

Pruebas diagnósticas: Lo positivo de los resultados negativos y viceversa

Diagnostic tests: The positive side of negative results and viceversa

Dr. Carlos Martínez-Duncker R,* Dr. Luis Mauricio Hurtado-López**

En medicina es aceptado generalmente que, cuando se desea confirmar la presencia de una enfermedad, se prefieren las pruebas diagnósticas más específicas (una prueba que, de ser positiva, confirme la enfermedad) mientras que, cuando se requiere descartar una enfermedad, se prefieren las pruebas diagnósticas más sensibles (un resultado que de ser negativo descarte la enfermedad).¹

Sin embargo, sucede que los médicos frecuentemente sobreestiman la probabilidad de una enfermedad al considerar su presencia en el enfermo por haber obtenido un resultado positivo en una prueba diagnóstica que tenga mayor sensibilidad que especificidad, lo que puede repercutir en una alta incidencia de falsos positivos y, por lo tanto, de pacientes que reciben tratamientos innecesarios. Así mismo, a lo largo del tiempo la comunidad médica ha mostrado mayor inclinación por realizar pruebas diagnósticas que confirmen la existencia de un padecimiento en sus pacientes que por aquellas pruebas diagnósticas que la descarten.

Hay estudios que revelan que las definiciones correctas de sensibilidad, especificidad y valores predictivos son bien conocidas por la mayoría de los médicos, sin embargo, pocos saben cómo aplicarlas en los padecimientos de sus pacientes.²

En muchos casos, los resultados negativos de una prueba diagnóstica muy sensible pueden ser tan importantes como aquellos resultados positivos de una prueba diagnóstica muy específica (ej. por las implicaciones terapéuticas, económicas y existenciales de un paciente, descartar la presencia de cáncer con alta seguridad puede ser tan importante como confirmar la existencia de cáncer con alta seguridad).

La finalidad de las pruebas diagnósticas es maximizar la información sobre un padecimiento específico. Sin embargo, todas las pruebas diagnósticas sufren de al-

gún tipo de error que resulta en información incompleta y, por lo tanto, en conclusiones diagnósticas potencialmente equivocadas.³ De hecho, las pruebas diagnósticas que sean al mismo tiempo 100% sensibles y 100% específicas se consideran virtualmente inexistentes (que de ser negativas descarte en el 100% de los casos la presencia de una enfermedad "X" y de ser positivas confirmen en el 100% de los casos la presencia de esa enfermedad "X"),⁴ ya que, por lo general, una prueba diagnóstica con una alta sensibilidad (que en el mejor de los casos puede llegar a ser hasta del 100%) sufre de una menor especificidad y viceversa. En otras palabras, lo frecuente es tener pruebas diagnósticas que son más sensibles que específicas o viceversa e infrecuentemente al 100% en cualquiera de los dos valores. Es por esta razón que en un gran número de pacientes es necesario realizar más de una prueba diagnóstica para llegar a una conclusión que permita descartar o confirmar la necesidad de tratamiento (y/o del tratamiento óptimo para la enfermedad del paciente).

A manera de ejemplo, cuando en presencia de la sospecha clínica de una enfermedad en particular obtenemos un resultado negativo de una prueba diagnóstica muy sensible (que descarta una enfermedad en particular en el 100% de los casos) y un resultado positivo en otra prueba diagnóstica no muy específica (que confirma esa misma enfermedad en menos del 100% de los casos), un resultado negativo de la primera podría ser más importante que un resultado positivo en la segunda.

Del anterior ejemplo, común en la clínica diaria, se deduce que la gran mayoría de las pruebas diagnósticas no son "perfectas", entendiendo por esto que sean capaces de descartar o confirmar la presencia de una enfermedad con un 100% de certeza (aun tratándose de pruebas distintas).

* Jefe del Departamento de Medicina Nuclear Molecular del Hospital Infantil de México Federico Gómez. Academia Mexicana de Cirugía.

