

GUÍA PARA LOS USUARIOS DE LA LITERATURA MEDICA

III. Cómo utilizar un artículo sobre pruebas diagnósticas

Tabla 1 - Evaluación y Aplicación de los Resultados de Estudios de Pruebas Diagnósticas.

<p>♦ A. Los resultados del estudio: ¿Son válidos? Guías primarias: ¿Hubo una comparación independiente y enmascarada con un patrón de referencia? ¿La muestra de pacientes incluyó un espectro adecuado de pacientes a los cuales la prueba diagnóstica será aplicada en la práctica clínica?</p> <p>Guías secundarias: ¿Los resultados de la prueba evaluada influyeron en la decisión de establecer el patrón de referencia? ¿Los métodos de ejecución de la prueba fueron descriptos con suficiente detalle como para permitir su reproducción?</p> <p>♦ B. Los resultados ¿Son útiles? Guía primaria ¿Se presentaron las razones de verosimilitud de la prueba o los datos necesarios para calcularlas?</p> <p>Guía secundaria Los resultados: ¿Me ayudarán en la atención de mis pacientes? ¿La reproducción de los resultados de la prueba y su interpretación serán satisfactorios en mi medio? ¿Los resultados son aplicables a mis pacientes? ¿Los resultados modificarán mis decisiones? ¿Los pacientes se verán beneficiados como resultado de la prueba?</p>

SITUACIÓN CLÍNICA

Usted es un médico de consulta al que un cirujano le pide que examine una paciente de 78 años, en el décimo día posterior a una cirugía abdominal, que presenta disnea en aumento durante las últimas 24 horas. Ella también ha estado experimentando lo que describe como una molestia en el tórax, que en ciertas ocasiones se incrementa con la respiración profunda. Los resultados anormales que se manifestaron en el examen físico, son una sensibilidad remanente en el abdomen y rales esparcidos en la base de ambos pulmones. La radiografía de tórax revela un pequeño derrame pleural derecho, pero ésta es la primera radiografía después de la operación. Los gases en la sangre arterial muestran un PO₂ de 70 mmHg, con una saturación del 92%. El electrocardiograma sólo muestra cambios no específicos.

Usted sospecha que la paciente, a pesar de recibir 5000 U de heparina dos veces al día, puede tener una embolia pulmonar (EP). Usted solicita un estudio de ventilación-perfusión (V/P) y el resultado transmitido a la enfermera por teléfono es de una “probabilidad intermedia” de embolia pulmonar (EP). Aún con cierta incertidumbre con respecto al diagnóstico, usted ordena una anticoagulación completa. A pesar de que usted ha usado frecuentemente el control de ventilación-perfusión en el pasado, y piensa que ha tenido una noción

bastante buena de cómo usar los resultados, se da cuenta de que su comprensión está basada en la intuición y en la práctica local, más que en las propiedades del control V/P provenientes de la literatura original. Por ello, en su camino al departamento de medicina nuclear para ver el estudio de V/P, usted se detiene en la biblioteca.

LA BÚSQUEDA

Su plan es encontrar un estudio sobre las propiedades del control V/P con respecto a las aplicaciones que usted utiliza en su práctica clínica en general y a este paciente en particular. Usted está familiarizado con el programa GRATEFUL MED y lo utiliza para su búsqueda. El programa le proporciona una lista de Medical Subject Headings (MeSh), y su primer elección es “embolia pulmonar”. Como hay 1749 artículos con ese encabezado MeSh de publicación entre 1989 y 1992 (el rango seleccionado), usted tendrá que recortar su búsqueda. Elige dos estrategias: levantar sólo los artículos que tengan “imágenes radionúclidas” como un subencabezado y que también tengan asociado el encabezado MeSH “estudio comparativo” (ya que usted necesitará un estudio comparativo del control V/P con algunos patrones de referencia). Esta búsqueda alcanza 31 artículos, de los cuales usted excluye 11 que evalúan nuevas técnicas de diagnóstico, 9 que se relacionan con el diagnóstico y tratamiento de trombosis venosa profunda, y uno que examina la historia natural de la embolia pulmonar. Los 11 restantes se refieren al control V/P en embolia pulmonar. Uno, sin embargo, es una editorial; cuatro se limitan en su objetivo (trabajando sólo con estudios de perfusión, en situaciones en las que la determinación del diagnóstico debería comenzar con una angiografía pulmonar, o con un simple defecto de perfusión). Del resto, el estudio de Investigación Prospectiva de Diagnóstico de Embolia Pulmonar uno (IPDEP) capta su atención, porque es una revista extensamente leída con la que usted se siente familiarizado y porque está referida en varios títulos de otras revistas. Usted imprime el resumen de este artículo y encuentra que incluye la siguiente información: entre los pacientes con resultado intermedio del control V/P, el 33% tuvo embolia pulmonar. Usted concluye que ha hecho una buena elección y recupera el artículo de los estantes de la biblioteca.

INTRODUCCIÓN

Regularmente los clínicos afrontan dilemas cuando ordenan e interpretan pruebas diagnósticas. La proliferación continua de la tecnología médica torna cada vez más importante la habilidad de los clínicos para evaluar los artículos sobre pruebas diagnósticas. Por este motivo, este artículo presenta los principios de una evaluación eficiente de los artículos sobre pruebas diagnósticas, y el uso óptimo de la información que ellos proporcionan.

Una vez que usted decide, como fue ilustrado en la situación clínica con el artículo sobre IPDEP, que un artículo es potencialmente relevante (esto es, que el título y el resumen sugieren que la información es directamente aplicable al problema del paciente que usted está tratando), usted puede plantearse las tres mismas preguntas que sugerimos en la Introducción y en los artículos de Terapia (**Tabla 1**).

Los resultados del estudio ¿Son válidos?

El hecho de que uno crea en los resultados de un estudio, está determinado por los métodos que se aplicaron cuando se llevó a cabo. Decir que los resultados son válidos implica que la exactitud de la prueba diagnóstica, tal como se reporta, se acerca suficientemente a la verdad como para justificar el próximo examen del estudio. Primero, usted debe determinar si puede confiar en los resultados del estudio considerando de qué forma los autores reunieron sus pacientes y cómo aplicaron la prueba y si tomaron un patrón de referencia apropiado para los pacientes.

¿Cuáles son los resultados del estudio?

Si usted decide que los resultados son válidos, el próximo paso es determinar la exactitud de la prueba diagnóstica. Esto se realiza examinando (o calculando usted mismo) las razones de verosimilitud de la prueba (referidas frecuentemente como las “propiedades” de la prueba).

¿Los resultados me ayudarán en la atención de mis pacientes?

El tercer paso es decidir cómo utilizar la prueba, tanto para el paciente individual, como para su práctica en general. Debemos preguntarnos:

¿Son generalizables los resultados del estudio? -es decir, ¿puede usted aplicarlos a este paciente en particular y al tipo de pacientes que usted atiende más frecuentemente?

