

# Evaluación de pruebas diagnósticas para especialistas en rehabilitación cardiopulmonar

Víctor Z. Rizo-Tello<sup>1</sup> , Ana M. Londoño-Espinel<sup>1</sup> , Carlos D. Páez Mora<sup>1</sup> 

## Resumen

**Introducción:** A diario, los profesionales de la salud se enfrentan a importantes retos diagnósticos cuando atienden a sus pacientes, por lo que se apoyan en exámenes paraclínicos que complementan su ejercicio semiológico, porque les permiten confirmar o descartar una enfermedad.

**Objetivo:** Explicar a los especialistas en rehabilitación cardiopulmonar los conceptos requeridos para interpretar de forma crítica los resultados de las pruebas diagnósticas.

**Materiales y métodos:** Revisión narrativa de la literatura, que expone conceptos actualizados, ejemplos y gráficas con enfoque académico y didáctico.

**Resultado:** Se realizó la actualización y explicación de los conceptos de sensibilidad, especificidad, valores predictivos positivo y negativo y los *likelihood ratio* positivo y negativo y su interpretación en el normograma de Fagan, a través de ejemplos del día a día del especialista.

**Conclusión:** Comprender los conceptos que acompañan la evaluación de pruebas diagnósticas ayuda a que los especialistas en rehabilitación cardiopulmonar analicen críticamente los resultados de las ayudas paraclínicas funcionales y estructurales que acompañan a sus pacientes y, de esta manera, puedan caracterizar bien el diagnóstico y seguimiento de las personas.

**Palabras clave:** rehabilitación cardiaca; medicina física y rehabilitación; pruebas diagnósticas de rutina; sensibilidad y especificidad; valor predictivo de las pruebas; epidemiología.

<sup>1</sup> Universidad Manuela Beltrán (Bogotá, Colombia).

**Autora de correspondencia:** Ana M. Londoño-Espinel. Correo electrónico: [anamarialondono17@gmail.com](mailto:anamarialondono17@gmail.com)

### Citar este artículo así:

Rizo-Tello VZ, Londoño-Espinel AM, Páez Mora CD. Evaluación de pruebas diagnósticas para especialistas en rehabilitación cardiopulmonar. Rev Investig Salud Univ Boyacá. 2021;9(2):173-193. <https://doi.org/10.24267/23897325.871>

## Evaluation of Diagnostic Tests for Professionals who Perform Cardiopulmonary Rehabilitation

### Abstract

**Introduction:** Health professionals face important diagnostic challenges daily when they care for their patients, which is why they rely on paraclinical tests that complement their semiological exercise by confirming or ruling out a disease.

**Objective:** Explain to cardiopulmonary rehabilitation specialists, the concepts necessary to critically interpret the results of diagnostic tests.

**Materials and methods:** narrative review of the literature was carried out, which exposes updated concepts, examples, and graphs with an academic and didactic approach.

**Result:** The concepts of sensitivity, specificity, positive and negative predictive values and the positive and negative Likelihood Ratio and their interpretation in the Fagan normogram were updated and explained, through examples of the specialists day-to-day.

**Conclusion:** understanding the concepts that accompany the evaluation of diagnostic tests help cardiopulmonary rehabilitation specialists to critically analyze the results of the functional and structural paraclinical aids that accompany their patients, and thus be able to properly characterize the diagnosis and follow-up. of people.

**Keywords:** cardiac rehabilitation; physical and rehabilitation medicine; diagnostic tests routine; sensitivity and specificity; predictive value of tests; epidemiology.

## Avaliação de testes de diagnóstico para especialistas em reabilitação cardiopulmonar

### Resumo

**Introdução:** Diariamente, os profissionais da saúde enfrentam desafios diagnósticos, recorrendo a exames que complementem a sua prática semiológica, permitindo-lhes confirmar ou excluir uma doença.

**Objetivo:** Explicar aos especialistas em reabilitação cardiopulmonar os conceitos necessários para interpretar criticamente os resultados dos testes de diagnóstico.

**Materiais e métodos:** Revisão narrativa da literatura, fornecendo conceitos atualizados, exemplos e gráficas com uma abordagem acadêmica e didática.

**Resultado:** Os conceitos de sensibilidade, especificidade, valores preditivos positivos e negativos e razões de verossimilhança positivas e negativas e a sua interpretação no normograma de Fagan foram atualizados e explicados com exemplos do trabalho diário do especialista.

**Conclusão:** A compreensão dos conceitos que acompanham a avaliação dos testes de diagnóstico ajuda aos especialistas em reabilitação cardiopulmonar na análise crítica dos resultados das ajudas clínicas funcionais e estruturais dos pacientes e, desta forma, caracterizar bem o diagnóstico e o acompanhamento dos indivíduos.

**Palavras-chave:** Reabilitação Cardíaca; Medicina Física e Reabilitação; Testes Diagnósticos de Rotina; Sensibilidade e Especificidade; Valor Preditivo dos Testes; epidemiologia.

## Introducción

A medida que avanza el tiempo, se han ido evidenciando grandes e importantes investigaciones centradas en comprender mejor la historia natural de la enfermedad y, con ello, la creación de nuevas formas de diagnosticarla. Los profesionales de la salud están en un constante proceso de aprendizaje y actualización de conocimientos, pues una de sus tareas más importantes cuando enfrentan la realidad de un paciente es la de dar un correcto diagnóstico teniendo en cuenta que los exámenes paraclínicos usados no tienen la mayor precisión diagnóstica y pueden presentar resultados falsos positivos y falsos negativos, que llevan a un desafío de la práctica clínica para brindar el mejor tratamiento disponible en el momento y de forma oportuna (1,2).

Por lo general, el ejercicio diagnóstico es atribuido al médico; sin embargo, poco se tienen en cuenta las actividades de los demás profesionales de la salud, puesto que al establecer la línea de base del estado clínico de sus pacientes, participan en el análisis estructurado de diagnóstico, a partir del cual se sustenta el plan de tratamiento. Así mismo, con frecuencia se cree que el personal dedicado a la rehabilitación —y, de manera estricta, la rehabilitación cardiopulmonar— no debe adquirir competencias en la interpretación de las pruebas diagnósticas. Por ello, en esta revisión narrativa se hace hincapié

en la comprensión transversal que deben tener los profesionales afines a la responsabilidad de integrar el estado clínico con imágenes, pruebas y laboratorios tomados, para que de forma crítica argumenten su conformidad o no con el diagnóstico registrado como antecedente en la historia clínica por parte del médico tratante y así aportar de manera integral al paciente en el proceso de reincorporación a la cotidianidad (1-3).