** Clínica de Tiroides Hospital General de México. Academia Mexicana de Cirugía

Recibido para publicación: 2 enero 2006

Aceptado para Publicación: 27 enero 2006

Correspondencia: Dr. Luis Mauricio Hurtado-López. Clínica de Tiroides, Servicio de Cirugía General Hospital General de México,

Dr. Balmis 148 Col. Doctores México D.F. 06726

E-mail: luismauriciohurtado@terra.com.mx

Lo anterior se debe a que la gran mayoría de las pruebas diagnósticas tienen resultados falsos negativos (pruebas que resultan normales aun en pacientes que sí tienen una enfermedad) y/o tienen resultados falsos positivos (pruebas que resultan anormales aun en pacientes sanos). Por ejemplo, una radiografía convencional de tórax puede demostrar la existencia de una anormalidad o lesión focal o difusa en un pulmón, pero por otro lado, la misma radiografía convencional de tórax puede ser normal en otros pacientes con la misma enfermedad pulmonar antes de que la lesión o enfermedad se consolide (resultados falsos negativos).

Es necesario recordar que las pruebas diagnósticas pueden no ser "perfectas" porque no estén diseñadas, elaboradas, desarrolladas, perfeccionadas, etcétera, para estudiar ciertos padecimientos, así como que la presencia de otro padecimiento, antecedente o tratamiento impida la correcta valoración del estado actual del sitio o padecimiento de interés actual.

Por ejemplo, es incorrecto creer que una tomografía PET con F-18-FDG esté indicada para valorar a un paciente con cáncer de próstata, ya que un resultado negativo (ausencia de lesiones) en la tomografía funcional PET con F-18-FDG no descartaría la existencia de células neoplásicas en ese paciente debido a particularidades biológicas de esa neoplasia. Un resultado negativo en el paciente de este ejemplo no significa que el procedimiento (PET con F-18-FDG) no sirva o que sea "imperfecto". Simplemente indica que esa prueba diagnóstica no está indicada en pacientes con cáncer de próstata, como puede estarlo en pacientes con cáncer de colon, pulmón, mama, melanoma, linfoma, etcétera. Dicho esto, es un error común en nuestro medio, considerar que el procedimiento del ejemplo (PET con F-18-FDG) es óptimo para valorar todas las enfermedades oncológicas, como es incorrecto también hacer comparaciones que no corresponden (ej. calificarlo de subóptimo o "inferior a la biopsia").

La literatura también indica que, como en el caso anterior, las pruebas diagnósticas de medicina nuclear con SPECT son superiores a las realizadas con PET y F-18-FDG en pacientes con tumores (neoplasias) que no concentran FDG (glucosa) como son los tumores neuroendocrinos (cáncer medular de tiroides, feocromocitomas, neuroblastomas, paragangliomas y otros que expresen receptores para somatostatina, como algunos tumores gástricos y pancreáticos, etcétera).⁵

Otro ejemplo, los estudios de resonancia magnética y tomografía computada, que ciertamente revolucionaron la medicina diagnóstica en el siglo pasado y que son de gran ayuda o inclusive indispensables para el correcto y oportuno diagnóstico de ciertos padecimientos en la actualidad, no deben utilizarse para realizar el diagnóstico diferencial de algunas enfermedades (ej. del cáncer tiroideo debido a su baja especificidad para ese fin en particular).⁶

En el caso de los resultados negativos debemos preguntarnos "si el resultado de la prueba es negativo ¿cuál es la probabilidad de que mi paciente no tenga la enfermedad?". En otras palabras, un paciente con un

resultado negativo en una prueba que tenga valor predictivo negativo del 100% tiene 0% de probabilidad de tener la enfermedad. Cuando se realiza una prueba diagnóstica cuyo valor predictivo negativo es del 100% no se requieren otras pruebas diagnósticas para descartar ni para confirmar esa enfermedad. Del mismo modo, cuando una prueba diagnóstica con un valor predictivo positivo del 100% (que confirma en todos los casos la presencia de una enfermedad cuando el resultado es positivo para la presencia de dicha enfermedad) es positiva, que demuestra la presencia de esa enfermedad, se requiere tratamiento, tampoco se necesitarían otras pruebas diagnósticas para descartar ni para confirmar esa enfermedad.

