

9



Lectura crítica de un artículo sobre diagnóstico

Enrique Ortín Ortín
Juan Antonio Sánchez Sánchez
Juan Francisco Menárguez Puche
Isabel M^a Hidalgo García.
Médicos de Familia y miembros del
Grupo de MBE de la Sociedad Murciana
de Medicina de Familia y Comunitaria.

1.INTRODUCCIÓN. OBJETIVOS DEL CAPÍTULO.

2.-EL PROCESO DE TOMA DE DECISIONES DIAGNOSTICAS.

- 2.1. Consideraciones generales sobre las pruebas diagnósticas.
- 2.2.¿Cuándo solicitar una prueba diagnóstica. Los Umbrales de Decisión.
- 2.3.Criterios de valoración de una prueba diagnóstica: Determinantes estadísticos.
- 2.4.Probabilidad y Odds de una prueba diagnóstica.
- 2.5.Validez Interna.
 - 2.5.1.Sensibilidad (S).
 - 2.5.2.Especificidad (E).
 - 2.5.3.Prevalencia de la enfermedad en la muestra estudiada.
 - 2.5.4.Prevalencia del test.
 - 2.5.5.Valor global.
- 2.6.Validez Externa.
 - 2.6.1.Valor predictivo positivo (VPP).
 - 2.6.2.Valor predictivo negativo (VPN).
 - 2.6.3.Probabilidad pretest (prevalencia).
 - 2.6.4.Razón o cociente de probabilidad o cociente de verosimilitud (RP).
 - 2.6.5.Probabilidad post-test.
 - 2.6.6.Normograma de Fagan.



3.OBTENIENDO Y VALORANDO INFORMACIÓN PARA UN PROBLEMA SOBRE DIAGNÓSTICO.

- 3.1.La pregunta clínica.
- 3.2.Buscando la información.
- 3.3. Lectura crítica de un artículo sobre pruebas diagnósticas.
- 3.4.Condiciones que debe reunir una publicación científica sobre una prueba diagnóstica.
- 3.5.Sesgos a tener en cuenta al realizar lectura crítica de un estudio sobre pruebas diagnósticas.
- 3.6.Un ejemplo concreto.
 - 3.6.1.Lectura crítica del artículo seleccionado.
 - 3.6.2.Ejemplo de aplicación de los determinantes estadísticos (realización de un CAT).

4.PUNTOS CLAVE.

5.-BIBLIOGRAFÍA.

En el sector Sanitario hay un importante desequilibrio entre la producción y el uso de la información científica: Esta situación que se ha hecho más ostensible durante los últimos años, obliga a los clínicos a desarrollar habilidades y estrategias de búsqueda de información útil: aquella que partiendo de las evidencias disponibles, aporta recursos para resolver las dudas y lagunas de conocimientos que habitualmente aparecen durante la práctica clínica. De esta forma puede disminuir el número de problemas no resueltos y mejorar la calidad de la asistencia.

En la práctica habitual, el médico se enfrenta a problemas clínicos concretos y es, tras el abordaje de la atención clínica de los pacientes, cuando las necesidades de información y el número de dudas que surgen son mayores. En varios estudios se ha podido comprobar que por cada 10 pacientes atendidos se genera una media entre 2 y 4 preguntas. Sin embargo solo un tercio de las preguntas que se plantean son solucionadas, quedando el resto sin respuesta.

La fuente de información más frecuentemente utilizada por los profesionales para resolver sus dudas suelen ser el vademécum y los libros de texto. Los médicos por otra parte solo realizan una estrategia de búsqueda sistematizada en muy contadas ocasiones (en algunos estudios sólo en 2 preguntas por cada 1100 planteadas). Esta situación puede ser debida en parte a que bases de datos como Medline no están preparadas para dar respuesta a preguntas clínicas, y a que se requiere tiempo y habilidades específicas para realizar búsquedas. Como consecuencia, el profesional tras una inversión importante de tiempo y esfuerzo, se encuentra ante una gran cantidad de información que le resulta irrelevante para responder a sus dudas clínicas.

Las decisiones clínicas son el resultado final de un proceso compuesto por una secuencia de acciones de la atención a un paciente. Se inicia a través de una adecuada anamnesis y exploración y con los datos recogidos se producirá el diagnóstico de sospecha o hipótesis diagnóstica. En ocasiones el proceso requerirá la solicitud de exploraciones complementarias, que ayuden a confirmar o descartar el cuadro clínico del paciente. El método clínico se desarrolla, por tanto, a través de una valoración de probabilidades diagnósticas basadas en los resultados de las pruebas diagnósticas (PD) realizadas.

9

→ 1

INTRODUCCIÓN. OBJETIVOS DEL CAPÍTULO

A partir de los recursos teóricos expuestos y de los ejemplos clínicos propuestos, el lector deberá ser capaz de:

- * Identificar e interpretar los parámetros que definen la validez y aplicabilidad de las pruebas diagnósticas (PD) en el proceso de atención clínica.
- * Buscar y realizar una lectura crítica básica de un artículo sobre validación de pruebas diagnósticas.

Supuestos clínicos

Ante un paciente que acude a consulta por dolor en la pantorrilla se pueden elaborar varias hipótesis diagnósticas según las características del paciente. En nuestro caso y por su especial relevancia, utilizaremos como ejemplo el trombo embolismo venoso profundo (TVP) en extremidades inferiores como cuadro clínico a confirmar:

Escenarios Clínicos

Caso A) Mujer de 39 años con Diabetes gestacional, a la que hace 20 días se le practicó una cesárea por un RN macrosómico con buen Apgar al nacer. Desde hace dos días nota dolor en la pantorrilla de la pierna izquierda. La exploración muestra dolor a la presión en anillo del soleo, edema en la zona del dolor y aumento del diámetro (>3cm) en relación con la otra pierna. Presenta calor local y no tiene circulación colateral.

Caso B) Hombre de 45 años que desde hace 7 días presenta dolor en pantorrilla derecha, que relaciona con una inusual caminata de 20 Km. Tiene dolor a la presión del gemelo con ligero aumento de la temperatura local, pero sin edema ni circulación colateral. Antecedente de varices previas.

El médico variará la solicitud de PD según las características de cada situación clínica, aunque, en ambos casos, el motivo de consulta sea un dolor en pantorrilla. Esta diferencia en el proceso de atención se debe al diferente diagnóstico de sospecha (probabilidad preprueba), que orientará el diagnóstico hacia pruebas que proporcionen información adicional útil, para confirmar o rechazar la sospecha (probabilidad postprueba). La probabilidad preprueba es la prevalencia de la enfermedad y define la frecuencia de la misma en un contexto clínico concreto.

En los casos enunciados, aunque los datos clínicos y la semiología de cada paciente no permitan confirmar o descartar un diagnóstico, son muy importantes para estimar las probabilidades de que la clínica en ambas situaciones sea producida o no por una TVP (prevalencia o probabilidad preprueba),

En las situaciones anteriores, el facultativo realizaría probablemente una revisión bibliográfica de la enfermedad, de la que obtendría los siguientes datos: para el caso A la prevalencia estimada de TVP es del 70% y para el B, la prevalencia estimada de TVP es del 15%.¹

La realización de otras PD para mejorar la información disponible, el conocimiento de la probabilidad preprueba (prevalencia) y de los cocientes de probabilidad (CP) de la prueba diagnóstica (Información sobre cuanto mayor es la probabilidad de encontrar un resultado positivo en un paciente con la enfermedad que en alguien sin ella), permiten calcular la probabilidad postprueba de tener esa enfermedad.

Por su importancia en el proceso de toma de decisiones, las pruebas diagnósticas son elementos esenciales para una práctica clínica de calidad, pero su utilización dependerá de la validez y fiabilidad de la información que proporcionen. Una PD es válida cuando mide lo que se pretende medir, o dicho de otra manera, cuando identifica de forma correcta la presencia o ausencia de enfermedad. La fiabilidad define la capacidad de la PD para obtener en un mismo contexto clínico el mismo resultado, independientemente de que se aplique en diferentes entornos. Por tanto, se deberán elegir aquellas pruebas con la mayor fiabilidad y validez posible.

Otro aspecto a considerar es que una misma prueba varía su rendimiento según el paciente. Si la aplicamos en situaciones de baja probabilidad preprueba, la posibilidad de obtener resultados falsos (positivos o negativos) aumenta proporcionalmente y por tanto la utilidad de la prueba disminuye.

En los supuestos clínicos la probabilidad preprueba de TVP es alta en caso clínico A y baja en el caso clínico B. Si realizáramos una eco-doppler diagnóstica, el resultado sería de gran utilidad en A, pero aportaría poca información útil en B, generando un coste alto y asumiendo el riesgo de tomar decisiones erróneas ante resultados poco concluyentes de la prueba.

9

→ 2

EL PROCESO DE TOMA DE DECISIONES DIAGNOSTICAS

2.1. CONSIDERACIONES GENERALES SOBRE LAS PRUEBAS DIAGNÓSTICAS

Una buena PD es aquella que permite discriminar entre pacientes sanos y enfermos en función del resultado de la misma. Por tanto, se deben elegir aquellas pruebas con mayor validez y fiabilidad posible, teniendo además en cuenta otros aspectos:

- * Accesibilidad a la PD en el medio de trabajo: Es la posibilidad de ser realizada en un determinado nivel asistencial.
- * Capacidad para diagnosticar correctamente a un paciente: La probabilidad de diagnóstico depende de la prevalencia de la enfermedad en esa población. Esta circunstancia hace necesario valorar la probabilidad de obtener resultados positivos antes de realizar la PD. La situación clínica del paciente y la prevalencia de la enfermedad (probabilidad preprueba) son por tanto el punto de partida para la toma de decisiones a la hora de solicitar una PD.
- * Los efectos indeseables de la prueba, con frecuencia, hacen injustificada la exploración. El paciente debe ser informado de ellos y mostrar su consentimiento del que debe quedar constancia escrita; requisito denominado en la literatura legal como: Consentimiento informado.
- * La exactitud de los resultados que proporcione la prueba, definidos por la sensibilidad, especificidad y valores predictivos, que más adelante se comentan.

Por último, se puede afirmar que la utilización de una PD sin un objetivo definido como ocurre con las "PD de rutina" induce errores frecuentes y genera costes muy elevados, haciendo improductivo el trabajo de los Servicios de apoyo: Radiología, Análisis y otros.

Volviendo a los casos anteriores; Antes de hacer una valoración clínica mas detallada, en el Hospital de referencia se podría realizar una ecografía doppler, la determinación sanguínea de Dímero D y una Flebografía venosa si fuera preciso.

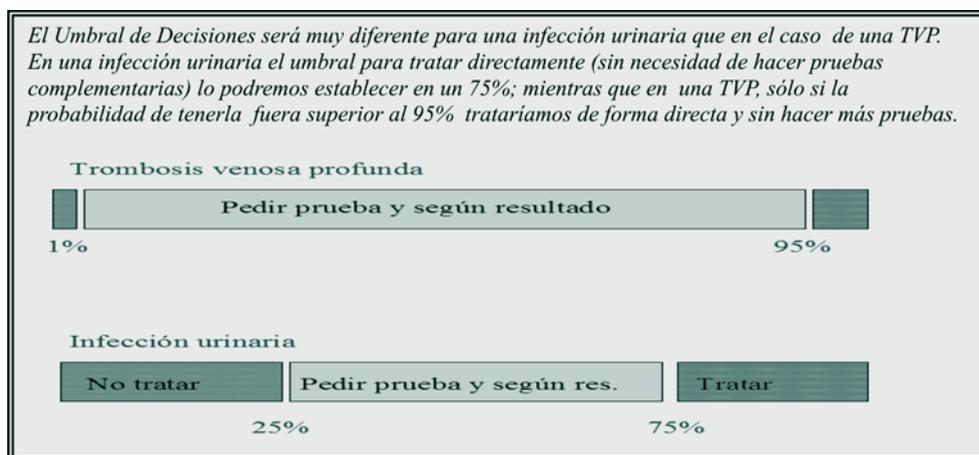
Ante el caso de un paciente en la puerta de Urgencias al que se debe orientar el diagnóstico será necesario conocer la fiabilidad de cada una de las pruebas, su seguridad y márgenes de error.

En un caso como este, resulta tan importante evitar un diagnóstico erróneo (tratar con anticoagulantes e ingresar una inexistente TVP) como obviar un diagnóstico (no tratar con anticoagulantes y aumentar el riesgo de padecer un embolismo pulmonar).

2.2.¿CUÁNDO SOLICITAR UNA PD?. LOS UMBRALES DE DECISIÓN

La aplicación de métodos explícitos y cuantitativos para la toma de decisiones en condiciones de incertidumbre define el análisis de umbrales de decisión. Utilizando la teoría estadística de Bayes se puede valorar la conveniencia de realizar o no una prueba diagnóstica a un paciente.

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.

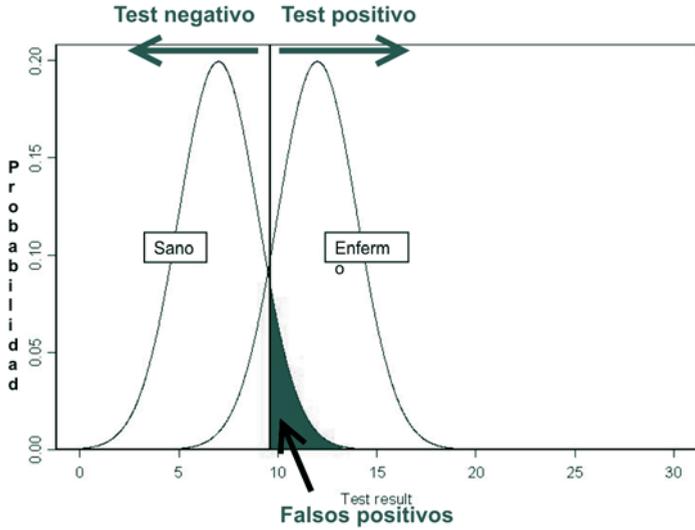


El Umbral de Decisión es el nivel de probabilidad a partir del que el beneficio previsible de tomar una decisión de intervenir iguala o mejora al beneficio de no tomarla. Para conocer el umbral de decisión se utilizan una serie de determinantes estadísticos que expresan la precisión y validez de la PD.

Una PD tendrá mayor validez en la medida que sea capaz de detectar muchos casos con la alteración buscada (sensibilidad alta), excluya a muchos casos sin la alteración (especificidad alta), el resultado positivo represente una alta probabilidad de presentar la alteración (valor predictivo positivo alto) y el resultado negativo refleje una alta probabilidad de la ausencia de la alteración (valor predictivo negativo).

Los puntos de corte son aquellos valores en los que el resultado de la PD se considera normal o alterado. La figura 1 muestra dos curvas de Gauss que representan las poblaciones de sujetos sanos y enfermos, que tiene un valor en el que habría un número de enfermos en que el test sería negativo (Falsos negativos) y un número de sanos considerados como enfermos (Falsos positivos).

Fig. 1



Dependiendo del nivel de corte elegido se obtendrán variaciones de los valores de sensibilidad y especificidad (figuras 2 y 3). Por tanto, si el punto de corte varía también lo harán los demás valores. De esta forma una PD puede mostrar diferentes valores de sensibilidad y especificidad dependiendo de la elección de los valores de los puntos de corte.

Fig 2

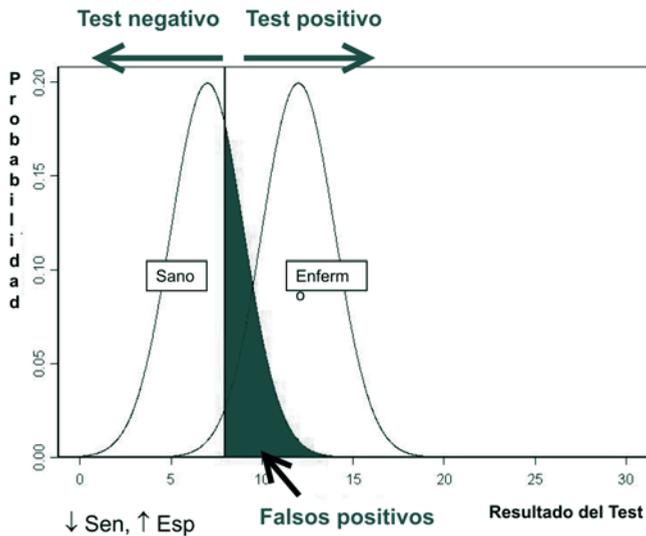
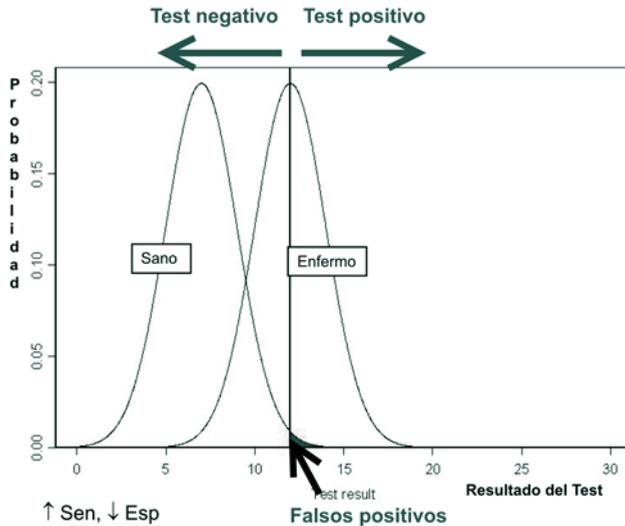


Fig 3



La elección de un umbral de decisión (punto de corte) para clasificar a los sujetos como positivos y negativos dependerá del nivel que consideremos aceptable en cuanto a: proporción de falsos positivos, verdaderos negativos, riesgo de realización de la prueba y riesgo de indicar un tratamiento inadecuado al prescribirlo a un paciente sano.

Para un punto de corte del Dímero D de 130 $\mu\text{g}/\text{dl}$, la sensibilidad es del 95% y la especificidad es del 30%.
 Para la ecodoppler la sensibilidad es del 95% y la especificidad del 95%

¿Pero como se aplican estos datos en nuestros dos pacientes?

Si se aplican estos datos a los dos pacientes con dolor en pantorrilla, se puede obtener el Valor predictivo positivo y negativo para cada una de las situaciones:

Con el dímero D:

Caso A) pasaría de un 70% antes de realizarse la prueba a una probabilidad del 76,5% si el resultado es positivo y de un 9% para un resultado negativo.

Caso B) pasaría de una probabilidad preprueba del 15% a la del 19,8% si el resultado es positivo y a la del 0,9% si el resultado es negativo

Los Umbrales de Decisión permitirían tomar una decisión sin realizar más pruebas (pedir eco-doppler) solo para el caso B con un resultado negativo de dímero D, en cuyo caso se podrían tomar la decisión de no tratar sin necesidad de hacer más pruebas. (Ver Figura)

2.3.CRITERIOS DE VALORACIÓN DE UNA PRUEBA DIAGNÓSTICA: DETERMINANTES ESTADÍSTICOS

Las pruebas diagnósticas permiten clasificar a los pacientes en diferentes categorías en relación con la presencia o ausencia de una enfermedad, característica que define la precisión de la prueba. A mayor precisión menor error aleatorio, expresado por un intervalo de confianza más estrecho. La precisión de una o varias PD dará lugar a la aplicación de un procedimiento terapéutico más o menos adecuado y a la estimación del correcto pronóstico del paciente.

Desgraciadamente algunos tests diagnósticos son incapaces de discriminar entre individuos sanos y enfermos, por lo que su realización puede inducir riesgos importantes para el paciente. Como consecuencia de todo ello se hace necesario utilizar diferentes métodos estadísticos que permitan evaluar y manejar con claridad la información proporcionada por los tests diagnósticos.

