

Análisis Crítico de la Literatura

Medwave 2012 Ago;12(7):e5465 doi: 10.5867/medwave.2012.07.5465

Análisis crítico de estudios de pruebas diagnósticas: I

Critical analysis of studies of diagnostic tests: I

Autor: Miguel Araujo Alonso⁽¹⁾

Filiación: ⁽¹⁾Hospital de Los Andes, Chile; Medwave, Santiago, Chile

E-mail: maraujo@medwave.cl

Citación: Araujo M. Critical analysis of studies of diagnostic tests: I. Medwave 2012 Ago;12(7):e5465 doi: 10.5867/medwave.2012.07.5465

Fecha de envío: 27/7/2012

Fecha de aceptación: 27/7/2012

Fecha de publicación: 1/8/2012

Origen: solicitado

Tipo de revisión: sin revisión por pares

Correspondencia a:

⁽¹⁾Medwave, Villaseca 21, Oficina 702, Ñuñoa, Santiago, Chile

Los aspectos metodológicos implicados en la evaluación de las pruebas diagnósticas se encuentran bastante menos estandarizados que en otras áreas, por ejemplo, que en la evaluación de los tratamientos farmacológicos. Tampoco los resultados de estos estudios se pueden interpretar de manera tan directa como aquellos que entregan los ensayos clínicos, o las investigaciones sobre factores de riesgo.

Sin duda, lo ideal sería que cada prueba diagnóstica confirmara positivamente que el paciente tiene la enfermedad en cuestión, o que está libre de ella. Por desgracia, ello no siempre se da. Las pruebas aportan datos que deben ser analizados en contexto con otras variables que también poseen valor diagnóstico, y que muchas veces resultan contradictorias en el terreno práctico: el paciente tiene un dolor sugerente, pero la ecografía es normal.

La información aportada por la prueba tampoco suele ser dicotómica, en términos de apoyar o descartar la presencia de enfermedad. Más bien nos señala que hay cierta probabilidad de que ello ocurra.

Tal como en otras formas de investigación clínica, los resultados de los estudios sobre pruebas diagnósticas se encuentran también expuestos a sesgos que usted deberá saber reconocer, y nuevamente aquí, la tendencia es que el error se manifieste a través de una sobreestimación de la exactitud de la prueba. ¿Qué implica ello? Que probablemente utilizaremos la prueba con mayor frecuencia de lo necesario, y que incurriremos en una mayor proporción de errores diagnósticos en nuestros pacientes, con todo lo que trae aparejado, por ejemplo:

- Efecto de etiquetado: las consecuencias de diagnosticar a un paciente una enfermedad inexistente, aunque sea transitoriamente. Demás está decir lo que puede significar para alguien que se le diagnostique un cáncer o una enfermedad crónica o estigmatizante (ejemplo: VIH) sin que realmente esté enfermo.
- Tratamientos innecesarios.
- Ejecución de nuevos test.
- Omisiones diagnósticas importantes.

Los desarrollos que ha experimentado la investigación de pruebas diagnósticas y las aproximaciones metodológicas a ella no siempre son bienvenidas por la comunidad clínica, que observa una excesiva racionalización en un tema que tiene mucho de "arte" e "intuición".

Como siempre, ningún extremo es bueno. Ambas vertientes del conocimiento tienen su espacio. Lo que importa en definitiva es que el error al clasificar la condición de un paciente específico sea el menor posible. Esta serie le entregará algunas de las herramientas derivadas de la epidemiología clínica para valorar la utilidad de una prueba. El resto corre por su cuenta.

Objetivos de los estudios de pruebas diagnósticas

Las distintas opciones de diseño existentes para evaluar las pruebas diagnósticas se relacionan, en primer lugar, con los objetivos que se plantea el autor del estudio. Esos objetivos pueden ir desde evaluar las características operacionales del test en condiciones de laboratorio, hasta conocer el impacto que el uso del test tiene sobre los desenlaces clínicos del paciente.

a) En las fases iniciales del desarrollo de la prueba, las preguntas se relacionan con su desempeño "técnico", por ejemplo, en términos de resolución espacial o su capacidad de discriminación *in vitro*.

b) Ya en el terreno clínico, una primera etapa es aquella en que uno se plantea si los resultados de la prueba serán distintos en los pacientes con enfermedad y en las personas sanas. El uso rutinario que hacemos de muchos exámenes, y el modo en que hemos "aprendido" acerca de ellos hace que normalmente no nos cuestionemos al respecto. No obstante, más de alguna vez usted se habrá preguntado si determinado hallazgo pudiera ser "normal", es decir, que quizás ocurre con similar frecuencia, tanto en sujeto enfermos o sintomáticos, como en aquellos que no lo están. Sabido es, por ejemplo, que una alta proporción de personas después de cierta edad presenta alteraciones radiológicas de columna que no se manifiestan por ningún síntoma o limitación funcional.

Para contestar este tipo de preguntas se debe reclutar un grupo de pacientes con enfermedad conocida, y otro de sujetos que, también es sabido, no la tienen.

Aoki Y, Saida T, Nakano I, Saito T, Ikeguchi K, Urabe T, Nishiguchi E, Suzuki H, Takahashi K, Mizuno Y. [Determination of medullasin levels for the diagnosis of multiple sclerosis] *Rinsho Byori*. 2002 Mar;50(3):308-12.

Gasparini C, Bertolotto M, Croce SL, Perrone R, Quiaia E, Tiribelli C. Evaluation of liver parenchymal blood flow with contrast-enhanced US: preliminary results in healthy and cirrhotic patients. *Acad Radiol*. 2003 Aug;10(8):869-76.

