



REVISIÓN DE TEMA

Likelihood ratio (razón de verosimilitud): definición y aplicación en Radiología



C. Silva Fuente-Alba* y M. Molina Villagra

Departamento de Imágenes, Clínica Alemana, Facultad de Medicina, Universidad del Desarrollo, Santiago de Chile, Chile

Recibido el 12 de abril de 2016; aceptado el 13 de noviembre de 2016

Disponible en Internet el 28 de diciembre de 2016

PALABRAS CLAVE

Prueba diagnóstica;
Rutina;
Razón de
verosimilitud;
Sensibilidad y
especificidad

KEYWORDS

Diagnostic tests;
Routine;
Likelihood ratio;
Sensitivity
and specificity

Resumen Los parámetros de medición de validez más conocidos en las pruebas diagnósticas son la sensibilidad, la especificidad y los valores predictivos (positivo y negativo). Las razones de verosimilitud o *likelihood ratio* conforman otra herramienta que resume la precisión de una prueba diagnóstica, permitiendo un uso clínico más intuitivo.

© 2016 Publicado por Elsevier España, S.L.U. en nombre de Sociedad Argentina de Radiología. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Likelihood ratios: Definition and uses in Radiology

Abstract The properties of a diagnostic test such as sensitivity, specificity and predictive values (positive and negative) are well known. Likelihood ratios are an alternative for summarising diagnostic accuracy, yet providing a more powerful approach to clinical interpretation and decision-making.

© 2016 Published by Elsevier España, S.L.U. on behalf of Sociedad Argentina de Radiología. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introducción

Las pruebas diagnósticas son herramientas de apoyo ampliamente utilizadas en la práctica clínica para sustentar

hipótesis diagnósticas. La diversa gama de exámenes imagenológicos disponibles nos obliga a analizar críticamente las propiedades de cada uno con el fin de escoger el más óptimo para el paciente en términos de seguridad, eficacia y costo-efectividad. Todo radiólogo debe conocer y dominar los distintos aspectos de una prueba diagnóstica, de forma de poder apoyar la toma de decisiones clínicas.

Toda hipótesis diagnóstica que se origina de un proceso de análisis clínico lleva asociada una probabilidad de ocurrencia

* Autor para correspondencia.

Correos electrónicos: csilvafa@alemana.cl,
drsilvafa@gmail.com (C. Silva Fuente-Alba).

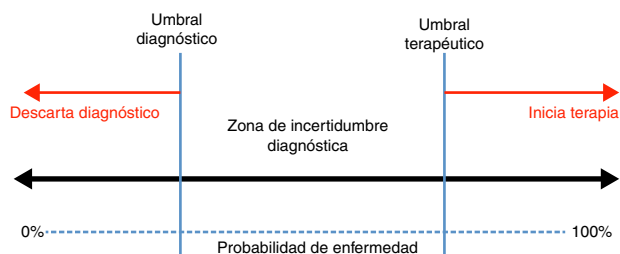


Figura 1 Umbrales de decisión presentes en el proceso diagnóstico.

denominada “probabilidad *pretest*”. Esta se logra a partir de un conjunto de elementos anamnésticos y del examen físico, que, junto con la experiencia, brinda una aproximación a una entidad diagnóstica. Si a juicio del médico esta probabilidad no supera el primer umbral, llamado “umbral diagnóstico”, se rechaza la hipótesis inicialmente planteada y se descarta el diagnóstico (pudiendo ser por elementos clínicos que hacen poco probable una patología en particular). Por el contrario, si su juicio le indica que la probabilidad *pretest* supera tanto el umbral de estudio adicional como un segundo umbral, denominado “umbral terapéutico”, el médico avala la hipótesis diagnóstica sin requerir otra confirmación (mediante imágenes, por ejemplo), dando paso a la indicación de la terapia pertinente^{1,2}.

