

# 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

### Lista de comprobación de estudios sobre precisión de pruebas diagnósticas: declaración STARD

González Rodríguez MP<sup>1</sup>, Velarde Mayol C<sup>2</sup>

<sup>1</sup>CS Barrio del Pilar. Área 5. Madrid. (España).

<sup>2</sup>CS de Segovia 1. Segovia. (España).

Correspondencia: María Paz González Rodríguez, [pazgonz@gmail.com](mailto:pazgonz@gmail.com)

---

**Palabras clave en inglés:** pruebas diagnósticas; estándares; listas guía; utilización; listas guía; métodos; listas guía; estándares

**Palabras clave en español:** diagnostic tests/routine standards; checklist; utilization; checklist; methods; checklist; standards

**Fecha de recepción:** 28 de mayo de 2012 • **Fecha de aceptación:** 29 de mayo de 2012

**Fecha de publicación del artículo:** 31 de mayo de 2012

---

Evid Pediatr. 2012;8:43.

#### CÓMO CITAR ESTE ARTÍCULO

González Rodríguez MP, Velarde Mayol C. Lista de comprobación de estudios sobre precisión de pruebas diagnósticas: declaración STARD. Evid Pediatr. 2012;8:43.

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=2012;8:43>

©2005-12 • ISSN: 1885-7388

# Lista de comprobación de estudios sobre precisión de pruebas diagnósticas: declaración STARD

González Rodríguez MP<sup>1</sup>, Velarde Mayol C<sup>2</sup>

<sup>1</sup>CS Barrio del Pilar. Área 5. Madrid. (España).

<sup>2</sup>CS de Segovia 1. Segovia. (España).

Correspondencia: María Paz González Rodríguez, pazgonz@gmail.com

Cuando un clínico solicita una prueba diagnóstica, se enfrenta a la labor de interpretar sus resultados, para elaborar un diagnóstico, un tratamiento y establecer un pronóstico<sup>1</sup>. Este proceso es de suma importancia para el clínico y sobre todo, para el paciente.

A la hora de evaluar una nueva prueba diagnóstica, es necesario analizar tanto su precisión como su calidad metodológica. Para evaluar la calidad se ha desarrollado la escala de valoración QUADAS<sup>2</sup>, descrita previamente en Evidencias en Pediatría<sup>3</sup>.

Para determinar su precisión, se comparan los resultados del estudio con los de la prueba de referencia en los mismos pacientes y se expresa en forma de sensibilidad, especificidad o combinaciones de estos índices, como los cocientes de probabilidad o los valores predictivos. En aquellas pruebas en que los resultados se pueden utilizar con distintos puntos de corte, el resultado se expresa como una curva ROC (curva de rendimiento diagnóstico), que refleja la sensibilidad y la especificidad para distintos puntos de corte.

La validez interna y externa de los resultados de una prueba de precisión diagnóstica se pueden alterar por la selección de los pacientes, la realización de la prueba, su interpretación entre otros aspectos<sup>4</sup>. A pesar ello, el desarrollo de herramientas para evaluar la precisión de los estudios diagnósticos ha sido posterior al desarrollo de las herramientas para valorar los estudios de tratamiento o las revisiones sistemáticas.

Con el objetivo de establecer una herramienta para valorar la precisión diagnóstica se desarrolló la declaración STARD (Standards for Reporting of Diagnostic Accuracy)<sup>5</sup>.

## La declaración STARD paso a paso, ítem a ítem

En 1999, el grupo de trabajo de la Cochrane de pruebas diagnósticas se reunió en Roma con el objetivo de mejorar el diseño de los estudios sobre precisión de una prueba diagnóstica. El grupo de trabajo siguió el ejemplo de la iniciativa CONSORT que evalúa los ensayos clínicos, descrito en la serie Fundamentos de Evidencias en Pediatría<sup>5</sup>. Al revisar las publicaciones y bases de datos más relevantes, el grupo de trabajo recuperó 33 listas, extrayendo 75 puntos potenciales. A continuación, se elaboró un documento a partir de una reunión de

consenso, en la que participaron los miembros del propio grupo de trabajo de STARD, investigadores, editores, metodólogos y organizaciones profesionales.

Se elaboró un listado de 25 preguntas y un diagrama de flujo, para que el diseño fuese el adecuado. En el mismo se tienen en cuenta los pasos necesarios, como la inclusión de los pacientes, el orden de la realización de la prueba, el número de pacientes que reciben la prueba y la prueba de referencia seleccionada.

La iniciativa STARD está dirigida a editores de revistas y a los autores de artículos, con el objetivo de poder valorar los sesgos potenciales del estudio (validez interna) y la generalización o aplicabilidad de los resultados (validez externa). La escala y el diagrama de flujo se pueden consultar en: <http://www.stard-statement.org/>

Los resultados se publicaron el 2003 en importantes revistas médicas como *Annals of Internal Medicine*, *Radiology*, *Clinical Chemistry* y *The British Medical Journal*<sup>6</sup>. Varias editoriales recomendaban a los autores y a los editores su aceptación. Su aprobación se basó en que la evaluación de la investigación depende de un informe completo y preciso y que la incorporación de STARD contribuiría a mejorar el proceso<sup>7,8</sup>. Su utilización fue impulsada por el Comité Internacional de Editores de Revistas Médicas, que en el año 2006 estableció su utilización como requisito para la publicación de los trabajos<sup>9</sup>.