La sensibilidad, especificidad y los valores predictivos se han usado como indicadores de la validez de las pruebas diagnósticas, aunque en los últimos años se han aplicado otros indicadores, como las curvas ROC y la razón de probabilidad que desarrollaremos más adelante, y que superan algunas limitaciones de los primeros.

Para analizar un test clínico dicotómico, en el que el resultado es sano o enfermo, se emplean las tablas de contingencia (Tabla 1) en la que se representan gráficamente los resultados de un test: positivo o negativo y las posibilidades en que se encuentra un paciente: sano o enfermo.

Tabla 1

	Enfermedad +	Enfermedad -	Total
Test +	a (verdaderos positivos)	b (falsos positivos) error tipo I o alfa	a + b
Test -	c (falsos negativos) error tipo II o beta	d (verdaderos negativos)	c + d
Total	a + c	b + d	a + b + c + d = n

a: corresponden al grupo de pacientes con enfermedad y que además son detectados por el test. Por tanto son los verdaderos positivos de la prueba.

d: corresponde al grupo de pacientes sanos en los que el test es negativo. Por tanto son los verdaderos negativos de la prueba.

El objetivo ideal de cualquier prueba diagnóstica sería que todos los pacientes enfermos tuvieran un resultado positivo (detectados) y que todos los pacientes sin enfermedad obtengan un test negativo (descartados).

c: Corresponde al grupo de pacientes con enfermedad y test negativo, constituyen el grupo de falsos negativos.

b: Corresponde al grupo de sanos con test positivo, constituyen el grupo de falsos positivos.

Los subgrupos b y c proporcionan idea de la imperfección de la prueba.

2.4. PROBABILIDAD Y ODDS DE UNA PRUEBA DIAGNÓSTICA.

Las probabilidades y las proporciones (también llamadas Odds) son las formas habituales de expresión de los resultados obtenidos al realizar una PD o una intervención terapéutica, resultando imprescindibles para comprender la utilidad de un test.

Odds: Es una forma de expresar las probabilidades de que un suceso ocurra mediante el uso de proporciones. Se expresan cuantificando las posibilidades a favor y las posibilidades en contra de un suceso. Ej.: Si en una caja hay cuatro bolas, tres amarillas y una roja, la Odds es de 3 a 1.

Probabilidad: Es el cociente entre los casos posibles de que ocurra un hecho dividido por el total de casos. El cociente se expresa mediante un valor entre 0 y 1. Siguiendo el ejemplo de las bolas amarillas y rojas, la probabilidad de sacar una bola roja de la caja sería de $1 / 4 = 0.25 = 25\%$.

Aplicándolo a un caso clínico, la probabilidad de que un paciente con dolor torácico sufra un infarto de miocardio P (D+) sería igual a la expresión: n° pacientes con dolor torácico e IAM [D+] / n° pacientes con dolor torácico e IAM [D+] + n° pacientes con dolor torácico sin IAM [D-];

$$P [D+] = D [+]/(D[+] + D[-]).$$

La probabilidad de que un paciente con un síntoma determinado sufra una enfermedad, aporta al médico clínico un dato de sospecha muy útil para orientar la toma de decisiones.

La expresión de los resultados mediante probabilidades o mediante Odds se puede transformar en una u otra según facilite la comprensión, utilizando las sencillas fórmulas expresadas a continuación:

$$\text{Odds} = \text{probabilidad} / (1 - \text{probabilidad})$$

$$\text{Probabilidad} = \text{Odds} / (\text{Odds} + 1)$$

2.5. VALIDEZ INTERNA

Se define como la capacidad de la prueba para medir lo que pretende medir, se cuantifica mediante la sensibilidad y la especificidad, junto con otros parámetros que más adelante se enuncian. La validez interna está condicionada por la adecuación metodológica del diseño y desarrollo del estudio que valora una PD. Garantiza que los resultados no sean sesgados y por tanto nos permitan una adecuada estimación de la utilidad potencial de la PD.

2.5.1. Sensibilidad (S)

Es la capacidad de la prueba para detectar como enfermos los que realmente lo están. Es por tanto la proporción de sujetos con enfermedad y test positivo. Utilizando la tabla 1 corresponde al cociente $[a / a+c]$, donde el denominador es el total de sujetos enfermos.

Si existiera un test perfecto tendría una sensibilidad del 100% y sería capaz de detectar todos los pacientes enfermos sin errores al considerar un paciente enfermo como sano (falsos negativos). En este caso beta sería igual a 0. La sensibilidad también es el complementario de la probabilidad de falsos negativos, por lo que también se puede definir como:

Sensibilidad = 1 - tasa de falsos negativos = 1 - Beta.

Expresado en otras palabras, un valor negativo en un test con alta sensibilidad señala una elevada probabilidad de estar sano.

Utilizar una prueba con elevado grado de sensibilidad es muy útil en la detección de enfermedades graves que no pueden pasar desapercibidas y en las que hay importantes repercusiones al errar el diagnóstico. Cuando un signo o síntoma tiene una elevada sensibilidad, un resultado negativo casi descarta el diagnóstico. Ejemplo: La sensibilidad de la pérdida de pulso de la vena central de la retina en el diagnóstico de la hipertensión intracraneal es del 100%. Así, la observación de pulso en la vena central de la retina, descarta incrementos importantes de la presión intracraneal.

2.5.2. Especificidad (E)

Es la capacidad que tiene una prueba para reconocer como sanos a los realmente sanos. Se mide mediante la proporción de sujetos sin la enfermedad y con test negativo. Utilizando la tabla 1: $[E = d / (b + d)]$, donde el denominador es el total de sujetos sanos. Por tanto, si existiera un test con especificidad del 100% señalaría una probabilidad absoluta de estar sano y no tendría errores (falsos positivos) al considerar como positivo del test a un sujeto sano. En este caso alfa sería igual a 0. La especificidad es el complementario de la probabilidad de falsos positivos, por lo que también se puede definir como:

Especificidad = 1 - la tasa de falsos positivos = 1 - alfa.

En otras palabras, un valor positivo de un test con alto grado de especificidad, señala una elevada probabilidad de enfermedad. Cuando un signo, síntoma o test tiene una elevada especificidad (alrededor del 95%) un resultado positivo casi confirma el diagnóstico.

Ejemplo: La especificidad de 3 ó más respuestas positivas en el test de CAGE confirma el diagnóstico de alcoholismo en el 99% de los pacientes en el ámbito de los ingresos hospitalarios de Medicina Interna. Por tanto, si una persona responde afirmativamente a 3 ó 4 preguntas en ese contexto clínico prácticamente confirmaría el diagnóstico de alcoholismo.

Sensibilidad y especificidad deben ser valorados según la estrategia de elección de la prueba. Así, para conseguir detectar la mayor proporción de individuos enfermos, minimizando los falsos negativos, el test debe tener una alta sensibilidad. Por el contrario, si se desea asegurar que cada test positivo sea verdadero, evitando los falsos positivos, el test debe tener una alta especificidad. En definitiva ambos determinantes definen las características de la prueba, aunque tienen la desventaja de que no proporcionan información relevante para tomar una decisión clínica ante un determinado resultado de la PD.

Ninguno de los dos parámetros está influido por la prevalencia de enfermedad. Ambos valores suelen estar inversamente relacionados, de forma que el aumento de uno supone la disminución del otro. Por tanto, cuando los resultados de una prueba diagnóstica se expresan mediante una escala numérica continua, la especificidad y la sensibilidad pueden variar dependiendo de los valores de corte.

2.5.3. Prevalencia de la enfermedad en la muestra estudiada

Se define como el número de enfermos que existe en una población determinada en un período de tiempo concreto. Es igual al total de enfermos de la muestra de estudio dividido por el total de la muestra:

$$\text{Prevalencia} = a + c / n.$$

2.5.4. Prevalencia del test

Es el número total de positivos obtenidos en el PD en relación al total de sujetos que componen la muestra estudiada:

$$\text{Prevalencia test} = a + b / n$$

2.5.5. Valor Global

Es la proporción de sujetos correctamente clasificados por el test en relación al total de sujetos que componen la muestra de estudio:

$$\text{Valor global} = a + d / n$$

2.6. VALIDEZ EXTERNA

La validez externa se refiere al grado en que los resultados de un estudio son generalizables a otros sujetos o poblaciones. En algunas publicaciones se le denomina también fiabilidad. Depende de la prevalencia de la enfermedad y se mide mediante los valores predictivos.

2.6.1. Valor predictivo positivo (VPP)

Es la proporción de sujetos con test positivo que están realmente enfermos. Utilizando la tabla 1 el VPP correspondería al resultado del cociente: $a / a + b$. A mayor especificidad mayor valor predictivo positivo.

La probabilidad de ser un verdadero positivo (pVP) será igual a la capacidad del test para reconocer a los enfermos (sensibilidad) en cada uno de los que realmente están enfermos (prevalencia):

$$\text{pVP} = \text{sensibilidad} \times \text{prevalencia}.$$

Se elige una prueba con elevado VPP cuando el tratamiento de los falsos negativos puede tener graves consecuencias.

2.6.2. Valor predictivo negativo (VPN)

Es la proporción de sujetos con test negativo que realmente no tienen la enfermedad. Utilizando la tabla 1, el VPN correspondería al resultado del cociente: $d / c + d$. Donde el denominador es el total de pacientes con test negativo. A mayor sensibilidad mayor valor predictivo negativo.

La probabilidad de ser un verdadero negativo (pVN) será igual a la capacidad del test para reconocer a los sanos (especificidad) en cada uno de los que realmente están sanos (complementario de la prevalencia = $1 - \text{prevalencia}$):

$$pVN = \text{Especificidad} \times (1 - \text{prevalencia}).$$

Expresando las probabilidades de resultado mediante una tabla de contingencia:

Tabla 2

	Enfermos	Sanos	Total
Prueba positiva	$pVP = \text{Sensibilidad} \times \text{prevalencia}$	$pFP = (1 - \text{Prevalencia}) \times (1 - \text{Especificidad})$ Error tipo I o alfa	Probabilidad de Positivos = $(P \times S) + ((1 - P) \times (1 - E)) = pVP + pFP$.
Prueba negativa	$pFN = \text{Prevalencia} \times (1 - \text{Sensibilidad})$ Error tipo II o beta	$pVN = \text{Especificidad} \times (1 - \text{Prevalencia})$	Probabilidad de negativos = $((P \times (1 - S)) + ((1 - P) \times E)) = pVN + pFN$
Totales	Prevalencia p (ser enfermo) $pVP + pFN$	$1 - \text{prevalencia}$ (ser sano). $pFP + pVN$	Población: n

VPP y VPN permiten evaluar la capacidad de un test para identificar una determinada condición o enfermedad, dando una idea de la utilidad del test. Sin embargo no permiten la comparación entre varios test ni la extrapolación de otros estudios a un caso concreto, ya que la prevalencia de la enfermedad en cada medio influye considerablemente sobre ambos parámetros. Así, cuando la prevalencia disminuye, el VPP también lo hace y el VPN aumenta, independientemente de la sensibilidad y especificidad del test. Por el contrario si la prevalencia aumenta, también lo hace el VPP mientras disminuye el VPN.

Los valores predictivos aportan importante información para la toma de decisiones y para la información del diagnóstico al paciente, aunque con la limitación de ser influidos por la prevalencia. Cuando es baja, un resultado negativo del test descartará la enfermedad con mayor seguridad, siendo el valor predictivo negativo mayor.

La conclusión de todo lo anterior, es que calcular los valores predictivos no tiene sentido en un estudio donde la prevalencia haya sido artificialmente estimada sin confirmación de la realidad de esa población

2.6.3. Probabilidad pretest (prevalencia)

Es la proporción de pacientes que tienen la enfermedad en la muestra estudiada, y cuyo diagnóstico es fehacientemente establecido por un estándar de referencia:

$$[a + c] / [a + b + c + d].$$

2.6.4. Razón o cociente de probabilidad o cociente de verosimilitud (RP)

Es la mejor medida de utilidad de una prueba diagnóstica. La razón de probabilidad es el cociente entre la probabilidad de obtener un determinado resultado en los individuos enfermos y la probabilidad de ese mismo resultado en individuos no enfermos⁹= Probabilidad de resultado en enfermos / Probabilidad de resultado en NO enfermos.

La razón de probabilidad aporta una medida directa de obtener el VPP o VPN sin necesidad de establecer cálculos que suponen un esfuerzo y tiempo añadido.

La RP informa de cuanto mayor es la probabilidad de encontrar un resultado positivo en un paciente con la enfermedad que en alguien sin ella. Su valor no está influido por la prevalencia de la enfermedad, característica que permite realizar comparaciones entre diferentes test diagnósticos.

Razón de probabilidad para un test positivo (RP+) o Likelihood ratio + (LR+)

Es igual a la expresión: sensibilidad / 1 - especificidad

La razón de probabilidad para un resultado positivo informa de la medida en que la probabilidad de la enfermedad aumenta cuando el test es positivo, es deseable que su valor sea mayor de 1. Resultados de una razón de probabilidad para un test positivo con valores entre 1-2 se consideran prácticamente inútiles para confirmar un diagnóstico; entre 2-10 se considera una prueba “moderada”; siendo un buen test cuando presenta unos valores entre 10 a 50 y un test excelente cuando es mayor de 50 .

Utilizando los datos de la tabla 1. Sería el resultado del cociente entre la fracción de verdaderos positivos (sensibilidad) y los falsos positivos (1 - especificidad): $a/a+c$ y $b/b+d$.

Razón de probabilidad para un test negativo (RP-) o Likelihood ratio - (LR-)

Es igual la expresión: (1 - sensibilidad) / especificidad

La RP para un test negativo informa de como la probabilidad de la enfermedad disminuye cuando el test es negativo, es deseable que su valor sea menor de 1. Resultados de una razón de probabilidad para un test negativo con valores entre 1-0.5 se consideran prácticamente inútiles para excluir prueba diagnóstica; entre 0.5 a 0.1 se considera una prueba “moderada”; siendo un buen test cuando presenta unos valores entre 0.1 a 0.02 y un test excelente cuando es menor de 0.02¹¹ .

Utilizando los datos de la tabla 1, sería el resultado del cociente entre la fracción de falsos negativos y los verdaderos negativos (especificidad): $c/c+a$ y $d/d+b$

Para realizar el cálculo de la Razón de Probabilidad se requiere conocer la Odds pretest: definida como los odds (proporción) de que un paciente tenga la enfermedad antes de que el test diagnóstico se efectúe.

Transformando las Odds ratio (proporciones) a probabilidades:

$$\text{Odds} = \text{probabilidad pre-test} / (1 - \text{probabilidad pre-test}).$$

El producto de la Odds pre-test por la razón de probabilidad será igual a la Odds post-test. Equivale a los odds (proporción) de que un paciente tenga la enfermedad una vez que se ha realizado el test diagnóstico:

$$\text{Odds post-test} = \text{Odds pre-test} \times \text{Razón probabilidad}.$$

2.6.5. Probabilidad post-test.

Es la proporción de pacientes con un resultado positivo del test y que tienen la enfermedad = (pos-test odds/[1 + pos-test odds]).

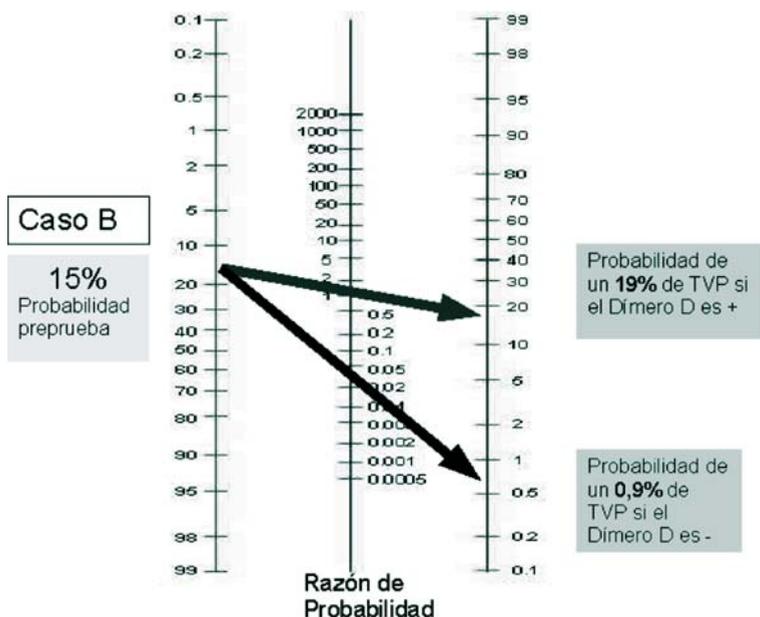
Una razón de probabilidad mayor de 1 da lugar a una probabilidad post-test más alta que la probabilidad pre-test. Por la misma razón una razón de probabilidad menor que 1 dará lugar a una probabilidad post-test menor que la probabilidad pre-test. Cuando el cociente de probabilidades para un resultado positivo o negativo es de valor 1, entonces dicho resultado no aporta información, por el contrario cuanto más alejado esté de 1, más información añade.

Por último, es conveniente recordar que todos estos parámetros de valoración de una PD se estiman partiendo de un muestreo y siempre habría que ofrecer y fijarse en los Intervalos de Confianza para tener una adecuada visión del significado de estos datos.

2.6.6. Normograma de Fagan (Figura 4)

Es un método de cálculo de la probabilidad post-test mediante la representación gráfica de una tabla. Permite conocer la probabilidad post-test mediante el trazo de una línea que una la probabilidad pre-test y la razón de probabilidad, el siguiente punto de intersección con la línea situada a la derecha, señala la probabilidad post-test.

Figura 4



En la dirección de Internet: <http://www.cebm.net/nomogram.asp> se encuentra un diagrama de Fagan que ilustra de manera automática el funcionamiento del nomograma para conocer la probabilidad post-test.

En nuestros ejemplos:

Si la $RP+$ de Dímero D es de 1,4 y la $RP-$ 0,05, con el Nomograma se pueden estimar, conociendo la probabilidad pre prueba, el valor de la probabilidad posprueba ($VP+$ o $VP-$) de forma directa y sin necesidad de realizar cálculos. Para la eco-doppler la $RP+$ es de 18.11 y la $RP-$ es de 0,09

Estos resultados permiten decidir de forma rápida entre diferentes opciones de pruebas diagnósticas.

Para una prevalencia baja (caso B) el dímero D, puede descartar la presencia de TVP si es negativo, Si el resultado es positivo hay que realizar la ecografía para confirmar o descartar el diagnóstico.

En el caso A, de una prevalencia estimada antes de realizar las pruebas del 70%, la realización del Dímero D no modifica la necesidad de realizar la ecografía. Si la ecografía es positiva, confirmaría el diagnóstico y si es negativa (todavía habría un 8% de probabilidades de tener un TVP!) habría que realizar la flebografía o repetir la ecografía pasada una semana.

2.6.7. Curvas de ROC (Receiver Operating Characteristic)

Para que una prueba diagnóstica pueda ser utilizada en la clínica es necesario seleccionar los valores a partir de los que se tomarán decisiones sobre el proceder con el paciente, valores a los que se denominan "puntos de corte" o "umbrales de decisión".

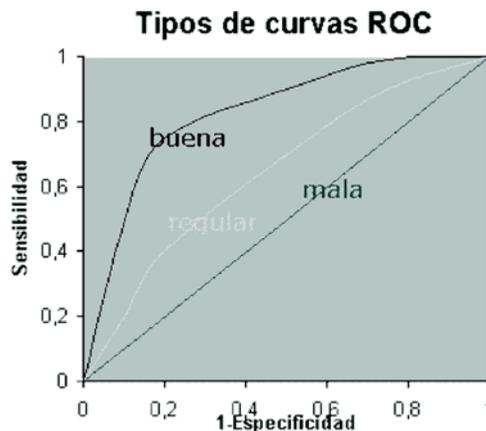
Muchos de los tests diagnósticos proporcionan resultados en una escala de datos continuos, como la glucosa o la hipertensión arterial, en estos casos es necesario establecer criterios que permitan transformar una información numérica en resultados dicotómicos de "normalidad" y "no normalidad". Para ello se establecen los distintos puntos de corte o niveles de decisión, que son los valores numéricos a partir de los cuales se modifica el criterio de normalidad.

