

¿Qué saber para optimizar el uso de medios diagnósticos en la clínica?

What to know for optimizing the use of diagnostic means in the clinical practice?

Teddy Osmin Tamargo Barbeito, Rosa Eugenia Jiménez Paneque, Tania Hidalgo Costa, Isabel Mora Díaz, Aisa Peña Casanovas, Ángela Rosa Gutiérrez Rojas

Hospital Clínico Quirúrgico "Hermanos Ameijeiras". La Habana, Cuba.

RESUMEN

Introducción: la eficacia de un medio diagnóstico se refiere a la capacidad para detectar al verdadero paciente enfermo. La sensibilidad, especificidad y valores predictivos positivo y negativo son indicadores de la validez de un medio diagnóstico. Otra manera de obtener la probabilidad de enfermar es a través de la estimación de las razones de verosimilitud positiva y negativa.

Objetivo: describir conceptos generales e interpretación de sensibilidad, especificidad, valores predictivos y umbrales diagnósticos.

Métodos: se realizó una revisión de artículos y libros sobre las propiedades básicas de un medio diagnóstico para poder interpretarlos de manera correcta y con ello optimizar su uso en el ámbito clínico.

Conclusiones: trabajo que servirá como guía para los profesionales de la salud que se debaten constantemente con la necesidad de hacer un diagnóstico correcto para tratar oportunamente a los pacientes y evitarles complicaciones y cuando sea posible evitar la muerte.

Palabras clave: medios diagnósticos; sensibilidad; especificidad; valor predictivo positivo y negativo.

ABSTRACT

Introduction: The effectiveness of a diagnostic means refers to the ability to detect the true sick patient. The sensitivity, specificity and positive and negative predictive values are indicators of the validity of a diagnostic means. Another way to obtain the probability of getting sick is through the estimation of positive and negative likelihood ratios.

Objective: To describe general concepts and interpretation of sensitivity, specificity, predictive values and diagnostic thresholds.

Methods: A review of articles and books on the basic properties of a diagnostic means was made in order to interpret them correctly and thereby optimize their use in the clinical setting.

Conclusions: Our paper will serve as a guide for health professionals who are constantly debating the need to make correct diagnosis to treat patients timely and to avoid complications and, to avoid death when possible.

Keywords: Diagnostic means; sensitivity; specificity; positive and negative predictive value.

INTRODUCCIÓN

El diagnóstico es un proceso complejo que requiere de una gran pericia del médico u otro personal de la salud que brinde atención a los pacientes enfermos.¹⁻⁵ Ochoa plantea que "el diagnóstico médico es un proceso dinámico en el que se intenta tomar decisiones idóneas en presencia de incertidumbre. En esencia, realizar el diagnóstico es asignar con razonable verosimilitud un paciente a una clase (o grupo) constituida por sujetos con una enfermedad o entidad nosológica".¹

El diagnóstico es un eslabón trascendental que junto con otras evidencias permitirá la toma de decisiones clínicas. En él intervienen una serie de componentes como la anamnesis, después le sigue un examen físico minucioso y por último, si es necesario, la realización de pruebas complementarias que servirán de ayuda para precisar el diagnóstico definitivo, el cual es la evidencia suficiente para indicar un tratamiento de forma oportuna y el paciente tenga menos riesgos de complicaciones. El diagnóstico es una manera de incorporar una evidencia para tomar decisiones en la clínica, incluso muchas veces no se tiene seguridad en el diagnóstico que se plantea y sin embargo ante la incertidumbre se impone un tratamiento médico que puede ser correctamente indicado.¹⁻³

En la actualidad con el desarrollo impetuoso que ha tenido la tecnología existen muchas pruebas complementarias a las que se les han denominado procedimientos diagnósticos, pruebas diagnósticas, test diagnósticos o medios diagnósticos (PD), ello ha traído consigo que en ocasiones los médicos indiquen análisis a los pacientes deliberadamente sin tener en cuenta un juicio científico basado en el método clínico como forma particular del método científico, esta conlleva a una crisis del método clínico.⁶

El médico muchas veces no le da importancia a la clínica, pues se necesita tiempo para poder hacer un buen interrogatorio al paciente y establecer una buena relación, examinarlo y después escribir en la historia clínica, además en países con sistemas de salud privado no les reporta una buena economía y porque muchos piensan que la tecnología puede sustituir al método clínico. Se ha deteriorado la relación médico paciente, no se hace una buena anamnesis y examen físico, así como se ha sobrevalorado la tecnología por el uso indebido e indiscriminado de una serie de procedimientos diagnósticos (PD) que son costosas, entrañan riesgos en muchas ocasiones, como la tomografía axial computadorizada (TAC) por la cantidad de radiaciones que recibe el paciente y además pueden ser cruentos o invasivos.

El avance tecnológico es bueno pues permite tener PD que ayuden a tener mayor certidumbre para plantear un diagnóstico adecuado, pero nunca se puede obviar un buen proceso de diagnóstico, los PD son complementarios y nunca podrán suplantar una buena valoración clínica del enfermo.⁶

Existen varias definiciones de PD, entre ellas la de Jiménez en el texto de la ayuda del programa estadístico Epidat versión 3.1 (Análisis epidemiológico de datos tabulados) programa que fue desarrollado por el servicio de información sobre salud pública de la Dirección General de Salud Pública de la Consejería de Sanidad (Junta de Galicia) en colaboración con el área de análisis de salud y sistemas de información de salud de la Organización Panamericana de la Salud (OPS-OMS), allí la mencionada autora refiere que un PD es "*cualquier proceso, más o menos complejo, más o menos invasivo, que pretenda establecer la presencia de cierto rasgo, supuestamente patológico, del paciente no susceptible de ser observado directamente*". Ochoa plantea que un PD es "*cualquier procedimiento realizado para confirmar o descartar un diagnóstico o incrementar o disminuir su verosimilitud*".¹

Un PD útil será aquel que sea lo menos invasivo posible, tenga buena confiabilidad (capacidad para producir los mismos resultados cada vez que se aplica en similares condiciones) que sus mediciones sean un reflejo exacto de lo que se quiere medir (validez o exactitud), que produzca un beneficio clínico es decir que sea ventajoso para tomar decisiones en la clínica y por último que no se incrementen los costos por concepto de su aplicación.