Por ejemplo, en los pacientes con lesiones tiroideas palpables (nódulos tiroideos) es necesario descartar cáncer. La gammagrafía de tiroides con Tc-99m-MIBI, que tiene un valor predictivo negativo del 100%, puede descartar la presencia de cáncer tiroideo en todos los pacientes que tengan un resultado negativo (captación nula de MIBI en el interior de la lesión tiroidea palpable o diagnosticada por otro medio). En otras palabras, si la prueba diagnóstica con Tc-99m-MIBI demuestra la ausencia de metabolismo en la lesión tiroidea, el paciente puede ser catalogado y tratado (para comodidad, tranquilidad y satisfacción del paciente y de su médico) confiablemente como un paciente con una lesión tiroidea benigna. No se requieren otras pruebas diagnósticas para confirmar ni para descartar la presencia de cáncer tiroideo, aunque podrían requerirse de otras para conocer la etiología de la lesión tiroidea benigna⁷.

Hemos visto en este ejemplo que una gammagrafía tiroidea con Tc-99m-MIBI con resultado negativo descarta la posibilidad de cáncer de tiroides en un nódulo tiroideo. Sin embargo, sería erróneo pensar que lo contrario, la captación de MIBI en la lesión tiroidea, fuera diagnóstico de cáncer. Esta prueba diagnóstica es capaz de descartar cáncer cuando el resultado es negativo, pero en los casos en que la prueba es positiva (captación de MIBI en la lesión tiroidea) el diagnóstico es inespecífico, ya que la captación de la radiomolécula en la lesión tiroidea (indicativa de actividad metabólica en la misma) sólo demuestra que existe una lesión tiroidea conformada por células metabólicamente activas, como pueden serlo tanto lesiones benignas (ej. adenomas, tiroiditis, etcétera) como lesiones malignas. Por ello, en este último caso se requieren otros métodos de diagnóstico para hacer el diagnóstico diferencial de la lesión (ej. biopsia por aspiración con aguja fina, estudio histopatológico transoperatorio, etcétera).

Existen muchos otros ejemplos de esta situación en el campo de la cirugía general, así la determinación de tripsinógeno urinario-2 puede descartar la presencia de pancreatitis aguda,⁸ la tomografía computada con enema de agua también puede descartar la presencia de una neoplasia colorrectal,⁹ este mismo estudio tomográfico descarta la presencia de íleo biliar,¹⁰ el ganglio centinela en cáncer de cavidad oral también descarta la presencia de metástasis regionales,¹¹ situación

similar, pero el 97.5% para el cáncer de mama¹² y ningún paciente con trauma abdominal fallecerá a consecuencia del mismo cuando tenga un ultrasonido negativo a detección de lesiones o colecciones,¹³ por tanto deberemos aprender a utilizar y a confiar en pruebas que "descartan" la enfermedad.

También existen pruebas diagnósticas distintas con alta especificidad y sensibilidad que confirman o descartan la presencia de un padecimiento con resultados similares, pero cuyos costos pueden ser muy diferentes (como puede ser la detección inicial de isquemia miocárdica [insuficiencia coronaria] mediante un SPECT cardiaco por aprox. \$900 dólares EUA o mediante un PET F-18-FDG de miocardio por aprox. \$2,500 dólares EUA o el seguimiento de un paciente con cáncer de tiroides con tiroidectomía y tratamiento ablativo con I-131 en el que un rastreo con I-131 puede ser diagnóstico de recurrencia o recidiva y suficiente para justificar y decidir administrar tratamiento con un costo aprox. de \$400 dólares *versus* un estudio inicial con PET F-18-FDG para valorar recurrencia o recidiva del cáncer de tiroides por un costo de aprox. \$2,500 dólares EUA).¹⁴ No obstante lo anterior, miles de pacientes con probable isquemia miocárdica (diagnóstico), con cáncer de próstata (valorar extensión) o cáncer de tiroides (sin haber realizado una gammagrafía de cuerpo entero con I-131) serán estudiados en nuestro medio con PET F-18-FDG.