Pero su aceptación ha sido variable<sup>10</sup>. Actualmente, los criterios de calidad de STARD se cumplen en un porcentaje bajo de publicaciones. Lo cual no quiere decir que los estudios sean de mala calidad, sino que es necesario describir el proceso, para que se pueda interpretar con transparencia los resultados<sup>11</sup>.

En estudios realizados acerca de algunas publicaciones biomédicas en España, encuentran que apenas la mitad de los artículos evaluados siguen las recomendaciones de la guía STARD<sup>12</sup>. La declaración STARD clasifica la evaluación en cinco secciones, que son las clásicas IMRD de cualquier artículo original (Introducción, Métodos, Resultados y Discusión), con una sección previa (Título y Resumen). Se puede consultar en la tabla I. Una descripción detallada de cada punto está disponible en la página web.

**Tabla 1.** Lista STARD para el informe de los estudios de precisión diagnóstica. Disponible en: STARD checklist

SECCIÓN Y TEMA		
<b>TÍTULO/RESUMEN/ PALABRAS CLAVE</b>	1	Identifica el artículo como un estudio de <b>precisión diagnóstica</b> (se recomienda el termino MeSH "sensibilidad y especificidad").
<b>INTRODUCCIÓN</b>	2	Especifica que las preguntas del estudio, o el objetivo del mismo, es <b>estimar la precisión diagnóstica</b> o comparar la precisión entre diferentes grupos y con otros test.
<b>MÉTODOS</b>		
Participantes	3	Describe la <b>población incluida</b> en el estudio: los criterios de inclusión y exclusión, lugar y localización dónde se realizó la recogida de datos.
	4	Describe cómo se seleccionaron los participantes: ¿la selección se basó en los síntomas iniciales, en los resultados de las pruebas previas, o en el hecho de que los participantes recibiesen la prueba de estudio y la prueba de referencia?
	5	Describe la selección de los participantes ( <b>el muestreo</b> ): ¿la población del estudio fue una serie consecutiva de participantes definida por los criterios de selección de los puntos 3 y 4? Si no es así, especifica cómo se seleccionaron los participantes.
Métodos de las pruebas	6	Describe la <b>recogida de datos</b> : si se planeó previamente antes de realizar la prueba de estudio y la de referencia (estudio prospectivo) o después (estudio retrospectivo).
	7	Describe la prueba de referencia y en que se basa (su justificación).
	8	Describe las cuestiones técnicas del material y métodos, incluyendo cómo y cuándo se tomaron las medidas, y/o cita referencias para la prueba en estudio y la prueba de referencia.
	9	Describe la definición y bases de las unidades, puntos de corte y/o categorías de los resultados de las pruebas en estudio y de la prueba de referencia.
	10	Describe el número, preparación y grado de conocimiento de las <b>personas que realizan e interpretan la prueba</b> de estudio y la de referencia.
Métodos estadísticos	11	Describe si la aplicación de la prueba y del test de referencia se realizó de forma ciega.
	12	Describe los métodos para calcular o comparar las medidas de precisión diagnóstica, y los métodos estadística utilizados para cuantificar la incertidumbre (por ejemplo los intervalos de confianza del 95%).
	13	Describe los métodos para calcular la reproductibilidad de la prueba, si se realizaron (es decir su fiabilidad).
<b>RESULTADOS</b>		
Participantes	14	Describe cuándo se realizó un estudio, incluyendo las <b>fechas</b> de inicio y finalización de inclusión de los participantes.
	15	Describe las <b>características clínicas y demográficas</b> de la población de estudio (por ejemplo edad, sexo, espectro de los síntomas, comorbilidad, tratamientos actuales, centros de reclutamiento).
	16	Describe el número de participantes que cumplen los criterios de inclusión que recibieron o no las pruebas de estudio y/o la prueba de referencia; describe por qué los participantes fallaron en recibir alguno de las pruebas (se recomienda un diagrama de flujo que identifique las pérdidas).
Resultados de las pruebas	17	Describe el <b>intervalo de tiempo</b> entre las prueba de estudio y la prueba de referencia, y si se ha administrado algún tratamiento entre medias.
	18	Describe la <b>distribución de la gravedad de la enfermedad</b> (criterios definidos) en aquellos con la enfermedad en estudio; otros diagnósticos en los participantes sin la enfermedad en estudio.
	19	Describe los resultados en una tabla (incluyendo los indeterminados y los no obtenidos) con los resultados de la prueba de estudio y la de referencia. En las variables continuas, la distribución de los resultados según los valores de la prueba de referencia.
	20	Describe cualquier <b>efecto adverso</b> que pueda haber originado el realizar la prueba de estudio o la de referencia.
Estimaciones	21	Describe las <b>estimaciones de precisión diagnóstica</b> y las medidas de incertidumbre estadística (por ejemplo el intervalo de confianza del 95%).
	22	Describe como se interpretaron los resultados indeterminados, las pérdidas y valores atípicos o extremos de la prueba de estudio.
	23	Describe la estimación de la variabilidad de la precisión diagnóstica entre subgrupos de participantes, investigadores o centros, si se hizo.
	24	Si procede, describe la <b>estimación de la reproductibilidad</b> (fiabilidad) de la prueba.
<b>COMENTARIO (DISCUSIÓN)</b>	25	Comenta la <b>aplicabilidad clínica</b> de los resultados del estudio.