En el presente artículo se pretende abordar una serie de conceptos generales e interpretación de sensibilidad, especificidad, valores predictivos y umbrales diagnósticos, propiedades básicas de los PD que le permitan a los galenos un mejor desenvolvimiento en el planteamiento de un diagnóstico adecuado. Asimismo, servirá de una guía práctica para que los investigadores y personal de la salud en general involucrado en esta difícil tarea, puedan leer y dilucidar con facilidad las diferentes aristas que tienen los indicadores acerca de la eficacia de un PD.

Propiedades básicas de un procedimiento diagnóstico

En la práctica médica todo PD no está exento de que se cometan errores, por tanto, es una necesidad imperiosa conocer sus principales atributos para caracterizarlo y a la vez someterlo a una evaluación para poder estar seguros de que ayudará a descartar, confirmar o pesquisar determinada enfermedad. La evaluación de dichos PD conlleva la realización de investigaciones, donde los resultados se publican de con mucha frecuencia en revistas médicas.⁷⁻¹⁰

*Manterola*¹¹ en un artículo hace referencia a un estudio realizado en la Universidad de Yale con una muestra probabilística estratificada de 300 médicos, el principal resultado fue que muy pocos utilizaban métodos adecuados para la evaluar la eficacia de un PD y la mayoría de las veces si utilizaban la sensibilidad y la especificidad, era de forma errónea.

Los PD, test o medios diagnósticos pueden tener diferentes resultados, algunos son dicotómicos, presencia o ausencia (ejemplo test de embarazo positivo o negativo) varios ordinales (bajo, moderado y alto) y otros de forma cuantitativa (ejemplo triglicéridos, ácido úrico, creatinina). Los últimos mencionados pueden convertirse en dicotómicos si se establecen puntos de corte a partir de los cuales se dan los resultados en positivos o negativos (ejemplo valor de triglicéridos mayores que 1,7 mmol/l).

Los indicadores básicos o atributos de un PD son la sensibilidad (S) y la especificidad (E), pero se necesita de un PD capaz de detectar la enfermedad en cuestión con la máxima certidumbre que se conozca, es decir, la más exacta posible, de manera que se cometan errores con muy baja probabilidad, por ejemplo una arteriografía coronaria para el diagnóstico de cardiopatía isquémica, es lo que se conoce con el nombre de prueba de verdad, estándar de oro o referencia (EO), entonces se comparan los resultados del PD con el EO. En ocasiones no se cuenta con un estándar de oro y es necesario construirlo mediante paneles de expertos y la evolución clínica por ejemplo en reumatología hay un grupo de enfermedades que en sus comienzos no tienen signos evidentes y entonces se siguen los pacientes durante un tiempo prudente que supuestamente se manifieste completamente la enfermedad entonces ese criterio clínico se toma como prueba de referencia o estándar de oro.¹²⁻²³

El EO además de ser adecuado tiene que ser independiente de la prueba que se está evaluando, una característica de esta independencia es que el EO no puede formar parte del PD que se evalúa. *Valenzuela y colaboradores*²² ejemplifica con un estudio que evaluó la utilidad de la medición de amilasa sérica y urinaria en el diagnóstico de pancreatitis aguda, considerando como EO más adecuado un listado de criterios dentro de los cuales se incluían la amilasa sérica y urinaria, lo que produce un falso aumento del poder discriminatorio del PD evaluado. Por otra parte si EO es apropiado e independiente, habría que preguntarse si los que obtienen el resultado del PD no conocían el resultado del EO y viceversa para evitar el sesgo de revisión del test y de revisión del diagnóstico.²³

Otra cuestión importante es que se les aplique el EO a todos los pacientes sin tener en cuenta el resultado del test sujeto a evaluación, puesto que puede suceder que las propiedades del test diagnóstico se alteren si el resultado de éste influye sobre la decisión de qué pacientes serán sometidos después al EO. En algunos estudios se aplica el EO sólo a los pacientes que presentan un resultado positivo en el test, especialmente si el EO es invasivo o presenta riesgos importantes; esto se denomina sesgo de verificación. Este sesgo hará inevitablemente aparecer con mayor poder diagnóstico al test al seleccionar pacientes con mayor probabilidad de tener un EO positivo. Un ejemplo de ellos es evaluar la eficacia del ultrasonido para el diagnóstico de cirrosis hepática, es obvio que solo se realice la biopsia a los pacientes que tiene un resultado alterado de dicho PD.²³⁻²⁶

Cuando se comparan los resultados de un PD con una prueba de referencia o EO, si el resultado de la prueba que se evalúa es binario o dicotómico se pueden obtener cuatro posibles resultados: verdaderos positivos, falsos negativos, verdaderos negativos y falsos positivos. Se presentan a partir de la siguiente tabla de contingencia dos entradas de la cual se dará el concepto de los diferentes indicadores y se explicará cómo se calculan e interpretan (tabla).