¿Con qué frecuencia los r resultados de la prueba proporcionan información valiosa? ¿La prueba, proporciona información adicional por encima y más allá de la historia clínica y el examen físico?

¿Es menos costosa o más fácilmente disponible que otras pruebas diagnósticas para la misma enfermedad en cuestión?

Por último, ¿los pacientes experimentaron una mejoría cuando se les aplicó la prueba?

En la primera parte (A), trataremos en detalle la primer pregunta, mientras que en la segunda parte (B) nos ocuparemos de la segunda y tercer pregunta. Utilizaremos el artículo de IPDEP para ilustrar el proceso.

En el estudio de IPDEP, a 731 pacientes que consintieron en ingresar al estudio, y de los cuales se sospechaba que podían padecer EP, se les realizó un estudio de V/P y una angiografía pulmonar. Se consideró que la angiografía pulmonar era el mejor camino para probar si el paciente padecía realmente EP y por ello fue el patrón de referencia. Cada angiografía se interpretó según los siguientes tres resultados: EP presente, EP incierta, o EP ausente.

La precisión del estudio V/P se comparó con la angiografía, y el resultado del estudio V/P se reportó según una de las siguientes categorías: probabilidad alta (de EP), probabilidad intermedia, probabilidad baja, o casi normal o normal. Las comparaciones de los estudios V/P y las angiografías se muestran en las Tablas 2 y 3. Las diferencias entre estas tablas las trataremos más adelante; por ahora aplicaremos la primer pregunta de las tres de este artículo.

Tabla 2 - Relación entre los resultados de la angiografías pulmonares y los resultados de los centellogramas de Ventilación/Perfusión en pacientes con angiografías exitosas.

Categoría del centellograma V/P	Angiografía	
	EP Presente	EP Ausente
Probabilidad alta	102	14
Probabilidad intermedia	105	217
Probabilidad baja	39	199
Normal/ Casi normal	5	50
TOTAL	251	480

Tabla 3- La relación entre los resultados de las angiografías pulmonares y los resultados de los centellogramas de Ventilación/Perfusión *

Categoría del centellograma V/P	Angiografía	
	EP Presente	EP Ausente
Probabilidad alta	102	14
Probabilidad intermedia	105	217
Probabilidad baja	39	273
Normal/ Casi normal	5	126
TOTAL	251	630

* Incluye 150 pacientes con controles V/P de probabilidad baja y normales o casi normales, sin angiografías (n: 136) o no interpretables (n: 14), y tromboembolismos no clínicamente importantes en el seguimiento.

♦ A.

LOS RESULTADOS DEL ESTUDIO ¿SON VÁLIDOS?

Guías Primarias

¿Hubo una comparación independiente y enmascarada con un patrón de referencia?

La mejor forma de determinar la precisión de una prueba diagnóstica es compararla con la “verdad”. Para ello, los lectores deben asegurarse de que se ha aplicado a todos los pacientes un patrón de referencia (como una biopsia, cirugía, autopsia, o un seguimiento a largo plazo) apropiado, junto con la prueba en investigación. En el estudio de IPDEP, las angiografías pulmonares se utilizaron como patrón de referencia, ya que pudieron realizarse sin sacrificar a los pacientes. En la lectura de artículos sobre pruebas diagnósticas, si usted no puede aceptar el patrón de referencia (entre las razones, es que ¡nada es perfecto!), entonces es improbable que el artículo le proporcione resultados válidos para sus propósitos.

Si usted acepta el patrón de referencia, la próxima pregunta es saber si los resultados de la prueba y del patrón de referencia fueron determinados independientemente unos de otros (es decir, que los encargados de determinar el resultado no tenían conocimiento de la otra prueba). Nuestra propia experiencia clínica, nos muestra el motivo por el cual esto es importante. Una vez que nos han mostrado un nódulo pulmonar en una tomografía computada, vemos la lesión antes desapercibida en la radiografía de tórax; una vez que conocemos el resultado del ecocardiograma, comenzamos a oír el murmullo cardíaco previamente inaudible. Cuanto mayor sea la probabilidad de que la interpretación de una prueba pueda estar influenciada por el conocimiento del resultado del patrón de referencia (o viceversa), mayor es la importancia de la interpretación independiente de ambas. Los investigadores de la IPDEP no establecieron explícitamente en el artículo que las pruebas hayan sido interpretadas a ciegas. Sin embargo uno podría deducir del esfuerzo puesto en asegurar la reproducción de lecturas independientes, que los intérpretes hayan estado de hecho enmascarados. Hemos confirmado a través de correspondencia con uno de los autores que así fue. Cuando estas cuestiones están en duda, a la mayoría de los autores les satisface clarificarlas si son contactados en forma directa.

¿La muestra de pacientes incluye un espectro apropiado de pacientes a los cuales la prueba diagnóstica será aplicada en la práctica clínica?

Una prueba diagnóstica es realmente útil, sólo en la medida que distinga la enfermedad en cuestión o estados, que de otra manera podrían ser confusos. Prácticamente todas las pruebas pueden distinguir el estado de salud en una afección severa; ésta habilidad no nos dice nada con respecto a la utilidad de la prueba. En

verdad, el valor práctico de una prueba se establece sólo en un estudio que refleja en forma cercana lo que sucede en la práctica clínica.

Un vívido ejemplo de cómo las esperanzas surgidas con la introducción de una prueba diagnóstica pueden destruirse en investigaciones subsecuentes, proviene de la historia del antígeno carcinoembrionario (ECE) en cáncer colorrectal. Los niveles de ACE, medidos en 36 personas con un cáncer avanzado confirmado de colon o recto, fueron altos en 35 de ellos. Al mismo tiempo, se encontraron niveles muy inferiores en personas normales y con una variedad de otras condiciones. Los resultados sugirieron que las mediciones de los niveles de ACE podrían ser útiles en el diagnóstico de cáncer colorrectal o aún en el rastreo de la enfermedad. Un estudio subsiguiente de pacientes con estados menos avanzados de cáncer colorrectal (y, por lo tanto, de una severidad menor de la enfermedad) y pacientes con otros cánceres u otros desórdenes gastrointestinales (y, por lo tanto, con desórdenes diferentes pero potencialmente confundentes), la exactitud de las mediciones de ACE cayó abruptamente, y se abandonó el uso de los niveles de ACE para el diagnóstico y screening de cáncer. El ACE se recomienda actualmente sólo como un elemento en el seguimiento de pacientes con cáncer colorrectal confirmado, pues disminuye con la extirpación del tumor.

En el estudio IPDEP, el espectro completo de pacientes de los cuales se sospechaba que padecían EP fue elegido y reclutado, incluyendo a aquéllos que entraron al estudio con sospecha clínica alta, media y baja de EP. Podemos concluir entonces, que se eligió la muestra de pacientes apropiada.