Cada punto de corte lleva consigo diferentes valores para la sensibilidad y especificidad de la exactitud de la prueba. Las curvas ROC proporcionan una representación gráficamente completa de la capacidad de una prueba para discriminar lo normal de lo que no lo es. Se examinan todos los umbrales de decisión mediante la presentación de los pares compuestos de sensibilidad y especificidad para todo el rango de resultados observados. En el eje "Y" se representa la sensibilidad o fracción de verdaderos positivos de una prueba, en el eje "X" se representan los falsos positivos, que es igual a la resta: $1 - \text{especificidad}$.

Cada punto en la figura representa un par de sensibilidad/especificidad correspondiente a un umbral de decisión particular. Una prueba diagnóstica con una discriminación perfecta dará lugar a una curva que pasa a través de la esquina superior izquierda, donde la fracción de verdaderos positivos es 1 ó 100 % (sensibilidad perfecta) y la fracción de falsos positivos es 0 (especificidad perfecta), por tanto cuanto más cerca esté la línea de la esquina superior izquierda, mayor será la precisión de la prueba.

La curva teórica de una prueba que no discrimina (distribución idéntica de los resultados para ambos grupos) es una línea diagonal de 45° desde la esquina inferior izquierda hasta la superior derecha. Debido a que no hay pruebas perfectas, la mayoría dan lugar a curvas intermedias entre las dos situaciones extremas (Figura 5).

Figura 5



La posición relativa de las líneas indica la precisión relativa de las pruebas. Así, una curva que cae encima y a la izquierda de otra indica mayor precisión.

Las curvas ROC son útiles para:

- Conocer el rendimiento global de una prueba mediante el Área bajo la curva ROC (*), que se expone el siguiente apartado.
- Comparar dos pruebas o dos puntos de corte. Comparación de dos curvas o de dos puntos de corte sobre una curva.
- Elegir el punto de corte apropiado entre los resultados positivos y negativos de una prueba.

* *Área bajo la curva ROC (ABC)*: El área bajo la curva ROC es un excelente indicador global de la precisión de una prueba diagnóstica. Hace posible expresar la utilidad de una prueba mediante un número simple. El valor del área bajo la curva oscila entre 1, que representa la máxima exactitud (discriminación perfecta) y 0,5 que es la mínima exactitud (no hay diferencias en la distribución de los valores de la prueba en los dos grupos). Si el valor fuera menor de 0,5 habría que invertir el criterio de positividad de la prueba. Como ejemplo sobre la interpretación del valor del área bajo la curva: Un área de 0,8 significa que un individuo seleccionado aleatoriamente del grupo de enfermos tiene un valor de la prueba mayor que uno seleccionado del grupo de sanos en el 80 % de ocasiones. Cuando se compara la exactitud en el diagnóstico de dos pruebas diferentes, se escoge aquella que presenta mayor área bajo curva. El cálculo se realiza mediante el contraste de hipótesis y la estimación del intervalo de confianza para el área. Rechazar la hipótesis de que las pruebas muestran igualdad en los resultados, donde el área teórica es igual a 0,5 ($p < 0,05$ y/o intervalo de confianza que no contiene el valor 0,5), demuestra que la prueba diagnóstica puede diferenciar con mayor precisión entre los 2 subgrupos: sanos y enfermos.

Desde un punto de vista estadístico el problema es valorar si la diferencia observada entre las ABC calculadas para dos pruebas distintas es debida a la variabilidad por el muestreo o es más bien atribuible a una diferencia real en la exactitud de ambas pruebas. Otro error posible puede ocurrir cuando dos curvas ROC presentan distinta morfología pero con igual o similar valor del ABC. En estos casos las diferencias pueden tener implicaciones clínicas importantes, ya que las diferencias serán las "Áreas Parciales de la Curva", que son aquellos valores en los que aumenta la sensibilidad o la especificidad de la prueba. Por tanto, dependiendo de la finalidad de la prueba se utilizarán aquellas pruebas que muestren partes de la curva que aseguren altos niveles de sensibilidad o especificidad.

9

→ 3

OBTENIENDO Y VALORANDO INFORMACIÓN PARA UN PROBLEMA SOBRE DIAGNÓSTICO

El manejo del gran volumen de información disponible en los diferentes repertorios bibliográficos para buscar respuesta a una cuestión clínica puede resultar difícil por las propias limitaciones del profesional para recuperar información, plantear el problema y sobre todo por falta de tiempo. Para hallar una opción basada en las mejores pruebas conocidas que permita resolver un problema o proceso asistencial concreto es imprescindible aplicar una metodología correcta. La calidad de la

información recuperada y el tiempo empleado en ello son directamente proporcionales al empleo de un método específico y sistemático, sin el cual el clínico puede perder tiempo y esfuerzo en una búsqueda estéril que no despeje sus dudas.

En muchas ocasiones las preguntas clínicas que nos formulamos en nuestro trabajo clínico cotidiano versarán sobre aspectos diagnósticos. Tras formular nuestra duda en forma de pregunta clínica y realizar una adecuada búsqueda bibliográfica, será preciso reflexionar sobre la calidad de la evidencia que nos ofrezcan los estudios que hayamos seleccionado para responder a nuestras dudas y poder excluir aquellos cuya calidad científica sea dudosa.

3.1.LA PREGUNTA CLÍNICA

El primer paso en la búsqueda de respuestas a dudas sobre diagnóstico consiste en transformar el problema o incertidumbre que se nos plantea en una pregunta clínica, siguiendo la metodología habitual. Deberán para ello quedar bien determinados y diferenciados el perfil de sujetos sobre los que se aplica la prueba diagnóstica, la prueba a valorar y los diferentes resultados posibles.

En general y para cualquier pregunta clínica, el tipo de recurso que más nos conviene utilizar dependerá del tipo de pregunta planteada y del tiempo que dispongamos (tabla 3)

Los diseños de estudio que utilizaremos para buscar respuestas a preguntas sobre diagnóstico serán por tanto estudios transversales, revisiones sistemáticas y Guías de práctica Clínica¹

Tabla 3: Tipo de pregunta y estudio a localizar

Pregunta	Tipo de estudio
	Estudio transversal
Diagnóstico	Revisión Sistemática Guía de Práctica Clínica (GPC)
Etiología	Estudios de Cohortes Caso-Control Revisión Sistemática GPC
Pronóstico	Estudios de Cohortes-Supervivencia Revisión Sistemática GPC
Intervención	Ensayo Clínico Aleatorizado y Controlado Revisión Sistemática GPC
Frecuencia	Estudios de Cohortes Estudio transversal Revisión Sistemática GPC

3.2.BUSCANDO LA INFORMACIÓN (RECOMENDABLE CONSULTAR EL CAPÍTULO DE BÚSQUEDA DE PUBMED)

Siempre es útil iniciar nuestra búsqueda con la adecuada elección de las palabras clave / descriptores (términos MeSH) que configuren el perfil de búsqueda y faciliten la exploración de las fuentes de información. Para ello podemos utilizar MeSH Database de PubMed (esta funcionalidad está descrita en el Manual de PubMed con detalle).

A partir de aquí, en función del tiempo disponible y de los conocimientos que el clínico tenga sobre búsqueda, recuperación y evaluación de la información científica, caben dos posibilidades:

1. Buscar la información científica acerca del tema en repertorios bibliográficos y evaluarla hasta depurar la evidencia que se desprenda de la misma. Ya descrito con detalle en el capítulo de búsquedas de la evidencia. Una vez recuperado los documentos que nos parecen relevantes seguiremos con:

- A. Evaluación crítica de los mismos.
- B. Síntesis de la Información.
- C. Resolución del escenario clínico.

2. Acceder a fuentes que ya hayan realizado este proceso y ofrezcan la información revisada y catalogada desde el punto de vista de la evidencia (publicaciones filtradas): Acceso a publicaciones de síntesis basadas en la evidencia, bases de Guías de Práctica Clínica (GPC), Revisiones sistemáticas, Agencias de Evaluación de Tecnologías Sanitarias (AETS), revistas secundarias o de resúmenes...

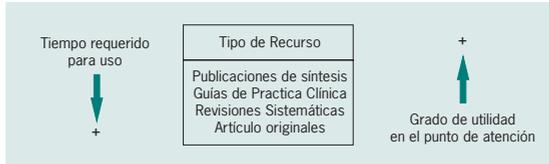
- a. Realización de la búsqueda.
- b. Recuperación de los documentos.
- c. Comparación con el contexto propio.
- d. Resolución escenario clínico.

¿Cuándo elegir una alternativa u otra? Indiscutiblemente, la segunda es la ruta más corta y certera, aunque no siempre hay revisiones sobre el tema que se está tratando. En esos casos habría que recurrir a generar la evidencia si es posible. Para ello se requiere conocimientos específicos sobre búsqueda, evaluación y síntesis de la literatura. Algo que los clínicos pocas veces poseen.

Respecto al tema de la búsqueda de información sobre PD se hace necesaria una reflexión. Aunque hay muchas publicaciones biomédicas, la información de calidad sobre validación de pruebas diagnósticas es relativamente escasa. Sin embargo, el proceso de buscar y leer lo mucho o poco que encontremos nos es de mucha utilidad para: a) reflexionar sobre la información objetiva en la que basamos nuestras decisiones diagnósticas y b) detectar posibles lagunas de información y por lo tanto de investigación aplicada.

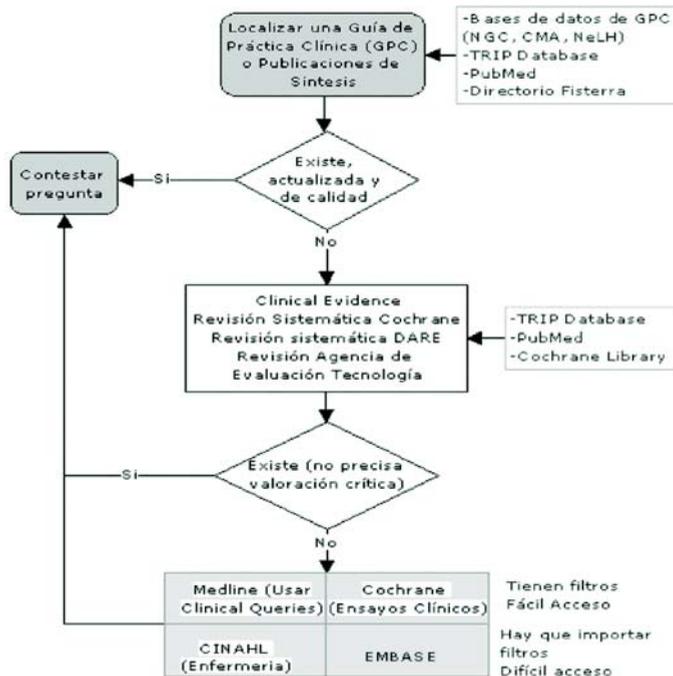
Iniciar la búsqueda por fuentes muy elaboradas, preparadas para responder preguntas en el puesto de consulta requiere menos formación en lectura crítica y es menos costosa en tiempo (figura 6).

Figura 6



Es por esta razón el importante desarrollo de publicaciones de síntesis finales (que en nuestro esquema situamos al nivel de las GPC) en apoyo a las decisiones en consulta. El proceso general que habitualmente se podría seguir en la búsqueda de respuestas a preguntas clínicas sería (Figura 7):

Figura 7



Si elegimos la vía de las publicaciones de síntesis para resolver nuestras dudas no será necesario ningún proceso adicional de análisis de la información recuperada. El análisis debe hacerse previamente, seleccionando las fuentes fiables por su rigor en la metodología de elaboración de sus documentos y su actualización.



En las búsquedas sobre diagnóstico puede ser de gran utilidad el uso de los filtros metodológicos. Se trata de estrategias de búsqueda muy elaboradas diseñadas por documentalistas expertos y que han demostrado ser las más eficientes para identificar determinado tipo de trabajos. Según el tema que se esté buscando permite diferentes opciones como:

- Seleccionar si se buscan aspectos de diagnóstico, tratamiento, pronóstico o etiología.
 - Elegir una búsqueda sensible, muy amplia para que no se escape ningún trabajo o más específica y que se ajuste mucho más a los criterios de búsqueda.
- En PubMed se puede encontrar como “Clinical Query.”

Otra buena alternativa sería el uso de un metabuscaor como TRIPdatabase, que nos permite localizar recursos específicos, muy útiles para recopilar información útil sobre problemas relacionados con el diagnóstico. Para finalizar la búsqueda podríamos ampliar la búsqueda en otras bases de datos más específicas sobre Guías de Práctica Clínica, ya descritas en otro capítulo de este libro.

3.3. LECTURA CRÍTICA DE UN ARTÍCULO SOBRE PRUEBAS DIAGNÓSTICAS

Lamentablemente los profesionales sanitarios no suelen utilizar de forma sistemática los métodos de valoración crítica de pruebas diagnósticas, mostrando un importante déficit de formación y desaprovechando una gran oportunidad de mejora en la atención al paciente.

De forma genérica y aunque cada tipo de estudio contemple algunas cuestiones específicas, hay tres grandes preguntas que debe superar cualquier artículo sometido a una evaluación crítica:

- 1.- ¿Cuáles son los hallazgos principales? (Mensaje del artículo)
- 2.- ¿Se justifican las conclusiones con la metodología empleada? (validez)
- 3.- ¿En que medida son aplicables a mis pacientes? (Utilidad)

Para estudios sobre diagnóstico la estrategia de búsqueda de soluciones a nuestras preguntas es menos eficiente que para estudios sobre tratamiento, porque la colaboración Cochrane y Clinical Evidence están enfocadas fundamentalmente a intervenciones terapéuticas. Las publicaciones de síntesis final (recursos filtrados) en relación a temas diagnósticos más importantes son EBM-Guidelines (www.ebm-guidelines.com) que contiene a fecha 20 de Diciembre de 2006 cerca de 1000 GPC basadas en la evidencia, además de 2711 resúmenes de evidencias, 919 imágenes y 78 registros de audio sobre auscultación y el texto completo de las revisiones de la Librería Cochrane citadas (Figura 8).

The screenshot shows the Wiley InterScience website. At the top, there's a navigation menu with 'HOME', 'ABOUT US', 'CONTACT US', and 'HELP'. Below that, the page title is 'EBM Guidelines: Evidence-Based Medicine'. The main content area includes a description: 'Easy to use clinical guidelines supported by sound scientific evidence'. There's also a sidebar on the right with 'SPECIAL FEATURES' and 'ALSO OF INTEREST' sections, including a link to 'The Cochrane Library'. The bottom of the page lists 'The collection includes' with bullet points about the number of guidelines and evidence summaries.

De extraordinaria utilidad también resulta una serie de JAMA sobre examen clínico racional que esta compendiada en un libro de texto.

Una vez comprobado que no hay información filtrada sobre la pregunta que nos hagamos y realizada una búsqueda, que nos permita localizar información que en teoría pueda responder a nuestra pregunta, deberíamos realizar un análisis crítico de la misma. Este análisis nos permitirá excluir aquellos artículos de dudosa calidad científica.

Para el proceso de lectura crítica hay diversas guías accesibles localizables a través de diferentes enlaces. Uno de ellos podría realizarse desde la Web del grupo de Urología Basada en Pruebas de la Asociación Murciana de Urología (<http://www.uroportal.net/uroevidencia/guiaslectura.htm>). La página con los diferentes enlaces puede observarse en la figura 9.

The screenshot shows a web browser window with the URL 'http://www.uroportal.net/uroevidencia/guiaslectura.htm'. The page header includes 'UroPortal.net' and 'Urología Basada en la Evidencia'. The main content area is titled 'Guías para la lectura crítica' and lists several resources: 'Users' Guides to Evidence-Based Practice', 'Users' Guides to the Medical Literature', 'How to read a paper', and 'Herramientas para la lectura crítica de CASP-españa'. Each resource has a brief description and a link to the full text.

El proceso de evaluación crítica está descrito en una serie de artículos denominados genéricamente como los "User's guides...", en los que se detalla cómo debe ser este procedimiento en función del tipo de estudio (tratamiento, diagnóstico, etc.). Como recurso muy útil en castellano disponemos de la ayuda de guías como la que se puede encontrar en la página Web de CASPe (www.redcaspe.org). Figura 10



En definitiva, en una revisión crítica se revisa la metodología del estudio en cuestión, cómo se seleccionaron los pacientes, si éstos culminaron el estudio y si fueron tratados homogéneamente y en el grupo al que fueron asignados... para comprobar tanto la validez interna como externa del estudio. En caso de que el estudio cumpla con estos dos requisitos, la información que aporta podría ser utilizada en nuestro medio para responder a nuestras preguntas clínicas.

3.4. Condiciones que debe reunir una publicación científica sobre una prueba diagnóstica

Quando se realiza la lectura crítica de un estudio, lo primero que debe ocuparnos es establecer su validez interna, definida por un adecuado diseño y desarrollo, para garantizar que sus resultados no sean sesgados y permitan una buena estimación de la eficacia real de la intervención estudiada. Si nuestras preguntas son sobre pruebas diagnósticas, las publicaciones que aportan mayor información al médico clínico serán aquellas realizadas con estudios transversales, que incluyan una muestra representativa de la población de interés y donde se analicen simultáneamente los resultados de la prueba estudiada y los resultados obtenidos con una prueba de referencia (Gold Standard) que previamente ha demostrado ser de utilidad. Generalmente buscaremos estudios que incluyan pacientes en estadios más precoces de enfermedad, ya que los estudios realizados sobre pacientes muy evolucionados ofrecerán información poco útil. Nos basaremos para comprobar si se cumplen esos requisitos en las plantillas de valoración crítica de CASPe para este tipo de publicaciones. Hay tres aspectos generales a valorar cuando se hace lectura crítica de un artículo:

- ¿Son válidos los resultados?
- ¿Cuáles son los resultados?
- ¿Son aplicables en tu medio?

Las 10 preguntas que presenta la plantilla CASPe disponible en <http://www.redcaspe.org/herramientas/lectura/10diagnostico.pdf>

“10 preguntas para dar sentido a un artículo de diagnóstico” están pensadas para responder a estos tres aspectos globales. Las tres primeras preguntas son de eliminación y se pueden responder con facilidad. Si su respuesta no es positiva podemos desechar el artículo y evitar pérdidas de tiempo al lector,. Sólo en caso de respuesta afirmativa a las tres primeras se garantizan unos mínimos metodológicos que hacen que merezca la pena continuar respondiendo a las 7 restantes.

Preguntas 1-3 (de eliminación): ¿Son válidos en lo esencial los resultados del estudio?

Pregunta 1: ¿Existió una comparación con una prueba de referencia adecuada? Esta primera pregunta define la necesidad de que haya una prueba de referencia o patrón oro (Gold Standard) adecuado para confirmar la presencia de la enfermedad cuyo diagnóstico estamos testando. Para responderla debemos plantearnos si es correcto o no el patrón oro para confirmar el diagnóstico. Habrá que tener en cuenta si se puede aplicar siempre el mismo patrón oro a todos los pacientes.

Pregunta 2: ¿Incluyó la muestra un espectro adecuado de pacientes? Es importante saber que casi cualquier prueba diagnóstica permite diferenciar entre personas absolutamente libres de enfermedad y otras con estadios muy avanzados de la misma. Para que una prueba diagnóstica resulte válida debe haber sido valorada en un espectro suficientemente amplio de pacientes con diferentes estadios de enfermedad. Solo así dispondremos de información suficiente sobre el comportamiento de la prueba, evitando sesgos en la interpretación. Para comprobar este punto deberemos reflexionar sobre si están adecuadamente descritos y pormenorizados los criterios de inclusión de los pacientes.