Por tanto, el objetivo de esta revisión narrativa es explicar los conceptos relacionados con la evaluación de pruebas diagnósticas, para que los especialistas en rehabilitación cardiopulmonar estén en capacidad de realizar un análisis crítico de la literatura científica publicada y de los reportes de las pruebas hechas a los pacientes, que permita orientar las decisiones clínicas.

## Materiales y métodos

Se llevó a cabo una revisión narrativa de la literatura sobre el tema, a partir de artículos publicados durante el periodo 2015-2022; sin embargo, se incluyó menos del 10% de la bibliografía con fecha de publicación más antigua, por su relevancia teórica. En la búsqueda se ayudó del metabuscador PubMed, además LILACS (para tener resultados de países de Latinoamérica y el Caribe), SciELO, Science Direct y la Biblioteca Virtual en Salud. Finalmente, se realizó la búsqueda en bola de nieve a través

de buscadores genéricos en internet. Los criterios de selección fueron: artículos de revisión, artículos epidemiológicos observacionales y experimentales originales, revisiones sistemáticas de literatura con o sin metanálisis y guías de manejo e implementación de enfermedades que abordaran el tema de evaluación de pruebas diagnósticas, conceptos, componentes e interpretación. No se tuvieron en cuenta artículos en otros idiomas diferentes al inglés y al español. En la búsqueda se obtuvieron un total de 66 documentos, de los cuales fueron eliminados 14, por duplicidad y falta de claridad en los conceptos. Así, quedaron un total de 52 documentos seleccionados y analizados para la presente revisión narrativa.

## Desarrollo

Una prueba diagnóstica se define como cualquier elemento utilizado para ayudar a confirmar o descartar un diagnóstico en los pacientes. Como ejemplo de ello se puede nombrar la anamnesis, el examen físico, las pruebas del laboratorio y las imágenes. Su función principal es disminuir la incertidumbre en el momento de tomar decisiones de abordaje de los pacientes, por lo cual su evaluación se fundamenta en la caracterización y la determinación de la funcionalidad (1-3).

En la práctica clínica asistencial se utilizan dos tipos de pruebas: las de tamizaje y las de diagnóstico, las cuales se diferencian porque las pri-

meras se utilizan en grandes poblaciones o a gran escala, normalmente en personas sanas o pacientes asintomáticos, con el objetivo de identificar la enfermedad tempranamente (en el modelo de prevención en salud-enfermedad corresponde a la prevención secundaria). Las segundas se aplican en pacientes en quienes previamente se tiene una impresión diagnóstica y se necesita tomar decisiones clínicas (3-5).

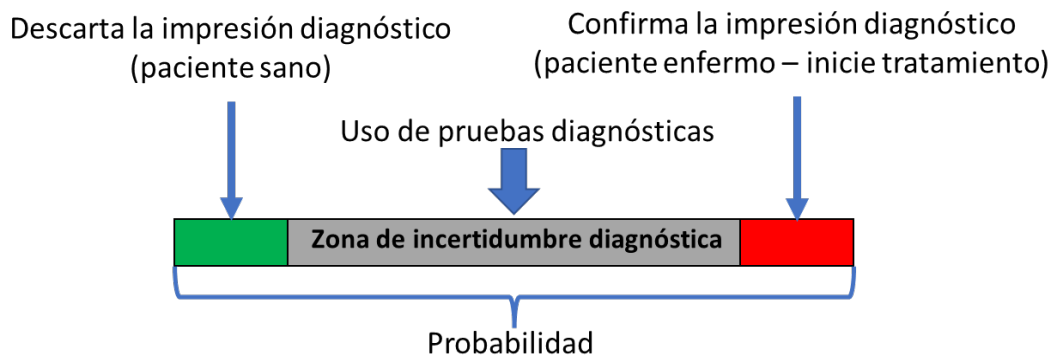
Las pruebas diagnósticas deben cumplir con algunas características generales, como su facilidad al aplicarse, la seguridad al usarlas, y deben ser económicas y fáciles de adquirir. En principio, se han identificado tres atributos importantes de una buena prueba: confiabilidad o reproducibilidad (de 100 veces que se repita la prueba cuántas van a dar el mismo resultado), seguridad y validez, las cuales son un principio fundamental de la precisión. La validez de las pruebas es el grado en el que se mida lo que debe medir o para lo que fue creada y, de esta manera, discrimine de manera apropiada los enfermos de los no enfermos (también conocida como exactitud) y que estos resultados se mantengan cuando se repita la prueba en diferentes momentos, escenarios y poblaciones bajo las mismas condiciones (6).

El uso de las pruebas diagnósticas es el complemento de un buen ejercicio diagnóstico por parte del rehabilitador cardiopulmonar tratante, quien debe realizar la anamnesis y el examen físico de

calidad, lo que ayuda a enfocar al paciente en el contexto de la probabilidad de ocurrencia o probabilidad pretest (probabilidad de que la enfermedad sospechada sea la que tiene el paciente antes de que se le realice la prueba diagnóstica). Si esa probabilidad previa no supera el umbral diagnóstico, entonces se procede a descartar el diagnóstico. Si ocurre de manera contraria, es decir, se supera el umbral, entonces el médico iniciará el tratamiento de la enfermedad sin requerir el uso de una prueba confirmatoria o la solicita como soporte legal para la historia clínica. En caso de que la probabilidad sea intermedia, el profesional debe apoyarse en la realización de pruebas diagnósticas, enfocado en una hipótesis diagnóstica, tal como se presenta en la figura 1 (7).

*Reflexión.* La evaluación pretest o hipótesis diagnóstica establecida no es el resultado de un análisis arbitrario o producto del azar; es el resultado del uso de todas las herramientas semiológicas (anamnesis, examen físico y todas las posibles relaciones sindrómicas o patológicas existentes). Esto lleva a hacer un uso racional de los recursos económicos y humanos, teniendo en cuenta que estos son finitos, y como ganancia secundaria ayudan a priorizar pacientes y a descongestionar laboratorios, servicios de urgencias y de imagenología.