Tabla. Relación entre el resultado de una prueba diagnóstica y la presencia o ausencia de una enfermedad

Procedimientos diagnósticos evaluados	Prueba de referencia (estándar de oro)		Total
	Positivo	Negativo	
Positivo	a (verdaderos positivos)	b (falsos positivos)	a + b
Negativo	c (falsos negativos)	d (verdaderos negativos)	c + d
Total	a + c	b + d	N

Prevalencia de la enfermedad (P): es la probabilidad o el porcentaje de enfermos en la población (o se en la muestra que sea valorada). Se calcula $P = a+c/N$.¹¹⁻²³

Probabilidad a priori o pretest: probabilidad de la enfermedad antes de la indicación de un PD por lo general en el contexto de un paciente individual y se puede estimar basado en la experiencia clínica o según la prevalencia de la enfermedad por medio del teorema de Bayes, fórmula que por su complejidad no se abordará en este artículo.¹¹⁻²³

La sensibilidad (S) es la capacidad de la PD para detectar a los verdaderos enfermos o lo que es lo mismo es la probabilidad de que si un paciente está enfermo el PD lo detecte como tal. Se calcula $a/a+c$ o verdaderos positivos/verdaderos positivos + falsos negativos.¹¹⁻²³

La tasa de falsos negativos (TFN) es la proporción de pacientes que el PD no detectó como enfermo y en realidad los estaban. Se calcula $1-S = c/a+c =$ falsos negativos/verdaderos positivos + falsos negativos.¹¹⁻²³

La especificidad (E) es la capacidad del PD de detectar la ausencia de la enfermedad, es decir, la probabilidad de que si un paciente no está enfermo el PD lo identifique como eso. Se calcula $d/b+d =$ verdaderos negativos/verdaderos negativos + falsos positivos.¹¹⁻²³

La tasa de falsos positivos (TFP) es la proporción de pacientes que la PD que se evalúa, identifica la presencia de la enfermedad sin embargo la prueba de referencia no lo manifiesta como tal. Se calcula $1-E = b/b+d =$ falsos positivos/falsos positivos y verdaderos negativos.¹¹⁻²³

La sensibilidad y especificidad son cualidades o atributos inherentes al PD que se evalúa y dependen de las características de los pacientes en las que ellas se están evaluando, es por ello que se recomienda escoger un amplio rango de enfermos y no enfermos para determinado propósito.¹¹⁻²³

En la clínica, cuando se evalúa la bondad de un PD, interesa a los médicos saber la exactitud o eficacia diagnóstica de ese procedimiento, este indicador se define como la probabilidad de que un individuo sea clasificado correctamente por la prueba. Se calcula $(a + d)/(a + b + c + d)$. En el numerador aparecen los enfermos con resultado positivo (a = verdaderos positivos) y los no enfermos con test negativo (d = verdaderos negativos); entonces en el denominador de la fórmula se coloca el número total de pacientes estudiados.¹¹⁻²³

A continuación con un ejemplo tomado de un artículo publicado por *Khooshideh y colaboradores*²⁷ los cuales evaluaron la eficacia de la alfa-microglobulina-1 placentaria para el diagnóstico de la rotura prematura de membranas (RPM) se hará una breve interpretación de los indicadores mencionados.

Los mencionados autores evaluaron la eficacia de la alfa-microglobulina-1 en el diagnóstico de la RPM. Su sensibilidad fue de 98,8 %, especificidad de 92,8 %, la TFP fue de 7,2 % y la TFN de 1,2 %. Es decir que el mencionado test es capaz de detectar al 98,8 % de las gestantes con RPM y al 92,8 % de las que no tienen dicha alteración. El porcentaje de pacientes que el test identifica con RPM y no es confirmado por el EO fue de 7,2 % y al 1,2 % el test no las detecta con RPM, sin embargo, la prueba de referencia si lo confirma.

¿Cómo funciona un PD en la clínica? ¿Qué probabilidad de enfermedad existe según el resultado de un PD?

La sensibilidad y especificidad son dos indicadores inherentes al PD y dependen de las características de los pacientes, pero en la práctica su capacidad de cuantificación de la incertidumbre es limitada.^{12,13,15-21,23}

El médico necesita más bien evaluar la medida en que sus resultados modifican realmente el grado de conocimiento que se tenía sobre el estado del paciente.²³

Concretamente, le interesa conocer la probabilidad de que un individuo para el que se haya obtenido un resultado positivo, sea efectivamente un enfermo; y lo contrario, conocer la probabilidad de que un individuo con un resultado negativo esté efectivamente libre de la enfermedad.²³ Las medidas o indicadores que responden a estas interrogantes se conocen como valores predictivos el positivo que también llamado probabilidad positiva postest y el valor predictivo negativo.¹¹⁻²³

El valor predictivo de una prueba positiva (VPP) equivale a la probabilidad condicional de que los individuos con una prueba positiva tengan realmente la enfermedad. Se calcula: $VP(+)= a/a+b= \text{verdaderos positivos}/\text{verdaderos positivos} + \text{falsos positivos}$.¹¹⁻²³

El valor predictivo de una prueba negativa (VPN) es la probabilidad condicional de que los individuos con una prueba negativa no tengan realmente la enfermedad. Se calcula: $VP(-)= d/c+d= \text{verdaderos negativos}/\text{falsos negativos} + \text{verdaderos negativos}$. La probabilidad negativa postest es igual al $1-VP-$.¹¹⁻²³

Se retoma el estudio de *Khooshideh y colaboradores*²⁷ en la evolución de la eficacia de la alfa-microglobulina-1 en el diagnóstico de la rotura prematura de membranas. El VPP fue de 99 % y el VPN de 93 %, es decir, dicho test es capaz de detectar la RPM en el 99 % de gestantes con un test positivo y al 93 % de gestantes sin RPM con un test negativo.