Así mismo, e independientemente de su costo, existen pruebas diagnósticas invasivas y pruebas diagnósticas no invasivas que pueden confirmar o descartar la presencia de alguna enfermedad con precisión diagnóstica similar, en cuyo caso, si ambas brindan la misma información, deberíamos considerar junto con el paciente tanto el costo económico como las molestias que sufrirá con una u otra prueba diagnóstica.

Por ejemplo, se han realizado literalmente millones de biopsias en lesiones de pacientes con cáncer cuyos resultados han sido negativos para cáncer (normales o indicativos de lesiones o enfermedades benignas), sólo para darse cuenta tiempo después (quizá demasiado tarde para un tratamiento oportuno) que la lesión era maligna (falsos negativos). Por otro lado, millones de estudios de laboratorio, patología o imagenología han sido diagnósticos de cáncer en pacientes con lesiones benignas y, por ello, han sido sometidos a cirugías u otros tratamientos incorrectos o innecesarios (falsos positivos).

Es por lo anterior que los médicos debemos conocer las indicaciones precisas de las pruebas diagnósticas, sus contraindicaciones y los casos en que se requieren complementaciones diagnósticas paralelas para realizar o descartar alguna opción diagnóstica, así como el costo o incomodidad para el paciente y el orden adecuado, óptimo, de su realización.

En cuanto a las pruebas diagnósticas paralelas, debido a que la mayoría no son 100% sensibles ni/100% específicas, en muchos casos el uso de dos pruebas diagnósticas incrementa la precisión diagnóstica considerablemente. Pero, como puede ser que nin-

guna de las dos pruebas diagnósticas sea 100% específica ó 100% sensible, es necesario recordar que especialmente en casos en que los resultados sean discrepantes, siempre hay una tercera opción, que ambas estén equivocadas (dos resultados equivocados no hacen un resultado correcto).¹⁵

En el tema de las pruebas paralelas o complementarias es importante considerar el orden adecuado de su realización para que una no interfiera con la otra, como puede ser el caso en la valoración funcional de la glándula tiroides mediante gammagrafía después de realizar una tomografía con medio de contraste o la valoración gammagráfica de la velocidad de filtración glomerular renal después de realizar una urografía excretora con medio de contraste. En ambos casos, las primeras deben ser realizadas antes de las segundas para evitar interferencia funcional secundaria al medio radiológico de contraste que podría alterar o anular los resultados de los estudios gammagráficos funcionales.⁶

Sin embargo, lo anterior no es fácil en nuestro medio debido a que no toda la tecnología médica de vanguardia está disponible y a que la que sí está, generalmente, se emplea (correcta o incorrectamente) al máximo, retrasando fechas de estudios que frecuentemente hacen que se soliciten otras pruebas alternativas cuyos resultados (en algunos casos menos sensibles y/o específicos) estarán antes que los primeros. Así mismo, ya que la tecnología médica (pruebas diagnósticas) ha avanzado a una velocidad superior a la de la enseñanza y formación de recursos humanos especializados, no hay muchos lugares o personas con experiencia calificados para hacer escuela y enseñar el uso adecuado y la optimización de dicha tecnología, a medida que se incorpora a nuestro arsenal diagnóstico. Existen numerosos artículos, consensos, etcétera, publicados en nuestro medio que contienen errores que promueven el descontento, la confusión y la falta de aceptación de ciertas pruebas diagnósticas por parte de los médicos. Pero esto no ocurre porque las pruebas diagnósticas "no sirvan", sino porque algunos consensos presentan indicaciones imprecisas, obsoletas, incompletas y/o equivocadas.

