

Medicina Basada en Evidencias

[Dr. Jaime Mañalich M., MSc](#)
[Departamento Medicina Interna, Clínica Las Condes](#)

La Medicina Basada en Evidencias (EBM) tiene ya 10 años de existencia. Se desarrolla a partir de los esfuerzos por tener pautas coherentes, reproducibles y transmisibles de "Evaluación Crítica de la Literatura Biomédica" en la Universidad de McMaster en Canadá, luego recogidos por el American College of Physicians formalmente a través del ACP Journal Club. Más tarde, se han desarrollado importantes grupos de trabajo y publicaciones especializadas en diversas universidades (1).

La disciplina desde donde nace la EBM es la Epidemiología Clínica, que es la aplicación de las herramientas propias de la Salud Pública al análisis de la práctica médica, en la que los pacientes son estudiados como un colectivo. De esta manera, el desarrollo de la EBM se fundamenta en la estadística, la aplicación de programas de computación, y el aporte de disciplinas como la sociología y psicología al estudio de las conductas de médicos y pacientes. La EBM nace además en un momento en que el uso de Internet, el acceso remoto a bibliotecas como la National Library of Medicine en Washington desde cualquier computador personal y sin costo, las publicaciones on-line y ciertos cambios en la forma en que se editan y publican artículos médicos* contribuyen sustancialmente al logro de sus objetivos.

** Uno de los mayores aportes de la EBM ha sido el que la mayoría de las revistas médicas adoptara el uso de "abstracts estructurados" como una forma de resumir los trabajos que publican.*

Hoy, EBM se enseña a nivel de pregrado en la mayoría de las Universidades de Norteamérica, existen programas de computación y sitios Web para ayudar a los usuarios. Se han escrito un sinnúmero de libros de texto sobre el tema y las autoridades en la materia se han erigido en los nuevos jueces que buscan orientar a la práctica de una Medicina efectiva, basada en datos, que disminuya la variabilidad en la conducta de los clínicos y que sea responsable ante la sociedad. Así, la EBM ha abierto una nueva etapa en el desarrollo de la Medicina, comparable en su influencia a aquella época en que la correlación anatómo-patológica gobernaba el espíritu de la enseñanza y la práctica o al de la fuerte influencia de la Fisiología Médica de los años cincuenta y sesenta.

La EBM hace un diagnóstico y se propone a sí misma como el tratamiento. El diagnóstico se puede sumarizar en:

1. La cantidad de información destinada a influir en la práctica médica se multiplica exponencialmente en el tiempo.
2. Es imposible hoy para un clínico estar al día en los temas a los que se dedica.
3. La práctica médica está determinada por tradiciones, estilos e informaciones

que muchas veces no son consistentes con el conocimiento actual.

4. La práctica médica suele ser ineficiente, desperdiciando recursos escasos que la comunidad ha puesto en manos de los clínicos.
5. El puro sustento anatómico, biológico o fisiológico no basta para la práctica médica.

Uno de los méritos de la aproximación sistemática a la literatura médica que ofrece la EBM reside en la jerarquización de la información. Así, se propone que para influir la práctica en temas relacionados con tratamiento, la información es más válida si proviene de ensayos randomizados doble ciego que si proviene de estudios fisiológicos. Un modelo de dicha priorización para trabajos relacionados con tratamiento se muestra en la Tabla 1 .

Tabla 1
Valoración de diferentes tipos de estudios sobre tratamiento
Valor decreciente 1-7

1. Ensayos Randomizados de N=1
2. Revisión sistemática de Ensayos Randomizados
3. Ensayo Randomizado individual
4. Revisión sistemática de Estudios Observacionales referentes a resultados importantes para los pacientes.
5. Estudio Observacional único referente a resultados importantes para los pacientes.
6. Estudios fisiológicos (Presión arterial, capacidad física, etc.)
7. Observaciones clínicas no sistematizadas.

Otra área de desarrollo significativo de la EBM ha sido el pre-proceso de la evidencia publicada (2). Como resulta difícil para un clínico en particular destinar tiempo suficiente para obtener todos los datos reportados respecto a un problema específico, ordenarlos según importancia y relevancia y entender rápidamente sus limitaciones metodológicas, se han generado diferentes fuentes de consulta para la búsqueda de información en la que estos procesos ya se han realizado por expertos. En el addendum están señaladas las principales fuentes que se pueden consultar para este propósito.