**Figura 1.** Probabilidad diagnóstica



Fuente: tomada de (6,7) y modificada por los autores.

### ***Prueba de referencia (gold standard) o patrón de referencia***

La prueba de referencia cuenta con la más alta precisión diagnóstica; en otras palabras, es la que dice sin equivocación que la enfermedad está presente o no. La prueba de referencia, en comparación con las otras opciones diagnósticas, es más probable que identifique correctamente a las personas con la enfermedad (es específica) y también a aquellas que no tienen la enfermedad (es sensible) (6). El rendimiento de toda prueba diagnóstica se basa en su comparación con estándar de oro (patrón de oro). Las pruebas de oro tienen consigo algunas desventajas, como: en ocasiones son difíciles de conseguir, pueden traer riesgos adicionales para el paciente, puede ser más costosa y de difícil acceso (6-8,9).

Siempre que se desee realizar la validación de una nueva prueba diagnóstica, esta debe ser comparada con el patrón de referencia, o con la que hasta el momento sea la de mayor precisión diagnóstica disponible (6,8).

El ejercicio de evaluación de pruebas diagnósticas contempla el cálculo de algunas medidas que se dividen en dos grandes grupos: las que establecen la propiedad discriminativa de la prueba (con las que se toman decisiones en el marco de las políticas de salud) y las que se utilizan para evaluar su capacidad predictiva (con las que se

mide la probabilidad de la enfermedad en los individuos) (10).

### ***Validez de la prueba***

Para hacer más fácil la comprensión de los conceptos, estos se explican a partir de la evaluación de la tomografía axial computarizada (TAC) simple de tórax (11,12), en comparación con la espirometría (patrón de referencia) (13), en el diagnóstico de la enfermedad pulmonar obstructiva crónica (EPOC). Los valores usados no corresponden a los reales y se emplean con fines académicos.

*Prueba de oro:* espirometría con respuesta pos-broncodilatador

*Nueva prueba diagnóstica:* tomografía de tórax simple

*Total de personas en el estudio:* 1049

La EPOC es de alta frecuencia, prevenible y tratable; comprende síntomas respiratorios persistentes y limitación del flujo del aire, secundarios a anomalías alveolares e inflamación de las vías aéreas por exposición a material particulado o gases. El diagnóstico definitivo se realiza con espirometría, evaluando el cociente entre el volumen espiratorio forzado en el primer segundo y la capacidad vital forzada ( $VEF1/CVF$ )  $< 0,7$  (en espirometría después de inhalar un broncodilatador) (14-17).

**Tabla 1.** Evaluación de pruebas diagnósticas (2 × 2)

		GOLD STANDAR		Total
		Positivo	Negativo	
PRUEBA NUEVA A EVALUAR	Positivo	(a) Verdaderos positivos	(b) Falsos positivos	a+b
	Negativo	(c) Falsos negativos	(d) Verdaderos negativos	c+d
Total		a+c Enfermos	b+d Sanos	n

Fuente: tomada de (23) y modificada por los autores.

Es de importancia recordar que las herramientas utilizadas en la evaluación de pruebas diagnósticas se derivan del diseño de estudio epidemiológico de corte transversal, y una de estas ayudas es la tabla 2 × 2 o de contingencia, con el fin de ilustrar los conceptos de la validez de una prueba diagnóstica. En la tabla 1 registra la ubicación del patrón de referencia en las columnas, que señala a los enfermos y a los no enfermos, y en las filas, la prueba diagnóstica nueva (18-22).

*Reflexión.* El clínico debe aprender los conceptos presentados aquí, pues mediante programas estadísticos se obtienen los cálculos de forma rápida; sin embargo, la función del clínico —en especial la de los rehabilitadores cardiopulmonares— va más allá de algo mecánico, y es entender el resultado, analizarlo y aplicarlo en un contexto de su paciente.

- Verdaderos positivos: son las personas que la prueba nueva detecta como positivos y en realidad están enfermos.

- Falsos positivos: personas que son etiquetadas erróneamente por la prueba nueva como positivos, pero en realidad están sanos.

- Falsos negativos: individuos detectados erróneamente como negativos por la prueba nueva, pero verdaderamente están enfermos.

- Verdaderos negativos: personas detectadas como sanas por la prueba nueva y verdaderamente están sanas (8-10,23,24).

Entre menos falsos positivos y falsos negativos tenga la prueba diagnóstica, mejor es su rendimiento



**Tabla 2.** Comparativo de 2 × 2: tomografía simple de tórax contra la espirometría en el diagnóstico de la enfermedad pulmonar obstructiva crónica

		Espirometría		Total
		Positivo	Negativo	
Tomografía de tórax simple	Positivo	540	5	545
	Negativo	22	643	665
Total		562	648	1210

### **Sensibilidad (S)**

La sensibilidad se define como la capacidad de una prueba para identificar acertadamente a individuos con la enfermedad; en otras palabras, se podría decir que es la probabilidad de que, en un individuo enfermo, la prueba resulte positiva. La sensibilidad se calcula en función de cuántas personas tienen la enfermedad (tabla 2). En realidad, lo que dice la fórmula no es más que la proporción de sujetos con resultado positivo entre el total de los enfermos y no se calcula sobre toda la población (8,25-27).

Se calcula así: sensibilidad = número de verdaderos positivos / (número de verdaderos positivos + número de falsos negativos) (8). En otras palabras, sensibilidad = verdaderos positivos / total de enfermos o  $S = a/a + c$  (6).

*Importante.* Si una prueba es altamente sensible, un resultado negativo descarta la enfermedad.

Entonces:

$S = 540 / 562 = 0,9609$ , que puede expresarse como un porcentaje así:

$$S = 0,9609 \times 100 = 96,1\%$$

La lectura e interpretación debe darse teniendo como base el concepto matemático de sensibilidad. A continuación, se enumeran algunas maneras para expresar la información:

- La tomografía simple de tórax es capaz de detectar al 96,1% de las personas que verdaderamente tienen la condición o enfermedad.