Guías Secundarias

Una vez que usted se ha convencido de que el artículo describe un espectro apropiado de pacientes que fueron sometidos a la comparación independiente y enmascarada de una prueba diagnóstica y un patrón de referencia, es más probable que los resultados representen una estimación sin sesgos de la exactitud real de la prueba -es decir, una estimación que no distorsione sistemáticamente la verdad- Sin embargo, usted puede reducir aún más las probabilidades de confundirse considerando otros aspectos.

¿Los resultados de la prueba influyeron en la decisión de realizar el patrón de referencia?

Las propiedades de una prueba diagnóstica se distorsionarán si sus resultados influyen la decisión de someter a los pacientes a la confirmación de un patrón de referencia. Esta situación, algunas veces llamada **“sesgo de verificación”** ó **“sesgo de elaboración”** se aplicaría, por ejemplo, cuando los pacientes con sospecha de padecer una enfermedad de las arterias coronarias y ergometrías positivas, tienen más probabilidad de ser sometidos a una angiografía (patrón de referencia), que aquellos con ergometrías negativas.

El sesgo de verificación fue un problema para el estudio IPDEP. Los pacientes en los que los estudios de V/P resultó más positivo (92%). Esto no es sorprendente ya que los clínicos serían renuentes a someter a pacientes con una probabilidad baja de EP, a los riesgos de una angiografía. Los resultados del estudio IPDEP restringidos a aquellos pacientes con una angiografía exitosa, se presentan en la Tabla 2.

La mayoría de los estudios deberían detenerse aquí, y los lectores deberían concluir que la magnitud del sesgo resultante de diferentes proporciones de pacientes con probabilidad alta y baja en los controles V/P, sometidos a angiografías adecuadas, es incierto pero quizás alto.

Sin embargo, los investigadores del IPDEP aplicaron un segundo patrón de referencia a los 150 pacientes con controles de probabilidad baja o normal/casi normal que no fueron sometidos a angiografía (136 pacientes) o a aquellos cuya interpretación de la angiografía fue incierta (14 pacientes): ellos deberían ser juzgados libres de EP, si continuaron en buen estado de salud sin la aplicación de ningún tratamiento.

Por consiguiente, los investigadores siguieron a todos los pacientes durante 1 año sin tratarlos con anticoagulantes. Ninguno de esos pacientes desarrolló evidencias clínicas de EP durante ese tiempo, por lo que podemos concluir que no existió una EP clínicamente (si definimos EP clínicamente importante a aquellas que requiere anticoagulación para prevenir eventos adversos subsiguientes), en el momento en que fueron sometidos al estudio V/P. Cuando esos 150 pacientes, juzgados libres de EP mediante este segundo patrón de referencia de buen pronóstico sin terapia anticoagulante, fueron agregados a los 480 pacientes con angiografías negativas en la Tabla 2, el resultado fue el que se muestra en la Tabla 3.

Nosotros esperamos que usted concuerde con nosotros que el mejor estimador de la precisión del estudio V/P proviene de la Tabla 3, que incluye los 150 pacientes que, en el seguimiento, no padecieron EP clínicamente importante. por consiguiente, utilizaremos aquellos datos en los cálculos subsiguientes.

Hubo aún 50 pacientes con estudios con probabilidad alta o intermedia que no se sometieron a una angiografía o cuyas angiografías fueron ininterpretables. Es posible que aquellos individuos pudieran sesgar los resultados. Sin embargo, ellos son una proporción relativamente pequeña de la población, y si sus características clínicas no son claramente diferentes de aquellos que se sometieron a la angiografía, es improbable que las propiedades de la prueba difieran sistemáticamente en la subpoblación. Por lo tanto, podemos tener cierta confianza en los resultados del IPDEP>

¿Los métodos de ejecución fueron descritos con suficiente detalle como para permitir su reproducción?

Si los autores concluyeron que usted debería utilizar la prueba diagnóstica, ellos deben decirle cómo usarla. La descripción debería cubrir todos los aspectos que son importantes en la preparación del paciente (dieta, drogas que deben ser evitadas, precauciones después de la prueba), la ejecución de la prueba (técnica, posibilidad de dolor), y el análisis e interpretación de sus resultados.

Una vez que el lector tiene confianza en que los resultados del artículo constituyen una estimación sin sesgo de las propiedades de la prueba, él puede determinar exactamente cuáles (y cuán útiles) son aquellas propiedades.

Aún sin ser impecables (los estudios casi nunca lo son) , nosotros podemos inferir firmemente que los resultados son una estimación válida de las propiedades del control V/P.

A continuación describiremos cómo interpretar y aplicar los resultados.

◆ B.

◆ LOS RESULTADOS ¿SON ÚTILES?

¿Cuáles son los resultados? ¿Me ayudarán para la atención de mis pacientes?

SITUACIÓN CLÍNICA

Usted está nuevamente en la situación en que lo ubicamos al comienzo de este artículo sobre pruebas diagnósticas, en la biblioteca, estudiando un artículo que lo guiará en la interpretación de los centellogramas pulmonares de ventilación/perfusión. Aplicando los criterios de la Tabla 1, usted ha decidido que la Investigación Prospectiva de Diagnóstico de Embolia Pulmonar (IPDEP) le proporcionará información válida. Casualmente llega una médica buscando un artículo que la ayude en la interpretación de los centellogramas V/P.

Su paciente es un hombre de 28 años que tuvo un acceso agudo de disnea y un vago dolor de tórax que comenzaron al finalizar un viaje en automóvil de 10 horas. El paciente había experimentado situaciones de molestias similares en el pasado, pero ninguno con esta severidad, y está muy preocupado por estos síntomas.

Efectuado un examen físico, un electrocardiograma y una telerradiografía de tórax, todos de características normales, con una medición de gases en sangre que muestra una PO_2 de 32 mmHg y una PO_2 de 82 mmHg, su colega ordenó un centellograma V/P.

Los resultados fueron informados como un control de probabilidad intermedia. Usted le comenta a su colega como utilizó el GRATEFUL MED para encontrar un artículo excelente sobre la precisión de los centellogramas V/P. Ella está complacida de que usted haya encontrado un artículo válido, y usted está de acuerdo en unir sus fuerzas a fin de aplicarlo a ambos pacientes.

En la situación clínica anterior, nosotros presentamos una propuesta para decidir si un artículo era válido y por consiguiente si se justificaba considerar sus resultados.

Al analizar esta nueva situación clínica, exploraremos los próximos pasos que envuelven la comprensión y el uso de los resultados de estudios válidos sobre pruebas diagnósticas.

¿cuáles fueron los resultados?

¿Se presentan las razones de verosimilitud para los resultados de la prueba diagnóstica o se incluyen los datos necesarios para su cálculo?

Probabilidad pre-prueba. El punto inicial para cualquier proceso de diagnóstico es el paciente, con su constelación de síntomas y signos.

Considere a los dos pacientes que dieron lugar a éste ejercicio - la mujer de 78 años, 10 días después de una cirugía abdominal y al hombre de 28 años angustiado, ambos con disnea y dolor de tórax no específico.