Pregunta 3: ¿Existe una adecuada descripción de la prueba? Cuando analizamos este aspecto estamos valorando la posibilidad de aplicación de la prueba en el contexto clínico del lector, la reproducibilidad en su medio. Se trata de conocer si se establece claramente la diferencia entre un resultado positivo y uno negativo. Hay una serie de aspectos que deben estar explícitos en un artículo sobre evaluación de pruebas diagnósticas:

- Preparación del paciente.
- Entrenamiento del profesional
- Variabilidad inter / intraobservador en pruebas que requieren una interpretación (pruebas de imagen, datos de examen físico...).
- Descripción detallada de cómo se realiza la prueba.

Todos estos aspectos nos informan de la caracterización de la prueba y nos permite decidir sobre si es adecuada en nuestro medio y para nuestros pacientes. Estas tres preguntas definen los mínimos de validez de un estudio y son básicos para interpretar si debemos seguir valorándolo.

Dos preguntas más de matiz (4,5) nos permiten valorar más detalladamente aspectos concretos sobre la validez interna del estudio:

Preguntas 4 y 5: ¿Hubo evaluación ciega e independiente de los resultados? Para valorar estos dos aspectos tendremos en cuenta que los resultados de una prueba no condicionaran la realización de la otra (independiente) y que la interpretación de cada una de las pruebas se hizo sin conocer los resultados de la otra (ciega).

La independencia no se producirá en los casos en que se incluyan preferentemente los resultados positivos en la prueba a evaluar o cuando se utilicen diferentes patrones oro en los positivos y negativos. Por ejemplo es probable que si estuviéramos valorando el diagnóstico de Cáncer de Próstata a través del PSA, no hiciéramos biopsia en todos los casos, tendiendo a practicarla más probablemente en los casos que la analítica haya sido claramente patológica.

Por otra parte podemos incluir en el sesgo de sospecha diagnóstica cuando la interpretación de una prueba se haga conociendo el resultado de la otra. Por ejemplo, en un estudio que contrastara le utilidad del CAGE para el diagnóstico de sospecha de alcoholismo ¿haría el médico la misma valoración del test AUDIT para confirmar el diagnóstico si supiera que el CAGE había sido positivo?

Estas 5 primeras preguntas pueden ayudarnos a resolver nuestras dudas acerca de la validez interna del estudio.

Para establecer si el estudio presenta resultados clínicamente relevantes, valoraremos cuales son los resultados a través de dos preguntas

Pregunta 6: ¿Se pueden calcular los cocientes de Probabilidad "likelihood ratios"? Para su cálculo habrá que valorar aspectos tales como si se han tenido en cuenta en el estudio los pacientes con resultados "no concluyentes" o si se pueden calcular los cocientes de probabilidad para los diferentes niveles de la prueba si procediera. En el ejemplo del PSA sería interesante que estuvieran disponibles los datos para calcular los CP (LR) en diferentes puntos de corte, como por ejemplo por encima de 10 ng/ml, de 4 ng/ml o por debajo de este nivel.

Aunque cada vez es más frecuente que se incluyan los datos sobre CP (LR) en los estudios sobre pruebas diagnósticas, su cálculo suele resultar sencillo si están disponibles unos datos elementales (sensibilidad y especificidad). El CP (+) (LR+) indicará cuanto mas probable es un resultado positivo en los enfermos que en los no enfermos y es deseable que sea mayor de 1. Por el contrario el CP (-) (LR-) indica cuanto mas probable es un resultado negativo en los enfermos que en los no enfermos. Por tanto es deseable que sea menor de 1. El cálculo resultará sencillo:

$$CP (+)(LR+) = Sen / 1 - Esp \quad CP (-)(LR-) = 1 - Sen / Esp$$

Como es lógico cuanto mayor de 1 sea el CP (+) o menor de 1 sea el CP (-) mayor será la utilidad de la prueba.

Pregunta 7: ¿Cuál es la precisión de los resultados? Habrá que localizar en el estudio los Intervalos de Confianza de los CP, o bien buscar los datos que nos permitan calcularlos. En un estudio en el que analizamos una muestra de la población, el proceso que nos permite a partir de los datos de esa muestra aproximarnos al conocimiento de la realidad se llama INFERENCIA. Cuanto más estrecho sea el Intervalo de Confianza mayor será la precisión con la que se estime el efecto en la población.

Las tres últimas preguntas permitirán sentar las bases para establecer si el estudio aporta resultados que puedan ser aplicables a nuestros pacientes.

Analizaremos con este tercer apartado la validez externa o generabilidad de los resultados. Para ello se considerarán las respuestas a las tres últimas preguntas:

Pregunta 8: ¿Serán satisfactorios en el ámbito del escenario la reproductividad de la prueba y su interpretación? Para contestar a esta pregunta el lector debe considerar si el ámbito en el que se estudió la prueba es demasiado diferente al escenario al que lo quiere aplicar el lector.

Pregunta 9: ¿Es aceptable la prueba en este caso? Para valorar este aspecto debe tenerse en cuenta la factibilidad de la prueba en el entorno del escenario, tanto desde el punto de vista de su disponibilidad, riesgos o molestias que pueda producir como de la posibilidad de asumir los costes que genere.

Pregunta 10: ¿Modificará los resultados de la prueba la decisión sobre cómo actuar? Se valorará que desde la perspectiva del escenario, si la actitud del clínico no va a cambiar la prueba será inútil. Otro aspecto a valorar al respecto de este punto será en que medida se modifica el umbral de acción y la posibilidad de enfermedad antes y después de la prueba.

3.5.SESGOS A TENER EN CUENTA AL REALIZAR LECTURA CRÍTICA DE UN ESTUDIO SOBRE PRUEBAS DIAGNÓSTICAS

Para poder realizar la lectura crítica de una publicación científica sobre la utilidad de una prueba diagnóstica hay que evaluar si el diseño del estudio es correcto, considerando que el mejor diseño es la comparación prospectiva y ciega de la PD con una prueba patrón que servirá de referencia, aplicándolo a una muestra de pacientes.

La selección de pacientes sobre los que se realiza el estudio puede producir sesgos en los resultados. Los más frecuentes son:

*Sesgo de verificación o de confirmación diagnóstica: El lector debe comprobar que los investigadores han intentado reducir este sesgo. Ocurre cuando el estudio se limita a los individuos sobre los que se realizó con anterioridad la prueba "Gold Standard", que suelen ser los de mayor probabilidad de padecer enfermedad. Los resultados sobreestimarán la sensibilidad al producir unos resultados con aumento de las pruebas positivas, además se producirá una disminución de los resultados de las pruebas negativas provocando infraestimación de la especificidad.

Controlar este sesgo requiere el uso de técnicas matemáticas complejas y la evaluación de la presencia o ausencia de enfermedad en los sujetos con resultado negativo utilizando otros métodos, como el seguimiento de pacientes, aunque sean menos adecuados que el de referencia ("gold Standard"). Con frecuencia es inevitable por razones éticas, como en el caso de las pruebas invasivas.



*Sesgo de interpretación de las pruebas: Las pruebas deben estar realizadas de forma independiente y a ciegas, de forma que el conocimiento de los resultados de una prueba no influyan en los resultados de la otra. Este sesgo es más frecuente cuando existe un componente de subjetividad en la interpretación de los resultados.

*Sesgo de resultados no interpretables: Generalmente, cualquier prueba diagnóstica es capaz de diferenciar entre personas muy sanas y otras muy enfermas, pero es la capacidad de discriminar entre situaciones intermedias y evitar errores, lo que determina la utilidad del test. Por ello, las muestras de pacientes deben ser similares incluyendo a pacientes con diferentes estadios de la enfermedad y con otras condiciones que plantean problemas de diagnóstico diferencial, y que sean similares a aquellos sobre los que se aplicaría el test en la práctica clínica.

El artículo debe definir explícitamente los criterios de selección ya que aunque la sensibilidad y especificidad son relativamente constantes independientemente de la prevalencia de la enfermedad, los valores predictivos si están muy influidos por ella.

*Sesgo por ausencia de “Gold Standard”: Los estudios sobre la utilidad de una prueba diagnóstica se realizan mediante comparación, pero no siempre es posible, ya que a veces no existe una prueba universalmente aceptada. En estos casos se desarrollan combinaciones de varios criterios que se utilizan como patrón de referencia. El sesgo se produce cuando entre los criterios de referencia se incluye la prueba de estudio.

*Sesgo en el espectro de la enfermedad: Una prueba diagnóstica debe tener distintos grados de exactitud dependiendo de los diferentes grados de severidad de la enfermedad. Por ello las características clínicas de los pacientes deben estar descritas en el estudio.

*Variabilidad inter-observador: Todas las pruebas diagnósticas requieren la valoración subjetiva del observador. Dos observadores pueden ser igualmente exactos pero uno más sensible o específico que el otro (operar en la misma curva ROC pero en puntos distintos). La exposición de los resultados debe indicar la concordancia entre diferentes observadores o entre diferentes interpretaciones de los resultados realizadas por un mismo observador.

*Sesgo de reproducibilidad: El estudio debe incluir la descripción detallada de las condiciones en que se han aplicado la prueba, así como las definiciones de los resultados positivos y negativos. Se deben incluir aspectos relacionados con la preparación del paciente (dieta, precauciones de la prueba y fármacos a emplear) y el análisis e interpretación de los resultados.

3.6. UN EJEMPLO CONCRETO

1. DESCRIPCIÓN DEL ESCENARIO CLÍNICO

Está de guardia de sábado en su Centro de Salud cuando acude una mujer de 65 años consultando por dolor e inflamación a nivel de MI derecho de 24 horas de evolución. No refiere traumatismo pero lleva varios días acatarrada. Por inspección aprecia unas varicosidades superficiales. La paciente nos pregunta: Doctor ¿no será una trombosis?

- Duda: ¿Tendrá la paciente una TVP?
- Pregunta: ¿Existe alguna escala o cuestionario clínicos que le permita estratificar el riesgo de padecer TVP?

¿Existen modelos fiables que orienten la probabilidad clínica de padecer TVP en pacientes que acuden a un servicio de urgencias?

Reflexionemos sobre lo que queríamos saber: si había modelos clínicos fiables que nos orienten sobre la probabilidad clínica de padecer TVP en pacientes que acuden a un servicio de urgencias por dolor en pantorrilla y con un contexto clínico sugestivo de dicho diagnóstico.

En este caso lo primero a hacer es transformar la duda en pregunta clínica: ¿En pacientes que acuden a un servicio de urgencias con sospecha clínica de TVP, el uso de un cuestionario de estratificación clínica de riesgo es válido para orientar de forma correcta el diagnóstico?

Para una primera búsqueda en PubMed se utilizaron los descriptores “venous thrombosis” AND “emergencias” como términos MESH, incorporando el filtro metodológico Clinical Query. Se solicitaron estudios sobre diagnóstico a través de una búsqueda sensible. Entre los 90 estudios identificados solo dos parecían poder responder a la pregunta. En ambos casos (ver mas abajo) el artículo original estaba en alemán por lo que se replanteó la búsqueda.

11: [Stiegler H.](#)

[Related Articles](#), [Links](#)



[Emergency in general practice. Possible deep venous thrombosis of the leg]
MMW Fortschr Med. 2002 Dec 12;144(50):45-8. Review. German. No abstract available.
PMID: 14610865 [PubMed - indexed for MEDLINE]

55: [Rudofsky G.](#)

[Related Articles](#), [Links](#)



[Problems in diagnosis and therapy of thrombotic venous diseases]
Vasa Suppl. 1991;33:67-8. German.
PMID: 1788768 [PubMed - indexed for MEDLINE]

En una búsqueda posterior empleando solo el descriptor “venous trombosis”, se utilizó el filtro “Clinical Query” para diagnóstico y se seleccionó una estrategia diferente para ser más específico. En este caso se identificaron 605 artículos entre los que parecía que varios podían responder a la pregunta. Como resultaba de interés conocer un contexto clínico similar al nuestro se seleccionaron por medio del “limit language spanish” los artículos cuya lengua original fuera el castellano, encontrando 5 estudios con estas características (figura 11). Uno de ellos (el número 3) parecía poder responder específicamente a la pregunta y estaba accesible a texto completo, por lo que se seleccionó para hacer una revisión crítica del mismo (figura 12).

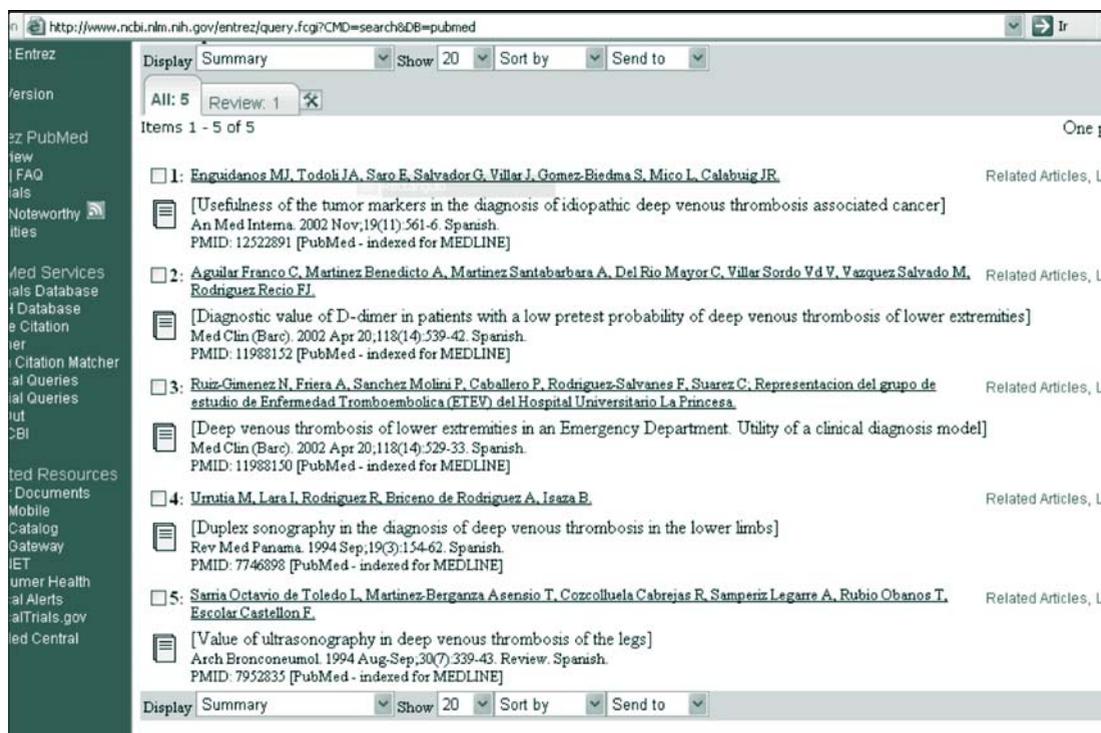


FIGURA 11

1: *Med Clin (Barc)*, 2002 Apr 20;118(14):529-33.

[Deep venous thrombosis of lower extremities in an Emergency Department. Utility of a clinical diagnosis model]

[Article in Spanish]

Ruiz-Gimenez N, Frieria A, Sanchez Molini P, Caballero P, Rodriguez-Salvanes E, Suarez C; Representación del grupo de estudio de Enfermedad Tromboembólica (ETE) del Hospital Universitario La Princesa.

Radiología, Hospital Universitario La Princesa, Universidad Autónoma de Madrid, Spain.

BACKGROUND: Deep vein thrombosis (DVT) is a difficult to diagnostic disease. The aim of this study was to determine the utility and accuracy of a risk stratification questionnaire and a diagnostic strategy, which were applied to patients with suspected DVT on lower extremities in an emergency department. **PATIENTS AND METHOD:** A prospective cohort study was performed in 569 outpatients with clinical suspected DVT during 14 months. The applied questionnaire stratified patients into three pre-test probability categories. Items included signs, symptoms, risk factors and potential alternative diagnosis, which were based on a modified Wells clinical model. DVT was diagnosed by the combined use of clinical model, compression ultrasonography (CUS) and follow-up CUS one week later in those moderate-high risk patients with an initial normal test. These patients were followed over three months for the development of venous thromboembolic complications. **RESULTS:** Two hundred three (35.7%) patients were classified as having a low, 186 (32.7%) moderate and 180 (31.6%) high clinical probability. Overall, DVT was diagnosed in 153 patients (26%; CI95%, 23.2-30.7%): 144 (96%) at the initial CUS, 6 (3.5%) at the second testing and 3 over the 3-month follow-up period. 22 patients had a low pretest probability (11%; CI95%, 7-16%), 43 (23%; CI95%, 17-30%) moderate, and 88 (49%;

Related Links

- ▶ Thrombosis in the emergency department: use of a clinical diagnosis model to safely [Arch Intern Med. 1999]
- ▶ Rapid D-dimer test combined a clinical model for deep vein thrombosis. Validation with [Thromb Haemost. 2004]
- ▶ Combined use of clinical assessment and d-dimer to improve the management of p [J Thromb Haemost. 2003]
- ▶ Prospective evaluation of a screening protocol to exclude deep vein thrombosis on [J Am Coll Surg. 2005]
- ▶ Excluding pulmonary embolism at the bedside without diagnostic imaging: management [Ann Intern Med. 2001]

▶ See all Related Articles...

FIGURA 12

En algunos casos las búsquedas conviene también ampliarlas con los denominados “digest sources”, que son sitios que “filtran” y seleccionan la información sanitaria de calidad. Entre ellos es especialmente interesante TRIPdatabase (Figura 13), que permite localizar documentos que, elaborados con una metodología de Medicina Basada en la Evidencia, han sido publicados en revistas médicas, así como GPC y libros. El acceso a los documentos depende de la publicación original, por lo que algunos no son gratuitos.

trip database

Home | About Trip | Contact Us

Turning Research Into Practice

EDM Links | Trip Learning | My Trip

Search Trip Database

venous thrombosis AND diagnosis

● Title ● Title & Text

Search

For search tips click here.

Show synonyms used Show Search History

Refine Your Search

Your search generated relatively few results. This may be because your search term was too specific. Repeat your search with a more generalised search term. Alternatively, repeat your search with 'title & text'

Results by Category

(Click on category to view results)

Evidence Based Synopses	5
Clinical Questions	1
Systematic Reviews	11
Guidelines	0
- North America	0
- Europe	0
- Other	0
E-Textbooks	0
Clinical Calculators	0
Medical Images	0
Patient Information Leaflets	0

MEDLINE Articles

Therapy	716
Diagnosis	656
Systematic Reviews	287
Prognosis	959
Rhology	799
NEJM/JAMA/Lancet/RM	768

Filter Medline by Specialisation

(Pick One)

Filter

FIGURA 13

En nuestro caso se realizó una búsqueda adicional introduciendo los términos venous thrombosis AND diagnosis. Se localizaron 11 revisiones sistemáticas de las que la n° 1 parecía poder responder a nuestra pregunta (Figura 14). Tras la lectura crítica del artículo originalmente seleccionado una posterior revisión de este segundo documento confirmó que las conclusiones de ambos eran similares (Figura 15). La restricción que TripDatabase mantenía, con una limitación de 3 búsquedas semanales cuando no existía una suscripción a este recurso, ha sido suprimida al ser, desde Septiembre del 2006, su acceso totalmente gratuito. Quizás lo ideal es empezar con un recurso en inglés de contrastada calidad (por ejemplo la página de National Guidelines Clearinghouse), buscando una GPC de reciente actualización que responda a la pregunta clínica. En ocasiones necesitaremos más de una búsqueda para encontrar lo que buscamos, para lo que exploraremos entre los diferentes sitios de calidad contrastada.