- La tomografía simple de tórax será positiva en el 96,1% de las personas que tengan la enfermedad.

- De 100 personas verdaderamente enfermas con EPOC. la tomografía simple de tórax podrá detectar a 96 de ellas.

- O, sencillamente, la sensibilidad de la tomografía simple de tórax para detectar EPOC es del 96,1%.

### **Especificidad (E)**

La especificidad se define como la capacidad de una prueba para identificar correctamente a los individuos que no tienen la enfermedad, es decir, es la probabilidad de que en una persona libre de la enfermedad, la prueba dé un resultado negativo. Si una prueba es altamente específica, un resultado positivo confirma la enfermedad (6,26).

Se calcula así: especificidad = verdaderos negativos / verdaderos negativos + falsos positivos; en otras palabras, verdaderos negativos / total de sanos o  $d/b + d$  (8,25).

En realidad, lo que dice la fórmula no es más que la proporción de sujetos con resultado negativo entre el total de los sanos. Continuando con el ejemplo:

$E = 643/648 = 0,9923$ , que puede expresarse como un porcentaje así:

$$E = 0,9923 \times 100 = 99,23\%$$

Lectura:

- La tomografía simple de tórax es capaz de detectar al 99,23% de las personas que verdaderamente no tienen EPOC.

- La tomografía simple de tórax será *negativa* en el 99,23% de las personas que no tienen EPOC.

- De 100 personas verdaderamente sanas, la tomografía simple de tórax podrá detectar al 99 de ellas, si esta tiene resultado negativo.

- Sencillamente, la especificidad de la tomografía simple de tórax es del 96,4% para el diagnóstico de EPOC.

### **Seguridad de las pruebas diagnósticas**

La seguridad de la prueba diagnóstica está dada a partir de los valores predictivos positivos y los valores predictivos negativos. Estos tienen utilidad en el escenario clínico, porque integran la probabilidad del individuo de estar enfermo o no, dependiendo del resultado de la prueba; mientras que la sensibilidad y la especificidad describen

muy bien la prueba, y lo verdaderamente importante es saber si se tiene o no la enfermedad (28).

### **Valor predictivo positivo**

El valor predictivo positivo (VPP) se define como la probabilidad que tiene un individuo de estar enfermo cuando el resultado de la prueba es positivo (9,26,28). Lo que pretende es identificar verdaderos positivos y no categorizar a algunas personas como portadoras de la enfermedad cuando en realidad no la tienen, es decir, se quieren evitar los falsos positivos (10,29-31).

Valor predictivo positivo (VPP) = verdaderos positivos / Total de los positivos de la prueba nueva o  $a/a + b$

Entonces:

$$\text{VPP} = 540 / 545 = 0,9908$$

$$\text{VPP} = 0,9908 \times 100 = 99,08\%$$

Lectura:

- Dado que el resultado de la tomografía simple de tórax es positivo, el paciente tiene el 99,1% de probabilidades de tener EPOC.

- De 100 personas en las que la tomografía simple de tórax es positiva, 99 de ellas verdaderamente EPOC.

- O, sencillamente, el VPP es del 99,1% para EPOC.

### **Valor predictivo negativo**

El valor predictivo negativo (VPN) se define como esa probabilidad de que un sujeto en quien le salió un resultado negativo no presente la enfermedad, o sencillamente la probabilidad que tiene un individuo de estar sano cuando el resultado de la prueba es negativo (9,26,28). Lo que pretende es identificar a los verdaderos negativos y no categorizar a algunas personas como sanas, cuando en realidad tienen la enfermedad, es decir, se quieren evitar los falsos negativos (10,29-31).

Valor predictivo negativo = verdaderos negativos / Total de los negativos de la prueba nueva o  $d/c + d$

Entonces:  $\text{VPN} = 643 / 665 = 0,9669$

Es preciso recordar que el VPN es una probabilidad y que las probabilidades se expresan como 0 o 1 (ausencia o presencia del evento); sin embargo, para hacer más fácil la comprensión, la convertimos en porcentaje multiplicando por 100.

$$\text{VPN} = 0,9669 \times 100 = 96,69\%$$

Lectura:

- Dado que el resultado de la tomografía simple de tórax es negativo, el paciente tiene el 96,69 % de probabilidades de no tener EPOC.
- De 100 personas en las que la tomografía simple de tórax fue negativa, 97 de ellas verdaderamente no tienen EPOC.
- Sencillamente, el VPP es del 96,69% para el diagnóstico de EPOC.

*Nota.* Los valores predictivos también dependen de la prevalencia de la enfermedad; es por esto por lo que si la prevalencia es alta, el VPP también será alto, y el VPN, bajo, y si la prevalencia es baja, entonces el VPP será bajo y el VPN será alto (32).

De la misma manera, es importante tener en cuenta lo siguiente: si aumenta la sensibilidad, entonces también aumenta el VPN, y si aumenta la especificidad, entonces aumenta el VPP.

### **Razón de verosimilitud o likelihood ratio**

La razón de verosimilitud (RV) o *likelihood ratio* (LR) también se conoce como el índice de eficiencia pronóstica. Se trata de un indicador fijo y utilizado cuando la prueba diagnóstica es cuantitativa y tiene un umbral o punto de corte. Son independientes de la prevalencia de la enfermedad

en una población. Indica que el resultado de una prueba diagnóstica elevará o reducirá la probabilidad de tener la enfermedad. Como lo expresan Felipe Salech et al. (9), es “la razón entre la probabilidad de tener determinado resultado de la prueba en la población con la condición versus tener el mismo resultado en la población sin la condición”.

Finalmente, lo que nos quiere mostrar es la magnitud y el sentido del cambio de la probabilidad pre- a postest, según el resultado de la prueba (33-35). El significado del LR tiene tres valores de importancia: cuando el valor está por encima de 1, cuando está por debajo de 1 y cuando es igual a 1. Por último, es importante que los clínicos expresen el significado de una LR dada usando palabras, en lugar de números y fórmulas que a veces pueden confundir y dar una percepción de dificultad en su comprensión (36).