**Tabla 2.** Diferencias entre la declaración STARD y la declaración QUADAS.

STARD	QUADAS-2
Para valorar la precisión de un estudio de pruebas diagnósticas.	Para evaluar la calidad de una prueba diagnóstica.
Aspecto a evaluar: <b>la precisión</b> , es decir la capacidad de una prueba para identificar a los pacientes con la enfermedad.	Aspecto a evaluar: <b>el rigor</b> de los estudios primarios.
Su uso es prospectivo: <ul style="list-style-type: none"> <li>• Para los investigadores en la fase de diseño del estudio.</li> <li>• Para los editores.</li> </ul>	Su uso es retrospectivo: <ul style="list-style-type: none"> <li>• Para las revisiones sistemáticas.</li> <li>• Para los metaanálisis.</li> </ul>
Evalúa principalmente la validez interna: sesgos y variaciones.	Evalúa principalmente su aplicabilidad.

### Diferencias entre STARD y QUADAS

En resumen, la escala STARD se utiliza para valorar la precisión diagnóstica es decir, para discernir la capacidad de una prueba de diferenciar entre pacientes que tienen la enfermedad de los que no la tienen. La herramienta STARD no es una herramienta para valorar la calidad metodológica de los estudios de precisión diagnóstica; para ello se debe utilizar otra lista de calidad como el QUADAS<sup>2,3</sup>. Ambas iniciativas coinciden en la búsqueda de un instrumento que detecte la variación y el sesgo de los estudios de pruebas diagnósticas utilizando la medicina basada en la evidencia.

Difieren entre sí en la intención del instrumento:

- STARD tiene como objetivo el proporcionar una lista que sirva de guía para la publicación de los estudios de precisión de pruebas diagnósticas. Es una herramienta que se utiliza de forma prospectiva para realizar un diseño adecuado de un estudio; por tanto, interesan a los investigadores en la fase de diseño del estudio y a los editores.
- QUADAS es una herramienta para valorar la calidad de los estudios primarios en las revisiones sistemáticas y metaanálisis. Se utiliza de forma retrospectiva para realizar un análisis crítico del rigor metodológico de un estudio de pruebas diagnósticas (Tabla 2). Algunos autores recomiendan que la combinación de ambas aportaría mayor rigor a la hora de validar los estudios de pruebas diagnósticas<sup>13</sup>.

### BIBLIOGRAFÍA

- Zamora J, Abaira V. Análisis de la calidad de los estudios de evaluación de pruebas diagnósticas. *Nefrología*. 2008; 28:42-5.
- QUADAS. A quality assessment tool for diagnostic accuracy studies. University of Bristol. Fecha de consulta 19 mayo 2012. Disponible en: <http://www.bris.ac.uk/quadas/>
- 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.
- Altman DG, Bossuyt PM. Estudios de precisión diagnóstica (STARD) y pronóstica (REMARK). *Med Clin*. 2005;125: 49-55.
- González de Dios J, Buñuel Álvarez JC, González Rodríguez P. Listas guía de comprobación de ensayos clínicos: declaración CONSORT. *Evid Pediatr*. 2011;7:72.
- Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwing LM et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *BMJ*. 2003;326:41-4.
- Rennie D. Improving reports of studies of diagnostic tests. The STARD initiative. *JAMA*. 2003;289:89-90.
- Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwing LM et al. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Ann Intern Med*. 2003;138:1-12.
- Equator network. Centro de recursos para la presentación correcta de informes sobre estudios de investigación sanitaria.
- Smidt N, Rutjes AW, van der Windt DAWM, Ostelo RW, Bossuyt PM, Reitsma JB et al. The quality of diagnostic accuracy studies since the STARD statement. Has it improved? *Neurology*. 2006;67:792-7.
- Selman TJ, Morris RK, Zamora J, Khan KS. The quality of reporting of primary test accuracy studies in obstetrics and gynaecology: application of the STARD criteria. *BMC Womens Health*. 2011;11:8.
- Gómez Sáez NG, Hernández-Aguado I, Lumberras B. Estudio observacional: evaluación de la calidad metodológica de la investigación diagnóstica en España tras la publicación de la guía STARD. *Med Clin*. 2009;133:302-10.
- Fernández de Oliveira MR, De Castro Gomes A, Toscano CM. QUADAS and STARD: evaluating the quality of diagnostic accuracy studies. *Rev Saúde Pública*. 2011;45(2).