Relación entre sensibilidad, especificidad y valores predictivos

Como ya se había mencionado la S y la E caracterizan al test propiamente dicho, dependen solamente de las definiciones de enfermo y no enfermo. Está claro que en un individuo enfermo la probabilidad que determinado test sea positivo es única. Igualmente pasa con los no enfermos.²³ Un PD con S y E de un 100 % posibilitaría el reconocimiento acertado de todos los enfermos y no enfermos en estudio.

En términos generales, cuando se estudia una variable de tipo cuantitativo continuo, puede observarse que existe una relación inversa entre el valor de la sensibilidad y el de la especificidad, de tal manera que, a medida que uno se incrementa, el otro disminuye.²⁸

Para un test con resultados cuantitativos para un valor determinado, es decir, un punto de corte que se define como aquel valor en el que el resultado de un PD se considera normal o alterado, existe una zona de posibles resultados de la prueba en la que la distribución de sujetos enfermos o no enfermos se solapan, es decir, habría un grupo de enfermos en el que el test sería negativo (falsos negativos) y otro en que un número de no enfermos serían positivos (falsos positivos). Para un test que la anormalidad va en el sentido del aumento de sus valores, si se desea aumentar la probabilidad de sujetos enfermos, lo que equivale a mover el punto de corte a la izquierda, entonces aumentan los falsos positivos. Si se desplaza el punto de corte hacia la derecha disminuirán los falsos positivos, pero aumentan los falsos negativos. En conclusión al mover el punto de corte cuando aumenta la sensibilidad, disminuye la especificidad y viceversa.²⁸

En el caso de los VP un valor positivo del test no constituye una verdad irrefutable sino solo el resultado de una prueba. Independientemente de su sensibilidad o especificidad, es natural que la probabilidad de que un test sea positivo dependa de la prevalencia de la enfermedad cuya presencia se intenta detectar. La probabilidad de resultados positivos depende de la suma de verdaderos positivos y falsos positivos. Para un nivel dado de S, a mayor prevalencia mayor cantidad de verdaderos positivos. El VPP está influenciado pues por la prevalencia de la enfermedad en cuestión, a medida que se incrementa la prevalencia de la enfermedad aumenta el VPP y disminuye el VPN y mientras menor sea la prevalencia de la enfermedad mayor es el VPN y menor el VPP, para un valor de S y E determinados. Esta limitación de los VP hace que no se puedan generalizar a todas las situaciones. Para cada valor de prevalencia o probabilidad pretest se deben calcular los VP.^{21,24,25}

Razones de verosimilitud

Las razones expuestas anteriormente provocan en el ámbito clínico algunos conflictos por lo que se necesitan otros indicadores que mantengan las fortalezas de la S, E y VP sin tener en cuenta sus limitaciones. Las razones de verosimilitud (RV) del inglés *Likelihood Ratio*, también llamadas cocientes o razones probabilidad, permiten estimar y cuantificar la magnitud y el sentido de la modificación del resultado obtenido previo al test. Se define como la razón entre la probabilidad de tener un determinado resultados del test en la población con la enfermedad y el mismo resultado en la que no tiene la enfermedad. Existen dos RV la RV+ y RV-, la primera es la razón entre la probabilidad de tener un test positivo en los enfermos y la probabilidad de obtener un test también positivo pero en los no enfermos y la segunda es la razón entre la probabilidad de tener un test negativo en los enfermos

y la de un test negativo en los no enfermos. Las fórmulas para el cálculo de cada una de ellas son las siguientes:^{11-23,29-33.}

$$RV+ = \frac{P(T+ / ENFERMO)}{P(T+ / NO ENFERMO)}$$

$$RV- = \frac{P(T- / ENFERMO)}{P(T- / NO ENFERMO)}$$

Sustituyendo en ambas fórmulas en términos de S y E sería:

$$RV+ = \frac{SENSIBILIDAD}{1 - ESPECIFICIDAD}$$

$$RV- = \frac{1 - SENSIBILIDAD}{ESPECIFICIDAD}$$

Un buen test debe tener una RV- cercana a 0 y una RV+ muy alta. Las RV tienen la ventaja de que no se ven afectadas por la prevalencia de la enfermedad en estudio, pueden calcularse para varios niveles del signo, síntoma o resultado de la prueba y constituyen una herramienta para reducir la lista de hipótesis diagnósticas.

Se verá ahora en un ejemplo cómo se interpretan las RV. *Romasan y colaboradores*³⁴ realizan una investigación para responder la interrogante ¿cuáles son los beneficios del uso del ultrasonido (US) en pacientes con sospecha clínica de salpingitis aguda y signos de la enfermedad inflamatoria pélvica (EIP)? y publican que la RV+ para la presencia de salpingitis aguda fue de 3,6 y la RV- de 0,23. Para la RV+ significa que son aproximadamente 4 veces más probables los hallazgos ultrasonográficos de salpingitis en las pacientes que la padecen con respecto a las que no la tienen. La RV- como se ha mencionado es la razón que permite conocer cuanto más probable es un resultado negativo en los enfermos que en los no enfermos y da un resultado menor que 1, lo cual es difícil de interpretar, lo que suele hacerse es dividir 1 entre el valor de la RV- y se obtiene un resultado que es equivalente a invertir la fórmula lo que antes era el denominador ahora es el numerador de la razón se había definido la RV- como 1-S/E ahora quedaría E/1-S. Por tanto $1/0.23=4,3$, se dice entonces que es 4 veces más probable no encontrar signos de salpingitis en US en aquellas pacientes que no la padecen que en las que realmente si están enfermas.

Una ventaja de las RV es que se pueden calcular para un resultado X cualquiera, es muy frecuente en la práctica clínica PD que se informan como un conjunto de resultados, signos o síntomas. Para un resultado Tx se puede calcular de la siguiente manera:

$$RV_x = \frac{P(T_x / Enfermo)}{P(T_x / No Enfermo)}$$

La fórmula anterior le da un poder a la RV, más generalizador y se obtiene la *RV de un determinado resultado*, lo cual permite conocer rápidamente si un resultado específico permite diferenciar entre enfermos y no enfermos. Si la RV es mayor que 1 apunta a la presencia de la enfermedad y si es menor que 1 hacia la ausencia.

