

# EVIDENCIAS EN PEDIATRÍA

Toma de decisiones clínicas basadas en las mejores pruebas científicas  
[www.evidenciasenpediatria.es](http://www.evidenciasenpediatria.es)

## Fundamentos de medicina basada en la evidencia

### Evaluación de la validez de las pruebas diagnósticas (I). Sensibilidad. Especificidad

Molina Arias M<sup>1</sup>, Ochoa Sangrador C<sup>2</sup>

<sup>1</sup>Servicio de Gastroenterología. Hospital Infantil Universitario La Paz. Madrid (España).

<sup>2</sup>Servicio de Pediatría. Hospital Virgen de la Concha. Zamora (España).

Correspondencia: Manuel Molina Arias, [mma1961@gmail.com](mailto:mma1961@gmail.com)

---

**Palabras clave en inglés:** errors; bias; confounding; effect modifiers.

**Palabras clave en español:** errores; sesgos; confusión; modificación de efecto.

**Fecha de recepción:** 1 de mayo de 2016 • **Fecha de aceptación:** 5 de mayo de 2016  
**Fecha de publicación del artículo:** 11 de mayo de 2016

---

Evid Pediatr. 2016;12:34.

#### CÓMO CITAR ESTE ARTÍCULO

Molina Arias M, Ochoa Sangrador C. Evaluación de la validez de las pruebas diagnósticas (I). Sensibilidad. Especificidad. Evid Pediatr. 2016;12:34.

Para recibir Evidencias en Pediatría en su correo electrónico debe darse de alta en nuestro boletín de novedades en  
<http://www.evidenciasenpediatria.es>

---

Este artículo está disponible en: [artículo está disponible en: http://www.evidenciasenpediatria.es/EnlaceArticulo?ref=2015;12:34](http://www.evidenciasenpediatria.es/EnlaceArticulo?ref=2015;12:34)  
©2005-16 • ISSN: 1885-7388

# Evaluación de la validez de las pruebas diagnósticas (I). Sensibilidad. Especificidad

Molina Arias M<sup>1</sup>, Ochoa Sangrador C<sup>2</sup>

<sup>1</sup>Servicio de Gastroenterología. Hospital Infantil Universitario La Paz. Madrid (España).

<sup>2</sup>Servicio de Pediatría. Hospital Virgen de la Concha. Zamora (España).

Correspondencia: Manuel Molina Arias, mma1961@gmail.com

Realizar diagnósticos es una de las actividades cotidianas de la práctica médica. Normalmente tenemos un paciente con un conjunto de síntomas que nos hace sospechar que padezca determinada enfermedad, por lo que realizaremos la prueba diagnóstica indicada para esa situación, cuyo resultado nos ayudará a definir si padece o no la enfermedad.

Clásicamente se aplicaba un abordaje diagnóstico simplista, según el cual el paciente se consideraba enfermo o sano en función de que el resultado de la prueba fuese positivo o negativo. Sin embargo, este abordaje debe considerarse erróneo. En Medicina nos hallamos inmersos en un mundo de incertidumbre en el que asignar un diagnóstico a una persona es equivalente a decir, con una razonable probabilidad, que pertenece a una clase de sujetos con una entidad nosológica determinada.

Nunca podremos estar seguros del diagnóstico al 100%, ya que ninguna prueba diagnóstica es infalible y todas pueden tener falsos positivos y falsos negativos. El paciente tiene una probabilidad preprueba de padecer la enfermedad en función de la población a la que pertenece, la sintomatología que refiere, etc., y esta probabilidad preprueba se ve modificada por el resultado de la prueba diagnóstica hasta un valor denominado probabilidad posprueba.

Por lo tanto, hemos de acostumbrarnos a movernos dentro del abordaje probabilístico: el paciente tiene una probabilidad mayor o menor de tener el diagnóstico.

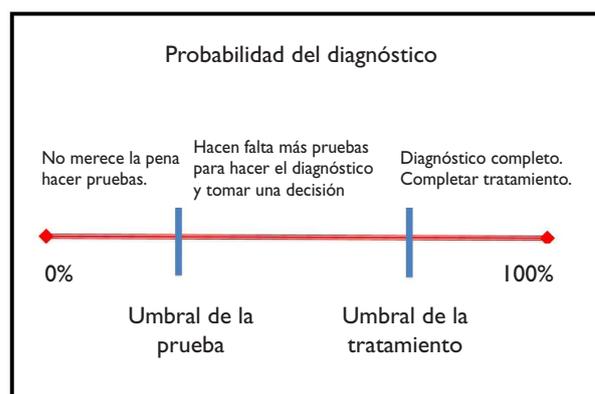
## ENTRE DOS UMBRALES

Habitualmente nos movemos entre los dos umbrales de probabilidad representados en la figura 1: el umbral de prueba o de diagnóstico y el umbral de tratamiento o terapéutico<sup>1</sup>.