### **Razón de verosimilitud positiva o likelihood ratio positivo**

La razón de verosimilitud positiva (RVP) o el *likelihood ratio positivo* (LR+) es una razón que expresa la probabilidad de que un individuo con la enfermedad tenga un resultado positivo en la prueba, dividido por la probabilidad de que un paciente no tenga una enfermedad, pero tenga un resultado positivo en la prueba. Compara los verdaderos positivos con los falsos positivos

(33,37-39). Informa que es muy probable que en un paciente enfermo la prueba resulte positiva, con respecto a un paciente sano, con el mismo resultado positivo. ¿Qué tanto más probable es encontrar una prueba positiva en alguien enfermo que en alguien sano? (9,35,38).

$LR+ = S$  (verdaderos positivos) / probabilidad de tener falsos positivos

$LR+ = S / (1-\text{especificidad})$  o

$LR+ = (a/a + c)/(b/b + d)$

*Recordemos.* Sensibilidad = 0,9609; especificidad = 0,9923

Entonces:

$LR+ = \text{Sensibilidad} / (1-\text{especificidad}) = LR+ = 0,9609 / (1-0,9923)$

$LR+ = 0,9609 / 0,0077 = LR+ = 124,79$

Entre más lejano a 1 sea el resultado, es mejor el comportamiento de la prueba para captar positivos, y lo interpretamos de la siguiente forma: por cada 125 verdaderos positivos de EPOC detectados por la prueba, va a existir un falso positivo.

### **Razón de verosimilitud negativa o likelihood ratio negativo**

La razón de verosimilitud negativa (RVN) o el *likelihood ratio negativo* ( $LR-$ ) es una razón que

expresa la probabilidad de que una persona sin la enfermedad con una prueba negativa, dividida por la probabilidad de que un paciente tenga una enfermedad, tenga un resultado negativo. Ello pone en evidencia la comparación entre los verdaderos negativos con los falsos negativos. En otras palabras: ¿qué tanto más probable es encontrar la prueba negativa en alguien enfermo que en alguien sano? (34-39).

$LR- = \text{probabilidad de tener falsos negativos} / E$  (verdaderos negativos)

$LR- = (1-\text{sensibilidad}) / (\text{especificidad})$  o  $LR- = (c/a + c)/(d/b + d)$

*Recordemos.* Sensibilidad = 0,9609; especificidad = 0,9923

$LR- = (1-0,9609) / 0,9923 = LR- = 0,0391 / 0,9923 = LR- = 0,039$

Entre más cerca de 0 sea el resultado, mejor es el comportamiento de la prueba para captar negativos. Para su interpretación debe sacarse el inverso a ese resultado  $1/0,039 = 25,57$ , lo que indica que por cada 26 personas verdaderas negativas para EPOC hay una persona falsa negativa (tabla 3).

**Tabla 3.** Interpretación del *likelihood ratio*

Likelihood Ratio +	Likelihood Ratio -	Utilidad
≥10	≤0,99	Altamente relevante
5-9,99	0,1-0,199	Buena
2-4,99	0,499-2	Regular
<2	≥0-5	Mala

Fuente: tomada (40) y modificada por los autores.

Teniendo en cuenta lo anterior, las razones de verosimilitud o *LR* ayudan al especialista clínico a adaptar la sensibilidad y especificidad de las pruebas a pacientes individuales. Los especialistas en rehabilitación cardiopulmonar pueden decidir realizar o interpretar una prueba u otra en particular, luego de una buena anamnesis y examen físico al paciente y de acuerdo con los hallazgos clínicos, como síntomas y signos del paciente, y los conocimientos derivados de la práctica profesional. De esta forma, puede sospechar cierto diagnóstico que sería necesario confirmar o descartar para realizar un mejor tratamiento al paciente.

La probabilidad de que un paciente tenga cierta enfermedad a partir de la valoración e historia clínica, de los conocimientos derivados de la práctica médica, de la prevalencia de esta enfermedad en la población y los resultados de investigación, antes de que se tenga el resultado de la prueba diagnóstica, se denomina *probabilidad preprueba* (41-44).

El objetivo de la realizar la prueba diagnóstica es, entonces, obtener información adicional de la enfermedad, con la consecuente modificación de la probabilidad preprueba, lo cual indica que una prueba positiva aumenta la probabilidad de tener la enfermedad; mientras que una prueba negativa la reduce. Adicional a esto, existe la probabilidad posprueba, la cual corresponde a la probabilidad de que el individuo tenga la enfermedad posterior a que se conozcan los resultados de la prueba. Esta es la que más debe interesar a los rehabilitadores cardiopulmonares y pacientes, para confirmar o descartar el diagnóstico y definir si se inicia el programa de rehabilitación o se reorienta para exámenes complementarios (42-45).

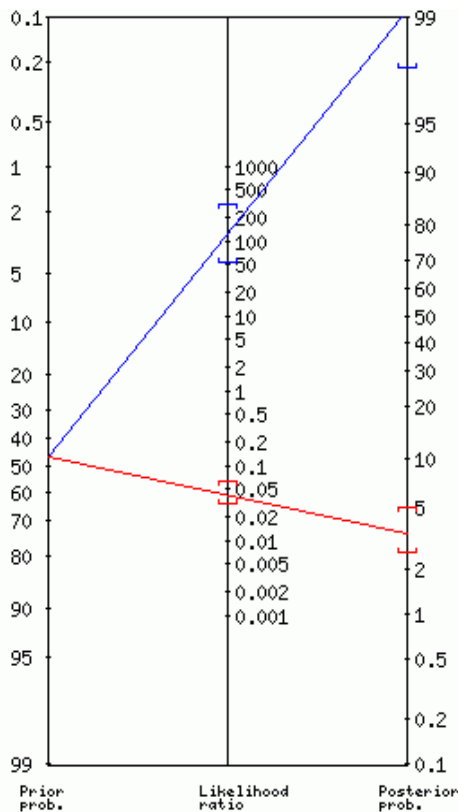
La probabilidad posprueba es sencilla de calcular, basándose en el teorema de Bayes, que multiplica la probabilidad preprueba por el *LR* así (41,42):

$$\text{Probabilidad posprueba} = \text{probabilidad preprueba} \times LR$$

Existe una herramienta gráfica que, en el ejercicio clínico rutinario es ampliamente usada, porque permite visualizar el cambio de la valoración preprueba posterior a tener los resultados de la prueba diagnóstica: el *normograma de Fagan* (46,47), que se representa en la figura 2. Este consiste en trazar una línea recta desde el valor de la probabilidad preprueba (eje izquierdo) cruzando por la RV de la prueba (eje central),

para finalmente proyectar la línea sobre el valor de la probabilidad de enfermedad posterior a la prueba (eje derecho) (48-51). Continuando con el ejemplo, la probabilidad preprueba es de 0,46 (562 /1210).