Existe un tercer escenario, situado entre el umbral diagnóstico y el terapéutico, en el que la probabilidad *pretest* impide aceptar o rechazar con certeza la hipótesis diagnóstica, estableciendo una zona de incertidumbre (fig. 1). Es en este escenario probabilístico que las pruebas diagnósticas presentan su mayor utilidad.

La ventaja de un examen diagnóstico radica fundamentalmente en poder modificar la probabilidad *pretest*, permitiendo pasar sobre el umbral terapéutico o por debajo del umbral de un estudio adicional para alejar la zona de incertidumbre.

Parámetros generales de prueba diagnóstica

Los términos de sensibilidad, especificidad y valores predictivos son los más utilizados para referirse al análisis de las propiedades de un test diagnóstico¹⁻³.

Los más conocidos son la sensibilidad y la especificidad, dos parámetros utilizados rutinariamente en la valoración de toda prueba diagnóstica. La sensibilidad de un test diagnóstico corresponde a la proporción de individuos enfermos, identificados como tal por el examen; mientras que la especificidad corresponde a la proporción de individuos sin la patología, registrados así por el referente. Ambas cuantificaciones tienen la fortaleza de ser propiedades inherentes a la prueba, independientemente de la prevalencia de la afección en la población estudiada. Además, entregan una aproximación con respecto a la capacidad de la técnica para identificar una entidad en particular.

No obstante, hay que considerar que estos parámetros están dados por el análisis de probabilidades reversas. Es decir, aquella posibilidad de tener un test positivo si el paciente efectivamente está enfermo, o de tener un test negativo en la eventualidad de que la persona esté sana.

Esto limita la posibilidad de hacer interpolaciones clínicas intuitivas para cada paciente³.

A su vez, los valores predictivos buscan responder la inquietud sobre la probabilidad de que la persona efectivamente esté enferma ante un test positivo. Sin embargo, sus resultados se encuentran altamente influenciados por la prevalencia de la enfermedad en el estudio que se analiza¹⁻³. Así, mientras mayor sea la prevalencia de la enfermedad, mayor es el valor predictivo positivo (VPP) y menor es el negativo (VPN) del test diagnóstico. Por ello, se hace necesario disponer de aproximaciones estadísticas alternativas para resumir la fortaleza diagnosticada de una prueba en particular y que arroje una utilidad clínica determinada.

Con el fin de dimensionar el beneficio clínico de un test diagnóstico en forma independiente de la prevalencia resulta de mayor utilidad la aplicación de los *likelihood ratio* (LR).

Likelihood ratio o razón de verosimilitud

Definición

Conocido también en español como razón de verosimilitud, el *likelihood ratio* (LR) se define como la razón entre la posibilidad de observar un resultado en los pacientes con la enfermedad en cuestión *versus* la posibilidad de ese resultado en pacientes sin la patología⁴.

El uso del LR constituye una herramienta de gran utilidad para la toma de decisiones clínicas frente a la solicitud de algún test diagnóstico, porque son valores inherentes a este e independientes de la prevalencia de la enfermedad. Si bien su cálculo deriva de probabilidades condicionadas en base al teorema de Bayes^{4,5}, se puede estimar en base a parámetros de sensibilidad y especificidad de la siguiente manera:

$$LR(+)=\frac{\text{Tasa de verdaderos positivos}}{\text{Tasas de falsos positivos}}=\frac{\text{Sensibilidad}}{1-\text{Especificidad}}$$

$$LR(-)=\frac{\text{Tasa de falsos negativos}}{\text{Tasa de verdaderos negativos}}=\frac{1-\text{Sensibilidad}}{\text{Especificidad}}$$

Esto refleja la capacidad de un test diagnóstico para cambiar una probabilidad *pretest* a una nueva probabilidad *posttest*. La aplicabilidad del LR es múltiple en la práctica radiológica, ya que es posible utilizarlo al analizar pruebas diagnósticas con resultados dicotómicos, en los que solo es posible determinar presencia o ausencia de enfermedad (negativo o positivo), o bien con resultados categóricos, por ejemplo mediante exámenes que tienen categorías de severidad (leve, moderada o severa).