Si la probabilidad preprueba de tener la enfermedad es inferior al umbral de prueba, la probabilidad del diagnóstico es tan baja que no merece la pena no ya tratar la enfermedad, sino tan siquiera preocuparse de hacer pruebas diagnósticas.

Si la probabilidad preprueba es superior al umbral terapéutico, la probabilidad de la enfermedad es tan alta que no hace falta realizar más pruebas y podemos comenzar el tratamiento específico de la enfermedad.

FIGURA 1. ESQUEMA REPRESENTANDO LOS UMBRALES DE PRUEBA Y TRATAMIENTO. MODIFICADO DE PAUKER ET AL.<sup>1</sup>.

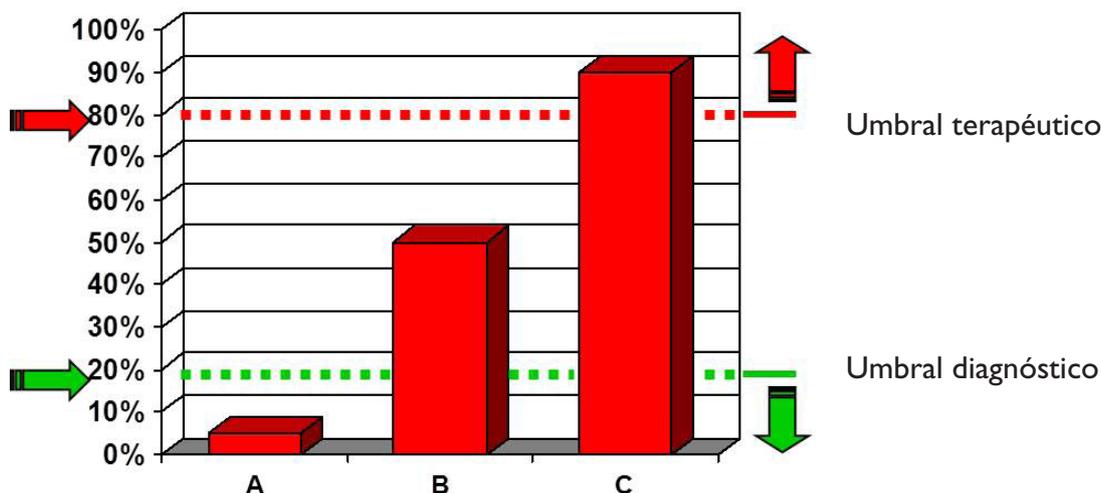


Quizás la situación más habitual es la de encontrarnos entre los dos umbrales. La probabilidad de enfermedad es suficiente como para preocuparnos con el tema, pero no tanto como para justificar el inicio del tratamiento. En este caso recurriremos a una prueba diagnóstica, con la que obtendremos una probabilidad posprueba que modificará nuestra posición entre los dos umbrales. Si caemos por debajo del umbral de prueba, descartaremos la enfermedad. Si subimos por encima del umbral terapéutico, iniciaremos el tratamiento. Si seguimos entre los dos umbrales, volveremos a iniciar el proceso.

Veamos un ejemplo en la figura 2. El paciente A es un escolar con fiebre de pocas horas de duración. La sospecha de, por ejemplo, una infección de orina es baja, situándonos por debajo del umbral de prueba, por lo que recomendamos una actitud expectante. El paciente B es el mismo, pero que ha desarrollado un síndrome miccional (llanto con la micción). En este caso la probabilidad preprueba de infección urinaria es mayor, pero no lo suficiente como para justificar el tratamiento. En consecuencia, obtenemos un cultivo de orina. El paciente C es el mismo, pero con un urocultivo positivo: no necesitamos más pruebas, estamos por encima del umbral terapéutico e iniciamos el tratamiento.

Como es lógico, los niveles en que se sitúan los umbrales diagnóstico y terapéutico dependerán de la frecuencia y relevancia de la enfermedad, de su pronóstico y respuesta al tratamiento, de la toxicidad del tratamiento, etc.

**FIGURA 2. REPRESENTACIÓN DE LOS UMBRALES DIAGNÓSTICO Y TERAPÉUTICO JUNTO CON LAS PROBABILIDADES PRE Y POSPRUEBA**



Es importante resaltar que el dato fundamental para la toma de decisiones en el proceso diagnóstico es la probabilidad posprueba, que depende de la probabilidad preprueba y de la capacidad de la prueba diagnóstica. El problema es que, como veremos más adelante, no es proporcionada de forma directa por los parámetros diseñados para evaluar el desempeño de las pruebas diagnósticas.

**SENSIBILIDAD Y ESPECIFICIDAD**

Sensibilidad y especificidad son dos parámetros bien conocidos por los clínicos, pero no por ello bien comprendidos<sup>2,3</sup>.