La EBM ha sido capaz de iluminar las diferentes etapas de la acción médica: la historia natural de la enfermedad, la prevención, el proceso de diagnóstico, la terapia, la toma de decisiones en Medicina, la diseminación de información científica, las guía o pautas clínicas, e incluso la docencia (3). Su gran mérito reside precisamente en esta iluminación, en hacer evidente los problemas y desafíos de cada uno de estos procesos, y de sistematizarlos en una manera no sospechada previamente. Sin embargo, las opiniones respecto a cuál ha sido el impacto final de la EBM en cambiar la práctica clínica en realidad son más discutibles. Quizás se requiera de un recambio generacional completo para que estos resultados sean evidentes. Por ahora, la EBM no ha simplificado la vida de quienes están en la práctica médica y es discutible el beneficio que ha representado para el paciente concreto.

Cabe señalar también algunas precauciones. La EBM no puede sustituir una enseñanza médica en que se desarrollen plenamente las capacidades clínicas del sujeto. Para

revisar la literatura y entender su aplicación a un paciente determinado hay que primero entender qué le pasa a ese paciente, cuál es su diagnóstico, cómo acepta su enfermedad. Por ello, hay que entender también que la práctica médica no requiere solo evidencia confiable para determinar cuál es el tratamiento adecuado en una determinada situación. La práctica médica se basa en valores. Lo que la "evidencia" indica como el tratamiento más adecuado para una condición, puede no ser apropiado para el paciente que está al frente. En Medicina, se ha tendido a denominar este conflicto como "momento ético". Es decir, cuando una acción determinada por la evidencia no es la más adecuada para una situación específica, y el médico debe hacer una recomendación calificada. Imaginemos por un momento la indicación de una Quimioterapia para el tratamiento de una leucemia aguda en un niño versus indicar el mismo tratamiento por idéntica condición en un anciano con un grave deterioro neurológico.

Por último, los instrumentos de la EBM son muchos más adecuados para el clínico experimentado que para el novato. Este último recurre habitualmente a una metodología de búsqueda no sistemática de la información. En cambio, el clínico con experiencia utiliza más frecuentemente una aproximación de "resolución de problemas", en la que un caso determinado lo obliga a informarse y a resolver una decisión bastante específica.

Addendum

ACPJournal Club	www.acponline.org/journals/acpjc/jcmenu.htm
Best Evidence	www.acpontine.org/catalog/electronic/best_evidence.htm
Cochrane Library	www.update-software.com/cochrane/cochrane-frame.html
UpToDate	www.uptodate.com
MEDLINE PubMed	www.ncbi.nlm.nih.gov/PubMed
Internet Grateful Med	igm.nlm.nih.gov
Other sources	www.medmatrix.org/info/medlinetable.asp
Scientific American Medicine	www.samed.com
Clinical Evidence	www.evidence.org
Harrison`s Online	www.harrisonsonline.com
emedicine	www.emedicine.com
Medscape	www.medscape.com/Home/Topics/homepages.html
Medical Matrix	www.medmatrix.org/index.asp
SchHARR Netting the Evidence	www.shef.ac.uk/~scharr/ir/netting/

Medical Word Search	www.mwsearch.com
Journal listings	www.nthames-health.tpmde.ac.uk/connect/journals.htm www.psigroup.com/dg/medjournals.htm
Clinical practice guidelines	www.guidelines.gov www.cma.ca/cpgs
MD Consult	www.mdconsult.com
Evidence-based Medicine Reviews (OVID)	www.ovid.com/products/clinical/ebmr.cfm (available through many medical libraries)

EBM aplicada a problemas de diagnóstico

Establecer un diagnóstico es básicamente descubrir una verdad que está oculta.

Para ello se recurre en primer lugar a métodos clínicos como son conocer el motivo de consulta, establecer una historia de los problemas recientes, el contexto familiar, social y laboral en que estos síntomas se desarrollan y ordenar procedimientos adicionales como exámenes de laboratorio, radiología, anatomía patológica y otros que ayuden a este proceso de establecer la verdad. Los clínicos han aprendido a formar en su mente en forma relativamente rápida ciertas hipótesis de trabajo para avanzar en este proceso (5). A pesar que la enseñanza clásica homologa estas a "hipótesis diagnósticas", no es exacto que sean diagnósticos lo que se utiliza como medios intermedios de trabajo. En algunos casos son "problemas"* y más auténticamente, agrupaciones o clusters que pueden referirse conjuntamente a signos, síntomas, tales como "hematuria silenciosa", "ictericia y fiebre", "retardo de crecimiento intrauterino", etc (6).