Nuestros clínicos intuyen la probabilidad de una embolia pulmonar (EP) como explicación de sus males, es decir, sus *probabilidades preliminares*. Estas son muy distintas para cada uno de ellos: la probabilidad de EP de la mujer mayor es alta en tanto que la del hombre joven es baja.

En consecuencia, aún si ambos centellogramas V/P muestran resultados compatibles con *probabilidad intermedia*, es muy probable que el manejo posterior de cada uno de ellos sea diferente. Uno podría tratar a la mujer mayor, pero ordenaría investigaciones adicionales para el paciente joven.

De esta línea de razonamiento emergen dos conclusiones.

Primero: cualquiera sea el resultado del estudio V/P, éste no nos dice si la EP está presente. Lo que ellos hacen es modificar la probabilidad de EP previa a la prueba, logrando producir una nueva probabilidad post-prueba.

La dirección y la magnitud de este cambio de pre-prueba a post-prueba, está determinada por las *propiedades de la prueba*, y la propiedad en la que nos concentraremos en esta sección es **la razón de verosimilitud (RV)**.

Segundo: La segunda conclusión que podemos inferir de nuestros pacientes, es que la probabilidad pre-prueba ejerce una influencia mayor en el proceso de diagnóstico.

Cada ítem de la historia clínica y del examen físico es una prueba diagnóstica que puede incrementar o disminuir la probabilidad de la enfermedad en cuestión.

Considere el joven paciente que se presentó a su colega. El hecho de que presente disnea aumenta la posibilidad de una EP. Asimismo, haber permanecido inmóvil durante un lapso de 10 horas incrementa esa probabilidad, pero su edad, falta de antecedentes de enfermedad, examen físico normal y los resultados de los estudios efectuados: ECG, Rx Tx y gases en sangre, todas disminuyen esa posibilidad.

Si conociéramos las propiedades de cada una de estas porciones de información (y algunas de ellas tenemos), nosotros nos podríamos manejar secuencialmente a través de ellas, incorporando cada porción de información a medida que la vamos obteniendo, recalculando continuamente la probabilidad de la enfermedad en cuestión.

Los clínicos proceden de esta forma, pero como las propiedades de los ítems individuales de la historia clínica y del examen físico usualmente no están disponibles, ellos deben confiar frecuentemente en la experiencia clínica y en la intuición para determinar las probabilidades pre-prueba que preceden a la orden de una prueba diagnóstica.

Para algunos problemas clínicos, incluyendo el diagnóstico de EP, su intuición les ha proporcionado una precisión sorprendente. Sin embargo, la limitada información sobre las propiedades de los ítems de la historia clínica y del examen físico, deriva frecuentemente en una variedad muy grande de estimaciones pre-prueba determinada por los clínicos.

Existe una gran cantidad de soluciones para este problema. En primer término los investigadores clínicos deberían estudiar la historia clínica y el examen físico para aprender más sobre las propiedades de estas pruebas diagnósticas. Afortunadamente, estas investigaciones se están volviendo comunes. Panzer y sus colegas han resumido mucha de esta información disponible en la forma de un texto médico, y se están publicando revisiones sistemáticas sobre la exactitud y precisión de las historias clínicas y exámenes físicos, concurrentemente con la serie de Guías para los Usuarios en el JAMA. sobre el Examen Clínico Racional.

Además, para algunas afecciones en cuestión., como la isquemia de miocardio, los análisis multivariados pueden proporcionar a los médicos formas de combinar información para generar probabilidades pre-prueba que sean precisas.

En segundo término, cuando no conocemos las propiedades de la historia clínica y del examen físico, podemos consultar a colegas sobre sus estimaciones de estas probabilidades. Es probable que una visión consensuada sea más precisa que nuestra intuición individual.

Finalmente, cuando permanecemos inseguros sobre la probabilidad pre-prueba, podemos asumir la probabilidad pre-prueba mas alta y la mas baja posibles, y ver si estas cambian el curso clínico de nuestra acción.

Ilustraremos en esta exposición, como podría realizarse esto último.

Tabla 2. Propiedades del centellograma de Ventilación /Perfusión

Embolia pulmonar					
Resultado del Centellograma V/P	Presente		Ausente		Razón de Verosimilitud
	N	Proporción	N	Proporción	
Probabilidad alta	102	102/251 = 0.406	14	14/630 = 0.022	18.3
Prob. intermedia	105	105/251 = 0.418	217	217/630 = 0.344	1.2
Probabilidad baja	39	39/251 = 0.155	273	273/630 = 0.433	0.36
Normal/ Casi normal	5	5/251 = 0.020	126	126/630 = 0.200	0.10
TOTAL	251		630		

Razones de verosimilitud

La utilidad clínica de una prueba diagnóstica se determina fundamentalmente por la exactitud con la que ella identifica la enfermedad en cuestión y la medida de exactitud en la que nos concentraremos es la *razón de verosimilitud (RV)*.

Miremos ahora la Tabla 2, completada con los resultados del artículo del estudio IPDEP. En el hubo 251 personas con EP comprobada mediante angiografía, y 630 personas cuyas angiografías o seguimientos

excluyeron la EP. Para todos los pacientes, los estudios V/P se clasificaron en cuatro niveles de probabilidad, desde probabilidad alta hasta normal o casi normal.

Cuán probable es obtener un centellograma V/P de probabilidad alta entre las personas con EP?

La tabla 2 muestra que 102 de 251 personas (ó 0.406) con EP tuvieron controles de probabilidad alta.

¿Cuán frecuente es el mismo resultado de la prueba, un decir un resultado del estudio con probabilidad alta, entre las personas que ha pesar de que se sospechaba EP, finalmente no lo padecían?

La respuesta es 14 de 630 ó 0.022. La razón de estas dos probabilidades se llama *razón de verosimilitud (RV)* y para un control de probabilidad alta es igual a $0.406/0.022 = 18.3$. En otras palabras, un control pulmonar de probabilidad alta es 18.3 veces más probable de obtenerse, *en un paciente con, en oposición a un paciente sin*, una EP.

De manera similar la razón de verosimilitud puede calcularse para cada nivel de resultado de la prueba diagnóstica. Cada cálculo envuelve la respuesta de dos preguntas:

- 1) - **Cuán probable es obtener un resultado dado (ej. un centellograma V/P de probabilidad baja) entre la gente con la enfermedad en cuestión (EP).**
- 2) - **Cuán probable es obtener el mismo resultado de la prueba (nuevamente un estudio con probabilidad baja), entre la gente sin la enfermedad en cuestión (sin EP).**

Para un control V/P de probabilidad baja, estas probabilidades son $39/251 = 0.155$ y $273/630 = 0.433$ y su razón de verosimilitud es 0.36. Como se ve en la Tabla 2 podemos repetir estos cálculos para el resto de los resultados del estudio.

¿Que significan todos estos números?

Las RV indican en cuánto, un resultado dado de la prueba diagnóstica aumentará o disminuirá la probabilidad pre-prueba de la enfermedad en cuestión.