Esto se ha simplificado a través del uso del nomograma de Fagan que consta de tres columnas paralelas entre sí, graduadas de forma de poder realizar una estimación de la probabilidad *posttest*, conocida la probabilidad *pretest* y un LR del test a ser probado.

El impacto de los valores del LR positivo y negativo se refleja en la tabla 1. Como término general, hay que recordar que un LR positivo mayor de 10 y un LR negativo menor de 0,1 indica un cambio relevante en la probabilidad *pretest*, lo cual determina con alta certeza un cambio de conducta clínica^{2,4,5}.

Con la finalidad de ejemplificar mejor estos conceptos, se ha seleccionado un artículo de la literatura

Tabla 1 Rangos de valores de *likelihood ratio* y su impacto en utilidad clínica

LR positivo	LR negativo	Utilidad
10	<0,1	Altamente relevante
5-10	0,1-0,2	Buena
2-5	0,5-0,2	Regular
<2	> 0,5	Mala

radiológica, donde se analiza la presencia o ausencia (resultado dicotómico) de hemorragia gastrointestinal en pacientes estudiados con angiografía por tomografía computada.

Ejemplo

Entre el año 2008 y el 2010, Martí *et al.*⁶ evaluaron la viabilidad y el desempeño de la angiografía por tomografía computada (angio-TC) en pacientes con hemorragia gastrointestinal aguda. Para ello, desarrollaron un estudio prospectivo en el que seleccionaron 47 pacientes que habían sido referidos para una colonoscopia convencional desde el servicio de urgencias.

Cada paciente fue sometido a una angio-TC según los parámetros estipulados en los materiales y métodos, comparando este examen con un estándar de referencia determinado por los autores. Los resultados mostraron que, de un total de 19 individuos enfermos (definidos así por el estándar de referencia), la angio-TC identificó correctamente a todos. De este modo, la sensibilidad del método para el diagnóstico de la hemorragia digestiva fue de $19/19 = 1$ (100%). En el subgrupo de individuos sin enfermedad, de un total de 28 la angio-TC detectó correctamente a 27 de ellos, por lo que la especificidad fue de $27/28 = 0,96$ (96%).

Si se interpretan estos resultados desde una óptica diferente, 19 pacientes fueron correctamente diagnosticados como enfermos por la angio-TC (verdaderos positivos), mientras que ninguno fue incorrectamente diagnosticado como sano (falso negativo). Por otro lado, 27 individuos fueron adecuadamente definidos como sanos por la angio-TC (verdaderos positivos), mientras que un individuo fue mal diagnosticado como enfermo por el método (falso positivo). Así, el VPP fue de $19/20 = 0,95$ (95%), que representa la probabilidad de que un individuo tenga hemorragia digestiva dado que el resultado de la angio-TC es positivo. Por su parte, el VPN fue de $27/27 = 1$ (100%), una cifra que traduce la probabilidad de que un individuo no tenga hemorragia digestiva dado que el resultado de la angio-TC fue negativo. Estos resultados se resumen en la [tablas 2 y 3](#).

Tabla 2 Resultado de la angiografía por tomografía computada y el estándar de referencia

		Estándar de referencia		
Angiografía por tomografía computada		+	-	
	+	19	1	20
	-	0	27	27
		19	28	

Tabla 3 Parámetros de la angiografía por tomografía computada

Sensibilidad	1
Especificidad	0,96
Valor predictivo positivo	0,95
Valor predictivo negativo	1