** Revísese por ejemplo toda la literatura referente a la "ficha médica orientada por problemas".*

Normalmente se utilizan simultáneamente más de un solo conjunto frente a un caso específico, y lo que ocurre en la mente del clínico es que toda la información adicional es contrastada contra ellos, en un pensamiento intuitivamente probabilístico. Si bien es cierto que un lenguaje matemático de probabilidades no es habitual en este razonamiento, la comunicación hace patente el uso de expresiones que conllevan esta forma de pensar (7). Frases como:

- *lo más probable es que este paciente tenga un abdomen agudo,*
- *yo descartaría una apendicitis,*
- *una posibilidad es un embarazo ectópico,*

reflejan una mezcla de incertidumbre y aproximación probabilística a los problemas que es habitual escuchar en reuniones clínicas o rondas médicas.

En esta forma de razonar, el pensamiento se concentra luego en definir cuanto más probable es que determinadas hipótesis diagnósticas expliquen lo que ocurre con el paciente (8). Así, la información adicional no hace sino desplazar esta probabilidad de diagnóstico hacia la certeza o descartarla (probabilidades de $p = 1$ y $p = 0$ respectivamente). De esta manera, la información adicional opera sobre un ámbito condicionado. Esto es, las nuevas probabilidades están condicionadas por la probabilidad que se tiene en mente al recurrir a la información adicional que se solicita. Imaginemos por ejemplo, la diferente importancia que tiene en la secuencia diagnóstica

de una cefalea con carácter de migraña en una mujer que refiere tener el síntoma 3 o 4 veces al mes desde hace 10 años, a un hombre de 55 años que relata nunca haber sufrido previamente de cefalea alguna. En este sentido, la información que da una test o examen diagnóstico, por ejemplo, tiene un poder diferente de modificar la probabilidad antes de aplicar dicho test (*p pre-test*) en relación a la que se obtiene al conocer su resultados (*p post-test*). Es decir, el contexto probabilístico en el que se utiliza una determinada herramienta de diagnóstico determina su utilidad. Esto se describe en la Figura 1, en que se tiene la probabilidad de enfermedad determinada por los medios clínicos y arbitrariamente fijada a manera de ejemplo en 0,33 o 33%, y que producto del examen diagnóstico se desplaza a aproximadamente 0,75. En este sentido, la manera convencional de analizar los exámenes de diagnóstico, en base a su Sensibilidad y Especificidad* se ve menos útil que aplicar los conceptos de *likelihood ratio* o razón de probabilidad (9). Esta razón para un test cuyo resultado es positivo o negativo, permite apreciar con mucha mayor propiedad el poder del resultado para desplazar la probabilidad pre-test.

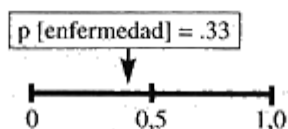
* *S*= probabilidad de un examen (+) o alterado en quienes padecen la enfermedad y *E*=probabilidad de un test (-) o un resultado normal en quienes no la padecen o están sanos.

La razón de probabilidad o LR debe leerse como *cuánto mayor es la probabilidad de un resultado positivo entre aquellos que sufren la condición o están enfermos en comparación a un resultado positivo entre los sanos, y cuánto mayor es la probabilidad de un resultado negativo entre los sanos en comparación a los enfermos* (Figura 2).

TEST DIAGNÓSTICOS

Impacto en cambiar la probabilidad de enfermedad

a) Apreciación clínica



b) Uso de un examen diagnóstico

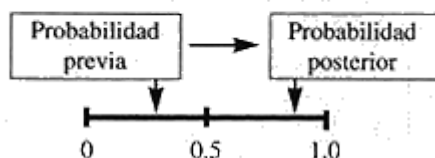


Figura 1

		ENFERMEDAD		
		Presente	Ausente	
T E S T	+	a	b	Razón de Probabilidad para un test (+) $\frac{\text{Sensibilidad}}{1 - \text{Especificidad}}$ $= \frac{[a/(a+c)]}{[b/(b+d)]}$
	-	c	d	
Razón de probabilidad para un test (-)		$\frac{1 - \text{Sensibilidad}}{\text{Especificidad}} = \frac{[c/(a+c)]}{[d/(b+d)]}$		