Una RV de 1 significa que la probabilidad post-prueba es exactamente la misma que la probabilidad pre-prueba. Razones de verosimilitud mayores de 1 incrementan la probabilidad de que la enfermedad esté presente, y cuanto más grande la RV mayor es este incremento. Recíprocamente RV menores de 1, disminuyen la probabilidad de la enfermedad en cuestión, y cuanto menor la RV, mayor será la disminución de la probabilidad y menor su valor final.

¿Cuán grande es una RV grande y cuán chica una chica?

El empleo sistemático de la Rvs en la práctica diaria, lo conducirá a su propio juicio de interpretación, pero como una guía preliminar:

- ◆ Razones de verosimilitud mayores que 10 y menores que 0.1 generan amplios y frecuentes cambios concluyentes de la probabilidad pre-prueba a la post-prueba.
- ◆ Razones de verosimilitud de 5 a 10 y de 0.1 a 0.2, generan cambios moderados de la probabilidad pre-prueba a la post-prueba.
- ◆ Razones de verosimilitud de 2 a 5 y de 0.5 a 0.2 generan pequeños (pero algunas veces importantes) cambios en probabilidad.
- ◆ Razones de verosimilitud de 1 a 2 y de 0.5 a 1 alteran la probabilidad en un pequeño (y raramente importante) grado.

Una vez que determinamos la magnitud y la importancia de la Rvs, ¿cómo las utilizamos para ir de las probabilidades pre-prueba a las post-prueba? Nosotros no podemos combinar las verosimilitudes directamente, en la forma en que podemos combinar probabilidades o porcentajes. Su uso formal requiere convertir las

probabilidades pre-prueba en ocurrencias, multiplicando el resultado de la RV, y convirtiendo la ocurrencia post-prueba consecuente en una probabilidad post-prueba.

Si bien no es muy dificultoso, este cálculo puede ser tedioso e incluye los siguientes pasos: La ecuación para convertir probabilidades en ocurrencias es (probabilidad / [1 - probabilidad]), que es equivalente a la probabilidad de padecer la enfermedad en cuestión / probabilidad de no padecer la enfermedad en estudio.

Una probabilidad de 0.5 representa una ocurrencia de $0.5 / 0.5$, o de 1 a 1; una probabilidad de 0.80 representa una ocurrencia de $0.80 / 0.20$, o de 4 a 1; una probabilidad de 0.25 representa una ocurrencia de $0.25/0.75$, o de 1 a 3, o de 0.33.

Una vez que usted tiene la ocurrencia pre-prueba, la ocurrencia post-prueba puede calcularse multiplicando la ocurrencia pre-prueba por la RV. La ocurrencia post-prueba puede volverse a convertir en probabilidad utilizando una fórmula: probabilidad = ocurrencia / (ocurrencia + 1).

Afortunadamente existe un camino más fácil. Un nomograma compuesto por Fagan (ver fig. 1) que hace todas las conversiones por nosotros y nos permite ir sencillamente de la probabilidad pre-prueba a la probabilidad post-prueba.

La primer columna de este nomograma representa la probabilidad pre-prueba, la segunda columna representa la RV, y la tercera muestra la probabilidad post-prueba.

Usted obtiene la probabilidad post-prueba ubicando una regla en la probabilidad pre-prueba y rotándola hasta que se alinie con la RV del resultado de la prueba obtenido.

Recuerde nuestra mujer mayor con sospecha de padecer EP después de una cirugía abdominal. La mayoría de los clínicos coincidirían con que la probabilidad de esta paciente de padecer EP es muy grande, de aproximadamente 70 %. Este valor representa entonces a la probabilidad pre-prueba. Suponga que su centellograma V/P se reportó como de probabilidad alta. Ubicando una regla en su probabilidad pre-prueba de 70 % y alineándola con la RV de 18.3 asociada con el control de probabilidad alta, su probabilidad post-prueba es de más del 97 %.

¿Que pasa si su estudio V/P produce un resultado distinto? Si el resultado de su centellograma V/P se reporta como de probabilidad intermedia (RV 1.2), la probabilidad de EP cambia apenas (a 74 %), mientras que un resultado casi normal adjudica una probabilidad post-prueba de 19 %.

Tabla 3 . Probabilidades pre-prueba, razones de verosimilitud (Rvs) de los resultados de los centellogramas de ventilación perfusión y probabilidades post-prueba en dos pacientes con embolia pulmonar.

Probabilidad pre-prueba % (Rango)	Resultado del estudio (RV)	Probabilidad post-prueba % (Rango)
• Mujer de 78 años con inicio repentino de disnea luego de una cirugía abdominal		
70 (60-80)	Probabilidad alta (18.3)	97 (96-99)
70 (60-80)	Probabilidad intermedia (1.2)	74 (64-83)
70 (60-80)	Probabilidad baja (0.36)	46 (35-59)
70 (60-80)	Normal / casi normal (0.1)	19 (13-29)
• Hombre de 28 años con disnea y dolor torácico atípico		
20 (10-30)	Probabilidad alta (18.3)	82 (67-89)
20 (10-30)	Probabilidad intermedia (1.2)	23 (12.34)
20 (10-30)	Probabilidad baja (0.36)	8 (4.6)
20 (10-30)	Normal / casi normal (0.1)	2 (1-4)

Hemos señalado que la probabilidad pre-prueba es una estimación, y que una manera de manejar la incertidumbre es examinar las implicancias de un rango posible de probabilidades pre-prueba. Asumamos que la probabilidad pre-prueba de este caso, es tan baja como el 60 %, o tan alta como el 80 %. Las probabilidades post-prueba que se obtendrían con estas probabilidades pre-prueba diferentes aparecen en la Tabla 3.

El mismo ejercicio puede repetirse para nuestro segundo paciente, el hombre joven con un dolor de tórax no específico e hiperventilación. Si se considera que su cuadro es compatible con una probabilidad del 20 % de EP, usando nuestro nomograma de probabilidad post-prueba, con un resultado del control de probabilidad alta es del 82 %, con una probabilidad intermedia es del 23 %, y con una probabilidad normal / casi normal es del 2 %.

Las probabilidades pre-prueba (con un rango de probabilidad. pre-prueba posible entre el 10 % y el 30 %), las Rvs y las probabilidades post-prueba asociadas con cada uno de los cuatro resultados posibles del control se ven en la Tabla 3.

Los lectores que han seguido la exposición hasta este punto entenderán la esencia de la interpretación de una prueba diagnóstica y pueden detenerse aquí. Sin embargo, deberían considerar la opción de la siguiente sección que trata sobre la sensibilidad y especificidad de una prueba.

La incluimos en forma ampliada porque los clínicos todavía encontrarán estudios que presentan sus resultados en términos de sensibilidad y especificidad, y pueden querer comprender esta forma alternativa de resumir las propiedades de una prueba diagnóstica.