La aplicabilidad práctica del *likelihood ratio* es modificar la probabilidad *pretest* que motivó la solicitud de esta prueba diagnóstica, tratando de abandonar la zona de incertidumbre. Si bien en el artículo publicado no se presentaron estos valores, disponemos de la información necesaria para su estimación. El cálculo de los LR en el estudio que utilizamos como ejemplo se puede realizar de la siguiente manera:

$$LR(+) = (\text{sensibilidad} / 1 - \text{especificidad}) = 1 / 1 - 0,96 = 25$$

$$LR(-) = (1 - \text{sensibilidad} / \text{especificidad}) = 1 - 1 / 0,96 = 0,04$$

En este caso, el LR (+) traduce que es 25 veces más verosímil que un paciente con hemorragia gastrointestinal obtenga una angio-TC positiva que un paciente sin hemorragia tenga el mismo resultado. Así, un *test* diagnóstico será de mayor utilidad en la medida en que su LR (+) sea de mayor magnitud, puesto que permite confirmar con certeza la presencia de enfermedad, y en que su LR (-) tenga un valor bajo, dado que descarta la enfermedad.

En el caso del estudio de Martí *et al.*⁶, el primer paso antes del cálculo de los LR, consistió en definir la probabilidad *pretest* clínica de los individuos a los que se les indicó una angiografía por tomografía computada. Esta se encontraba limitada al grupo con un cuadro de hematemesis, melena o hematoquezia, que fue evaluado por un médico y al que se le pidió este *test* diagnóstico, siendo mayor que la prevalencia en la población general o en aquella que consulta en un servicio de urgencia.

Para el efecto de este estudio, la prevalencia de la enfermedad en la población analizada (40%) se homologó con la probabilidad *pretest* y, junto con esto, fue posible obtener probabilidades después de la angio-TC (*posttest*) aplicando el nomograma de Fagan ([fig. 2](#)).

Los LR en el estudio de Martí *et al.*⁶, con una sospecha *pretest* del 40% en la situación de una angio-TC positiva para hemorragia gastrointestinal, aumentó a un 95% la posibilidad *posttest* de tener esta entidad. Por el contrario, en caso de *test* negativo, disminuye la posibilidad *posttest* a menos de un 2%. En otras palabras, la angio-TC sería una excelente herramienta para descartar o confirmar una hemorragia gastrointestinal y ayudar al clínico a abandonar la zona de incertidumbre diagnóstica.

Con estos resultados, podemos reforzar la hipótesis de los autores del artículo. Ellos establecen que el desempeño de la angio-TC es excelente, con una alta utilidad para la evaluación de pacientes con hemorragia gastrointestinal, en concordancia con otros trabajos que han reportado resultados similares. El análisis estadístico sugirió que la angio-TC permite descartar o confirmar fehacientemente la presencia de hemorragia, y que se asocia con un excelente correlato en la localización del sitio de sangrado. Esta información

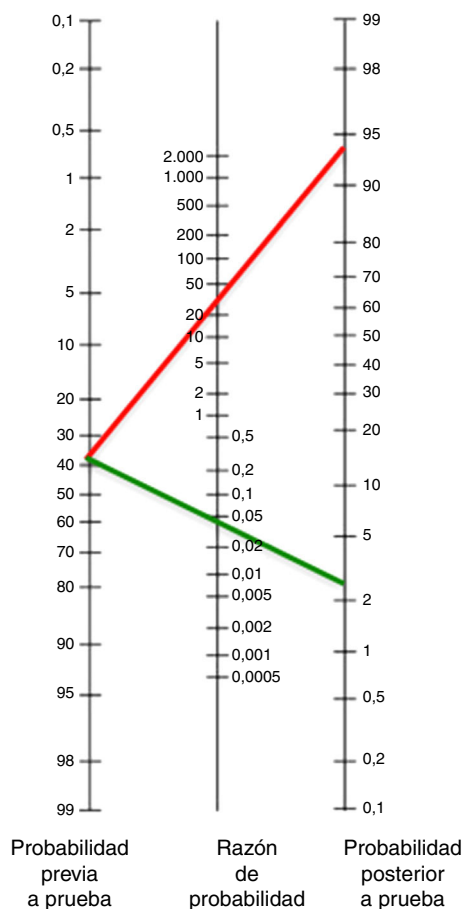


Figura 2 Nomograma de Fagan con representación gráfica de las probabilidades pre y *postest* de referencia⁶. La línea roja corresponde al resultado de la prueba positiva y la línea verde corresponde al resultado de la prueba negativa. Las probabilidades se expresan en porcentajes (%).