Figura 2

En este contexto de probabilidad, la información diagnóstica actúa de una manera tal que permite desplazarse **cruzando umbrales**. Estos umbrales son de izquierda a derecha, o desde una menor a una mayor probabilidad, el Umbral de un examen adicional (U Test) y el Umbral del tratamiento (U Tratamiento) (Figura 3). Esto puede ejemplificarse de la siguiente manera (10):

Se presenta al Servicio de Urgencia una paciente de 23 años, con un dolor de 12 horas de evolución en la fosa ilíaca derecha. Dice que su última menstruación fue hace 5 semanas. La t° es de $37,5^{\circ}$ axilar y $38,5^{\circ}$ rectal. El pulso 98 por min. Está en buen estado general. Es evidente que la probabilidad pre-test en esta enferma para dos diagnósticos que compiten entre sí, como son el de una apendicitis aguda y el de un embarazo ectópico es lo suficientemente alta como para no enviarla a su domicilio y solicitar información adicional.

Se practica un recuento de blancos, que se informan en 18.000, un hematocrito de 44%, una fracción beta en límites normales.

Es probable que con esta información, el médico de urgencia decida atravesar [U tratamiento], e intervenirla por una apendicitis aguda. Nótese que la probabilidad no es de 1,0 cuando se toma esta decisión.

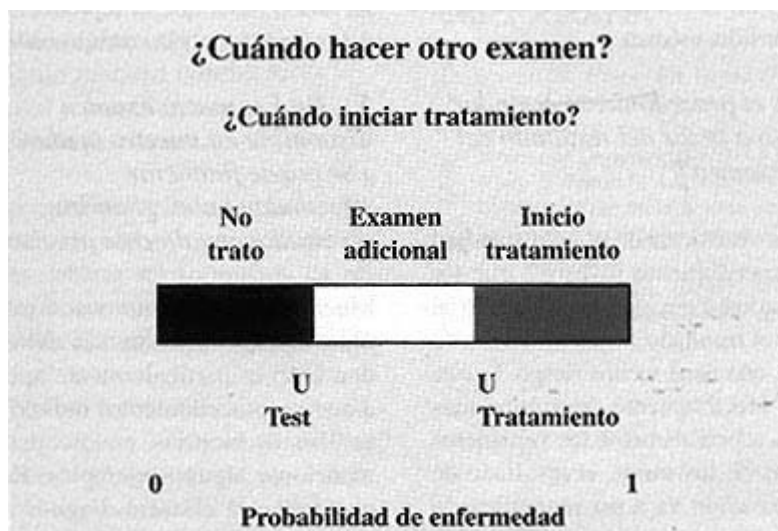


Figura 3

Evidentemente, estos umbrales varían de acuerdo a criterios como: cuál es el riesgo o costo del examen o la terapia que debe hacerse, cuál es la eficacia del examen para mover la probabilidad pre-test, cuál es el riesgo de no establecer el diagnóstico rápidamente, etc. Así, para iniciar una Quimioterapia, nadie estará dispuesto a hacerlo a menos que la probabilidad de la enfermedad en cuestión sea muy cercana a 1,0. Por otro lado, en el caso de un paciente que se presenta con síntomas propios de un resfrío común, se requerirá que entre en competencia con una probabilidad elevada el diagnóstico diferencial con una neumopatía antes de hacer una radiografía de tórax.

Se ha sistematizado el impacto de estos LR hasta el punto de publicar nomogramas que permiten determinar la variación desde la [p pre-test] a la [p posttest] simplemente fijando con una regla la primera, haciéndola pasar por el LR correspondiente y determinando de esta manera la segunda (11) (Figura 4).

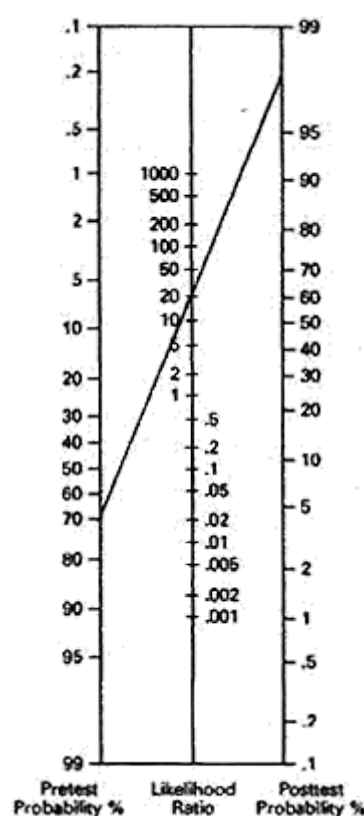


Figura 4

Pautas para evaluación de literatura sobre exámenes de diagnóstico

Con el avance de la tecnología, se publican todos los años cientos de artículos referentes a nuevas aproximaciones diagnósticas a enfermedades. Sin embargo, la validez de estos estudios debe ser analizada cuidadosamente antes de generar inversiones, introducir cambios en la práctica habitual o reemplazar procedimientos y estrategias que están adecuadamente validados.