SENSIBILIDAD Y ESPECIFICIDAD

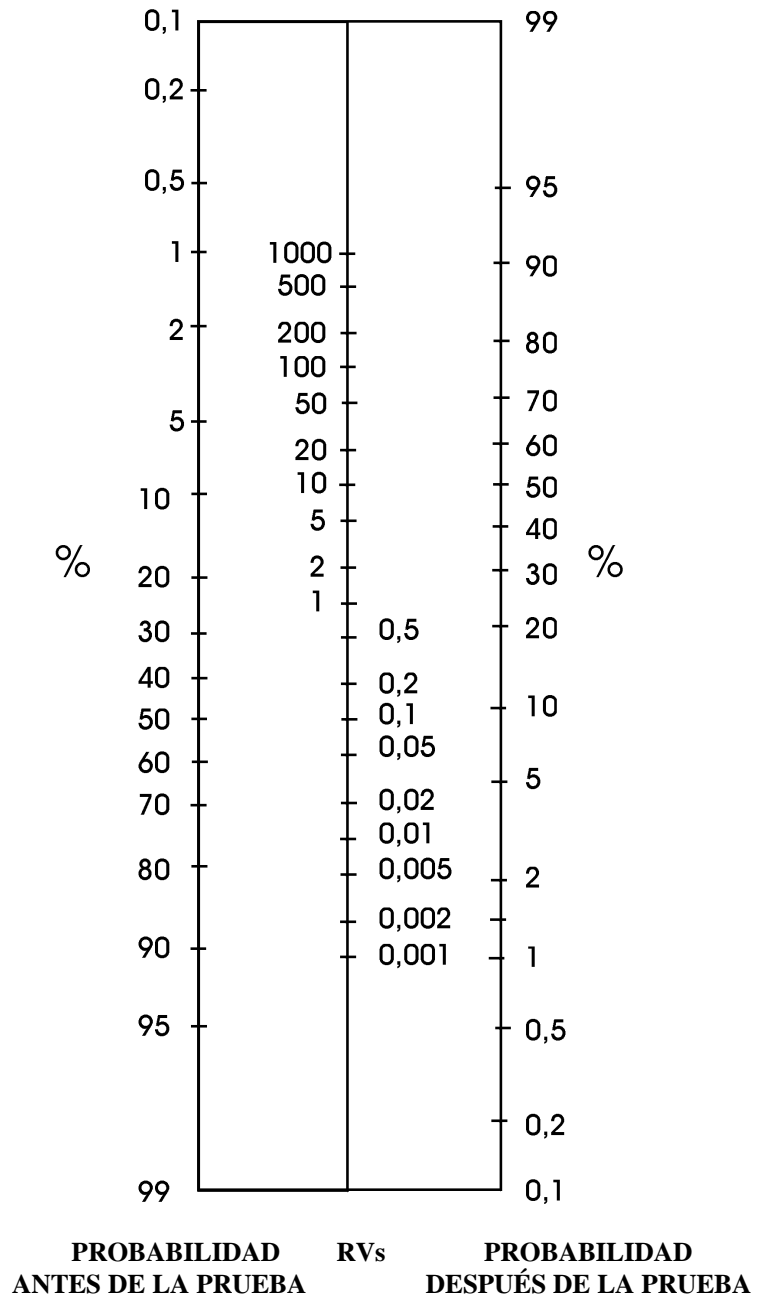
Usted puede haber notado que en nuestra exposición sobre Rvs no incluimos ningún comentario sobre pruebas normales y anormales. En su lugar, presentamos cuatro interpretaciones distintas de los exámenes V/P, cada uno con su correspondiente RV. Sin embargo, esta no es la forma en que los investigadores presentaron sus resultados. Ellos utilizaron los antiguos - pero menos útiles- criterios de sensibilidad y especificidad.

Tabla 4.- Comparación de los resultados de una prueba diagnóstica y un patrón de referencia

Resultado de la prueba	Patrón de referencia	
	Enfermedad presente	Enfermedad ausente
Enfermedad presente	Verdadero positivo (a)	Falso positivo (b)
Enfermedad ausente	Falso negativo (c)	Verdadero negativo (d)

- ◆ **Sensibilidad** = $a / (a+b)$
- ◆ **Especificidad** = $d / (b+d)$
- ◆ **RV para un resultado positivo de la prueba** = $\{ a / (a+c) \} / \{ b / (b+d) \}$
- ◆ **RV para un resultado negativo de la prueba** = $\{ c / (a+c) \} / \{ d / (d+b) \}$

NOMOGRAMA PARA LA INTERPRETACION DE LOS RESULTADOS DE LAS PRUEBAS DIAGNOSTICAS



La sensibilidad es la proporción de gente con la enfermedad en cuestión en la cual la prueba resulta positiva, y la especificidad es la proporción de gente sin la enfermedad en cuestión para la cual la prueba resulta negativa.

Para emplear estos conceptos debemos dividir los resultados de la prueba en normales y anormales, en otras palabras debemos crear una tabla de 2 x 2.

La forma general de la tabla de 2 x 2 que empleamos para comprender la sensibilidad y la especificidad se presenta en la Tabla 4. Mire nuevamente la tabla 2 y observe que podríamos transformar nuestra tabla 4 x 2 en cualquiera de las tres tablas de 2 x 2 que se forman según a que llamamos normal y anormal (ó a que llamamos resultado positivo y negativo de la prueba).

Asumamos que llamamos anormal solo a los resultados de probabilidad alta (o positivos). La tabla de 2 x 2 resultante se presenta en la Tabla 5.

Tabla 5. -Comparación de los resultados de una prueba diagnóstica (Centellograma V/P) con los resultados del patrón de referencia (angiografía pulmonar), asumiendo que solo los controles de probabilidad alta son positivos (verdaderamente anormales).*

Categoría del centellograma V/P	Angiografía	
	Embolia pulmonar presente	Embolia pulmonar ausente
Probabilidad alta	102	14
Otras	149	616
TOTAL	251	630

- ◆ *Sensibilidad: 41 %, especificidad: 98%, RV de un resultado de la prueba de probabilidad alta: 18.3
- ◆ RV de otros resultados: 0.61

Tabla 6. -Comparación de los resultados de una prueba diagnóstica (Centellograma de V/P) con los resultados del patrón de referencia (angiografía pulmonar) asumiendo que solo los controles normales/casi normales son negativos (verdaderamente anormal) *

Categoría del centellograma V/P	Angiografía	
	Embolia pulmonar presente	Embolia pulmonar ausente
Prob. alta, intermedia y baja	246	504
Normal /casi normal	5	126
TOTAL	251	630

- ◆ *Sensibilidad: 98 %, especificidad: 20 %, RV de una probabilidad alta, intermedia y baja: 1.23
- ◆ RV de normal/casi normal: 0.1

Para calcular la sensibilidad de los datos de la tabla 2, miramos el número de personas con EP comprobada (251) que se diagnosticaron como positivos por el control V/P: 102 - sensibilidad de 102/251, o aproximadamente 41 % ($a/\{a+c\}$). Para calcular la especificidad miramos la cantidad de personas sin la enfermedad en cuestión (630) que fueron diagnosticados como normales por el control V/P: 616 - especificidad de 616/630, o 98 % ($d/\{d+b\}$). También podemos calcular las Rvs para los resultados del estudio positivos y negativos usando este punto de corte, 18.3 y 0.6 respectivamente.