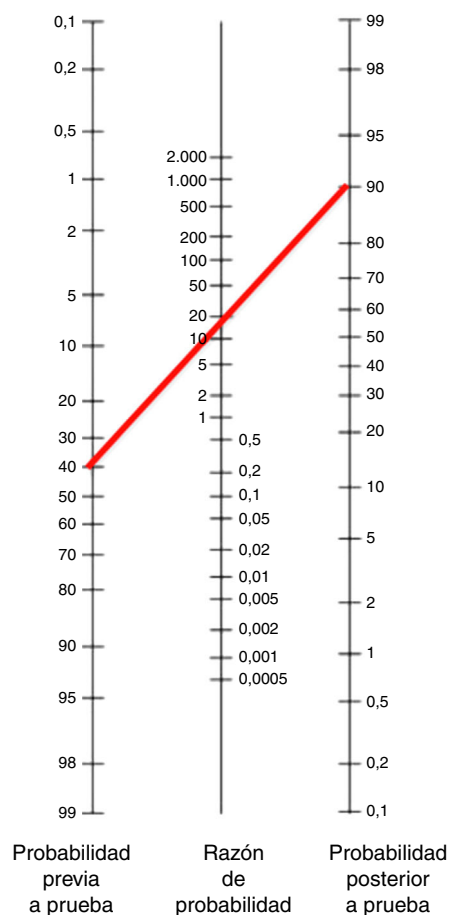


Figura 3 Nomograma de Fagan con representación gráfica de las probabilidades pre y *postest* con *likelihood ratio* conjunto derivado desde referencia 7. La línea roja corresponde al resultado obtenido por el *test* positivo múltiple con LR individual combinado. Las probabilidades se expresan en porcentajes (%).

es de gran utilidad para el clínico con el fin de direccionar la terapia y la eventual intervención hemostática, sea endoscópica, angiográfica o quirúrgica, de forma de evitar acciones innecesarias y riesgo de complicaciones.

Si bien los resultados de LR no fueron calculados por los autores, cualquier lector puede estimarlos desde los parámetros "clásicos" de sensibilidad y especificidad, aplicando probabilidades clínicas para cada paciente, lo que permite transportar su aplicabilidad directamente a la práctica clínica habitual^{1,4}.

Otras aplicaciones

En otras situaciones, un *test* puede tener un LR que no permite cambios relevantes por sí solos, pero su uso combinado o en serie con otras pruebas diagnósticas permite obtener un LR combinado que mejora el resultado^{1,2,5}. Es posible aplicar en serie tantas pruebas diagnósticas como el clínico o el radiólogo estimen convenientes, refinando secuencialmente la probabilidad *pretest* mediante la multiplicación de los LR de cada uno de los exámenes.

La selección de las pruebas en serie estará dada por los beneficios, riesgos, costos y disponibilidad de cada técnica, con la finalidad de realizar un diagnóstico acertado. Esto es lo que intuitivamente todo médico clínico lleva a cabo al agregar elementos al examen físico, anamnesis y estudios de laboratorio. Los hallazgos individuales de cada signo o síntoma pueden no ser suficientes para establecer un diagnóstico, pero su combinación determina una entidad o hipótesis diagnóstica más certera.