La escuela de EBM propone los siguientes criterios, en la forma de preguntas, para realizar esta evaluación.

1. ¿Hubo una comparación independiente y ciega, con una referencia estándar o "gold standard" de diagnóstico?

Los elementos claves que se deben cumplir para responder afirmativamente a esta pregunta son:

- Los pacientes fueron sometidos a ambos medios de diagnóstico, el convencional y el nuevo método que se está probando.
- Hubo una evaluación independiente de cada resultado. Esto es, el resultado de un test no era influido por el conocimiento del resultado del otro (12).

2. ¿Fue el test diagnóstico evaluado en un contexto apropiado de pacientes?

Esta pregunta apunta a si los pacientes evaluados se asemejan a los de la práctica donde se pretende aplicar el examen. Así por ejemplo, si el examen fue evaluado en un segmento de pacientes seleccionado, como podría ser el de un hospital de referencia para una determinada patología, su generalización a la población general será difícil. Recuérdese lo que se comentaba de los LR como la capacidad que tenían los exámenes para variar la [p pretest]. Si la prevalencia o [p pre-test] de una enfermedad

es de 0,1/1000 en la población general, y de 100/1000 en la clínica especializada, el comportamiento del examen será muy diferente en ambos escenarios (13). Un ejemplo apropiado es el la alfa-feto proteína en Cáncer de Colon. Un examen apropiado para el seguimiento de los pacientes con la enfermedad, que predice recurrencias, no tiene utilidad como screening de Cáncer de Colon en la población general (14). Por último, se puede mencionar el uso de la Proteína C reactiva (PCR) para determinar la naturaleza bacteriana de un supuesto cuadro infeccioso. Esto, que resulta apropiado en el contexto de pacientes de alto riesgo de sufrir un cuadro séptico, como es una Unidad de Tratamiento Intensivo, pierde validez al utilizarlo en la población que concurre a una consulta médica.

3. ¿Fue el procedimiento estándar aplicado a pesar del resultado del nuevo examen?

Ocurre a veces cuando el gold standard es un procedimiento invasivo, que los investigadores deciden no utilizarlo al conocer el resultado negativo del nuevo examen, que tiene menos riesgo. Si este examen efectivamente discrimina más o menos adecuadamente los verdaderos enfermos de los sanos, el resultado de la comparación va a ser magnificar la concordancia entre la aproximación convencional y el examen que se está probando. Ejemplos recientes son el uso de Resonancia Magnética para el diagnóstico de enfermedad coronaria. Será muy difícil someter a un paciente asintomático a una Coronariografía si tiene una RM negativa para calcificaciones coronarias. Complementariamente, al hacer angiografías en aquellos pacientes que si tienen estas calcificaciones, la evaluación del método no invasivo mejorará.

4. ¿Fue el procedimiento diagnóstico validado en una segunda serie, independiente de la primera, de sujetos?

Voluntariamente o involuntariamente, los investigadores tienen sesgos de derivación de pacientes para los estudios que pueden limitar la generalización de los resultados (15). No es sorprendente ver en la literatura médica que luego de un entusiasmo inicial con un nuevo procedimiento, los resultados son mucho más pobres en estudios sucesivos, e incluso, una iniciativa promisoriosa queda abandonada. Por ello, la aproximación más apropiada es que desde el inicio se someta el nuevo examen a un segundo "test", que permita sacar conclusiones sobre la generalización. Si ello no existe, se debe ser conservador, hasta que la reproducibilidad del método haya sido comprobada.

5. ¿Está el nuevo examen disponible en nuestro medio? ¿Se puede financiar adecuadamente? ¿Son los resultados igualmente precisos?