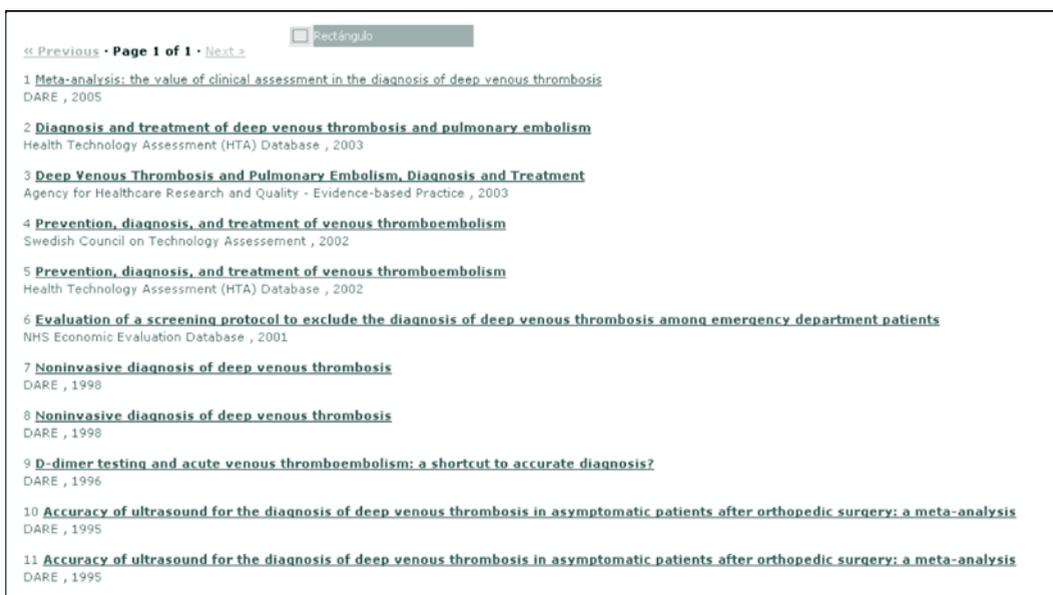


FIGURA 14

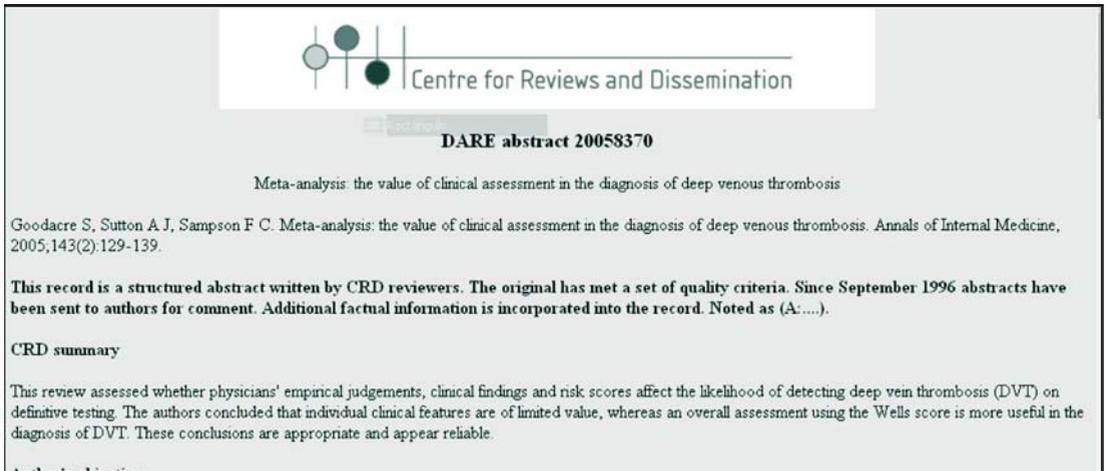


FIGURA 15

3.6.1. Lectura crítica del artículo seleccionado

El artículo finalmente seleccionado tras la búsqueda para responder a nuestra pregunta y ser valorado críticamente fue:

Ruiz-Gimenez N, Frieria A, Sánchez Molini P, Caballero P, Rodríguez-Salvanes F, Suarez C. Trombosis venosa profunda en miembros inferiores en un servicio de urgencias. Utilidad de un modelo clínico de estratificación de riesgo.

Med Clin (Barc) 2002; 118(14): 529-533.

Para la lectura crítica de un artículo se puede utilizar la plantilla CASPe previamente descrita.

Otra opción es la utilización de herramientas de lectura crítica (Critically Appraisal Topic o CAT) que pueden ser usadas mediante aplicaciones informáticas en forma de programas (CatMaker), actualmente disponibles a través de MurciaSalud. Las diferentes fases del proceso a través del CatMaker pueden observarse en las figuras 16-21.



FIGURA 16

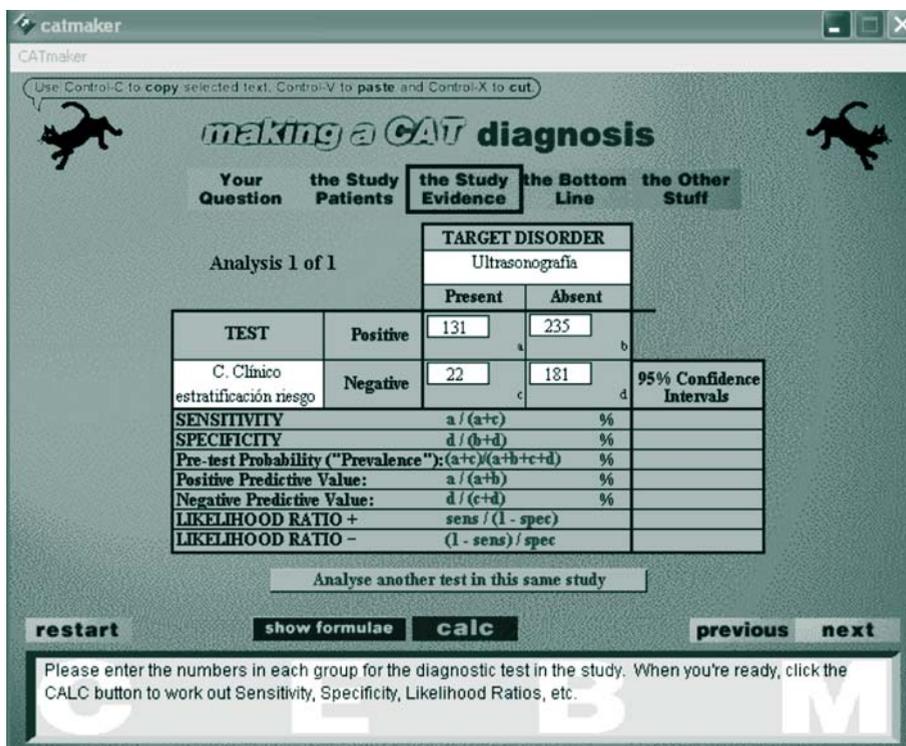


FIGURA 17

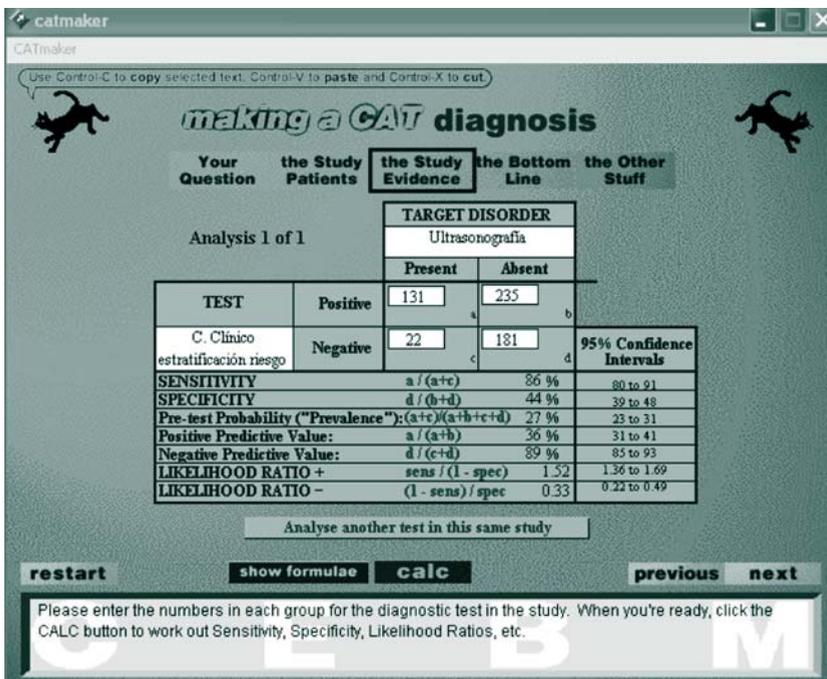


FIGURA 18

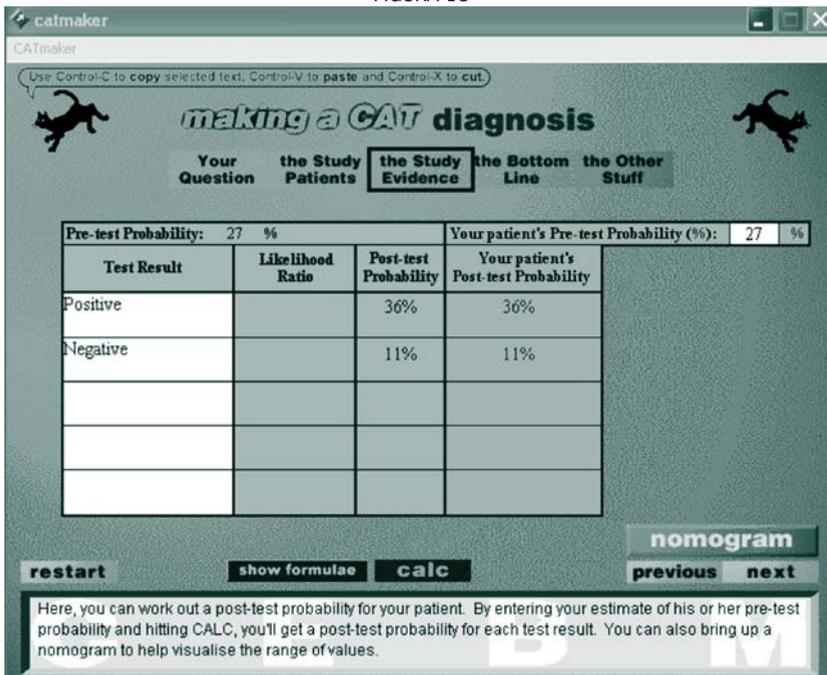


FIGURA 19
267

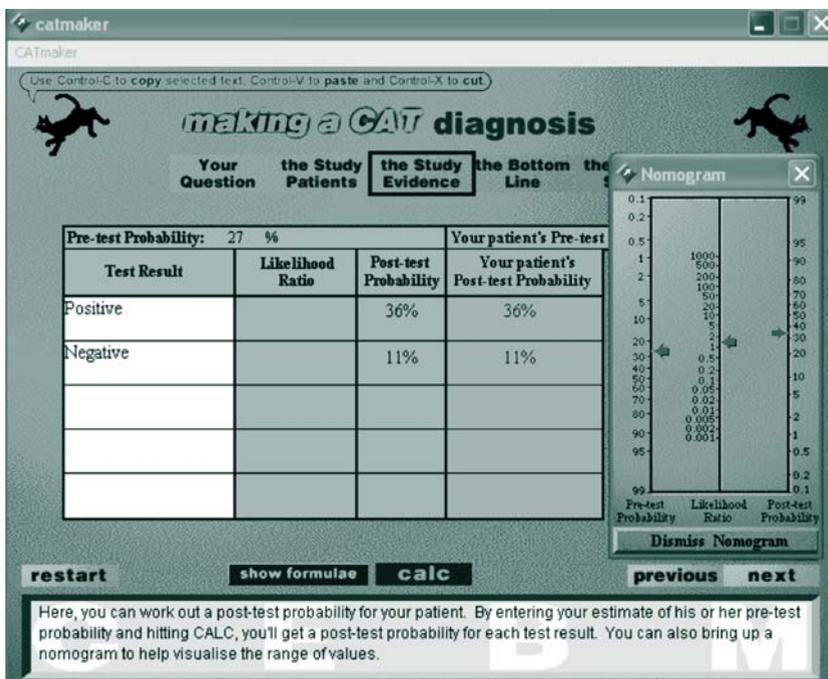


FIGURA 20

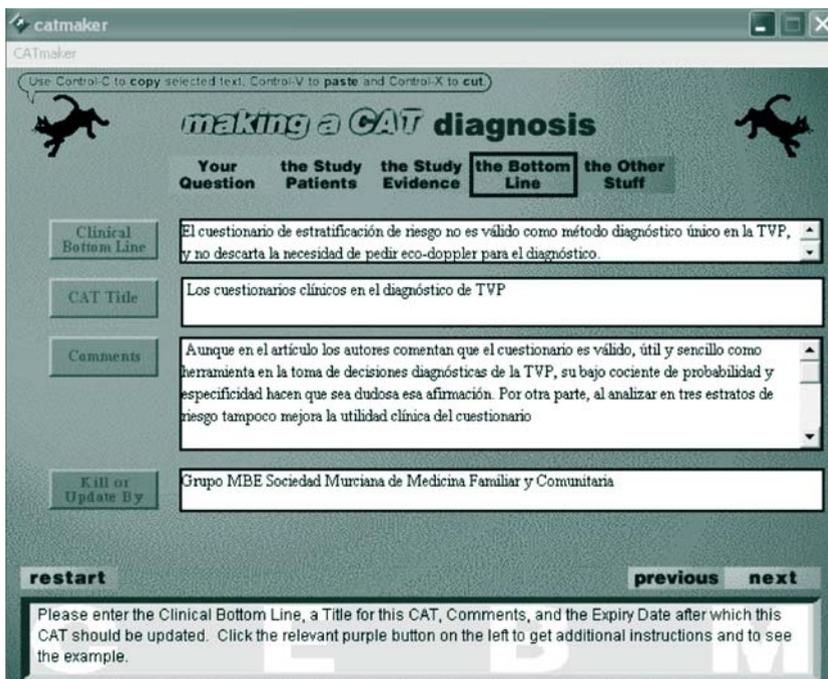


FIGURA 21
268

El método de los CAT se basa en que cuando para resolver dudas los clínicos aplicamos la metodología de la MBE aplicamos secuencialmente un proceso que consta de 5 pasos:

- 1.- Traducimos las preguntas a dudas concretas que pueden ser respondidas
- 2.- Buscamos en bases de datos la mejor evidencia para responder a las preguntas clínicas formuladas
- 3.- Valoramos esa evidencia y su aplicabilidad en un contexto clínico concreto
- 4.- Integramos la evidencia a nuestra experiencia clínica y la aplicamos.
- 5.- Evaluamos la puesta en práctica.

Un CAT es un resumen breve de las pruebas científicas más relevantes en respuesta a una pregunta clínica concreta. Para realizarlo se desarrollan los 3 primeros pasos en un resumen escrito y con formato preestablecido. El formato resumen del CAT puede observarse más abajo:

3.6.2. Ejemplo de aplicación de los determinantes estadísticos (realización de un CAT)

Los determinantes estadísticos expuestos en este capítulo deben ser aplicados a ejemplos concretos, para ello Internet puede ser un instrumento de gran ayuda. Aquí se reproduce como ejemplo el desarrollo de un CAT (Critically Appraisal Topic): No existen modelos fiables que orienten la probabilidad clínica de padecer TVP en pacientes que acuden a un servicio de urgencias, sobre los datos obtenidos en la publicación: Ruiz-Gimenez N, Frieria A, Sánchez Molini P et al. Trombosis venosa profunda en miembros inferiores en un servicio de urgencias. Utilidad de un modelo clínico de estratificación de riesgo. Med Clin (Barc) 2002; 118 (14): 529 533.

El cuestionario de estratificación de riesgo no es válido como método diagnóstico único en la TVP y no descarta la necesidad de pedir eco-doppler para el diagnóstico.

Cita bibliográfica:

Ruiz-Gimenez N, Frieria A, Sánchez Molini P.

Trombosis venosa profunda en miembros inferiores en un servicio de urgencias. Utilidad de un modelo clínico de estratificación de riesgo.

Med Clin (Barc) 2002; 118 (14): 529 533.

Cuestión clínica:

¿En pacientes que acuden a un servicio de urgencias con sospecha clínica de TVP, el uso de un cuestionario de estratificación clínica de riesgo es válido para orientar de forma correcta el diagnóstico, identificando aquellos casos en los que habría que solicitar eco-doppler para confirmarlo?

Términos de búsqueda:

Se utilizó el descriptor “venous trombosis” empleando el filtro “Clinical Query” para diagnóstico y seleccionando la estrategia especificidad. Se seleccionaron por medio del “limit” language spanish los artículos cuya lengua original fuera el castellano.

Pacientes del estudio:

	Sí	No	Sin determinar
Independiente?	X		
Ciego?		X	
Aplicación de valores estándar a los resultados del test	X		
Espectro apropiado?	X		

Enfermedad diana y prueba patrón (Gold Standard):

La Evidencia:

	Ultrasonografía				LR
	Presente		Ausente		
Questionario clínico estratificación riesgo	Num.	Prop.	Num.	Prop.	LR
Positivo	131	a	235	b	1,52
Negativo	22	c	181	d	0.33

Sensibilidad 86%; CI: 80 a 91

Especificidad 44%; CI: 39 a 48

Prevalencia: 27%; CI: 23 a 31

Valor Predictivo Positivo 36%; CI: 31 a 41

Valor Predictivo Negativo 89%; CI: 85 a 93

Comentario:

Aunque en el artículo los autores comenten que el cuestionario es válido, útil y sencillo como herramienta para la toma de decisiones diagnósticas en TVP, su bajo coeficiente de probabilidad y especificidad hacen que esa afirmación sea dudosa. Por otra parte al analizar en tres estratos de riesgo tampoco mejora la utilidad clínica del cuestionario.

Aplicación a mis pacientes:

Pre-test probabilidad: 27%

Test Resultado Post-test probabilidad

Positivo 36%

Negativo 11%

Otros ejemplos pueden obtenerse en recopilaciones bibliográficas de publicaciones tan importantes como: la revista British Medical Journal disponible en la dirección http://bmj.bmjournals.com/cgi/collection/diagnostics_tests y en la iniciativa Stats: <http://www.cmh.edu/stats/library/diagnostic.asp>

9

→ 4 PUNTOS CLAVE

*Por su importancia en el proceso de toma de decisiones, las pruebas diagnósticas (PD) son elementos esenciales para una práctica clínica de calidad. Una buena PD es aquella que permite discriminar entre pacientes sanos y enfermos en función del resultado de la misma. Por tanto, se deberán elegir aquellas pruebas con la mayor fiabilidad y validez posible.

*Además de la validez y fiabilidad al solicitar una prueba diagnóstica deberá tenerse en cuenta: la accesibilidad a la PD en el medio de trabajo, su capacidad para diagnosticar correctamente, sus efectos indeseables y la exactitud de los resultados.

*La sensibilidad, especificidad y los valores predictivos se han usado como indicadores de la validez de las pruebas diagnósticas, aunque en los últimos años se han aplicado otros indicadores, como las curvas ROC y la razón de probabilidad.

*La Validez Interna se define como la capacidad de la prueba para medir lo que pretende medir, se cuantifica mediante la sensibilidad y la especificidad. La Sensibilidad y especificidad tienen la desventaja de que no proporcionan información relevante para tomar una decisión clínica ante un determinado resultado de la PD.

*La validez externa se refiere al grado en que los resultados de un estudio son generalizables a otros sujetos o poblaciones. Depende de la prevalencia de la enfermedad y se mide mediante los valores predictivos y la Razón o cociente de probabilidad o cociente de verosimilitud (RP)

*La Razón de probabilidad es la mejor medida de utilidad de una prueba diagnóstica. Viene definida como el cociente entre la probabilidad de obtener un determinado resultado en los individuos enfermos y la probabilidad de ese mismo resultado en individuos no enfermos. Presenta la gran ventaja que su resultado no depende de la prevalencia de la enfermedad.