**Figura 2.** Normograma de Fagan para tomografía simple vs. espirometría



Fuente: calculado con los valores del ejercicio en <http://araw.mede.uic.edu/cgi-bin/testcalc.pl>

La línea azul grafica la probabilidad posprueba si el resultado es positivo, teniendo en cuenta un  $LR+$  de 124,79; entre tanto, la línea roja representa la probabilidad posprueba si el resultado es negativo, teniendo en cuenta el  $LR-$  de 0,039. En ese normograma se evidencia cómo hay un cambio importante en la probabilidad posprueba en relación con la presencia de la enfermedad.

## Discusión

En el mundo, el volumen de publicaciones científicas en el área de la salud aumenta significativamente día a día, y ello fuerza a los profesionales clínicos a mejorar su capacidad analítica e investigativa (4). La calidad en la toma de las decisiones en el ámbito clínico, hospitalario o ambulatorio depende de la habilidad con la que se interpreten los resultados de los estudios epidemiológicos y la evidencia científica (4,5). A diario, la tecnología diagnóstica evoluciona a efectos de detectar con mayor precisión la enfermedad en los pacientes; por ello, varios autores dedicados a la medicina basada en la evidencia o, en general, a la práctica basada en la evidencia refieren que esta actividad no debe ser un proceso mecánico de únicamente interpretar las medidas necesarias para identificar la precisión y el rendimiento diagnóstico, sino que el clínico debe saber calcular y ejecutar las fórmulas, pues estas son la base para conocer la ruta en el surgimiento de la información que favorece la fácil aplicación en el entorno asistencial.

Y este es el motivo por el cual se justificó la elaboración de este trabajo investigativo enfocado en una metodología didáctica (10-13).

### **Conclusión**

Interiorizar la importancia de entender los conceptos que acompañan la evaluación de pruebas diagnósticas, les ofrece a los especialistas en rehabilitación cardiopulmonar la posibilidad de analizar críticamente los resultados de las ayudas paraclínicas no solo en el contexto de los parámetros normales de clasificación de salud o enfermedad, sino en el escenario del rendimiento, precisión y probabilidad diagnóstica que acompañan a sus pacientes, para que, de esta manera, sean capaces de caracterizar bien el diagnóstico, el tratamiento y el seguimiento de las personas.

Del mismo modo, es importante que los rehabilitadores cardiopulmonares interpreten los resultados de las pruebas diagnósticas no solo en el contexto de los parámetros normales de clasificación de salud o enfermedad, sino en el escenario del rendimiento, precisión y probabilidad diagnóstica, lo que los llevará a elegir de una manera más segura la que mejor los ayude en su ejercicio asistencial.

### **Limitaciones**

El acceso exclusivo para la búsqueda en las siete bases de datos revela evidencia científica localizada

que puede racionalizar la información; no obstante, el análisis de los artículos deja ver que la literatura socializa este tema de forma fragmentada y que no existía un artículo con enfoque didáctico con el paso a paso y la interpretación, y que integrara todos los conceptos. Sin embargo, desde esta revisión podrán derivarse otras que amplíen su búsqueda con enfoques investigativos académicos.

### **Agradecimientos**

Los autores agradecen a la Universidad Manuela Beltrán y a su equipo docente de la Especialización de Rehabilitación Cardiopulmonar, y a la Universidad de Boyacá, por los espacios académicos que brinda para el fomento de la investigación en el campo de la salud.

### **Financiación**

Los autores no reciben financiación de ninguna institución o empresa; asumen los costos totales de la elaboración y publicación del trabajo de investigación.

### **Conflictos de intereses**

Los autores informan que no tienen conflictos de intereses. Este trabajo es con fines académicos.



## Referencias

1. Leeflang MMG, Allerberger F. How to: evaluate a diagnostic test. *Clin Microbiol Infect.* 2019 Jan;25(1):54-59. <https://doi.org/10.1016/j.cmi.2018.06.011>
2. Leeflang MM, Deeks JJ, Gatsonis C, Bossuyt PM; Cochrane Diagnostic Test Accuracy Working Group. Systematic reviews of diagnostic test accuracy. *Ann Intern Med.* 2008 Dec 16;149(12):889-97. <https://doi.org/10.7326/0003-4819-149-12-200812160-00008>
3. Sitch AJ, Dekkers OM, Scholefield BR, Takwaingi Y. Introduction to diagnostic test accuracy studies. *Eur J Endocrinol.* 2021 Feb;184(2):E5-E9. <https://doi.org/10.1530/EJE-20-1239>
4. Gogtay NJ, Thatte UM. Statistical Evaluation of Diagnostic Tests (Part 1): Sensitivity, Specificity, Positive and Negative Predictive Values. *J Assoc Physicians India.* 2017 Jun;65(6):80-84.
5. Maxim LD, Niebo R, Utell MJ. Screening tests: a review with examples. *Inhal Toxicol.* 2014 Nov;26(13):811-28. <https://doi.org/10.3109/08958378.2014.955932>. Epub 2014 Sep 29. Erratum in: *Inhal Toxicol.* 2019 Jun;31(7):298. <https://doi.org/10.3109/08958378.2014.955932>
6. Kumar R. Evaluation of diagnostic tests. *Clin Epidemiol Global Health.* 2016;4(2):76-79. <https://doi.org/10.1016/j.cegh.2015.12.001>
7. Medina C. Generalidades de las pruebas diagnósticas, y su utilidad en la toma de decisiones médicas. *Rev Colomb Psiquiatr.* 2011;40(4):787-97. [https://doi.org/10.1016/S0034-7450\(14\)60165-7](https://doi.org/10.1016/S0034-7450(14)60165-7)
8. Swift A, Heale R, Twycross A. What are sensitivity and specificity? *Evid Based Nurs.* 2020 Jan;23(1):2-4. <https://doi.org/10.1136/ebnurs-2019-103225>. Epub 2019 Nov 12. Erratum in: *Evid Based Nurs.* 2022 Apr;25(2):e1. <https://doi.org/10.1136/ebnurs-2019-103225>
9. Salech F, Mery V, Larrondo F, Rada G. Estudios que evalúan un test diagnóstico: interpretando sus resultados. *Rev Méd Chile.* 2008 Sep;136(9):1203-8. <https://doi.org/10.4067/S0034-98872008000900018>
10. Simundic AM. Diagnostic accuracy-Part 1 Basic concepts: sensitivity and specificity, ROC analysis, STARD statement [internet]. 2009. Disponible en: <https://acutecaretesting.org/-/media/acutecaretesting/files/pdf/diagnostic-accuracy--part-1-basic-concepts--sensitivity-and-specificity-roc-analysis-stard-statement.pdf>