Muchos procedimientos son extraordinariamente dependientes del operador. Esto es particularmente aplicable a nuevos procedimientos radiológicos, análisis de biopsias, endoscopias, por mencionar algunos ejemplos. Resulta obvio que la eficacia diagnóstica de un operador entrenado durante años, con miles de casos a su haber, y dedicado exclusivamente al procedimiento en análisis, va a mostrar resultados muy distintos al de un novicio. En otro orden, el tema del financiamiento es también crucial. La adquisición de algunas tecnologías implica cuando el volumen de enfermos analizados es pequeño, un costo por caso prohibitivo, que hace la aplicación del procedimiento sencillamente imposible.

6. ¿Podemos generar una estimación de las probabilidades pre-test de nuestros pacientes?

El énfasis de este punto ya se ha señalado. Sin embargo, su aplicación es difícil. La definición de [p pretest] puede basarse en el recuerdo, Idealmente documentado, de la propia experiencia del Centro. Se puede recurrir también a la información epidemiológica local, sea este regional o nacional, publicado por la autoridad sanitaria. Tercero, se pueden rescatar los datos de una población distante; pero que tenga en el atributo en estudio, características similares a la propia. Se puede utilizar también la [p pretest] del estudio analizado si hay convencimiento que el muestreo fue adecuado y que los riesgos de generalización son bajos.

7. ¿Los resultados post test afectarán nuestra conducta habitual y beneficiarán

a nuestros pacientes?

Esta pregunta es quizás la más importante en término de generalización. Quizás debería ser formulada en primer lugar, como un condicional: si los resultados fueran válidos y aplicables en nuestro medio, ello cambiará la conducta de los clínicos y beneficiará a los pacientes? Con el gigantesco avance de la tecnología, hoy se asiste no pocas veces a procedimientos diagnósticos "en busca de una enfermedad". En la evaluación de la literatura se debe tener en cuenta la presión de la industria por incorporar nuevas técnicas.

Si un escáner de abdomen permite evitar las antiguas laparotomías exploradoras, o si un eco-doppler obstétrico permite evitar una amniocentesis, la aceptación será rápida y la conducta cambiará acorde. Pero existen innumerables ejemplos de procedimientos caros o de riesgo que no aportan información útil alguna. Afortunadamente, ellos son defendidos por pocos, y lo más probable es que luego se olviden.

Bibliografía

1. Sackett DL, Haynes RB, Guyatt GH, Tugwell P *Clinical Epidemiology, A Basic Science for Clinical Medicine*. 2nd ed. Boston: Little, Brown and Company; 1991: 53-57
2. Irwig L, Tosteson AN, Gatsonis C, et al. *Guidelines, for meta-analyses evaluating diagnostic tests*. *Ann Intern Med*. 1994; 120: 667-676.
3. Sox HC, Blatt MA, Higgins MC, Marton KI *Medical Decision Making*. Boston: Butterworths; 1988.
4. Schnridt HG, Norman GR, Boshuizen HP. *A cognitive perspective on medical expertise: theory and implication*. *Acad Med*. 1990; 65: 611-621.
5. Bordage G. *Elaborated knowledge: a key to successful diagnostic thinking*. *Acad Med* 1994; 69: 883-885.
6. Regehr G, Norman GR. *Issues in cognitive psychology: implications for professional education*. *Acad Med*. 1996; 71: 988-1001.
7. Glass RD. *Diagnosis: A Brief Introduction*. Melbourne: Oxford University, Press; 1996.
8. Barondess JA, Carpenter CCJ, eds. *Differential Diagnosis*. Philadelphia: Lea & Febiger; 1994.
9. Choi BC. *Sensitivity and specificity of a single diagnostic test in the presence of work-up bias*. *J Clin Epidemiol*. 1992; 45: 581-586.
10. Pauker SG, Kassirer JP. *The threshold approach to clinical decision-making*. *N Engl J Med*. 1980; 302: 1109-1117.
11. Fagan TJ. *Nomogram for Bayes's theorem*. *N Engl J Med*. 1975; 293:257
12. Lijmer JG, Mol BW, Heisterkamp S, et al. *Empirical evidence of design-related bias in studies of diagnostic tests*. *JAMA* 1999; 282: 1061-1066.
13. Sox HC, Hickam DH, Marton KI, et al. *Using the patient's history to estimate the probability of coronary artery disease: a comparison of primary care and referral practices*. *Am J Med*. 1990; 89: 7-14.
14. Bates SE. *Clinical applications of serum tumor markers*. *Ann Intern Med*. 1991; 115: 623-638.

15. Ransohoff DF, Feinstein AR. Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. *N Engl J Med.* 1978; 299: 926-930.