Veamos como actúa la prueba si decidimos determinar los positivos y los negativos en forma diferente en la Tabla 2. Por ejemplo, llamemos sólo a los controles V/P normales y casi normales resultados negativos. Esta tabla de 2 x 2 (Tabla 6) muestra que la sensibilidad es ahora 246/251, o 98 % (entre 251 personas con EP, fueron diagnosticadas 246 por los controles V/P, pero ¿que ha pasado con la especificidad?. Entre las 630 personas sin EP, sólo 126 tuvieron controles negativos (especificidad del 20 %). Las correspondientes RV son 1.23 y 0.1. Observe que con este corte no solamente perdemos información de diagnóstico asociada con el resultado de probabilidad alta del control, sino que también interpreta los resultados de probabilidad intermedia y baja como si incrementaran la verosimilitud de EP, cuando de hecho ellos disminuyen esa probabilidad. Usted mismo puede generar la tercer tabla de 2 x 2 colocando el punto de corte en el medio - si su sensibilidad y su especificidad son 82% y 63 %, respectivamente, y las Rvs asociadas de una prueba negativa y una positiva son 2.25 y 0.28, usted lo tiene bien. (Si usted crea un gráfico en el que el eje vertical representa la sensibilidad {o la proporción verdadero positivo}, para diferentes puntos de corte y en el eje horizontal, muestre 1 - especificidad {o la proporción de falsos positivos} para los mismos cortes, y usted une los puntos generados por los distintos puntos de corte, obtendrá lo que se denomina una curva de características funcionales (Receiver Operating Characteristic ROC). Una curva de estas puede utilizarse para comparar formalmente el valor de diferentes pruebas, examinando el área bajo la curva, pero una vez que uno ya no necesita un punto de corte, esta no tiene otra aplicación clínica directa.

Usted puede observar que utilizando la sensibilidad y la especificidad usted debe dejar de lado importante información o recalcular la sensibilidad y la especificidad para todos los puntos de corte. Nosotros recomendamos el enfoque de la RV porque es más sencillo y más eficiente.

De este modo hemos establecido que los resultados son probablemente ciertos para la personas incluidas en el estudio IPDEP, y determinamos las Rvs asociadas con los diferentes resultados de la prueba. ¿Cuán útil es la prueba en nuestra práctica clínica?

¿Me ayudarán los resultados en la atención de mi paciente?

¿La reproducción de los resultados de la prueba y su interpretación serán satisfactorios en mi medio?

El valor de cualquier prueba depende de su capacidad para proporcionar el mismo resultado cuando se aplica a los pacientes habituales. Una reproductibilidad pobre puede dar como resultado problemas con la prueba en sí misma (ej. variaciones en los reactivos en las cajas de radioinmunoensayos para determinar los niveles hormonales). Una segunda causa para obtener resultados diferentes en los pacientes habituales surge cuando una prueba requiere interpretación (Ej, la medida de elevación del segmento ST en un electrocardiograma) Lo ideal es que un artículo sobre una prueba diagnóstica le diga a los lectores cuán reproducible se puede esperar que sean los resultados de la prueba. Esto es especialmente importante cuando se requiere pericia para la ejecución o la interpretación de la prueba (y esto lo puede confirmar recordando las discrepancias clínicas que surgen cuando usted y uno o más colegas examinan el mismo electrocardiograma, ultrasonido, o tomografía computada, aún cuando todos ustedes son expertos.

Si la reproducibilidad de una prueba en el medio en estudio es mediocre y las discrepancias entre los observadores son comunes, y sin embargo la prueba aún discrimina bien a aquellos que tienen la enfermedad de aquellos que no la tienen, la prueba es muy útil. Bajo estas circunstancias, es probable que la prueba pueda ser fácilmente aplicada en su ambiente clínico. Si la reproducibilidad de una prueba diagnóstica es muy alta y la variación entre los observadores muy baja, la prueba es muy simple y no ambigua, o aquellos que la interpretan son muy hábiles. Si se trata de la última situación, intérpretes menos habilidosos en su práctica clínica puede no hacerlo tan bien.

En el estudio IPDEP los autores no solo proporcionaron una descripción detallada de sus criterios diagnósticos para la interpretación de los centellogramas V/P, sino que también informaron sobre la coincidencia entre sus dos observadores independientes. Las discrepancias clínicas entre los centellogramas de probabilidad intermedia y baja fueron comunes (25% y 30 % respectivamente), y en esos casos recurrieron a la opinión de un panel de expertos.

¿Los resultados son aplicables a mi paciente?

Este tema consiste en ver si la prueba tendrá la misma exactitud entre sus pacientes que la que se reportó en el artículo. Las propiedades de la prueba pueden cambiar con una mezcla de grados de severidad diferentes de la enfermedad o una distribución diferente de las condiciones competentes. Cuando los pacientes con la afección tienen todos una enfermedad severa, las Rvs se alejarán del valor 1 (se incrementa la sensibilidad). Si los pacientes sin la enfermedad en cuestión tienen unas condiciones competentes que mimetizan los resultados de la prueba vistos en pacientes que padecen la enfermedad en cuestión, las Rvs se acercarán a 1 y la prueba parecerá menos útil. En un grupo clínico diferente, en el que menos pacientes sin la enfermedad tengan estas condiciones competentes, las Rvs se alejarán de 1 y la prueba parecerá más útil.

El fenómeno de diferentes propiedades de la prueba para subpoblaciones diferentes fue notablemente demostrado por la ergometría en el diagnóstico de enfermedad de las arterias coronarias. Por ejemplo cuanto más pronunciada es la severidad de la enfermedad de arterias coronarias, más grandes son las Rvs de ergometrías anormales para las arterias coronarias con estenosis angiográficas. Otro ejemplo proviene del diagnóstico de tromboembolismo venoso, donde la compresión por ultrasonido para la trombosis de venas proximales ha probado ser mejor en pacientes ambulatorios sintomáticos que en pacientes postoperatorios sin síntomas.

Algunas veces, las pruebas fallan en aquellos pacientes en los que uno espera que funcione mejor. La RV de una prueba de tira diagnóstica para el diagnóstico rápido de infección de las vías urinarias, es aproximadamente 0.2 en pacientes con síntomas claros y por lo tanto una alta probabilidad de infección del tracto urinario, pero es de más de 0.5 en aquellos pacientes con baja probabilidad, esto la vuelve de poca ayuda para descartar la infección en estos últimos.