Los signos y síntomas pueden tener diferente peso, importancia o valor para la realización del diagnóstico, algunos con su presencia son patognomónicos y, por tanto, su peso o valor es muy grande, otros sugieren el diagnóstico, pero como ocurren en muchas otras enfermedades, son inespecíficos, por lo que tienen un peso escaso, es decir, tan sólo con ellos es muy difícil hacer un diagnóstico.²⁸ Asumiendo la misma notación de la fórmula expresada anteriormente se explica según lo referido por *Santana-Chávez y colaboradores*²⁹ el significado de las RV para este caso particular. Los mencionados autores plantean que la RV permite identificar la importancia que tienen los signos y los síntomas para hacer un buen diagnóstico, ya que brindan información de cuanto más probable es que en presencia de un síntoma o signo "X", un paciente tenga una enfermedad "A" en comparación con el paciente "B" que no tiene el signo o síntoma "X".

Se retoma el ejemplo de la investigación publicada por *Romasan y colaboradores*.³⁴ Las pacientes con EIP pueden no solo tener salpingitis sino también endometritis, cervicitis, quistes ováricos, torsiones de las trompas, apendicitis, entre otros diagnósticos. El US podía revelar la presencia de determinados signos como masas anexiales bilaterales y masas anexiales bilaterales adyacentes al ovario. La RV+ para el primer signo en pacientes femeninas con salpingitis fue de 4,8 y la RV- de 0,22. Es aproximadamente 5 veces más probable la presencia de masas anexiales bilaterales en las pacientes con salpingitis que en las que tienen otros diagnósticos de la cavidad pélvica y con el mismo recurso explicado en con anterioridad para la RV-, es aproximadamente 5 veces más probable ($1/0,22 = 4,5$) que no exista ese signo en las pacientes con otros diagnósticos con respecto a las que poseen la enfermedad en cuestión.

Después de haber abordado una serie de elementos importantes acerca de los cuatro indicadores básicos de la validez de un PD se exponen algunos aspectos esenciales de un PD en la clínica descritos por *Cuadrado y colaboradores*.³⁵

1. La validez de una prueba diagnóstica se define como su capacidad para señalar correctamente la presencia o ausencia de la enfermedad.
2. Se expresa matemáticamente en índices: sensibilidad, especificidad, valor predictivo positivo y valor predictivo negativo.
3. Parte del análisis de pacientes sospechosos de padecer la enfermedad, en los que se realiza la prueba problema.
4. Sus resultados se comparan con la distribución real de la enfermedad: obtenidos en el mismo grupo de pacientes mediante una prueba de referencia.

¿Cuándo indicar o no un MD? ¿Cómo se decide qué conducta tomar? ¿Cómo se manobra la incertidumbre?

*Capurro*³ expone que cada vez que aplicamos un test adicional, lo hacemos por una de dos razones:

- Porque creemos necesario aumentar la probabilidad del diagnóstico a un nivel suficientemente alto, en el cual consideramos que vale la pena tomar las conductas de manejo y tratamiento concordantes con este.

- Porque queremos llevar la probabilidad a un nivel suficientemente bajo como para descartar.

Lo que habitualmente se hace es lo siguiente, si la PD da positiva se adjudica la enfermedad y si es negativa no, pero eso es erróneo porque las PD no son perfectas y pueden haber errores (falsos positivos y falsos negativos). Se debe adoptar otra posición más real, cada paciente tiene una probabilidad a priori de tener una enfermedad lo cual depende de las características higiénicas epidemiológicas del contexto que se trate. Cuando se aplica el PD, si es positivo dicha probabilidad se incrementará y si es negativo decrecerá. Existe una nueva probabilidad que es la posprueba (valores predictivos) la cual renovará la incertidumbre diagnóstica, pero no posibilitará posesionarse del diagnóstico inmediatamente, a no ser que esa probabilidad sea cercana al 100 % o a 0 %.^{1,36}

En el escenario clínico constantemente el médico de asistencia se mueve entre dos umbrales, el diagnóstico y el terapéutico, es lo que algunos autores han denominado enfoque de umbral para la toma de decisiones.^{1,36,37} *Ortín y colaboradores*²⁸ definen que "el concepto en que se basan los umbrales de decisión es que realizar un test diagnóstico, resulta inútil si no sirve para modificar la estrategia terapéutica o de manejo clínico global del paciente".

Para comprender dicho enfoque se necesita tener bien claro los términos de prevalencia de la enfermedad, de sensibilidad, especificidad, probabilidad pretest y postest abordados previamente. Es posible que la probabilidad a priori de determinada enfermedad sea muy baja como para no indicar ningún PD o muy alta, entonces si se indicará algún PD. Si la probabilidad a priori supera el umbral terapéutico, se asumirá determinado diagnóstico y si se indicará el tratamiento al paciente. De semejante forma si la probabilidad a priori no supera al umbral diagnóstico se descartará el diagnóstico. *Ochoa y colaboradores*¹ refieren que estos "umbrales son diferentes para cada escenario clínico y se situarán en niveles de probabilidad más a menos alejados respectivamente del 100 % y 0 % en función de la gravedad del diagnóstico, de su frecuencia esperada y del coste de los falsos positivos y negativos, tanto en términos clínicos (riesgo de la demora diagnóstica, interés del tratamiento precoz, etc.) como económicos".