A manera de ejemplo, con fines constructivos y no menospreciando el esfuerzo que la realización de un consenso lleva implícito, en el apartado de gammagrafía pulmonar de unas guías y recomendaciones sobre el diagnóstico, estratificación y tratamiento de la tromboembolia pulmonar aguda¹⁶ y a diferencia de lo que la literatura de primer nivel ha descrito y actualizado constantemente en varias versiones desde hace ya varios años,¹⁷ los autores mencionan la administración de aproximadamente "60,000 microesferas o macroagregados", cuando las guías de primer nivel mencionan que deben ser más de 100,000 y hasta 200,000 macroagregados de albúmina (MAA) para no obtener imágenes subóptimas que pueden llevar a errores diagnósticos o inutilidad diagnóstica de la prueba realizada. Así mismo, se menciona que dichas "microesferas" son marcadas con "Tc-99", cuando en realidad lo son con

"Tc-99m". Otra limitación de la guía es que solamente mencionan la gammagrafía de ventilación pulmonar con Xe-133, siendo que debido a la falta de disponibilidad de dicho gas radiactivo en nuestro medio (como lo es la falta en nuestro medio del Kr-81m que tampoco se menciona), las gammagrafías de ventilación pulmonar se realizan, debido a su disponibilidad, con aerosol marcado con Tc-99m. Esto es importante porque los criterios de interpretación que se mencionan son distintos a los necesarios cuando se utilizan gases nobles radiactivos. Por lo anterior, puede ser que un médico especialista con un paciente con probable tromboembolismo pulmonar solicite la realización de una gammagrafía de "ventilación con Xe-133" y una de perfusión con "microesferas marcadas con Tc-99", cuando lo que puede hacerse en México es una gammagrafía de ventilación pulmonar con "aerosol marcado con Tc-99m" y una de perfusión con "MAA marcados con Tc-99m".

De manera similar, un consenso sobre enfermedad tromboembólica venosa¹⁸ al que también se cita en la guía mencionada, recomienda reclasificar los resultados del gammagrama pulmonar como "normal", "no contribuye" o "diagnóstico", cambiando la nomenclatura vigente e internacionalmente aceptada de "probabilidad alta", "probabilidad intermedia" o "probabilidad baja" de tromboembolismo pulmonar (TEP). Cabe señalar que la nomenclatura internacional ha sido modificada por los expertos en el tema basados en el análisis repetido a lo largo de los años (desde 1983) de bases de datos que contienen la información de miles de pacientes. Además, una "probabilidad intermedia" de tromboembolismo (lo que en el consenso mencionado propusieron cambiar a "no contribuye") no significa que la evaluación no proporcione información importante. Una gammagrafía pulmonar de probabilidad intermedia de tromboembolismo significa que existe una probabilidad de TEP de entre 20 y 79%, significa que el paciente tiene uno o más defectos de perfusión pulmonar aun en ausencia de alteraciones en la radiografía convencional de tórax, lo que aun cuando no sea diagnóstico (de alta probabilidad) de TEP, tampoco puede ser "descartado" ni catalogado como "no contribuye", sino que debe ser considerado como información adicional importante en el análisis médico específico del paciente y en la justificación y/o necesidad de realizar otras pruebas diagnósticas o administrar tratamiento. Además, los autores mencionan que "numerosos pacientes con embolia pulmonar no muestran datos anormales en el gammagrama", cuando la gran cantidad de literatura médica de los últimos 40 años y hasta la fecha indica precisamente lo contrario, que la sensibilidad de la gammagrafía pulmonar en la detección de TEP es muy alta cuando se realiza e interpreta apropiadamente. Una vez más, quizá una confusión entre los términos de sensibilidad y especificidad.

Por lo anterior, reconociendo los esfuerzos realizados, pero haciendo énfasis en los que faltan, es indispensable que toda guía diagnóstica y médica en práctica profesional utilice estudios diagnósticos adecuados

en el paciente adecuado y en el momento adecuado y que se analice a profundidad cómo se realizan éstos en el mundo y cómo se realizan en nuestro medio, qué indicaciones y contraindicaciones precisas tienen y, en consecuencia, cómo son y cómo deberían ser interpretados, dando su justo valor tanto a los resultados positivos como a los negativos.