Este tipo de estudio nos puede permitir, además, identificar condiciones específicas del resultado del test que podrían ser las que realmente discriminan desde el punto de vista diagnóstico (por ejemplo, qué signos radiológicos o patrones de respuesta son los aparentemente más valiosos, sobre o bajo qué valor podríamos considerar el resultado como anormal, etc.).

Shafik A, Shafik AA, El-Sibai O, Ahmed I. Electrocavernosogram in erectile dysfunction: a diagnostic tool. *Arch Androl*. 2004 Sep-Oct;50(5):317-25.

c) Si está suficientemente claro que el test tiene un resultado diferente en los sujetos sanos y enfermos, ya sea porque existe un estudio publicado al respecto, o porque es obvio o conocido que así es, podemos preguntarnos entonces cuál es su desempeño si se aplica en un grupo de pacientes con sospecha de enfermedad. Es decir, bajo condiciones similares a las que uno enfrentaría en la práctica.

Este tipo de preguntas se responden con el diseño "clásico" de los estudio de pruebas diagnósticas, contrastando los resultados del test (o prueba índice) con los de un patrón de oro o referencia. Más adelante en estas series nos dedicaremos de lleno a analizar las condiciones que otorgan –o restan– validez a sus resultados, que se expresan mediante las conocidas

sensibilidad, especificidad, valores predictivos, y otras medidas de rendimiento diagnóstico.

Froehling DA, Elkin PL, Swensen SJ, Heit JA, Pankratz VS, Ryu JH. Sensitivity and specificity of the semiquantitative latex agglutination D-dimer assay for the diagnosis of acute pulmonary embolism as defined by computed tomographic angiography. *Mayo Clin Proc*. 2004 Feb;79(2):164-8.

d) Si el test efectivamente permite avanzar en el diagnóstico, la pregunta que sigue es si esa información afecta de modo significativo el manejo del paciente. ¿Cambian las decisiones terapéuticas con ese diagnóstico más preciso? Imagine el caso extremo de un paciente con un accidente vascular cerebral con compromiso neurológico profundo y fuera de alcance quirúrgico. Conocer las características más específicas de la lesión, ¿modifica la conducta terapéutica?

van Lankveld MA, Peeters PH, van Eijkeren MA, Koot VC, Witteveen PO, Mali WP. The value of abdominal CT scans in decision-making during chemotherapy in ovarian cancer. *Med Oncol*. 2004;21(1):41-8.

e) La pregunta final es si la prueba modifica o no el pronóstico del paciente, es decir, si por el hecho de ejecutarla puede esperarse en definitiva un mejor desenlace clínico. En muchos casos esto es obvio y no se requiere hacer un estudio para demostrarlo.

Por la vía de inducir al uso de tratamientos infectivos o peligrosos, un test con excelente capacidad diagnóstica podría resultar en definitiva dañino para los pacientes. Piense por ejemplo en el efecto psicológico que el hallazgo de un sutil "infiltrado pulmonar crónico" (no detectable en la radiografía convencional) podría tener sobre una paciente con disnea psicógena secundaria a un trastorno de ansiedad. O en cuántos pacientes fueron sometidos en exceso a cirugía por una hernia de núcleo pulposo cuando estas lesiones comenzaron a ser fácilmente detectables con las técnicas de imágenes actuales.

Este tipo de preguntas (d y e) se aborda dando al test la categoría de intervención, y por lo tanto deben ser respondidas idealmente a través de un estudio experimental, es decir, un ensayo clínico aleatorizado, en el que un grupo es asignado a recibir la intervención (en este caso la prueba diagnóstica), y otro a una intervención alternativa (otra prueba) o ninguna. Las opciones de diseño y las fortalezas y debilidades de los estudios que evalúan la efectividad de las intervenciones en salud serán discutidos más adelante.

Nascimento MA, Lira RP, Soares PH, Spessatto N, Karajose N, Arieta CE. Are routine preoperative medical tests needed with cataract surgery? Study of visual acuity outcome. *Curr Eye Res*. 2004 Apr;28(4):285-90.

Lidman C, Burman LG, Lagergren A, Ortvist A. Limited value of routine microbiological diagnostics in patients hospitalized for community-acquired pneumonia. *Scand J Infect Dis.* 2002;34(12):873-9.

Westwood ME, Kelly S, Berry E, Bamford JM, Gough MJ, Airey CM, Meaney JF, Davies LM, Cullingworth J, Smith MA. Use of magnetic resonance angiography to select candidates with recently symptomatic carotid stenosis for surgery: systematic review. *BMJ.* 2002 Jan 26;324(7331):198.

Desde otra perspectiva, debemos tener presente que los objetivos de las pruebas no se restringen simplemente a establecer o descartar la presencia de enfermedad, sino que abarcan una gama mucho más compleja de aplicaciones clínicas, por ejemplo:

- Pueden constituir un pre-requisito para, o apoyar la ejecución de un tratamiento (por ejemplo, la coronariografía no sólo indica que existe una estenosis, también permite planificar la cirugía o procedimiento de revascularización).
- Permiten definir el estado de avance de la enfermedad, es decir, aportan información pronóstica, o que sirve de línea basal para el posterior seguimiento del paciente.
- Los test diagnósticos permiten evaluar la respuesta al tratamiento. El hecho de medir la temperatura corporal -quizás la prueba diagnóstica más utilizada de todas- es un ejemplo. ¿Ha oído hablar de los "exámenes de control"?

Por último, la investigación puede estar dirigida a evaluar un solo test, o evaluar dos o más pruebas entre sí (frente a un estándar de oro).



Esta obra de Medwave está bajo una licencia Creative Commons Atribución-No Comercial 3.0 Unported. Esta licencia permite el uso, distribución y reproducción del artículo en cualquier medio, siempre y cuando se otorgue el crédito correspondiente al autor del artículo y al medio en que se publica, en este caso, Medwave.