

# EVIDENCIAS EN PEDIATRÍA

Toma de decisiones clínicas basadas en las mejores pruebas científicas

[www.evidenciasenpediatria.es](http://www.evidenciasenpediatria.es)

## Artículos Valorados Críticamente

### Las estimaciones de validez de las pruebas diagnósticas varían según la prevalencia en los estudios

Ruiz-Canela Cáceres J<sup>1</sup>, García Vera C<sup>2</sup>

<sup>1</sup>CS Virgen de África. Sevilla (España).

<sup>2</sup>CS Sagasta-Ruiseñores. Zaragoza (España).

Correspondencia: Juan Ruiz-Canela Cáceres, [jruizcanela@gmail.com](mailto:jruizcanela@gmail.com)

**Palabras clave en inglés:** diagnostic techniques and procedures; sensitivity and specificity; research design.

**Palabras clave en español:** técnicas y procedimientos diagnósticos; sensibilidad y especificidad; proyectos de investigación.

Fecha de recepción: 16 de octubre de 2013 • Fecha de aceptación: 28 de octubre de 2013

Fecha de publicación del artículo: 18 de diciembre de 2013

Evid Pediatr.2014;10:8.

#### CÓMO CITAR ESTE ARTÍCULO

Ruiz-Canela Cáceres J, García Vera C. Las estimaciones de validez de las pruebas diagnósticas varían según la prevalencia en los estudios. Evid Pediatr. 2014;10:8.

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: <http://www.evidenciasenpediatria.es/EnlaceArticulo?ref=2014;10:8>.

©2005-14 • ISSN: 1885-7388

# Las estimaciones de validez de las pruebas diagnósticas varían según la prevalencia de los estudios

Ruiz-Canela Cáceres J<sup>1</sup>, García Vera C<sup>2</sup>

<sup>1</sup>CS Virgen de África. Sevilla (España).

<sup>2</sup>CS Sagasta-Ruiseñores. Zaragoza (España).

Correspondencia: Juan Ruiz-Canela Cáceres, jruizcanela@gmail.com

Referencia bibliográfica: Leeflang MM, Rutjes AW, Reitsma JB, Hooft L, Bossuyt PM. Variation of a test's sensitivity and specificity with disease prevalence. CMAJ. 2013;185:E537-44.

## Resumen

**Conclusiones de los autores:** los factores que hacen que la prevalencia sea diferente en determinadas poblaciones también provocan que varíen la especificidad y la sensibilidad de sus pruebas diagnósticas.

**Comentario de los revisores:** con datos de 23 metanálisis diferentes (que agrupan 416 estudios), los autores demuestran que las diferencias en la prevalencia de las enfermedades pueden condicionar cambios significativos en la sensibilidad y especificidad de las pruebas que las diagnostican.

**Palabras clave:** técnicas y procedimientos diagnósticos; sensibilidad y especificidad; proyectos de investigación.

## Estimates of the validity of diagnostic tests vary with the prevalence of the studies

### Abstract

**Authors' conclusions:** the factors, that are responsible for the differences in the prevalence between certain populations, also vary the specificity and sensitivity of their diagnostic tests.

**Reviewers' commentary:** with data from 23 different meta-analyses (grouping 416 studies) the authors prove that differences in the prevalence of diseases may lead to significant changes in the sensitivity and specificity of the tests that are used for their diagnostic.

**Palabras clave:** diagnostic techniques and procedures; sensitivity and specificity; research design.

## RESUMEN ESTRUCTURADO

**Objetivo:** documentar la relación entre la prevalencia y la sensibilidad (Sn) y especificidad (Es) en los estudios de pruebas diagnósticas.

**Diseño:** análisis de revisiones sistemáticas sobre estudios de validez diagnóstica.

**Fuente de datos:** revisiones sistemáticas (RS) publicadas entre enero de 1999 y abril de 2002 en MEDLINE, EMBASE, DARE y la MEDION database of the University of Maastricht. Los descriptores utilizados fue-

ron: metanálisis, revisión, sensibilidad y especificidad, diagnóstico por imagen, pruebas diagnósticas de rutina. Se excluyeron los trabajos no publicados en inglés, alemán, francés y holandés.