Estos umbrales no son iguales para todas las situaciones clínicas y se encuentran en valores de probabilidades que oscilan entre 0 % y 100 % según la gravedad de la enfermedad, de su frecuencia esperada, de la tasa de falsos positivos y negativos, tanto en términos clínicos (la demora en hacer el diagnóstico, beneficio del tratamiento temprano precoz) como económicos. Podrá notarse que los elementos anteriormente referidos entrañan cierta subjetividad, por lo que los umbrales variarán en dependencia de las personas que los evalúe, incluso no podrían determinarse (*Fig. 1*).^{1,36,37} Cuando la probabilidad a priori se encuentra entre los dos umbrales diagnóstico y terapéutico, se necesita mayor información para aumentar el grado de certidumbre. Es por eso que se apela a la indicación del PD. Si el PD que se indica es válido, su resultado admitirá incrementar o reducir la estimación de la probabilidad de traspasar alguno de los dos umbrales (diagnóstico o terapéutico). En esa situación el juicio sobre el diagnóstico será correcto, si ocurre lo contrario, habrá que indicar nuevas PD o si no están disponibles, a diagnósticos presuntivos y pruebas terapéuticas debido a la incertidumbre (*Fig. 1*).^{1,36,37}

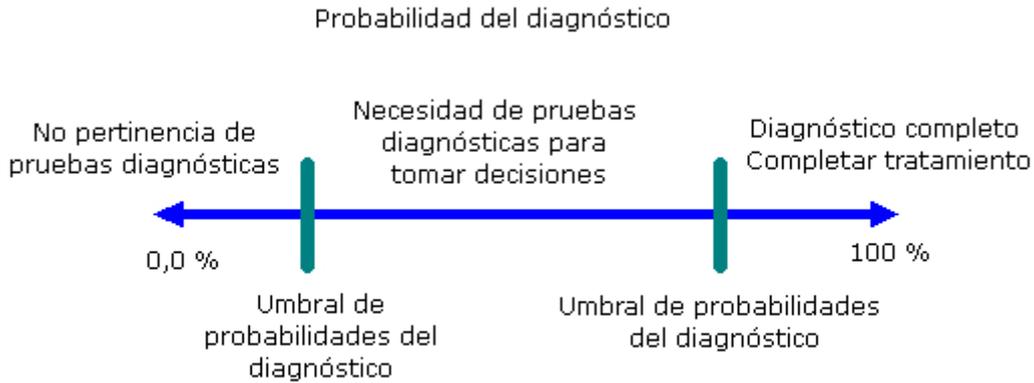


Fig. 1. Umbrales diagnóstico y terapéuticos.

Molina³⁶ en un artículo expone un ejemplo que permite comprender estos conceptos de umbrales diagnósticos y terapéuticos. A continuación se relata textualmente:

"Pensemos que nos llega un paciente de dos años de edad a consulta por presentar diarrea de características motoras. A pesar de que creamos que tiene una diarrea crónica inespecífica, decidimos pedirle unos anticuerpos antitransglutaminasa por aquello de que pueda ser celíaco (al fin y al cabo, es una enfermedad frecuente y la clínica no siempre es muy característica). Si los anticuerpos son negativos descartaremos la enfermedad con una seguridad razonable y decidiremos no hacer más pruebas diagnósticas: habremos traspasado el umbral de diagnóstico (también llamado umbral negativo de prueba). Pero si los anticuerpos son positivos ¿qué ocurre? Que la probabilidad de que sea celíaco es ahora más alta, pero nos parece insuficiente para retirar el gluten, así que decidimos biopsiar para confirmar el diagnóstico. Si en la biopsia vemos una atrofia vellositaria grave, la probabilidad de ser celíaco habrá aumentado un poco más, esta vez lo suficiente para la retirada del gluten de la dieta: habremos cruzado el umbral terapéutico (la sistemática del diagnóstico se ha simplificado para que sea más ilustrativa)".

CONCLUSIONES

Este artículo servirá como una guía para los profesionales de la salud que se debaten constantemente con la necesidad de poder hacer un diagnóstico correcto para poder tratar oportunamente a los pacientes y evitarles complicaciones y cuando sea posible evitar la muerte. Nos referimos a la importancia de los PD y su uso correcto. Se comentan cuestiones que se producen en la práctica y el médico necesita saber cuándo indicar o no un PD, cómo se decide qué conducta tomar y cómo se maneja la incertidumbre a través de los conceptos de umbrales diagnóstico y terapéuticos.

CONFLICTOS DE INTERESES

El autor declara no presentar conflictos de intereses en la preparación y publicación de este artículo y no tener otras relaciones o actividades que podrían influir sobre la publicación del mismo.