Referencias

1. Lee WC. Selecting diagnostic tests for ruling out or ruling in disease: the use of the Kullback-Leibler distance. *Int J Epidemiol* 1999; 28: 521-525.
2. Espallardo N. Decisions on diagnosis in family practice: Use of sensitivity, specificity, predictive values and likelihood ratios. *Asia Pacific Family Medicine* 2003; 2: 229-232.
3. Bowyer A. Quantitative information of specific diagnostic tests. *J Med Syst* 1988; 22: 3-13.
4. Dendukuri N, Rahme E, Belisle P, Joseph L. Bayesian sample size determination for prevalence and diagnostic test studies in the absence of a gold standard test. *Biometrics* 2004; 60: 388,397.
5. Schmidt M, Eschner W, Dietlein M, Theissen P, Schicha H. Established nuclear medicine techniques for tumour diagnosis (tumour SPECT): can they still compete with (18)F-FDG-PET? *Nuklearmedizin* 2005; 44: 37-48.
6. AACE/AME Task Force on Thyroid Nodules. American Association of Clinical Endocrinologists and Associazione Medici Endocrinologi medical guidelines for clinical practice for the diagnosis and management of thyroid nodules. *Endocr Pract* 2006; 12: 63-102.
7. Hurtado-López LM, Arellano-Montaña S, Torres-Acosta EM, Zaldivar-Ramírez FR, Duarte Torres RM, Alonso-de-Ruiz P, et al. Combined use of fine needle aspiration biopsy, MIBI scans and frozen section biopsy offers the best diagnostic accuracy in the assessment of the hypofunctioning solitary thyroid nodule. *Eur J Nucl Med Mol Imaging* 2004; 31: 1273-1279.
8. Matull WR, Pereira SP, O'Donohue JW. Biochemical markers of acute pancreatitis. *J Clin Pathol* 2006; 59(4): 340-344.
9. Pilleul F, Bansac-Lamblin A, Monneuse O, Dumortier J, Milot L, Valette PJ. Water enema computed tomography: diagnostic tool in suspicion of colorectal tumor. *Gastroenterol Clin Biol* 2006; 30: 231-234.
10. Yu CY, Lin CC, Shyu RY, Hsieh CB, Wu HS, Tyan YS, et al. Value of CT in the diagnosis and management of gallstone ileus. *World J Gastroenterol* 2005; 11: 2142-2147.
11. Civantos FJ, Moffat FL, Goodwin WJ. Lymphatic mapping and sentinel lymphadenectomy for 106 head and neck lesions: contrasts between oral cavity and cutaneous malignancy. *Laryngoscope* 2006; 112(3 Pt 2 Suppl 109): 1-15.
12. Gallegos-Hernandez JF, Chavez-Garcia M. [Usefulness of lymphatic mapping with sub-areolar injection of patent blue dye, in the typification of breast cancer]. *Rev Invest Clin* 2003; 55: 407-11.
13. Porter RS, Nester BA, Dalsey WC, O'Mara M, Gleeson T, Pennell R, Beyer FC. Use of ultrasound to determine need for laparotomy in trauma patients. *Ann Emerg Med* 1997; 29: 323-330.
14. Ruiz Franco-Baux JV, Borrego Dorado I, Gomez Camarero P, Rodriguez Rodriguez JR, Vazquez Alberto RJ, Navarro Gonzalez E, et al. [F-18-fluorodeoxyglucose positron emission tomography on patients with differentiated thyroid cancer who present elevated human serum thyroglobulin levels and negative I-131 whole body scan]. *Rev Esp Med Nucl* 2005; 24: 5-13.

15. Miller WC. Bias in discrepant analysis: when two wrongs don't make a right. *J Clin Epidemiol* 1998; 51: 219-231.
16. Jerjes-Sánchez CD, Elizalde González JJ, Sandoval Zárate J, Gutiérrez Fajardo P, García de León MS, Ramirez Rivera A, et al. Diagnóstico, estratificación y tratamiento de la tromboembolia pulmonar aguda: Guías y Recomendaciones del Capítulo de Circulación Pulmonar de la Sociedad Mexicana de Cardiología. *Arch Cardiol Mex* 2004; 74(S3): S547-S585.
17. Parker JA, Coleman RE, Hilson AJW, et al. *Society of Nuclear Medicine Procedure Guideline for Lung Scintigraphy Version 3.0, approved February 7, 2004.*
18. Consenso mexicano sobre enfermedad tromboembólica venosa. *Cir Cir* 2000; 68(S6): S7-S29.