**Selección de estudios:** las RS seleccionadas debían tener como objetivo la estimación de la validez de una prueba diagnóstica, incluir al menos diez estudios originales sobre la misma prueba, no excluir estudios primarios en función de las características del diseño del estudio y ofrecer la posibilidad de reproducir las tablas de 2 x 2 de los estudios originales. Por la imposibilidad de calcular la prevalencia de la enfermedad, se excluye-

ron los estudios de casos y controles. La selección se realizó de forma independiente por dos de los autores. Para evaluar la calidad de los estudios, se estableció una lista de nueve criterios en función de las posibles fuentes de sesgo o variación.

**Extracción de datos:** para cada uno de los metaanálisis elegibles se estableció el modelo para generar las estimaciones resumen de Sn y Es. Para evaluar la asociación entre prevalencia y Sn y Es, se incluyó la prevalencia como covariable continua en el modelo. Para el análisis de heterogeneidad, se usó el modelo bivariante que permite correlacionar Sn y Es en el estudio y entre los diferentes estudios. Se agrupan los efectos a través del método de la varianza inversa. Al modelo se le añaden los siguientes factores relacionados: ubicación, patrón de referencia sobre el paciente (síntomas o resultados de alguna prueba), reclutamiento consecutivo, exclusión de pacientes con aspectos de difícil diagnóstico, verificación diagnóstica y verificación parcial. Se comparó la bondad de ajuste del modelo con la prevalencia como covariable con otros modelos que incluyeran una covariable diferente para los mismos datos.

**Resultados principales:** finalmente, se trabaja sobre 23 metaanálisis de diferentes pruebas diagnósticas que recogían un total de 416 estudios primarios. En 8 de los 23 metaanálisis (35%) se observa una asociación significativa entre prevalencia y *logit* de Sn, *logit* de Es o de ambos. En todos los casos, a mayor prevalencia se derivaba menor Es (por cada punto porcentual de aumento de prevalencia la Es disminuiría entre 0,1 y 0,5 puntos porcentuales). En los dos metaanálisis con una asociación significativa entre prevalencia y Sn, la mayor prevalencia se relacionó con mejor Sn de la prueba, aunque en el global de estudios no hubo una asociación estadísticamente significativa entre prevalencia y Sn. En seis de las ocho revisiones en las que la prevalencia se asoció al *logit* de Es, la bondad de ajuste fue mejor con la prevalencia como covariable que con otras covariables.

**Conclusión:** la Es y la Sn de una prueba varían con la prevalencia de la enfermedad, porque los factores que determinan el cambio de prevalencia (por ejemplo, las características de la población) también influyen en aquellas. Los clínicos deberían utilizar la prevalencia como guía para hacer inferencia de estudios de pruebas diagnósticas a su población.

**Conflicto de intereses:** no se declaran, pero no parece haber posibilidad de que los haya ni de que influyan.

**Fuente de financiación:** dos autoras reciben apoyo económico de The Netherlands Organization for Scientific Research (NWO) (proyecto 916.10.034 y

beca de investigación número de registro 945-10-012). La entidad financiadora no participa en ninguna de las etapas del estudio desde el diseño hasta su publicación.

## COMENTARIO CRÍTICO

**Justificación:** la validez diagnóstica es fundamental en la evaluación de las pruebas diagnósticas. La validez de una prueba puede ser expresada como Sn y Es, o valores de cocientes de probabilidades (CP) positivo y negativo. Los CP son más útiles en la clínica que la Sn y la Es<sup>1</sup>, ya que permiten estimar la probabilidad postprueba para escenarios con prevalencias diferentes. No obstante, la prevalencia de la enfermedad puede influir en las estimaciones de estos indicadores de validez, porque sus variaciones pueden reflejar diferencias en los criterios de clasificación de la enfermedad o sesgos de selección (exclusión de pacientes dudosos y sesgos de verificación) o clasificación (sesgos de incorporación). El estudio que analizamos intenta demostrar que la Sn y la Es pueden variar dependiendo de la prevalencia de la enfermedad.