REFERENCIAS BIBLIOGRÁFICAS

1. Ochoa Sangrador C. Aprender a entender e interpretar las pruebas diagnósticas. Herramientas y aplicaciones. En AEPap ed. Curso de Actualización Pediatría 2015. Madrid: Lúa Ediciones 3.0; 2015[citado 28 ago 2015]:255-63. Disponible en: <https://www.aepap.org/sites/default/files/cursoaepap2015p255-263.pdf>
2. Medina MC. Generalidades de las pruebas diagnósticas, y su utilidad en la toma de decisiones médicas. Rev. Colomb. Psiquiat. 2011[citado 28 ago 2015];40(4):787-97. Disponible en: <http://www.redalyc.org/articulo.oa?id=80622345014>
3. Capurro D, Rada G. El proceso diagnóstico. Rev Méd Chile. 2007[citado 28 ago 2015];135:534-8. Disponible en: <http://www.scielo.cl/pdf/rmc/v135n4/art18.pdf>
4. Díaz Novás J, Gallego Machado B, León González A. El diagnóstico médico: bases y procedimientos. Rev Cubana Med Gen Integr. 2006[citado 25 ago 2015];22(1):11. Disponible en: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0864-21252006000100007&lng=es
5. Kostopoulou O, Oudhoff J, Nath R. Predictors of diagnostic accuracy and safe management in difficult diagnostic problems in family medicine. Med Decis Making. 2008;28:668-80. Disponible en: <http://mdm.sagepub.com/content/28/5/668.full.pdf>
6. Cisnero Álvarez Y, Hernández Castellanos G, Jiménez Leyva M, Hernández Castellanos G. Valoraciones sobre la crisis del método clínico en el nuevo milenio. Correo Científico Médico. 2013;17(1):64-9. Disponible en: <http://scielo.sld.cu/pdf/ccm/v17n1/ccm09113.pdf>
7. Shan X, Wang X, Yuan Q, Zheng Y, Zhang H, Wu Y, et al. Evaluation of the diagnostic accuracy of nonstructural protein 1 Ag-based tests for dengue virus in Asian population: a meta-analysis. BMC Infect Dis. 2015;21(15):360. doi:10.1186/s12879-015-1088-4. Disponible en: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4545931/pdf/12879_2015_Article_1088.pdf
8. Peng L, Bian XW, Li DK, Xu C, Wang GM, Xia QY, et al. Large-scale RNA-Seq Transcriptome Analysis of 4043 Cancers and 548 Normal Tissue Controls across 12 TCGA Cancer Types. Sci Rep. 2015[citado 28 ago 2015];21(5):18. Disponible en: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4544034/pdf/srep13413.pdf>
9. mMcCormick N, Bhole V, Lacaille D, Avina Zubieta JA. Validity of Diagnostic Codes for Acute Stroke in Administrative Databases: A Systematic Review. PLoS One. 2015[citado 31 ago 2015];10(8):26. Disponible en: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4546158/pdf/pone.0135834.pdf>
10. Renz PB, Cavagnolli G, Weinert LS, Silveiro SP, Camargo JL. HbA1c Test as a Tool in the Diagnosis of Gestational Diabetes Mellitus. PLoS One. 2015[citado 1 sept 2015];10(8):11. Disponible en: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4546239/pdf/pone.0135989.pdf>

11. Manterola C. Cómo interpretar un artículo sobre pruebas diagnósticas. Rev. Med. Clin. Condes. 2009;20(5):708-17. Disponible en: <http://www.scielo.cl/pdf/rchcir/v62n3/art18.pdf>
12. Burgos ME, C. Cómo interpretar un artículo sobre pruebas diagnósticas. Rev Chil Cir. 2010;62(3):301-8. Disponible en: <http://www.scielo.cl/pdf/rchcir/v62n3/art18.pdf>
13. Whiting PF, Davenport C, Jameson C, Burke M, Sterne JA, Hyde C, et al. How well do health professionals interpret diagnostic information? A systematic review. BMJ Open. 2015[citado 9 sep 2015];5(7):9. Disponible en: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4521525/pdf/bmjopen-2015-008155.pdf>
14. Avalos O. Las pruebas diagnósticas. Su aplicación en los estudios epidemiológicos. Nefrología. 2000;20(5):403-7. Disponible en: <http://www.revistanefrologia.com/es-publicacion-nefrologia-articulo-las-pruebas-diagnosticas-su-aplicacion-los-estudios-epidemiologicos-X0211699500012909>
15. Salech F, Victoria M, Larrondo F, Rada G. Estudios que evalúan un test Diagnóstico: interpretando sus resultados. Rev Méd Chile. 2008;136:1203-08. Disponible en: <http://www.scielo.cl/pdf/rmc/v136n9/art18.pdf>
16. Altman DG, Bland JM. Statistics Notes: Diagnostic tests 1: sensitivity and specificity. BMJ 1994;308:1552. Disponible en: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2540489/pdf/bmj00444-0038.pdf>
17. Castro Bouza M. Algo básico sobre los instrumentos de medida: validez, fiabilidad, sensibilidad y especificidad. Revista Gallega de Terapia Ocupacional TOG. 2005[citado 28 ago 2015];2:16. Disponible en: <http://www.revistatog.com/num2/pdfs/num2art3.pdf>
18. Sánchez NA. ¿Son la sensibilidad y la especificidad medidas obsoletas para determinar la bondad de una prueba diagnóstica? Rev. Fac. Nac. Salud Pública. 2002;20(1):149-59. Disponible en: <http://www.redalyc.org/articulo.oa?id=12020113>
19. Molinero LM. Valoración de pruebas diagnósticas. Asociación de la Sociedad Española de Hipertensión. Liga Española para la lucha contra la Hipertensión Arterial. 2002[citado 28 ago 2015]: 7. Disponible en <http://www.seh-lelha.org/pdiagnos.html>
20. Campillo Soto A, Ramón Lirón Ruiz R, Martín Lorenzo JG, Aguayo Albasini JL. Algunas consideraciones sobre la interpretación de resultados de estudios de pruebas diagnósticas: estudio DINO. Rev Esp Cardiol. 2009;62(10):2. Disponible en: http://pdf.revespcardiol.org/watermark/ctl_servlet? f=10&pident_articulo=13141825&pident_usuario=0&pcontactid=&pident_revista=25&ty=133&accion=L&origen=cardio&web=www.revespcardiol.org&lan=es&fichero=25v62n10a13141825pdf001.pdf
21. Escrig Sos J, Martínez Ramos D, Miralles Tena JM. Pruebas diagnósticas: nociones básicas para su correcta interpretación y uso. Cirugía Española. 2006;79(5):267-73. Disponible en: www.elsevier.es/es-revista-cirugia-espanola-36-pdf-13087491-S300

