limits of the parameters that originated them.

Example

In order to clarify the possible use of the proposed calculator we employed it to assess the diagnostic performance of BNP assay reported in a classical paper[16]. The authors demonstrated that properly choosing the cut-off of BNP assay in patient with suspected heart failure improves a lot the diagnostic power of the test. Really, Figure 1 shows that BNP concentration cut-off of 17.9 ng/L yields a very poor performance; for example the Post test odds is 0.614 (95% confidence interval: 0.349-1.079) and the number needed to diagnose is 4.72 (i.e. almost five assays are needed for a single diagnosis) making the test practically useless. On the contrary, adopting a different cut-off (76 ng/L) (Fig. 2) the performance of the assay improves; Post test odds rise to 2.36 (95% confidence interval: 0.9-6.23) and the number needed to diagnose lowers to 1.91 (i.e. only two assays are needed for a diagnosis).

Conclusions

EBLM will become really essential in the modern laboratory medicine practice only if its methods and tools will be translated into the daily practice and the "interaction" with the clinicians. Concepts such as Sensitivity, Specificity, reference intervals, decision limits, cut-off must be coupled with much more informative and clinical-oriented concepts such as the likelihood ratios that essentially convert the pre-test to the post-test probability. The Fagan's nomogram can be really helpful in this regard and has had an enduring popularity, even if it has a limited accuracy, although generally sufficient for bedside calculations, and usually shows pre-test and post-test probability spectra inadequate for some screening test.[17] Information Technology can assist in this field since most pathologists and physicians daily use desktop or laptop computer and handheld devices and could easily carry out the calculations and obtain the graphics needed for the EBLM[18]. We are convinced that EBLM did not find its momentum until now for the lack of proper tools for applying EBLM theory to daily practice[19]. We agree with Buckminster Fuller when he says "If you want to teach people a new way of thinking, don't bother trying to teach them. Instead give them a tool, the use of which will lead to new ways of thinking".

References

1. The Lewin Group. The Value of Diagnostics Innovation, Adoption and Diffusion into Health Care. Washington: Advamed 2005.
2. Institute of Medicine. Crossing The Quality Chasm. Washington; National Academy Press, 2001.
3. Graber AM, Mulrow C, Sox HC. Articles on diagnostic tests: a new series and an invitation to authors. Ann Intern Med 2002;137: 288-9.
4. Straus SE. Reporting diagnostic tests. BMJ 2003; 326:3-4.
5. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PS, Irwig LM et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative Clin Chem 2003; 49:1-6.
6. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM et al.The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. Clin Chem 2003; 49:7-18.
7. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PS, Irwig LM et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. Ann Intern Med 2003; 138:40-4.
8. Bossuyt PM, Reitsma JB, The STARD initiative. Lancet 2003; 361:71.
9. Mulherin SA, Miller WC. Spectrum bias or spectrum effect? Subgroup variation in diagnostic test evaluation. Ann Intern Med. 2002;137: 598-602.
10. Fagan TJ. Nomogram for Bayes theorem N Engl J Med 1975; 293:257.
11. Altman DG, Machin D, Bryant TN, Gardner MJ. Statistics with confidence. BMJ books, 2000.
12. Anonymous. Physician use of IT in the rise. Clin Lab News 2002; 28:1.
13. Ebell MH, Messimer SR, Barry HC. Putting computer-based evidence in the hands of clinicians. JAMA 1999; 281:1171-2.
14. Straus SE, Sackett DL. In response: Putting computer-based evidence in the hands of clinicians. JAMA 1999; 281:1171-2.
15. Wilson EB. Probable inference, the law of succession, and statistical inference. J Am Stat Assoc 1927; 22:209-12.
16. Sackett DL, Haynes RB. The architecture of diagnostic tests. In: Knottnerus JA, ed. London: BMJ Books, 2002.
17. Christenson RH. Evidence-Based Laboratory Medicine- a guide for critical evaluation of in vitro laboratory testing. Ann Clin Biochem 2007; 44: 111-130.
18. Glasziou P. Which methods for bedside Bayes? EBM 2001; 6: 164-5.
19. Price C. Application of the principles of evidence based-medicine to laboratory medicine. Clin Chim Acta 2003; 333: 147-54.