Cuando queremos estudiar el rendimiento de la prueba diagnóstica lo que hacemos es comparar el resultado de esa prueba en un grupo de individuos con el resultado en esos mismos individuos de otra prueba que se considera el patrón de referencia (*gold standard*). Con estos resultados se construye la tabla de contingencia (tabla 1), que nos muestra la distribución de sanos y enfermos y el resultado de la prueba de estudio. Clasificamos así los pacientes como verdaderos positivos, verdaderos negativos, falsos positivos (prueba de referencia negativa con prueba de estudio positiva) y falsos negativos (prueba de referencia positiva con prueba en estudio negativa). Será esta tabla de contingencia la que utilizaremos para calcular los diferentes parámetros que estudian el comportamiento de la prueba que estamos valorando.

Para entenderlo mejor, vamos a definir estos parámetros utilizando un ejemplo ficticio. Supongamos que tenemos una población de 100 escolares con síndrome miccional en los que queremos valorar la utilidad de la tira reactiva para el diagnóstico de la infección urinaria. En este caso seleccionamos el cultivo de orina como el patrón de referencia (asumiendo que el cultivo no falla nunca a la hora de clasificar sanos y enfermos) y realizamos ambas pruebas en los 100 niños. Como resultado, obtenemos un cultivo de orina positivo en 30 de los 100, con lo que establecemos la prevalencia de infección en nuestra población en el 30% (0,3). Esta es la probabilidad preprueba en nuestra población: si seleccionamos uno de los niños al azar, la probabilidad de que tenga una infección urinaria será de 0,3 (figura 3). Sin embargo, tal como se muestra en la figura 4, estos resultados no coinciden plenamente con los obtenidos al realizar la tira reactiva, que da positiva en 16 niños sin infección urinaria y negativa en cuatro niños con infección de orina (tabla 2).

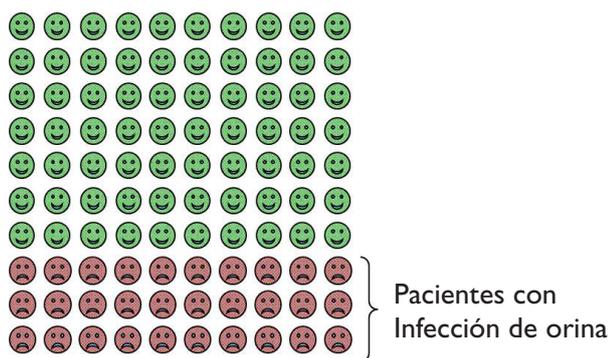
Definamos ahora el primer parámetro. La sensibilidad representa la probabilidad de clasificar correctamente a los enfermos o, lo que es lo mismo, la proporción de verdaderos positivos. Es una proporción en la que en el denominador se sitúa el total de enfermos y en el numerador los positivos verdaderos:

$$\text{Sensibilidad} = \text{verdaderos positivos} / \text{total de enfermos}$$

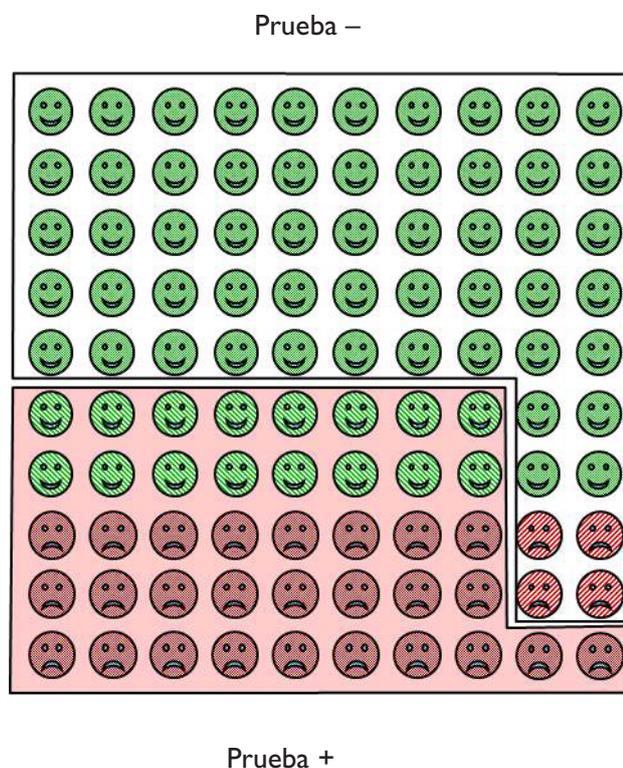
**TABLA 1. TABLA DE CONTINGENCIA CON LA COMPARACIÓN DE LOS RESULTADOS OBTENIDOS CON EL PATRÓN DE REFERENCIA Y LA PRUEBA EN ESTUDIO**