11. Singhvi D, Bon J. CT Imaging and comorbidities in COPD: beyond lung cancer screening. *Chest*. 2021 Jan;159(1):147-53. <https://doi.org/10.1016/j.chest.2020.08.2053>
12. Bakker JT, Klooster K, Vliedenthart R, Slebos DJ. Measuring pulmonary function in COPD using quantitative chest computed tomography analysis. *Eur Respir Rev*. 2021 Jul 13;30(161):210031. <https://doi.org/10.1183/16000617.0031-2021>
13. Schnieders E, Ünäl E, Winkler V, Dambach P, Louis VR, Horstick O, Neuhaus F, Deckert A. Performance of alternative COPD case-finding tools: a systematic review and meta-analysis. *Eur Respir Rev*. 2021 May 25;30(160):200350. <https://doi.org/10.1183/16000617.0350-2020>
14. Global Initiative for Chronic Obstructive Lung Disease-GOLD. Global strategy for prevention, diagnosis and management of COPD: 2022 report [internet]. 2022 Jan 24 [citado 2022 jun 13];1(1):1-56. <https://goldcopd.org/2022-gold-reports-2/>
15. Christenson SA, Smith BM, Bafadhel M, Putcha N. Chronic obstructive pulmonary disease. *Lancet*. 2022;399:2227. [https://doi.org/10.1016/S0140-6736\(22\)00470-6](https://doi.org/10.1016/S0140-6736(22)00470-6)
16. D'Urzo KA, Mok F, D'Urzo AD. Variation among spirometry interpretation algorithms. *Respir Care*. 2020 Oct;65(10):1585-90. <https://doi.org/10.4187/respcare.07294>
17. Hoesterey D, Das N, Janssens W, Buhr RG, Martinez FJ, Cooper CB, Tashkin DP, Barjaktarevic I. Spirometric indices of early airflow impairment in individuals at risk of developing COPD: spirometry beyond FEV1/FVC. *Respir Med*. 2019 Sep;156:58-68. <https://doi.org/10.1016/j.rmed.2019.08.004>
18. Kesmodel US. Cross-sectional studies - what are they good for? *Acta Obstet Gynecol Scand*. 2018 Apr;97(4):388-93. <https://doi.org/10.1111/aogs.13331>
19. Cataldo R, Arancibia M, Stojanova J, Papuzinski C. General concepts in biostatistics and clinical epidemiology: observational studies with cross-sectional and ecological designs. *Medwave*. 2019 Sep 25;19(8):e7698. <https://doi.org/10.5867/medwave.2019.08.7698>
20. Buka SL, Rosenthal SR, Lacy ME. Epidemiological study designs: traditional and novel approaches to advance life course health development research. En: Halfon N, Forrest CB, Lerner RM, et al., editores. *Handbook of life course health development*. Cham (CH):

- Springer; 2018. [https://doi.org/10.1007/978-3-319-47143-3\\_22](https://doi.org/10.1007/978-3-319-47143-3_22)
21. Campbell JM, Klugar M, Ding S, Carmody DP, Hakonsen SJ, Jadotte YT, White S, Munn Z. Diagnostic test accuracy: methods for systematic review and meta-analysis. *Int J Evid Based Healthc*. 2015 Sep;13(3):154-62. <https://doi.org/10.1097/XEB.0000000000000061>
  22. Korevaar DA, Gopalakrishna G, Cohen JF, et al. Targeted test evaluation: a framework for designing diagnostic accuracy studies with clear study hypotheses. *Diagn Progn Res*. 2019;3(22). <https://doi.org/10.1186/s41512-019-0069-2>
  23. Heuer C, Stevenson MA. Diagnostic test validation studies when there is a perfect reference standard. *Rev Sci Tech*. 2021 Jun;40(1):261-70. <https://doi.org/10.20506/rst.40.1.3223>
  24. Mouliou DS, Gourgoulialis KI. False-positive and false-negative COVID-19 cases: respiratory prevention and management strategies, vaccination, and further perspectives. *Expert Rev Respir Med*. 2021 Aug;15(8):993-1002. <https://doi.org/10.1080/17476348.2021.1917389>
  25. Aamir A, Hamilton RG. Predictive value model for laboratory tests: diagnostic sensitivity, diagnostic specificity, positive and negative predictive value, efficiency, likelihood ratio ([positive and negative]), incidence and prevalence. En: Mackay IR, Rose NR, Ledford DK, Lockey RF, editores. *Encyclopedia of medical immunology*. New York: Springer; 2014. [https://doi.org/10.1007/978-1-4614-9194-1\\_312](https://doi.org/10.1007/978-1-4614-9194-1_312)
  26. Monaghan TF, Rahman SN, Agudelo CW, Wein AJ, Lazar JM, Everaert K, Dmochowski RR. Foundational statistical principles in medical research: sensitivity, specificity, positive predictive value, and negative predictive value. *Medicina (Kaunas)*. 2021 May 16;57(5):503. <https://doi.org/10.3390/medicina57050503>
  27. Leeflang MM, Rutjes AW, Reitsma JB, Hooft L, Bossuyt PM. Variation of a test's sensitivity and specificity with disease prevalence. *CMAJ*. 2013 Aug 6;185(11):E537-44. <https://doi.org/10.1503/cmaj.121286>
  28. Ying GS, Maguire MG, Glynn RJ, Rosner B. Calculating sensitivity, specificity, and predictive values for correlated eye data. *Invest Ophthalmol Vis Sci*. 2020 Sep 1;61(11):29. <https://doi.org/10.1167/iovs.61.11.29>
  29. Staffa SJ, Zurakowski D. Statistical evaluation of diagnostic tests: a primer for pediatric surgeons. *J Pediatr Surg*. 2019 Apr;54(4):799-804. <https://doi.org/10.1016/j.jpedsurg.2018.06.010>