*Cuando se realiza la lectura crítica de un estudio de validación de pruebas diagnósticas, las publicaciones que aportan mayor información al médico clínico serán aquellas realizadas con estudios transversales, que incluyan una muestra representativa de la población de interés y donde se analicen simultáneamente los resultados de la prueba estudiada y los resultados obtenidos con una prueba de referencia (Gold Standard) que previamente ha demostrado ser de utilidad.

1. Wells P, Anderson DR, Bormanis J. Value of assessment of deep-vein thrombosis in clinical management. *Lancet* 1997; 350: 1795-1798.
2. Ruud Oudega, Karel GM Moons, Arno W. Hoes Limited value of patient history and physical examination in diagnosing deep vein thrombosis in primary care. *Family Practice* 2005; 22: 8691
3. Abraira V. glosario de terminos. Unidad de Investigación Hospital 1º de Octubre. Disponible en: http://www.hrc.es/investigación/mbe_glosario.html
4. Rodríguez Jiménez, Mª Esther. Técnicas de Imagen en Atención Primaria. Criterios orientativos de remisión de pacientes de Atención Primaria a los servicios de radiodiagnóstico. *Rev. Sociedad Madrileña de Medicina Familiar y C.* Nº2 .Vol 5. Agosto 2003
5. Ahlbom A. Norell S. Fundamentos de epidemiología. Ed. Siglo XXI. Madrid 2000
6. Wells PS, Owen C, Doucette S, Fergusson D, Tran H. Does this patient have deep vein thrombosis? *JAMA.* 2006 Jan 11;295(2):199-207
7. Watkins C, Daniels L, Jack C, Dickinson H, Van Der Broek M. Accuracy of a single question in screening for depression in a cohort of patients after stroke: comparative study. *BMJ,* 2001; 323: 1159.
8. Argimon Pallás, J.M. Epidemiología. En Martín Zurro A Cano Pérez J. Atención primaria. Conceptos, organización y práctica clínica. 5ª edición. Madrid. Elsevier.2003
9. Fagan TJ. Nomogram for Baye´s Theorem. *New England Journal Medicine* 1975; 293: 257.
10. Aguilar Franco C. El dímero D en pacientes con sospecha de tromboembolia venosa. *Formación Continuada.* 2005; 12:246-249
11. Phillips B. Towards evidence based medicine for paediatricians. *Archives of Disease in Childhood* 2003;88:82-83
12. Domínguez E. González Suárez R. Análisis de las curvas receiver-operating characteristic: un método útil para evaluar procederes diagnósticos.*Rev Cubana Endocrinol* 2002;13 (2):173-80. disponible en http://www.bvs.sld.cu/revistas/end/vol13_2_02/end10202.htm#cargo

13. López de Ullibarri Galparsoro I, Pita Fernández, S. Curvas ROC. *Aten Primaria* 1998; 5 (4): 229-235.
14. Muir Gray JA. *Atención sanitaria basada en la evidencia*. Madrid: Churchill Living Stone 1997.
15. Haynes RB. Of studies, syntheses, synopses, and systems: the "4S" evolution of services for finding current best evidence. *ACP J Club* 2001; 134 (2): A11-A13.
16. Ruiz-Canela Cáceres J, Louro González A. *Medicina Basada en la evidencia*. Donde y como buscar la información necesaria. Disponible en http://www.fisterra.com/mbe/mbe_temas/13/buscar_informacion.htm.
17. How doctors use tests?. *Bandolier* Mar 1999: 61-3 Consultado en <http://www.jr2.ox.ac.uk/Bandolier/band61/b61-3.html> el 10 de Junio del 2006
18. Black ER, Bordley DR, Tape TG, Panzer RJ. *Diagnostic Strategies*. American College of Physician-American Society of Internal Medicine ed. Philadelphia. 1999.
19. *Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice*. Gordon Guyatt, Drummond Rennie (Edit). Chicago: American Medical Association; 2002
20. Lijmer JC, Moll BW, Heisterkamp S et al. Empirical evidence of design related bias in studies of diagnostic tests. *JAMA* 1999; 282:1061-1066.
21. Richardson WS, Wilson MC, Guyatt GH, Cook DJ, Nishikawa J. Users' guides to the medical literature: XV. How to use an article about disease probability for differential diagnosis. *JAMA* 1999; 281 (13): 1214-9.
22. Ruiz-Gimenez N, Frieria A, Sánchez Molini P, Caballero P, Rodríguez-Salvanes F, Suarez C. Trombosis Venosa Profunda en miembros inferiores en un servicio de urgencias. Utilidad de un modelo clínico de estratificación de riesgo. *Med Clin* 2002; 118 (14): 529-33.



Acotado Automático de Términos: Brevemente, consiste en el enlace automático de una palabra o frase (que no es un término MeSH) al término MeSH que mejor lo describe. Una descripción más completa en el apartado correspondiente. Sinónimos (Sin): Automatic Term Mapping, ATM, Mapeo automático de términos.

Aleatorización: En un estudio en el que se comparan dos o más modalidades de tratamiento, es conveniente que los grupos formados sean parecidos en todas las características pronósticas excepto en el tratamiento recibido, de modo que cualquier diferencia en el curso clínico que se registre podrá ser atribuida a los diferentes tratamientos administrados (y sólo a ellos). La asignación aleatoria consiste en distribuir a cada participante de uno de los grupos de tratamiento por un método disciplinado de azar, de modo que cada sujeto tenga exactamente las mismas probabilidades de formar parte de uno u otro grupo de tratamiento. Sin: Randomización.

Alertas: Método usado para la difusión de información mediante mensajes electrónicos a ordenadores, agendas electrónicas e incluso teléfonos móviles. Representa una forma de actualización rápida y eficaz.

Análisis coste/beneficio: Medición de los costes y beneficios de una intervención en términos comensurables, generalmente monetarios. Permite determinar en primer lugar si una intervención concreta ofrece un beneficio global neto de bienestar, y en segundo lugar cómo se compara la ganancia de bienestar obtenida con esta intervención, con la obtenida con intervenciones alternativas. Al valorar todos los costes y beneficios en las mismas unidades, el análisis coste-beneficio compara diversas intervenciones con el criterio del beneficio neto (véase). Análisis económico en el que se consideran los costes de la asistencia sanitaria y la pérdida de ganancias netas debida a muerte o incapacidad, traduciendo todos los costes y beneficios a unidades monetarias.

Análisis coste/efectividad: Modalidad de estudio económico en la que las consecuencias de diferentes intervenciones pueden ser diferentes pero pueden ser medidas en unidades naturales idénticas, Mide los beneficios sanitarios en términos de unidades naturales, como años de vida salvados o mejorías del estado funcional. Dado que los costes no se expresan en las mismas unidades que los beneficios, su razón (ratio coste-efectividad) constituye una herramienta que permite valorar la eficiencia relativa. Análisis económico que compara los costes médicos incrementales y los resultados sobre la salud de programas sanitarios alternativos. Al contrario que el análisis de coste beneficio, los efectos sobre la salud se expresan en diversas unidades, como vida-años ganados, días libres de síntomas, casos evitados, etc., en lugar de en unidades monetarias

Análisis coste-utilidad: Adaptación del análisis coste-efectividad que mide el efecto de una intervención a la vez sobre los aspectos cuantitativos y cualitativos de la salud, mediante el uso de una variable denominada AVAC -Años de Vida Ajustados por Calidad- o QALYs . También evalúa la eficiencia relativa con una razón incremental, que en este caso se denomina ratio coste-utilidad.



Análisis de decisiones (AD): Es la aplicación de métodos explícitos y cuantitativos que cuantifican e integran los efectos del tratamiento, el pronóstico y los valores del paciente para analizar una decisión en condiciones de incertidumbre.

Análisis de minimización de costes: Modalidad de análisis coste-efectividad donde los resultados o beneficios de las alternativas se considera que son equivalentes o intercambiables para el mismo tipo de pacientes. A igualdad de resultados, la atención de los investigadores se centra en la diferencia de costes asociados.

Análisis de sensibilidad: En economía de la salud, técnica por la que se repiten las comparaciones entre recursos necesarios y resultados, pero con sucesivas modificaciones de las suposiciones en las que se basan los cálculos. De este modo, el análisis de sensibilidad pone a prueba la robustez de las conclusiones mediante variaciones de los ingredientes sobre los que existe incertidumbre.

Análisis de regresión: En un conjunto de datos sobre la variable dependiente y sobre una o más variables independientes, x_1, x_2, \dots, x_i , consiste en determinar el modelo matemático más ajustado que describa y como una función de las x o para predecir y a partir de las x . Término general para los métodos de análisis que se ocupan de estimar los parámetros de alguna posible relación entre una variable respuesta y una o más variables explicativas

Análisis de la varianza (ANOVA): Conjunto de técnicas estadísticas para conocer el modo en que el valor medio de una variable es afectado por diferentes tipos de clasificaciones de los datos. Con el análisis de la varianza se pueden ajustar las estimaciones del efecto de un tratamiento según otros factores como sexo, edad, gravedad, etc.

Análisis estratificado: Análisis de los resultados de un ensayo clínico o un estudio observacional, con el fin de examinar si hay diferencias de respuestas asociadas a alguna característica (por ej., edad o sexo).

Análisis factorial: Procedimiento estadístico empleado para la validación de constructo que permite identificar grupos o conglomerados de variables relacionadas.

Análisis multivariante: Cualquier método analítico que permite el estudio simultáneo de dos o más variables dependientes. Término genérico para los métodos que están diseñados para determinar las contribuciones relativas de diferentes causas a un solo suceso o resultado

Análisis por intención de tratar: En un ensayo clínico, análisis de los datos según el tratamiento asignado inicialmente por distribución aleatoria, independientemente de que el paciente no recibiera después este tratamiento.

Árbol de decisiones: Expresión gráfica de un análisis de decisiones.

Artículo científico: informe escrito que comunica por primera vez los resultados de una investigación. Los artículos científicos publicados en miles de revistas científicas componen la literatura primaria de la ciencia

Artículo de revisión: artículos de síntesis que resumen el conocimiento de un tema determinado y durante un periodo de tiempo definido. Junto con los libros, los artículos de revisión constituyen la literatura científica secundaria.

Asiento: Conjunto de información que, en una base de datos, se refiere a un mismo elemento. En Bases de Datos de bibliografía, es el conjunto de información que se refiere a una misma publicación. Sin: Registro, Record.

ATM: Ver Acotado Automático de Términos.

Atributos: Propiedades o características del instrumento que merecen consideraciones separadas e independientes en la evaluación.

Auditoria: Examen o estudio que evalúa e informa sobre la medida en que una condición, proceso o desempeño se ajusta a estándares o criterios predeterminados.

Automatic explosion: Ver Explode.

Automatic Term Mapping: Ver Acotado Automático de Términos.

AVAC: Índice que combina la supervivencia con la calidad de vida. La medida de la calidad de vida no está estandarizada y suele variar de un estudio a otro, según la enfermedad, el tratamiento evaluado y las preferencias de los autores. Para calcular un AVAC se multiplica la duración del estado de salud (en años) por un factor que representa la calidad (“utilidad”) de este estado de salud. Sin: QALY (Quality Adjusted Life Years).

Azar: Se denomina así al proceso que ocurre debido a la casualidad, de forma imprevisible, sin planeamiento ni influencia consciente.

Base de datos: Conjunto de datos homogéneos y ordenados, que se presentan normalmente, de forma legible, por ordenador y que se refieren a una materia determinada. Colección estructurada de datos que está organizada de tal manera que se puede acceder a ella con una amplia variedad de programas de aplicación.

Beneficio neto: Diferencia entre los costes y los beneficios de una intervención.

Bioestadística: Aplicación de la teoría y los métodos de la estadística a las ciencias biológicas. En sentido estricto, rama de la ciencia que aplica los métodos estadísticos a los problemas biológicos, aunque ahora se utiliza la mayoría de las veces para referirse a la estadística aplicada a la medicina y las ciencias de la salud.

Blinding: Ver Enmascaramiento.

Buscadores: Motores de búsqueda de información en Internet que identifican la información contenida en bases de datos formadas a partir de publicaciones previamente indexada. Ver Indización.

Calidad de Vida Relacionada con la Salud (CVRS): Impacto que una enfermedad y el tratamiento consiguiente tienen sobre la percepción del paciente de su satisfacción y bienestar físico, psíquico, social y espiritual.

Campo: Cada una de las unidades de información en que se divide un asiento. Sin: Field.

Carga de administración: Se define como el grado de trabajo y esfuerzo que deben realizar quienes administran un instrumento de medida (cuestionario), y depende del tiempo precisado para su administración, formación requerida del entrevistador y recursos necesarios para su aplicación y puntuación.

Carga de respuestas: Se entiende como el tiempo, esfuerzo y otros aspectos que se exigen a quienes se les administran un instrumento de medida (cuestionario).

CAT: Herramienta de lectura crítica consistente en un resumen breve de las pruebas científicas más relevantes en respuesta a una pregunta clínica concreta.

Causalidad: Indica que las asociaciones encontradas entre las exposiciones y los efectos son realmente de tipo causal y no consecuencia de sesgos ni del efecto de factores de confusión.

Cociente de probabilidad: Véase Razón de probabilidad.

Coefficiente de correlación: Índice que cuantifica la relación lineal entre un par de variables. Medida de asociación que indica el grado en el que dos variables continuas x e y poseen una relación lineal ($y = a + bx$). Su valor puede situarse entre -1 y $+1$. Cuando su valor es 0 no existe ningún grado de correlación. El coeficiente de este tipo que más se utiliza es el de Pearson

Coefficiente de Cronbach: Fórmula utilizada para la comprobación de la consistencia interna de un test que presupone que si varios ítems están midiendo el mismo atributo o dimensión, estos ítems deberían estar intercorrelacionados. Esta correlación la mediremos mediante el coeficiente alfa de Cronbach.

Coefficiente kappa: índice corregido por el azar de la concordancia entre, por ejemplo, juicios o diagnósticos que hacen 2 ó más observadores. El coeficiente adopta desde el valor 1 (cuando hay una concordancia perfecta) hasta el valor 0 cuando la concordancia observada es igual a la debida al azar.

Confusión: Situación en la que la estimación del resultado estudiado está sesgada por algún factor relacionado con la exposición y el resultado.

Consentimiento informado: Procedimiento formal para aplicar el principio de autonomía. Debe reunir por lo menos tres elementos: voluntariedad, información y comprensión. La voluntariedad implica que los sujetos puedan decidir libremente si quieren participar en un estudio, sin que haya persuasión, manipulación ni coerción. La información debe ser comprensible y debe incluir el objetivo del estudio y su procedimiento, los beneficios y riesgos potenciales, las posibles incomodidades derivadas de la participación y la posibilidad de retirarse del estudio en cualquier momento sin que ello de lugar a perjuicio, en este contexto es básica la comprensión de la información, que no sólo depende de esta, sino también de la competencia de quien la recibe, es decir de su capacidad para comprender.

Consistencia interna: Grado en que los ítems de una dimensión se correlacionan entre ellos mismos. Ver: Coeficiente de Cronbach.

Constructo: En psicometría, concepto abstracto no medible directamente que se quiere convertir en variable operativa medible. Generalmente, se utiliza para un concepto que existe en teoría, pero que no se puede observar directamente. Esencialmente, es una variable latente.

Correlación ítem-total:(Correlación total de las preguntas): Verifica la homogeneidad. Proceso mediante el cual comprobamos si cada ítem mide lo mismo que los demás y por lo tanto es sumable en una puntuación total que supuestamente mide el rasgo y que es la que después interpretamos y utilizamos.

Coste: Es el resultado del producto de la cantidad total del recurso consumido y el valor monetario de la unidad de dicho recurso.

Coste oportunidad: En economía de la salud, valor de las oportunidades perdidas por usar recursos en un servicio. Generalmente es cuantificado considerando el beneficio que se obtendría invirtiendo los mismos recursos en la mejor alternativa posible.

Criterios: Condiciones utilizadas como un estándar para juzgar los atributos.

Criterios de selección: Características sociodemográficas, clínicas, etc. que determina que los pacientes puedan ser o no incluidos en un estudio. Nos van a servir para definir la población de estudio.

Curva de Gauss: En matemáticas, la campana o curva de Gauss es la representación gráfica de la ecuación matemática que corresponde a una Distribución normal. Tiene forma de campana y debe su nombre al matemático alemán Carl Friedrich Gauss.

Curvas ROC: Representación gráfica completa de la capacidad de una prueba para discriminar lo normal de lo que no lo es.

Delimitador de campo: Ver Etiqueta.

DeCS - Descriptores en Ciencias de la Salud: Tesauro que puede considerarse, hasta cierto punto, como una versión en español del MeSH - Medical Subject Headings de la U.S. National Library of Medicine. Fue desarrollado a partir de él con el objetivo de permitir el uso de terminología común para búsqueda en tres idiomas (español, inglés y portugués), proporcionando un medio consistente y único para la recuperación de la información independientemente del idioma. Además de los términos médicos originales del MeSH contiene áreas específicas de Salud Pública y Homeopatía.

Descriptor: Término del lenguaje documental, carente de variabilidad, sinonimia y polisemia, escogido para representar el contenido de un documento, que puede ser utilizado en el análisis y recuperación de la información. Relacionado con Palabra clave.

Descriptor mayor: Ver Mayor Topic.

Descriptores en Ciencias de la Salud: Ver DeCS

Descuento: Procedimiento para calcular los costes o beneficios que ocurren en diferentes fechas a su valor en el momento presente, llamado valor actual.

Dimensiones: Aspectos relevantes que engloba el constructo o concepto que se mide a través del cuestionario.

Diseño: Conjunto de estrategias y procedimientos que se seguirán para dar respuestas al problema y comprobar la hipótesis. Engloba al selección del tipo de estudio, la población a estudiar, la definición de las variables y escalas de medición y los métodos de recogida, procesado y análisis de los datos.

Docencia virtual (“on-line” o “e-learning”): Definida como la utilización de las nuevas tecnologías multimedia y de internet para mejorar la calidad del aprendizaje, facilitando el acceso de recursos y servicios; así como los intercambios y la colaboración a distancia.

Dosis Diaria Definida (DDD): Dosis de mantenimiento promedio en la principal indicación de un fármaco en pacientes adultos. Se usa para cuantificar el consumo de medicamentos, sobre todo cuando se compara el consumo de diversos fármacos de un mismo grupo, o cuando se realizan comparaciones internacionales. No se ve afectada por diferencias de precios o de tamaño de las presentaciones. El consumo se suele expresar en número de DDD por 1.000 habitantes y por día, y esta medida da una idea del número promedio de personas tratadas diariamente con un fármaco o grupo de fármacos en una región o país.

Efectividad: Grado en que una determinada intervención origina un resultado beneficioso en las condiciones de la práctica clínica habitual, sobre una población determinada.

Efecto indeseado: Ver Reacción Adversa.

Eficacia: Grado en que una determinada intervención origina un resultado beneficioso en condiciones ideales de aplicación.

Eficiencia: Relación entre los efectos o resultados finales que se alcanzan con una intervención y coste que supone la misma (recursos humanos, materiales y de tiempo).

Encabezamientos de Materias Médicas: Ver MeSH.

Enmascaramiento: Procedimiento por el que se asegura que los sujetos participantes en un ensayo clínico, los observadores o ambos, no conocen el tratamiento u otra característica que pudiera sesgar los resultados. Se denomina simple ciego cuando el sujeto, pero no el observador ignora cuál de los posibles tratamientos recibe; doble ciego es cuando sujeto y observador ignoran el tratamiento recibido; y triple ciego es cuando el sujeto participante, el investigador observador y el investigador que analiza los datos ignoran el tratamiento recibido.

Ensayo clínico: “Toda investigación efectuada en seres humanos para determinar o confirmar los efectos clínicos, farmacológicos y/o demás efectos farmacodinámicos, y/o de detectar las reacciones adversas, y/o de estudiar la absorción, distribución, metabolismo y excreción de uno o varios medicamentos en investigación con el fin de determinar su seguridad y/o su eficacia” (Real Decreto de 16 de febrero de 2004).