Si usted practica en un grupo similar a aquel de la investigación y su paciente cumple con los criterios de inclusión y no viola ninguno de los criterios de exclusión, usted puede estar confiado de que los resultados son aplicables. De otra manera se requiere una evaluación. Como en las intervenciones terapéuticas, usted se debería preguntar si existen razones apremiantes para las cuales los resultados no deberían ser aplicados a sus pacientes, ya sea por la severidad de la enfermedad de sus pacientes, o debido a que la mezcla de condiciones competentes sea tan distinta que la generalización no pueda garantizarse. El tema de la generalización puede resolverse si usted encuentra una revisión sistemática que combine los resultados de un cierto número de estudios

Los pacientes del IPDEP eran una muestra representativa de pacientes con sospecha de EP, provenientes de una cantidad de hospitales generales grandes. Por consiguiente, los resultados son realmente aplicables a la mayoría de los pacientes en América del Norte. Existen grupos para los cuales uno debería ser reacio a generalizar los resultados, como los pacientes críticamente enfermos (quienes fueron excluidos del estudio, y que es más probable que tengan un espectro de condiciones competentes, que otros pacientes).

¿Los resultados modificarán mis decisiones?

Es útil que la toma, enseñanza y comunicación de las decisiones de conducta, sean vinculadas explícitamente con la probabilidad de la enfermedad en cuestión. Por lo tanto, para cualquier enfermedad existen ciertas probabilidades bajo las cuales un clínico descartaría un diagnóstico y no ordenaría más pruebas (*un límite de prueba*). En forma similar, existen probabilidades sobre las cuales un clínico consideraría que el diagnóstico está confirmado, e interrumpiría las pruebas para comenzar con el tratamiento (*un límite de tratamiento*). Cuando la probabilidad de la enfermedad en cuestión se encuentra dentro de los límites de tratamiento y de prueba, se requieren más pruebas.

Una vez que se decide cuáles son nuestros límites de prueba y tratamiento, las probabilidades post-prueba tienen implicancias directas de tratamiento. Supongamos que estamos dispuestos a tratar aquellos con una probabilidad de EP del 80% o más, sabiendo que trataremos a un 20 % de nuestros pacientes innecesariamente. Además, supongamos que estamos dispuestos a descartar el diagnóstico de EP, en aquellos cuyas probabilidades post-prueba sean de 10 % o menos. Usted puede querer aplicar distintos números, aquí los límites de tratamiento y prueba son materia de juicio, y cambian para diferentes condiciones dependiendo de los riesgos de la terapia (si es riesgosa, usted querrá estar seguro de su diagnóstico) y del peligro que genera la enfermedad si permanece sin tratamiento (si el peligro de ignorar la enfermedad es alto, como en el caso de la EP, usted querrá considerar una probabilidad post-prueba muy baja antes de abandonar la búsqueda del diagnóstico).

En nuestro hombre joven, un centellograma de probabilidad alta generará una probabilidad post prueba de 82 % y puede indicar tratamiento (ó, al menos, más investigación), un centellograma de probabilidad intermedia (una probabilidad post-prueba de 23 %) indicará más pruebas -quizás una venografía de miembros inferiores, una pletismografía seriada por impedancia o ultrasonido, o una angiografía de pulmón-, mientras que un examen de probabilidad baja o normal (probabilidad < del 10 %) permitirá la exclusión del diagnóstico de EP.

En la mujer mayor, un centellograma de probabilidad alta indica tratamiento (probabilidad post-prueba de EP de 97 %), un resultado intermedio (probabilidad post-prueba de 74 %) puede ser compatible con el tratamiento o con más estudios (probablemente una angiografía pulmonar), mientras que cualquier otro resultado indica más investigación.

Si la mayoría de los pacientes tienen resultados con Rvs cercanas a 1, la prueba no resultará muy útil. Por lo tanto la utilidad de una prueba diagnóstica está fuertemente influenciada por la proporción de pacientes bajo sospecha de padecer la enfermedad en cuestión, cuyos resultados de la prueba tengan Rvs muy altas o muy bajas, y por lo tanto el resultado de la prueba cambiará sus probabilidades de enfermedad a través de un límite de enfermedad o de tratamiento. En los pacientes bajo sospecha de padecer EP de nuestro ejemplo de centellograma V/P, una revisión de la Tabla 2 nos permitirá determinar la proporción de pacientes con resultados extremos, (ya sea con una probabilidad alta y una RV de más de 10, ó con exámenes normales / casi normales y una RV de 0.1) La proporción puede calcularse como $(102 + 14 + 5 + 126) / 881$ ó $247/881 = 28\%$.

Los clínicos que se han frustrado repetidas veces por frecuentes resultados de probabilidades intermedias o bajas, en sus pacientes con sospecha de EP ya sabrán que esta proporción (28%) está lejos de ser óptima. Por lo tanto, a pesar de la alta RV asociada a un examen de probabilidad alta, y la baja RV asociada con un resultado normal/casi normal, el centellograma V/P tiene una utilidad limitada en pacientes con sospecha de EP.

Un comentario final tiene que ver con el uso secuencial de pruebas. Hemos demostrado como cada ítem de la historia clínica, o cada hallazgo de la revisión física, representa una prueba diagnóstica. Nosotros generamos probabilidades pre-prueba que modificamos con cada nuevo hallazgo. En general, también podemos usar las pruebas de laboratorio u otros procedimientos que imaginemos, de la misma manera. Sin embargo si dos pruebas están muy estrechamente relacionadas, la aplicación de la segunda prueba puede proporcionar muy poca o ninguna información, y la aplicación secuencial de las Rvs producirá resultados erróneos.

Por ejemplo, una vez que uno tiene los resultados de la prueba de laboratorio más potente para definir la deficiencia de hierro, *-ferritina sérica-*, pruebas adicionales como hierro sérico ó saturación de transferrina, no agregan ninguna información.

¿Los pacientes se verán beneficiados como resultado de la prueba?

El último criterio para determinar la utilidad de una prueba diagnóstica es comprobar si ella agrega alguna información diferente más allá de la disponible, y si esta información conduce a un cambio en las decisiones que posteriormente produce un beneficio al paciente. El valor de una prueba exacta, no será discutido, cuando la enfermedad en cuestión, si permanece no diagnosticada, es peligrosa, la prueba posee riesgos aceptables; y existen tratamiento efectivos. Los centellogramas de V/P de probabilidad alta o normal/casi normal pueden eliminar adecuadamente la necesidad de seguir investigando y como resultado se darán o se descartarán los anticoagulantes apropiadamente. Cualquiera de estos recursos de acción tienen una influencia substancial en el resultado del paciente.

En otras situaciones clínicas, las pruebas pueden ser exactas y el comportamiento puede modificarse como resultado de su aplicación, pero su impacto en el pronóstico final del paciente puede estar muy lejos de la certidumbre. Los ejemplos incluyen el cateterismo de corazón para muchos enfermos críticos, o el valor incremental de la resonancia magnética por imágenes sobre la tomografía computada para una gran variedad de problemas.