30. Trevethan R. Sensitivity, specificity, and predictive values: foundations, pliabilitys, and pitfalls in research and practice. *Front Public Health*. 2017 Nov 20;5:307. <https://doi.org/10.3389/fpubh.2017.00307>
31. Kotowski M, Szydłowski J. Radiological diagnostics in nasal dermoids: Pitfalls, predictive values and literature analysis. *Int J Pediatr Otorhinolaryngol*. 2021 Oct;149:110842. <https://doi.org/10.1016/j.ijporl.2021.110842>
32. Vali Y, Yang B, Olsen M, Leeflang MMG, Bossuyt PMM. Reporting of test comparisons in diagnostic accuracy studies: a literature review. *Res Synth Methods*. 2021 May;12(3):357-367. <https://doi.org/10.1002/jrsm.1469>
33. Shreffler J, Huecker MR. Diagnostic testing accuracy: sensitivity, specificity, predictive values and likelihood ratios. 2022 Mar 9. En: *StatPearls*. Treasure Island (FL): StatPearls Publishing; 2022.
34. Baratloo A, Safari S, Elfil M, Negida A. Evidence based emergency medicine part 3: positive and negative likelihood ratios of diagnostic tests. *Emerg (Tehran)*. 2015;3(4):170-1.
35. Fierz W. Likelihood ratios of quantitative laboratory results in medical diagnosis: the application of Bézier curves in ROC analysis. *PLoS One*. 2018 Feb 22;13(2):e0192420. <https://doi.org/10.1371/journal.pone.0192420>
36. McGinn T, Jervis R, Wisnivesky J, Keitz S, Wyer PC; Evidence-based Medicine Teaching Tips Working Group. Tips for teachers of evidence-based medicine: clinical prediction rules (CPRs) and estimating pretest probability. *J Gen Intern Med*. 2008 Aug;23(8):1261-8. <https://doi.org/10.1007/s11606-008-0623-z>
37. Parikh R, Parikh S, Arun E, Thomas R. Likelihood ratios: clinical application in day-to-day practice. *Indian J Ophthalmol*. 2009 May-Jun;57(3):217-21. <https://doi.org/10.4103/0301-4738.49397>
38. Fierz W, Bossuyt X. Likelihood ratio approach and clinical interpretation of laboratory tests. *Front Immunol*. 2021 Apr 16;12:655262. <https://doi.org/10.3389/fimmu.2021.655262>
39. Grimes DA, Schulz KF. Refining clinical diagnosis with likelihood ratios. *Lancet*. 2005 Apr 23-29;365(9469):1500-5. [https://doi.org/10.1016/S0140-6736\(05\)66422-7](https://doi.org/10.1016/S0140-6736(05)66422-7)
40. Silva Fuente-Alba C, Molina Villagra M. Likelihood ratio (razón de verosimilitud): definición y aplicación en radiología. *Rev Argent Radiol*. 2017 Sep;81(3):204-8. <https://doi.org/10.1016/j.rard.2016.11.002>

41. Akobeng AK. Understanding diagnostic tests 2: likelihood ratios, pre- and post-test probabilities and their use in clinical practice. *Acta Paediatr.* 2007 Apr;96(4):487-91. <https://doi.org/10.1111/j.1651-2227.2006.00179.x>
42. Weatherall M. Information provided by diagnostic and screening tests: improving probabilities. *Postgrad Med J.* 2018 Apr;94(1110):230-5. <https://doi.org/10.1136/postgradmedj-2017-135273>
43. Habibzadeh F, Habibzadeh P. The likelihood ratio and its graphical representation. *Biochem Med (Zagreb).* 2019 Jun 15;29(2):020101. <https://doi.org/10.11613/BM.2019.020101>
44. Uy EJB. Key concepts in clinical epidemiology: estimating pre-test probability. *J Clin Epidemiol.* 2022 Apr;144:198-202. <https://doi.org/10.1016/j.jclinepi.2021.10.022>
45. van Norman ER, Klingbeil DA, Nelson PM. Posttest probabilities: an empirical demonstration of their use in evaluating the performance of universal screening measures across settings. *School Psychol Rev.* 2017 Dec 1;46(4):349-62. <https://doi.org/10.17105/SPR-2017-0046.V46-4>
46. Fagan TJ. Letter: Nomogram for Bayes theorem. *N Engl J Med.* 1975 Jul 31;293(5):257. <https://doi.org/10.1056/NEJM197507312930513>
47. Abushouk AI. Evolution of Fagan's nomogram; a commentary. *Emerg (Tehran).* 2016 Summer;4(3):114-5.
48. Caraguel CG, Vanderstichel R. The two-step Fagan's nomogram: ad hoc interpretation of a diagnostic test result without calculation. *Evid Based Med.* 2013 Aug;18(4):125-8. <https://doi.org/10.1136/eb-2013-101243>
49. Safari S, Baratloo A, Elfil M, Negida A. Evidence based emergency medicine; part 4: pre-test and post-test probabilities and Fagan's nomogram. *Emerg (Tehran).* 2016 Winter;4(1):48-51.
50. Ma Z, Chong HY, Liao PC. Real-time safety inspection and planning: a first application of the Fagan nomogram. *Can J Civil Eng.* 2020;47(4):438-49. <https://doi.org/10.1139/cjce-2018-0500>
51. Huang J, Liu M, He W, Liu F, Cheng J, Wang H. Use of the A2DS2 scale to predict morbidity in stroke-associated pneumonia: a systematic review and meta-analysis. *BMC Neurol.* 2021;21(33). <https://doi.org/10.1186/s12883-021-02060-8>



Esta obra está bajo una licencia de Creative Commons Reconocimiento-NoComercial 4.0 Internacional