Para ejemplificar, tomamos el artículo de Soardi *et al.*⁷, y lo aplicamos en la siguiente situación clínica: la tomografía computada de tórax solicitada por disnea en un paciente de 73 años, fumador activo de 30 paquetes/año, reveló un nódulo de 17 mm y contornos mal definidos en el lóbulo superior derecho. Antes del estudio, el médico tratante había estimado la probabilidad de neoplasia pulmonar en un 40%, y el artículo recogía los valores de LR (+): edad 70-79 años de 1,616, fumador < 40 paquetes/año de 1,216, diámetro de 16-20 mm de 2,308 y contornos mal definidos de 3,710.

Los resultados por separado no impactaban significativamente en la probabilidad *postest*; sin embargo, siendo una expresión de probabilidades condicionadas, en el que cada parámetro mencionado modifica en serie

la probabilidad *pretest* del siguiente, el resultado final equivale a multiplicar cada LR individual para arrojar un LR conjunto ($1,616 \times 1,216 \times 2,308 \times 3,710 = 16,826$). Esto eleva la probabilidad *posttest* de nódulo pulmonar maligno a un 90% (fig. 3).

Por ello, los LR proveen una valiosa información aplicada a las variadas características de los pacientes, en tanto permiten entregar probabilidades ajustadas a las particularidades del caso, utilizando la información obtenida en grandes estudios poblacionales. Sin dudas, el rol de la radiología moderna debe ser orientar y dirigir las diferentes técnicas imagenológicas con el fin de colaborar en el proceso diagnóstico.

Es importante que todo radiólogo conozca y use los LR como forma de relacionarse con el accionar clínico, así como es altamente recomendable que todos los *test* diagnósticos presenten sus LR o entreguen los datos necesarios para su cálculo, con el fin de colaborar adecuadamente con los médicos tratantes para la mejor toma de decisiones.

Conclusión

Los LR son herramientas de importancia para poder asistir en el proceso de tomas de decisiones clínicas. El conocimiento de estos parámetros orienta al médico clínico en el impacto que puede tener la realización de un examen en particular, tanto para aumentar la probabilidad de enfermedad (en caso de resultado positivo) como para descartar la patología (si el resultado es negativo).

Responsabilidades éticas

Protección de personas y animales. Los autores declaran que para esta investigación no se han realizado experimentos en seres humanos ni en animales.

Confidencialidad de los datos. Los autores declaran que han seguido los protocolos de su centro de trabajo sobre la publicación de datos de pacientes.

Derecho a la privacidad y consentimiento informado. Los autores declaran que en este artículo no aparecen datos de pacientes.

Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses.

Bibliografía

1. Florkowski CM. Sensitivity, specificity, receiver-operating characteristic (ROC) curves and likelihood ratios: communicating the performance of diagnostic tests. *Clin Biochem Rev.* 2008;29 Suppl 1:S83-7.
2. Busel D, Silva C. Radiología basada en la evidencia: estrategia conceptual focalizada para la práctica de la imagenología. *Rev Chil Radiol.* 2004;10:109-17.
3. Loong TW. Understanding sensitivity and specificity with the right side of the brain. *BMJ.* 2003;327:716-9.
4. Deeks JJ, Altman DG. Diagnostic tests 4: likelihood ratios. *BMJ.* 2004;329:168-9.
5. Hayden SR, Brown MD. Likelihood ratio: a powerful tool for incorporating the results of a diagnostic test into clinical decision making. *Ann Emerg Med.* 1999;33:575-80.
6. Martí M, Artigas JM, Garzón G, Alvarez-Sala R, Soto JA. Acute lower intestinal bleeding: feasibility and diagnostic performance of CT angiography. *Radiology.* 2012;262:109-16.
7. Soardi GA, Perandini S, Motton M, Montemezzi S. Assessing probability of malignancy in solid solitary pulmonary nodules with a new Bayesian calculator: improving diagnostic accuracy by means of expanded and updated features. *Eur Radiol.* 2015;25:155-62.