Ensayo clínico controlado: Ensayo clínico en el que los sujetos son asignados a dos grupos: uno (grupo experimental) recibe el tratamiento que se está probando y el otro (grupo de comparación o control) recibe un tratamiento alternativo. En el ensayo clínico controlado aleatorizado la asignación a un tratamiento experimental o control se realiza al azar de forma que ni el sujeto ni el investigador responsable de su selección o tratamiento puedan influir en su asignación.

Entry Term: Relacionado con Palabra clave (véase).

Epidemiología: Estudio de la distribución y los determinantes de los estados o acontecimientos relacionados con la salud de poblaciones. Ciencia que investiga las causas y la historia de las enfermedades y el desarrollo de éstas en relación con determinadas intervenciones o exposiciones.

Epistemología: Doctrina de los fundamentos y métodos del método científico. La función de la estadística en el modelo empírico analítico.

Equivalencia: capacidad del instrumento de proporcionar los mismos valores en las mismas condiciones clínicas cuando las mediciones las hacen distintas personas.

Error aleatorio: Las diferencias de medidas entre la realidad y nuestra muestra y son debidas al azar.

Escala tipo Likert: La respuesta del individuo se da mediante una escala numérica, a la cual se le atribuyen las propiedades de las escalas de intervalo, permitiendo por tanto la realización de análisis estadísticos más complejos. Escalas que se utilizan con frecuencia en estudios de actitudes en los que las puntuaciones crudas se basan en respuestas alternativas graduadas a cada una de una serie de preguntas.

Especificidad: Referida a una prueba diagnóstica, es la proporción de personas verdaderamente no enfermas que son catalogadas como tales por medio de la prueba.

Estabilidad: Ver Reproducibilidad.

Estabilidad del test: Véase Método test-retest.

Estadística: Recopilación y organización de hechos o datos numéricos, sobre todo cuando hay muchos. Conjunto de métodos que ayudan a la toma de decisiones frente a la incertidumbre. La estadística no ofrece pruebas ni certidumbre, sino probabilidades.

Estadístico: Medida estadística (media, mediana, etc...) calculada sobre los datos de una muestra.

Estadístico kappa: Medida del acuerdo, corregida por el azar, aplicada a datos binarios u ordinales. Se calcula como la proporción de acuerdo, aparte del que ya sería de esperar por azar, que ha sido observado entre dos repeticiones del mismo instrumento.

Estudios ambispectivos: Estudios que se inician como los retrospectivos (Ver: Estudios retrospectivos) para el análisis de un fenómeno ya ocurrido y, posteriormente, continuamos con un seguimiento prospectivo hasta la aparición de un efecto a largo plazo (Ver: Estudios prospectivos).

Estudios analíticos: Estudios diseñados para examinar asociaciones, cuyo objeto final suele ser identificar o medir los efectos de factores de riesgo o de intervenciones específicas sobre la salud.

Estudios de caso-control: Estudios en los que la selección de la población a estudiar se hace en función de la presencia o ausencia de una enfermedad o efecto, eligiéndose un grupo de individuos que tienen la enfermedad o efecto (casos), y otro que no los presenta (controles).

Estudios de cohortes: Tipo de estudios en los que los sujetos del estudio se seleccionan a partir presencia o ausencia de la exposición que se trata de analizar y de la que se sospecha es responsable del efecto observado.

Estudio descriptivo: Estudio cuyo objetivo es constatar hechos, observar lo que ocurre en nuestra población. Permiten generar hipótesis.

Estudio experimental: Estudio en el que el investigador asigna y controla el factor de estudio.

Estudios de incidencia: Tipo de estudios en los que seguimos a una población determinada a lo largo del tiempo para analizar la aparición o no de un fenómeno de interés. La unidad de análisis es el individuo.

Estudios longitudinales: Estudios en los que la recogida de los datos de interés hace referencia a dos momentos separados en el tiempo (se analizan los sujetos en varios momentos a lo largo del tiempo).

Estudios observacionales: Estudios en los que el investigador no participa directamente en la intervención o exposición (factor de estudio), se limita a observar (registrar) lo que ocurre en la realidad.

Estudios pragmáticos: Estudios cuyo objetivo principal es la evaluación de la adecuación de la práctica asistencial en condiciones normales.

Estudios pre y post-intervención: Estudios en los que se registran la situación basal de la muestra y los cambios que experimenta tras su exposición a una determinada intervención.

Estudios de prevalencia: Estudios en los que se describe una serie de variables en una población determinada y en un momento determinado. La unidad de análisis es el individuo.

Estudios prospectivos: Tipo de estudios en los que los hechos a estudiar no han tenido lugar aun y los sujetos son seguidos hasta obtener el resultado.

Estudio retrospectivo: Estudio en el cual los hechos a estudiar ya han tenido lugar al iniciar el estudio.

Estudio transversal: Tipo de estudio en el que se estudia la exposición y el desenlace al mismo tiempo.

Etiqueta: Abreviatura del nombre de un campo que puede utilizarse en su lugar para dirigir la búsqueda de información. Sin: Tag, Delimitador de campo.

Evaluación crítica: Véase Lectura crítica.

Evaluación económica: Análisis comparativo de opciones alternativas en términos de sus costes y consecuencias. Los estudios de evaluación económica pueden ser parciales, cuando sólo se describen los costes, los resultados o ambos a la vez pero de una única opción o que comparan dos o más alternativas pero sólo en función de sus costes o de sus consecuencias sin valorar ambos a la vez, o completos, cuando se analizan comparativamente dos o más alternativas mediante la valoración de los costes y de las consecuencias de cada una de ellas.

Evaluación de tecnologías sanitarias: Proceso de análisis e investigación dirigido a estimar el valor y contribución relativos de cada tecnología sanitaria (véase) en la mejora de la salud individual y colectiva, teniendo en cuenta su impacto económico, social e, incluso, implicaciones éticas.

Explode: Obtención de las referencias que contienen el descriptor especificado o cualquiera de los descriptores de su mismo campo semántico pero de significado más restringido. Sin: Automatic explosion, Exploding, Explosion, Explosión, Explosión automática.

Exploding: Ver Explode.

Exploración complementaria: Conjunto de pruebas diagnosticas realizadas con la finalidad de confirmar una hipótesis diagnostica o hacer el seguimiento de la patología de un paciente.

Explosión: Ver Explode.

Explosión automática: Ver Explode.

Factibilidad: Indica las posibilidades reales y práctica de que un instrumento de medida pueda ser contestado, completado y analizado. Está relacionad con el nivel de aceptabilidad del instrumento y se refleja en la proporción de rechazos, no respuesta y datos perdidos.



Factor de confusión: Variable que está simultánea e independientemente asociada al factor de riesgo estudiado y a la enfermedad estudiada, que puede modificar el resultado del estudio. Tal variable debe ser identificada y se debe evitar su influencia antes de la recogida de datos (mediante el emparejamiento o la restricción) o bien en la fase de análisis, mediante el análisis estratificado y la regresión múltiple.

Factor de impacto: Indicador bibliométrico, elaborado por el ISI (véase) y publicado en el Journal Citation Reports, que mide el uso de las revistas. Suele calcularse dividiendo el número de veces que son citados los artículos publicados en una revista durante un plazo de tiempo (habitualmente dos años), entre el número de artículos publicados por esa misma revista y en ese mismo plazo de tiempo. Puede modificarse para calcular el impacto de autores, instituciones... Sin: Impact Factor.

Factor de riesgo: Factor relacionado con el riesgo de que ocurra un suceso (p. ej. enfermar).

Falso negativo: Referido a una prueba diagnóstica, número de enfermos en que el resultado sería negativo.

Falso positivo: Referido a una prueba diagnóstica, número de sanos en que el resultado sería positivo.

Farmacovigilancia: Programas que proporcionan de forma continuada una evaluación sobre la seguridad de los fármacos. El factor clave en este sistema son los profesionales sanitarios que participan en el programa de notificación espontánea de reacciones adversas.

Fiabilidad: Grado de consistencia con que mide el atributo para el que está diseñado. La consistencia implica la ausencia de variabilidad o desacuerdo en repetidas observaciones o mediciones. Grado en que las mismas mediciones de los individuos que se obtienen en diferentes condiciones dan resultados similares.

Fiabilidad inter-observador (inter-ratio): Mide la variación que ocurre entre diferentes observadores al medir a un mismo individuo con un mismo instrumento .

Field: Ver Campo.

Filtros metodológicos: Son estrategias de búsqueda que nos ayudan a encontrar, en bases de datos como Medline, diferentes tipos de estudio sobre tratamiento, diagnosis, etiología o pronóstico. En Pubmed se denominan "Clinical Queries". Para aplicar uno de estos filtros sólo tenemos que combinar nuestra búsqueda con AND al filtro que nos interesa.

Forest/ Forest Plot: Tipo de gráfico ampliamente utilizado para representar los resultados de los diversos estudios seleccionados para una revisión sistemática. El Forest Plot ayuda a decidir si los resultados de los estudios individuales son suficientemente similares para poder ser combinados en una sola estimación.

Generabilidad: Véase Validez externa.

Gold Standard: Es el criterio que define quién tiene "realmente" la enfermedad (o condición) en estudio.

Grado o nivel de significación: En las pruebas de significación estadística, es el valor de p, el cual, en sentido estricto, en un ensayo clínico debe ser preespecificado en la fase de diseño, el nivel aceptado con mayor frecuencia es de 0.05. Ver: P (probabilidad), Significación clínica y Significación estadística.

Grupo de comparación: distintos grupos que intentamos comparar en un estudio analítico, ya sean estudios caso-control, estudios de cohortes o ensayos clínicos.

Grupo control: En un ensayo clínico, el grupo que recibe placebo o bien un tratamiento activo ya conocido, aceptado y establecido. Sin: Grupo de referencia.

Grupo intervención: En un ensayo clínico, el grupo que recibe el tratamiento que se desea evaluar.

Guía farmacoterapéutica: Documento que contiene una relación limitada de medicamentos recomendados para la prescripción en un ámbito determinado, seleccionada a partir de la oferta farmacéutica en función de unos criterios previamente establecidos, con la participación y el consenso de los profesionales a los que va destinada.

Guía de practica clínica: Conjunto de recomendaciones diseñadas para ayudar a los profesionales sanitarios y a los usuarios a seleccionar las opciones diagnósticas y/o terapéuticas más adecuada en el abordaje de una condición clínica específica.

Hipótesis: Suposición, fundada en observaciones o reflexiones, que puede conducir a predicciones refutables. También, conjetura elaborada de forma que pueda ser comprobada y refutada.

History: Archivo de todas las consultas bibliográficas realizadas a lo largo de una sesión en PubMed. También designa al subgrupo de registros que provienen de artículos sobre historia de la medicina.

Holística: Relativa al "holismo" que es la doctrina que propugna la concepción de cada realidad como un todo distinto de la suma de las partes que lo componen.

Impact Factor: Ver Factor de impacto.

Incremento absoluto del riesgo (IAR): Cuando el tratamiento experimental aumenta la probabilidad de un acontecimiento adverso, diferencia absoluta entre las tasas de acontecimientos adversos.

Incremento relativo del riesgo (IRR): Cuando el tratamiento experimental aumenta la probabilidad de un acontecimiento adverso, aumento proporcional de la tasas de acontecimientos adversos en el grupo experimental relativa a la tasa en el grupo de referencia.

Indicador de adhesión a la Guía Farmacoterapéutica: Indicador genérico de prescripción que indica el porcentaje de Dosis Diaria Definida (DDD, véase) de medicamentos prescritos incluidos en la Guía Farmacoterapéutica respecto del total de DDD de medicamentos prescritos.



Índice Bibliográfico Español en Ciencias de la Salud (IBECS): Base de Datos de la Biblioteca Nacional de Ciencias de la Salud (BNCS) del Instituto de Salud Carlos III (ISCIII), contiene referencias de artículos de revistas científico sanitarias editadas en España e incluye contenidos de las diferentes ramas de las ciencias de la salud tales como Medicina (incluyendo Salud Pública, Epidemiología y Administración Sanitaria), Farmacia, Veterinaria, Psicología, Odontología y Enfermería.

Índice Médico Español (IME): Base de datos referencial bibliográfica, creada y gestionada por el Instituto de Historia de la Ciencia y Documentación (IHCD), centro mixto de la Universidad de Valencia y del Consejo Superior de Investigaciones Científicas (CSIC). Cubre 321 publicaciones periódicas editadas en España especializadas en Biomedicina (Administración sanitaria, Farmacia clínica, Medicina experimental, Microbiología, Psiquiatría, Salud pública).

Indización: Proceso por el que el documentalista asigna unos descriptores que describen el contenido conceptual de documento con el fin de facilitar su recuperación

Index: Listado de todos los términos (unitérminos o compuestos) con significado que están incluidos en la Base de Datos. Esta opción permite hacer una búsqueda seleccionando los términos directamente del listado.

Indexar: Ver indización.

Ineficiencia: Situación en la cual los recursos no son empleados de la forma en que se incrementarían los beneficios producidos.

Inferencia: Proceso de pasar de las observaciones y axiomas a la generalizaciones. En estadística, desarrollo de la generalización a partir de los datos de la muestra, generalmente con un margen de incertidumbre.

Instrumentos de medida de CVRS específicos: Los instrumentos específicos incluyen dimensiones de la CVRS de una determinada afección que se quiere estudiar, capacidad funcional, o de un determinado grupo de pacientes.

Instrumentos de medida de CVRS genéricos: Los instrumentos de calidad de vida genéricos no están relacionados con ningún tipo de enfermedad o especialidad en particular. Son aplicables tanto a la población general como a grupos específicos de pacientes.

Interpretabilidad: grado en el que uno puede asignar un significado comprensible a las puntuaciones cuantitativas del instrumento.

Intervalo de confianza: Margen o recorrido comprendidos entre dos valores extremos (el límite inferior y el superior), entre los que cabe esperar el valor medio real de la población con un margen grado prefijado de seguridad. Ver: Nivel de confianza.

Intervención: Cualquier acción (preventiva, diagnóstica, sintomática o curativa) aplicada con el fin de modificar el curso de la situación previa. Cualquier tratamiento o actuación sanitaria que queramos evaluar

ISI: El Institute for Scientific Information fue creado por Eugene Garfield en 1958 con el objeto de publicar el primer Current Contents que cubría unas 200 revistas. En 1971, cubría más de 4000 revistas, publicaba más de 650 páginas de contenidos a la semana, y era consultado por más de un millón de científicos. En 2001, se desarrolló el ISI Web of KnowledgeSM, que integra literatura científica, tanto revistas (journals) como actas (proceedings), patentes, compuestos químicos, y hasta secuencias genómicas. El ISI fue adquirido por The Thomson Corporation, que lo convirtió en el Thomson Institute for Scientific Information. Es, por tanto, una institución privada que produce diversas Bases de Datos de primer orden en calidad y utilidad en la investigación científica, pero que sobre todo ha obtenido un notable éxito por la difusión de lo que se conoce como Impact Factor o Factor de Impacto.

ISSN: International Standard Serial Number, número internación único que identifica de manera inequívoca cada revista.

Ítem: Cada uno de los elementos, preguntas o cuestiones que constituyen un test.

Key word: Ver Palabra clave.

Lectura crítica: Es el proceso de evaluar e interpretar la evidencia aportada por la literatura científica, considerando sistemáticamente los resultados que se presenten, su validez y su relevancia para el trabajo propio. Sin: Evaluación crítica.

Lenguaje documental: Lenguaje artificial creado específicamente para el análisis y recuperación de la información. Contrapuesto al Lenguaje natural, se diferencia de este especialmente por carecer de variabilidad, sinonimia y polisemia.

Lenguaje natural: Expresión de ideas o de conceptos con los términos y con las reglas propios de un idioma o de un grupo especial de hablantes (p.e. sanitarios). Se contrapone a Lenguaje artificial y, específicamente en documentación, a Lenguaje documental.

Límites: Opción que ofrecen varias bases de datos, como Medline, que permite limitar/restringir una búsqueda de acuerdo a unas características como p.ej. artículos publicados en un idioma determinado, tipo de publicación (revisión, ensayo clínico), o grupo de edad, sexo, etc.

Likehood Ratio: Véase Razón de probabilidad.

Literatura inédita: Literatura que no ha sido publicada.

Mapeo Automático de Términos: Ver Acotado Automático de Términos.

Mayor Topic: Materia principal de un artículo. Sin: Descriptor mayor.

Médical Subject Headings: Ver MeSH

Medicina Basada en la Evidencia (MBE): Práctica de la medicina basada en la integración de la maestría clínica de cada médico (es decir, el dominio del conocimiento y el juicio adquirido por la experiencia) con las mejores pruebas clínica externas disponibles. También ha sido definida como la utilización consciente y juiciosa de las mejores demostraciones provenientes de la investigación clínica para la toma de decisiones en el cuidado de cada paciente.

Medicamentos genéricos (Especialidad Farmacéutica Genérica, EFG): Especialidad con la misma forma farmacéutica e igual composición cualitativa y cuantitativa en sustancias medicinales que otra especialidad de referencia, cuyo perfil de eficacia y seguridad está suficientemente establecido por su continuado uso clínico. La EFG debe demostrar la equivalencia terapéutica con la especialidad de referencia mediante los correspondientes estudios de bioequivalencia.

Medicamentos huérfanos: Cuando un fármaco no es desarrollado porque el detentador de la patente considera que no recuperará los costes de este desarrollo (en el caso de enfermedades raras y otras enfermedades no raras pero que ocurren en países pobres en los que no existe un “mercado” potencial), se lo denomina medicamento huérfano.

MeSH (Médical Subject Headings, Encabezamientos de Materias Médicas): Tesauro (diccionario que permite conocer la equivalencia entre los términos del lenguaje natural y los del lenguaje documental) de Medline y de PubMed.

MeSH Subheading: Subencabezamiento propio de MeSH.

MeSH Term: Descriptor propio de MeSH.

Meta-análisis: Integración estructurada y sistemática de la información obtenida en diferentes estudios sobre un problema determinado. Consiste en identificar y revisar los estudios controlados sobre un determinado problema, con el fin de dar una estimación cuantitativa sintética de todos los estudios disponibles.

Método matricial multirasgo-multimétodo (multitrait-multimethod matrix method, MTMM): procedimiento empleado para la validación de constructo aplicando para ello los conceptos de convergencia (validez convergente) y discriminabilidad (validez discriminante).

Método test-retest: Método utilizado para la comprobación de la fiabilidad de un test. Correlación entre la puntuación de dos administraciones de una prueba a los mismos individuos. De esta manera se mide uno de los conceptos incluidos en la fiabilidad: la estabilidad del test.

Modelo de medida: Se refiere a las escalas y subescalas de un cuestionario y a los procedimientos seguidos para crear las puntuaciones.

Muestra: Fracción seleccionada de una población determinada, idealmente representativa de la población de referencia y de tamaño suficiente.

Muestreo: Selección de los sujetos de una población que tomarán parte en un estudio. El muestreo aleatorio implica que cada participante potencial de la población estudiada tiene las mismas probabilidades de ser incluido en la muestra.

NEAR: Operador booleano utilizado en las búsquedas en Librería Cochrane Plus. Buscará documentos donde las dos palabras que enlaza NEAR estén separadas entre sí por un máximo de seis palabras en cualquier orden.

NEXT: Operador booleano utilizado en las búsquedas en Librería Cochrane Plus. Recupera documentos donde las dos palabras están juntas en ese orden.

Nivel de confianza: Se expresa en términos probabilísticos y se interpreta como la probabilidad de que el verdadero parámetro poblacional esté dentro del intervalo de confianza. Ver: Intervalo de confianza.

Normograma de Fagan: Método de cálculo de la probabilidad posttest mediante la representación gráfica de una tabla. Permite conocer la probabilidad posttest mediante el trazo de una línea que una la probabilidad pretest y la razón de probabilidad, el siguiente punto de intersección con la línea situada a la derecha, señala la probabilidad posttest.