	Referencia positiva	Referencia negativa	
Prueba positiva	Verdaderos positivos	Falsos positivos	Total de positivos
Prueba negativa	Falsos negativos	Verdaderos negativos	Total de negativos
	Total de enfermos	Total de sanos	Total de pacientes

**FIGURA 3. RESULTADOS DEL CULTIVO DE ORINA EN NUESTRA POBLACIÓN**



**FIGURA 4. COMPARACIÓN DE LOS RESULTADOS DE LA TIRA REACTIVA Y EL UROCULTIVO EN NUESTRA POBLACIÓN**



,Por su parte, la especificidad representa la probabilidad de clasificar correctamente a los sanos o, lo que es lo mismo, la proporción de verdaderos negativos. Es una proporción en la que en el denominador figuran el total de sanos y en el numerador los negativos verdaderos:

$$\text{Especificidad} = \text{verdaderos negativos} / \text{total de sanos}$$

En nuestro ejemplo la sensibilidad sería de  $26/30 = 0,86$  (o del 86%) y la especificidad de  $54/70 = 0,77$  (o del 77%). Esto quiere decir que un niño con infección tiene un 86% de probabilidades de tener un cultivo positivo, mientras que uno sin infección tiene un 77% de probabilidades de tener un cultivo negativo.

Ambos parámetros, sensibilidad y especificidad, son propiedades intrínsecas de la prueba diagnóstica que prácticamente no se influyen por la prevalencia de la enfermedad. De esta manera, si realizamos la misma prueba en las mismas condiciones en poblaciones distintas, los valores de sensibilidad y especificidad son similares en las dos poblaciones, aunque la prevalencia de la enfermedad sea diferente. Lo que sí influye en los valores de los parámetros es la gravedad de la enfermedad. Lógicamente, es más fácil diagnosticar los casos más graves de enfermedad, por lo que la sensibilidad puede ser más alta en los casos graves que en los leves. Por esto se comprende la necesidad de que en los estudios de pruebas diagnósticas se incluyan enfermos en distintos estadios de la enfermedad que se diagnostica.

¿Cuál es el problema? El problema es que la sensibilidad nos dice la probabilidad de clasificar correctamente al enfermo **una vez que sabemos que está enfermo**. Por su parte, la especificidad nos dice la probabilidad de clasificar correctamente al sano pero **una vez que ya conocemos que está sano**. Sin embargo, en la práctica esto es lo que desconocemos, el estado de salud o enfermedad del paciente. Es por esto que el conocer los valores de sensibilidad y especificidad de una prueba no nos permiten por separado estimar la probabilidad del diagnóstico y por lo tanto saber si hemos traspasado los umbrales diagnóstico o terapéutico, incluso con valores de sensibilidad o especificidad próximos al 100%. Pensemos, por ejemplo, en una prueba que siempre dé resultado positivo. Su sensibilidad será del 100% (siempre clasifica a los enfermos como tales), pero su utilidad para el diagnóstico será nula<sup>4</sup>. Lo mismo ocurriría con una prueba que siempre dé negativa: tendrá una especificidad del 100% pero no modificará en absoluto el valor de la probabilidad preprueba que teníamos antes de realizarla.

**TABLA 2. COMPARACIÓN DE LOS RESULTADOS OBTENIDOS MEDIANTE TIRA REACTIVA Y CULTIVO DE ORINA EN NIÑOS CON SOSPECHA DE INFECCIÓN URINARIA.**

	Cultivo positivo	Cultivo negativo	
Tira positiva	26	16	42
Tira negativa	4	54	58
	30	70	100

Cuando una prueba diagnóstica tiene tanto la sensibilidad como la especificidad cercanas al 100% se comporta como una prueba de referencia y por lo tanto sus resultados serán casi siempre válidos. Sin embargo, esta circunstancia es excepcional, por lo que para estimar la probabilidad posprueba del diagnóstico necesitaremos otros parámetros que describiremos en próximas entradas de esta serie de Fundamentos de Medicina Basada en la Evidencia, como son los valores predictivos y los cocientes de probabilidades.

## BIBLIOGRAFÍA

---

1. Pauker SG, Kassirer JP. The threshold approach to clinical decision making. *N Engl J Med.* 1980;302:1109-17.
2. Dermirdjian G, Berlín V, Rowensztein H. Pediatría basada en la evidencia. Estudios de diagnóstico (I.ª parte). *Arch Argent Pediatr.* 2009;107:527-35.
3. Sensibilidad y especificidad. En: Argimón JM, Jiménez J (eds.). *Métodos de investigación clínica y epidemiológica*, 3ª ed. Elsevier España SA. Madrid, 2004. p. 335-40.
4. Loong TW. Understanding sensitivity and specificity with the right side of the brain. *Br Med J.* 2003;327:716-9.