22. Valenzuela L, Cifuentes L. Validez de estudios de test diagnósticos. Rev Med Chile. 2008;136:401-4. Disponible en: <http://www.scielo.cl/pdf/rmc/v136n3/art18.pdf>
23. Abraira V. Sesgos en los estudios sobre pruebas Diagnósticas. SEMERGEN. 2006;32(1):24-6. Disponible en: <http://www.elsevier.es/es-revista-semergen-medicina-familia-40-articulo-sesgos-los-estudios-sobre-pruebas-13083814>
24. Ochoa Sangrador C, González de Dios J, Buñuel Álvarez JC. Evaluación de artículos científicos sobre pruebas diagnósticas. Evid Pediatr. 2007[citado 28 de ago 2015];3:24:9. Disponible en: <http://www.evidenciasenpediatria.es/files/41-10675RUTA/Evaluaci%C3%B3n%20de%20art%C3%ADculos%20cient%C3%ADficos%20sobre%20prueb%20.pdf>
25. Marcos R. de Sousa MR, Ribeiro AL. Revisión Sistemática y Metaanálisis de Estudios de Diagnóstico y Pronóstico: una Guía. Arq Bras Cardiol. 2009[citado 28 ago 2015];92(3):235-45. Disponible en: http://www.scielo.br/pdf/abc/v92n3/es_13.pdf
26. Quispe Ramos D, Quispe Mamani J. Test Diagnostico. Rev. Act. Clin. Med. 2013;33:1695-99. Disponible en: <http://www.revistasbolivianas.org.bo/pdf/raci/v33/v33a07.pdf>
27. Khooshideh M, Radi V, Hosseini R, Hosseini L. The accuracy of placental alpha-microglobuline-1 test in diagnosis of premature rupture of the membranes. Iran J Reprod Med. 2015[citado 28 ago 2015];13(6):355-60. Disponible en: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4555055/pdf/ijrm-13-355.pdf>
28. Ortín Ortín E, Sánchez Sánchez JA, Menárguez Puche JF, Hidalgo García IM. Capítulo 9. Lectura crítica de un artículo sobre diagnóstico. 2009: En: Atención sanitaria basada en la evidencia. Su aplicación a la práctica clínica Comité editorial grupo ASBE: Sanchez López MC, Navarro Mateu F, Castillo Bueno MD, Menarguez Puche JF, Sanchez Sánchez JA. Integrar Magenta. Murcia. 2007:233-73. Disponible en: http://www.murciasalud.es/recursos/ficheros/136615-capitulo_9.pdf
29. Santana Chávez LA, Esparza Pérez. Razonamiento con los signos y los Síntomas. Med Int Méx. 2014[citado 28 ago 2015];30:442-50. Disponible en: http://www.cmim.org/boletin/pdf2014/MedIntContenido04_11.pdf
30. Cerda J Cifuentes L. Uso de tests diagnósticos en la práctica clínica (Parte 1): Análisis de las propiedades de un test diagnóstico. Rev. Chil. Infectol. 2010[citado 28 ago 2015];27(3):205-08. Disponible de: <http://www.scielo.cl/pdf/rci/v27n3/art04.pdf>
31. Cifuentes L, Cerda J. Uso de tests diagnósticos en la práctica clínica (Parte 2): Aplicación clínica y utilidad de un test diagnóstico. Rev. Chil. Infectol. 2010[citado 28 ago 2015];27(4):316-19. Disponible de: <http://www.scielo.cl/pdf/rci/v27n4/art05.pdf>
32. Zamora J, Abraira V. Análisis de la calidad de los estudios de evaluación de pruebas diagnósticas. Nefrología 2008[citado 28 ago 2015];28(Supl 2):42-45. Disponible de: <http://www.revistanefrologia.com/es-publicacion-nefrologia-articulo-analisis-calidad-los-estudios-evaluacion-pruebas-diagnosticas-X0211699508031945>

33. Ruiz de Adana Pérez R. Eficacia de una prueba diagnóstica: parámetros utilizados en el estudio de un test. *Jano*. 2009[citado 9 sep 2015]; 1736: 30-32. Disponible en:
http://www.jano.es/ficheros/sumarios/1/0/1736/30/00300032_LR.pdf
34. Romosan G, Bjartling C, Skoog L, Valentin L. Ultrasound for diagnosing acute salpingitis: a prospective observational diagnostic study. *Hum Reprod*. 2013[citado 28 ago 2015]; 28(6):1569-79. Disponible de:
<http://humrep.oxfordjournals.org/content/28/6/1569.full.pdf+html>
35. Cuadrado Cenzual MA, Ortega Madueño I, Arroyo M. Utilidad de las pruebas diagnósticas en la práctica clínica: Medicina de laboratorio basada en la Evidencia. Asociación Española de Biopatología Médica, Madrid. 2001[citado 3 sep 2015]; 687-14. Disponible en:
<http://www.aebm.org/formacion%20distancia/distancia%202011-2012/Actualizaciones/monografias%202011/3.-%20MBE.pdf>
36. Molina Arias M. Lectura crítica en pequeñas dosis. Características de las pruebas diagnósticas. *Rev Pediatr Aten primaria*. 2013; 15: 169-73. Disponible en:
http://scielo.isciii.es/pdf/pap/v15n58/lectura_critica.pdf
37. Skaltsa K, Jover LI, Carrasco JL. Estimation of the diagnostic threshold accounting for decision costs and sampling uncertainty. *Biometrical Journal*. 2010; 52(5): 676-97. Disponible de:
<http://onlinelibrary.wiley.com/doi/10.1002/bimj.200900294/pdf>

Recibido: 20 de abril de 2017.
Aprobado: 31 de julio de 2017.

Teddy Osmin Tamargo Barbeito. Hospital Clínico Quirúrgico "Hermanos Ameijeiras".
La Habana, Cuba.
Correo electrónico: teosmin@infomed.sld.cu