Número necesario para dañar (number needed to harm, NNH): Número de pacientes que deben tomar el tratamiento experimental para que se produzca un caso adicional de un determinado acontecimiento adverso, en comparación con los pacientes que reciben el tratamiento de referencia.

Número necesario para tratar (NNT): Es una medida de la eficacia de un tratamiento. Cuando el tratamiento experimental incrementa la probabilidad de un acontecimiento favorable (o cuando disminuye la de un acontecimiento adverso), número de pacientes que hay que tratar para dar lugar a un paciente más con mejoría (o para prevenir un acontecimiento adverso adicional). Se calcula como $1/\text{RAR}$.

Odds: Se llama odds a la razón en la que el numerador es la probabilidad de que ocurra un suceso y el denominador es la probabilidad de que tal suceso no ocurra.

Odds post-test: Equivale a los odds (proporción) de que un paciente tenga la enfermedad una vez que se ha realizado el test diagnóstico.

Odds pre-test: Definida como los odds (proporción) de que un paciente tenga la enfermedad antes de que el test diagnóstico se efectúe.

Odds ratio (OR): Resultado de dividir la odds del grupo de interés (expuesto) por la odds del grupo de referencia (control). Si el $\text{OR} = 1$ el riesgo o probabilidad de que ocurra el evento es el mismo para ambos grupos (el efecto del tratamiento experimental es el mismo que el del tratamiento control); si el $\text{OR} > 1$ se interpretará como que el tratamiento experimental favorece la aparición del evento y si el $\text{OR} < 1$ diremos que protege de la aparición del mismo.

Operadores booleanos: Partículas habitualmente expresadas en inglés (AND, OR, NOT) que indican operaciones lógicas entre conceptos. Las operaciones son las de intersección (AND = recupera los documentos que contengan ambos términos), adición (OR = recupera los documentos que contengan tanto un término como otro, o la suma de los dos) o exclusión (NOT = sirve para excluir documentos que tienen un determinado término).

P: Seguía de la abreviatura n.s. (no significativa) o del símbolo $<$ o una cifra decimal, indica la probabilidad de que la diferencia observada en una muestra haya ocurrido puramente por azar siendo los grupos comparados realmente semejantes, es decir bajo la hipótesis nula.



Palabra clave: Término propio del lenguaje natural, sin otras normas que las del habla de la cual procede, que puede utilizarse en la búsqueda de información. Con frecuencia se utiliza como sinónimo de Descriptor, pero este uso no es correcto y conviene evitarlo. En MeSH las palabras clave son el equivalente a los términos de entrada (Entry Terms) de un diccionario de idiomas. Sin: Key word.

Palabra de texto: PubMed considera “palabras de texto” a todos los números y términos con significado que aparecen, entre otros menos importantes, en los campos título (title [TI]), resumen (abstract [AB]), descriptores (MeSH terms [MH]), subencabezamientos (MeSH Subheadings [SH]), así como los nombres de sustancias químicas (Substance Name [NM]) y los nombres de personas (Personal Name [PS]) cuando unas y otras son la materia del artículo. Sin: Text Word.

Panel de expertos: Reunión de un grupo de profesionales de prestigio en una determinada materia que abordan, exponen, contrastan ... sus opiniones y experiencias sobre la actuación en un tema preocupante de interés para el que se busca la mayor uniformidad de criterios que no es posible obtener por otro tipo de estudios más eficientes.

Perdidas de seguimiento: Aquellos pacientes que son perdidos durante el estudio y en los cuales no podemos valorar los resultados finales.

Población: Grupo individuos, objetos, mediciones, etc., con determinadas características demográficas, de donde se obtiene la muestra y a la que se quiere hacer inferencia de los resultados de un estudio.

Positivismo lógico: También denominado neopositivismo. Movimiento filosófico contemporáneo que destaca en filosofía la importancia del análisis del lenguaje y de la metodología científica.

Pretest cognitivo: Conjunto de estrategias y procedimientos que se utilizan para determinar si el cuestionario funciona tal y como lo ha concebido el investigador. Debe reproducir exactamente las condiciones en las que se responderá el cuestionario definitivo y, por tanto, tiene que realizarse con el cuestionario terminado y empleando el mismo procedimiento de recogida de información, puesto que la única diferencia entre el pretest y el trabajo de campo es el número de cuestionarios contestados.

Prevalencia: Número de casos de una enfermedad en una población y en un momento dados.

Prevalencia del test: Número total de positivos obtenidos al realizar una prueba diagnóstica en relación al total de sujetos que componen la muestra estudiada.

Probabilidad: Es el cociente entre los casos posibles de que ocurra un hecho dividido por el total de casos.

Probabilidad pretest o preprueba: Proporción de pacientes que tienen la enfermedad en la muestra estudiada, y cuyo diagnóstico es fehacientemente establecido por un estándar de referencia. Es la prevalencia de la enfermedad.

Probabilidad postest o posprueba: Probabilidad de tener una enfermedad tras la realización de una prueba diagnóstica.

PubMed: Base de datos de bibliografía sobre ciencias de la salud. Puede obtenerse de ella información acerca de publicaciones de carácter científico y de contenido médico (en su sentido más amplio, incluyendo enfermería, odontología, veterinaria...).

Randomización: Ver Aleatorización.

Razón de probabilidad: Probabilidad de un resultado en los enfermos dividido por la probabilidad de ese mismo resultado en no enfermos. Sin: Cociente de probabilidad, Likelihood Ratio.

Reacciones adversas a medicamentos (RAM): Según la OMS, “reacción nociva y no deseada que se presenta tras la administración de un fármaco, a dosis utilizadas habitualmente en la especie humana, para prevenir, diagnosticar o tratar una enfermedad, o para modificar cualquier función biológica”. Las reacciones adversas se clasifican en dos grandes grupos: RAM tipo A, de tipo cuantitativo o predecibles, y RAM tipo B, de tipo cualitativo o no predecibles. Sin: Efecto indeseado.

Record: Ver Asiento.

Reducción absoluta del riesgo (RAR): Diferencia aritmética absoluta entre la tasa de acontecimientos (riesgo) en el grupo experimental y la tasa de acontecimientos (riesgo) en el grupo control. También se puede denominar: Diferencia de riesgos.

Reducción relativa del riesgo (RRR): Cociente entre la reducción absoluta del riesgo (RAR) y el riesgo en el grupo control. A la reducción relativa del riesgo se le llama también “fracción atribuible” o “diferencia relativa del riesgo”.

Registro: Ver Asiento.

Regresión logística: Tipo de análisis de regresión (véase) en el que la variable dependiente sólo puede tomar dos valores. En el análisis de regresión la variable dependiente puede ser o bien categórica o bien continua (véase).

Repetibilidad: Referida a una prueba, es el grado en que los resultados son idénticos o estrechamente semejantes cada vez que aquella se lleva a cabo por los mismos profesionales y los mismos emplazamientos.

Reproducibilidad: Grado en que los resultados son idénticos o estrechamente semejantes cada vez que una prueba se lleva a cabo por distintos profesionales y es distintos ámbitos. Sin: Estabilidad.

Retroalimentación: Método para mantener la acción o eficacia de un sistema mediante la continua revisión de los elementos del proceso y de sus resultados con el fin de realizar las modificaciones necesarias.

Reunión educacional: Visita individualizada de un educador a un profesional de la salud en su ámbito de trabajo; en ella se suelen combinar técnicas de auditoría y retroalimentación.

Revisión: Se denomina así a todo intento que sintetice los resultados y las conclusiones de dos o más publicaciones relacionadas con un tema dado.

Revisión bibliográfica: recopilación de literatura realizada sobre un tema concreto en un intervalo de tiempo concreto del que se realiza un resumen y en el que no quedan explícitos los criterios de búsqueda y se destacan las conclusiones a criterio del autor.

Revisiónes narrativas: Se basan en el análisis que hace un experto de los estudios originales existentes sobre un determinado tema o problema para posteriormente subrayar aquellos resultados que considera más importantes.

Revisiónes sistemáticas: Revisión en las que la evidencia sobre un tema ha sido sistemáticamente identificada, criticada y resumida de acuerdo a unos criterios predeterminados.

Riesgo: Resultado predecible o incierto desfavorable de una actividad, intervención o exposición. Probabilidad de que ocurra un determinado fenómeno adverso.

Riesgo relativo (RR): Cociente entre el riesgo de padecer un acontecimiento entre los individuos expuestos a un factor de riesgo determinado y el riesgo de padecerlo entre los no expuestos. También se puede denominar: Cociente de riesgos.

Seguimiento: Periodo de duración de un estudio durante el cual se mantienen contactos periódicos con los participantes en el mismo con el fin de evaluar las variables de estudio.

Sensibilidad: Referida a una prueba diagnóstica, es la proporción de personas verdaderamente enfermas que ha sido catalogadas como tales mediante dicha prueba.

Sensibilidad al cambio (de un instrumento): se refiere a la medida en que detecta y mide de manera fiable las variaciones más pequeñas de un atributo, rasgo, dimensión o constructo.

Serie de casos: Tipo de estudio que describe una serie de características poco conocidas de un proceso o tratamiento y facilitan el conocimiento del mismo y su evolución en el tiempo.

Sesgo: Error sistemático producido en la estimación de los resultados, a causa de defectos en el diseño del estudio, recogida de datos o seguimiento.

Sesgo de desgaste: Diferencias sistemáticas que se producen entre los grupos de un ensayo clínico por la pérdida de participantes durante el estudio.

Sesgo de información: Sesgo relacionado con la forma en la que se obtuvieron los datos. Entre ellos se incluyen los errores cometidos en la valoración del estado de los pacientes al inicio, durante o al final del estudio; también los que ocurren cuando los datos se obtienen de fuentes sin que se conozca la calidad de las mismas. Algunos de los sesgos de información serían: el sesgo del recuerdo o de memoria, sesgo de sospecha diagnóstica, sesgo del entrevistador...

Sesgo de interpretación de las pruebas: Error sistemático que surge debido a la realización de las pruebas de forma no independiente y a ciegas con la consecuencia de que el conocer los resultados de una prueba influye en los de la otra.

Sesgo de publicación: Sesgo que surge debido a la tendencia a publicar sólo estudios con resultados positivos.

Sesgo de selección: Error sistemático consecuencia de que las características de los sujetos incluidos en un estudio son diferentes de las características de los no incluidos, de modo que la muestra no es representativa de la población de referencia. Se incluyen en este grupo el sesgo de los voluntarios, el de prevalencia-incidencia, el de los emigrantes, el de selección no aleatorizada, el de Berkson, el de pertenencia a grupo, el del trabajador sano...

Sesgo de sospecha diagnóstica: Sesgo que ocurre cuando la interpretación de una prueba se haga conociendo el resultado de otra.

Sesgo de reproducibilidad: Sesgo debido a la ausencia de información detallada de las condiciones en que se ha aplicado la prueba, así como las definiciones de los resultados positivos y negativos.

Sesgo de resultados no interpretables: Error sistemático consecuencia de no incluir pacientes en diferentes estadios de la enfermedad.

Sesgo de verificación o de confirmación diagnóstica: Sesgo que ocurre cuando el estudio se limita a individuos sobre los que se realizó con anterioridad la prueba "gold Standard", que suelen ser los de mayor probabilidad de padecer enfermedad.

Sesgo en el espectro de la enfermedad: ¿

Sesgo por ausencia de "gold Standard": Error sistemático que se produce cuando entre los criterios de referencia que se emplean en la comparación de la utilidad de una prueba diagnóstica (en ausencia de "gold Standard") se incluye la prueba de estudio.

Sesión bibliográfica: Herramienta docente consistente en reuniones entre profesionales que realizan una lectura crítica de un artículo científico.

Significación clínica: Probabilidad de que una diferencia observada tenga una repercusión sobre el curso del problema o enfermedad tratados que sea relevante para un paciente dado o para un conjunto de pacientes.

Significación estadística: Probabilidad de que una diferencia observada sea resultado de la casualidad y no de los determinantes causales en un estudio. El hallazgo de una significación estadística no implica necesariamente significación clínica.

Síntesis: Fuentes de información que recogen resúmenes estructurados de revisiones sistemáticas de alta calidad o de artículos originales.

Síntesis: Fuentes de información que incluyen revisiones sistemáticas de alta calidad metodológica.



Sistema: Fuente de información que se caracteriza por abarcar una gran variedad de temas, proporcionar un resumen de los resultados de revisiones de alta calidad y ser actualizado con frecuencia.

Subencabezamiento: Término que delimita el significado del descriptor al que acompaña. P.e. “Drug therapy” (tratamiento farmacológico) como subencabezamiento unido al descriptor Hypertension nos informaría que un artículo trata, entre todos los aspectos posibles de la hipertensión, su tratamiento farmacológico. Sin: Subheading.

Subheading: Ver subencabezamiento.

Tag: Ver Etiqueta.

Tamaño de la muestra: Número mínimo necesario de sujetos que participan en un estudio, pero a su vez suficiente para proporcionar una respuesta fiable al objetivo final de dicho estudio. La determinación de dicho tamaño mínimo es de naturaleza estadística y depende de varios factores.

Técnica de grupos conocidos: Procedimiento utilizado para la validación de constructo según la cual el instrumento se aplica a los grupos que, según se presume, diferirán respecto del atributo crítico a causa de una característica conocida.

Tecnología sanitaria: Conjunto de medicamentos, aparatos, procedimientos médicos y quirúrgicos usados en la atención médicos y los sistemas organizativos y de soporte con los que se presta la atención sanitaria.

Teoría estadística de Bayes o Teorema de Bayes: Establecimiento inductiva de la probabilidad de una hipótesis a partir de algunas observaciones. Según el teorema de Bayes, la probabilidad posterior (o a posteriori) de una hipótesis es igual a una constante, multiplicada por la verosimilitud o posibilidad de la hipótesis (definida como la probabilidad de observar unos acontecimientos determinados si la hipótesis considerada fuera de hecho cierta) y por la probabilidad anterior de la hipótesis.

Término MeSH: Ver MeSH Term

Término de entrada: Relacionado con Palabra clave (véase).

Tesaurus: Es un repertorio de términos con expresión de las relaciones semánticas que los unen. Habitualmente consta de dos partes, una sección categorizada, que expresa las relaciones jerárquicas, genérico-específicas, entre los términos normalizados (descriptores), lo que nos permite conocer el campo semántico (significado) que abarcan los distintos términos. La sección alfabética es una ordenación de todos los términos, normalizados (descriptores) o no normalizados (palabras clave), que de forma similar a un diccionario de idiomas, permite conocer las relaciones de equivalencia entre unos y otros términos. Sin: Thesaurus.

Test Chi cuadrado (2): Prueba estadística utilizada para determinar el grado de asociación entre variables categóricas (véase).

Test de McNemar: Prueba de significación estadística para probar la hipótesis nula de inexistencia de cambios en la proporción de sujetos que experimentan un acontecimiento, cuando cada individuo es evaluado dos veces (en condiciones diferentes) y los datos están emparejados.

Test T de Student: Prueba de significación estadística paramétrica (véase) para contrastar la hipótesis nula respecto a la diferencia entre dos medias.

Test U de Mann-Whitney: Prueba de significación estadística no paramétrica (véase) para probar la hipótesis nula de que el parámetro de localización (generalmente la mediana) es el mismo cuando se comparan dos grupos independientes, cualquiera que sea el tipo de distribución de la variable (distribución normal o de otro tipo).

Test de Wilcoxon: El contraste de Wilcoxon es la técnica no paramétrica paralela al test T de Student para muestras apareadas.

Text Word: Ver Palabra de texto.

Thesaurus: Ver Tesoro.

Transferabilidad: Término acuñado en investigación cualitativa sugiere que en la medida en que el investigador ha respetado el criterio de maximizar la diversidad en la obtención de la muestra y realice una detallada descripción del contexto y participantes, los hallazgos pueden ser aplicables en contextos similares.

Triangulación: Técnica empleada en investigación cualitativa para controlar sesgos. Consiste en obtener distintas perspectivas del fenómeno estudiado utilizando diferentes métodos de investigación (cualitativos y cuantitativos). La triangulación puede ser simultánea o secuencial.

Truncado: El término truncar significa en español "cortar el extremo de una cosa". Aplicado a un término de búsqueda consiste en separar del término su terminación y sustituirla por un símbolo comodín (en PubMed es un asterisco *-). El resultado es la selección de todos los términos que contienen la raíz superviviente.

Umbral de decisión: Nivel de probabilidad a partir del cual el beneficio previsible de toma una decisión de intervenir iguala o mejora al beneficio de no tomarla

Validez: Capacidad de un instrumento para medir aquello para lo que ha sido diseñado, denota el grado en que el instrumento mide lo que se supone que debe medir.

Validez aparente o facial: Forma especial de la validez de contenido que se refiere al hecho de que el instrumento parezca, de la impresión a las personas evaluadas, de que es adecuado para medir lo que se pretende medir.

Validez concurrente: Denota la capacidad de un instrumento de distinguir a los individuos que difieren en su condición actual respecto de un criterio.

Validez convergente: Aspecto de la validez de constructo que se refiere al grado en que diferentes formas de medir el mismo constructo (o dimensiones de un constructo) se correlacionan la una con la otra.

Validez de constructo: Grado en que una prueba o instrumento mide un determinado rasgo, característica o constructo.

Validez de contenido: Determina si el instrumento contiene ítems representativos de todas las dimensiones que forman la definición del concepto o constructo en estudio.

Validez de criterio: Concordancia entre los resultados de la medida del instrumento y los resultados de la medida realizada mediante un criterio o patrón de referencia (gold standard) de lo que se quiere medir.

Validez discriminante o divergente: Aspecto de la validez de constructo que demuestra que una medida no se correlaciona con otras medidas cuya finalidad es medir rasgos, dimensiones o constructos diferentes a los que ella mide.

Validez externa: Grado en que las conclusiones obtenidas con la muestra de población que participa en un estudio pueden ser generalizadas a su población de referencia o a otras poblaciones, lugares, momentos e investigaciones.

Validez interna: Grado en que los resultados de un estudio miden la realidad para la muestra de población que participa en un estudio. Grado en que un estudio está libre de sesgos.

Validez predictiva: Se refiere a la idoneidad de un instrumento para diferenciar entre el desempeño o las conductas de los individuos respecto de un criterio futuro.

Valor global de un test: Proporción de sujetos correctamente clasificados por el test en relación al total de sujetos que componen la muestra estudiada.

Valor predictivo negativo: Referido a una prueba diagnóstica, la probabilidad de que una persona con un resultado negativo sea un verdadero negativo.

Valor predictivo positivo: Referido a una prueba diagnóstica, la probabilidad de que una persona con un resultado positivo sea un verdadero positivo.

Variabilidad: Variación entre individuos debido a determinantes genéticos, ambientales, patológicos... (variabilidad interindividual) u oscilación de variables biológicas en un mismo individuo según las circunstancias (variabilidad intraindividual).

Variable: Cualquier atributo, fenómeno o hecho que pueda tener diferentes valores.

Variable categórica: Se dice de la variable que presenta interrupciones. Una variable es categórica si, entre varios valores potencialmente observables, hay un valor que no puede ser observado. Ejemplo: variables binarias.

Variable continua: Se dice de la variable que no presenta interrupciones. Una variable es continua si, entre dos valores observables, siempre existe la posibilidad de que haya otro valor observable. Ejemplos: edad, peso, altura.

Verbatim: Su significado es: Literalmente, o palabra por palabra. Anglicismo utilizado en investigación cualitativa para denominar a la transcripción exacta, en la fase de exposición de resultados, de las expresiones surgidas durante el grupo de discusión de los participantes, es decir conservado la expresión literal grabada.

Verdaderos negativos: Grupo de pacientes sanos en los que el test es negativo.

Verdaderos positivos: Grupo de pacientes con enfermedad y que además son detectados por el test.