**Validez:** se trata de una RS de pruebas diagnósticas. Esta revisión se basa en una pregunta clínica claramente definida. La metodología utilizada para la selección y la búsqueda sistemática es adecuada y se describe de forma más amplia en una publicación previa<sup>2</sup>. La evaluación de los estudios individuales está bien descrita y es correcta. Los resultados de la RS están bien sintetizados y descritos, así como la metodología utilizada en el análisis de heterogeneidad. Las conclusiones están justificadas y son útiles. Se trata de un estudio de buena calidad<sup>3</sup>. Sin embargo, los estudios incluidos datan de enero de 1999 a abril de 2002, y podrían tener criterios actualmente no vigentes, mientras que la heterogeneidad entre los distintos escenarios valorados limita el origen de las variaciones observadas en la validez de los estudios.

**Importancia clínica:** en este estudio se aporta que el *logit* de Es decrece una media de 0,02 unidades (intervalo de confianza del 95% -0,03 a -0,01), por cada punto de incremento de la prevalencia. Esto se corresponde con una disminución de la Es del 0,1-0,5%. El efecto es mayor para Es cercanas al 50% y menor en las cercanas al 95%. Las variaciones observadas pueden ser cuantitativamente importantes, especialmente para algunas pruebas diagnósticas. Los clínicos que trasladen a la práctica los valores de los CP para calcular las probabilidades postprueba han de tener en cuenta la prevalencia (probabilidad preprueba) no solo para la realización de los cálculos, sino también para la selección de los CP más apropiados. Existen estudios publicados que aportan explicaciones fundamentalmente teóricas

en función del espectro de pacientes, sin tomar en consideración la prevalencia<sup>4</sup>. El hallazgo de que la prevalencia puede influir en las estimaciones de validez de las pruebas diagnósticas ofrece una pista para identificar defectos metodológicos en los estudios de evaluación de dichas pruebas.

**Aplicabilidad:** estos resultados tienen implicación para los clínicos que deben seleccionar los indicadores de validez a emplear en la toma de decisiones, y además para los investigadores que vayan a evaluar pruebas diagnósticas que tendrán que considerar si los pacientes presentan diferentes estadios o patrones de enfermedad para realizar una valoración de la Sn y la Es diferenciadas o ajustadas. Siempre se deberá tratar de localizar los estudios de valoración de pruebas diagnósticas de mayor calidad<sup>5</sup> y que más se adecuen a nuestra práctica clínica.

**Conflicto de intereses de los autores del comentario:** no existe.

## BIBLIOGRAFÍA

1. Ochoa Sangrador C, González de Dios J, Buñuel Álvarez JC. Evaluación de artículos científicos sobre pruebas diagnósticas. *Evid Pediatr.* 2007;3:24.
2. Rutjes AW, Reitsma JB, Di Nisio M, Smidt N, van Rijn JC, Bossuyt PM. Evidence of bias and variation in diagnostic accuracy studies. *CMAJ.* 2006;174:469-76.
3. Macaskill P, Gatsonis C, Deeks JJ, Harbord RM, Takwoingi Y. Chapter 10: Analysing and Presenting Results. En: Deeks JJ, Bossuyt PM, Gatsonis C (ed.), *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy* Version 1.0. The Cochrane Collaboration, 2010.
4. Mulherin SA, Miller WC. Spectrum bias or spectrum effect? Subgroup variation in diagnostic test evaluation. *Ann Intern Med.* 2002;137:598-602.
5. González Rodríguez MP, Velarde Mayol C. Listas guía de comprobación de estudios sobre pruebas diagnósticas incluidos en las revisiones sistemáticas: declaración QUADAS. *Evid Pediatr.* 2012;8:20.