

Hospital de Pediatría  
**Garrahan**

## ***Guía para la lectura crítica de artículos científicos***

- 1 - Estudios sobre tratamientos o intervenciones ..... pág 1***
- 2 - Estudios sobre métodos diagnósticos ..... pág 5***
- 3 - Estudios sobre factores de riesgo ..... pág 9***
- 4 - Estudios sobre factores pronósticos ..... pág 13***
- 5 - Estudios sobre meta-análisis ..... pág 17***

***Dirección Asociada de Docencia e Investigación***

***Dr. Hernán Rowensztein - Coordinador de Formación de RR.HH.***

**PLANILLA PARA LECTURA CRÍTICA: ESTUDIOS SOBRE TRATAMIENTO O INTERVENCIÓN**

**Validez interna:** ¿Son válidos los resultados del estudio?

<b>Criterios principales:</b>	
<b>1) ¿Fue la asignación de los pacientes al tratamiento aleatoria?</b>	
<b>¿Qué es mejor?</b>	<b>¿Dónde encuentro esta información?</b>
<p>El ensayo clínico aleatorizado es un diseño de estudio en el que los sujetos son asignados por azar al menos a dos grupos, uno (grupo experimental) recibe el tratamiento que se está probando, y el otro (grupo control) recibe un tratamiento alternativo. La aleatorización es la distribución que más garantía otorga que los grupos sean similares en los aspectos que conocemos pero también en los desconocidos, para esto, cada sujeto debe tener la misma chance de ser asignado a cualquiera de los dos grupos. El ensayo clínico aleatorizado (ECA) es el tipo de estudio con mayor grado de evidencia científica para estudios de intervención, establece la mejor evidencia causa-efecto y evita sesgos por confundidores.</p>	<p>Ya desde el <b>Título</b> del trabajo puede indicarse si la asignación fue al azar. La declaración CONSORT establece que en el título debe consignarse si fue randomizado. Luego en <b>Métodos</b> debería decir cómo fue la metodología de la aleatorización, si ésta fue oculta o no y detalles de cualquier restricción (como el tamaño de los bloques si los hay).</p>
<p>Este artículo: Si <input type="radio"/> No <input type="radio"/> No está claro <input type="radio"/></p> <p>Comentarios:</p>	
<b>2) ¿Se tomaron en cuenta a todos los pacientes que ingresaron al ensayo? ¿Fueron analizados en el grupo al cual fueron aleatorizados?</b>	
<b>¿Qué es lo mejor?</b>	<b>¿Dónde encuentro esta información?</b>
<p>Las pérdidas durante el seguimiento deberían ser mínimas -de preferencia menos del 20%- . Sin embargo, si pocos pacientes presentan el desenlace de interés, entonces incluso pequeñas pérdidas pueden sesgar los resultados. Una forma de comprobar si la pérdida es despreciable o no, es el <b>análisis del peor escenario</b> que consiste en asumir que todos los pacientes perdidos tuvieron mala evolución y volver a comparar los grupos para ver si se mantienen las mismas conclusiones.</p> <p>Por otra parte, los pacientes deberían ser analizados en el grupo al que fueron asignados, lo cual se denomina “análisis por intención de tratar”, es decir que si fueron asignados al grupo experimental pero por algún motivo no cumplieron con el tratamiento, igual serán analizados en ese grupo como forma de respetar la aleatorización.</p>	<p>En la sección <b>Métodos</b> debe decir cuáles fueron los criterios de inclusión, exclusión y de eliminación. Estos últimos son los pacientes que ingresaron al estudio, fueron asignados a algún grupo y luego eliminados en el transcurso del ensayo clínico, siendo fundamental que se justifique el por qué.</p> <p>La sección de <b>Resultados</b> debe decir cuántos pacientes fueron aleatorizados (por ejemplo, tabla de características basales) y cuantos pacientes fueron realmente incluidos en el análisis. Es necesario leer la sección de resultados para clarificar la razón y el número de pérdidas.</p>
<p>Este artículo: Si <input type="radio"/> No <input type="radio"/> No está claro <input type="radio"/></p> <p>Comentarios:</p>	

**Criterios secundarios:**

**1) ¿Fueron las mediciones objetivas o se mantuvieron “ciegos” al tratamiento que estaban recibiendo a los pacientes y tratantes?**

¿Qué es lo mejor?

¿Dónde encuentro esta información?

Lo ideal es que el estudio sea al menos “doble ciego” - esto es, tanto pacientes como investigadores no saben a qué tratamiento fueron asignados-. Si el desenlace es objetivo (por ejemplo, muerte) entonces el cegamiento es menos crítico. Si el desenlace es subjetivo (por ejemplo síntomas o función) entonces el cegamiento del evaluador del desenlace es crítico. No siempre se puede cegar la intervención (una cirugía, kinesiología, etc.), esto es posible cuando el programa de estudio consiste en administrar una droga y el de control es una medicación alternativa. En todos los casos en los que no se ha hecho la asignación ciega, se debe contemplar la posibilidad de error o sesgo.

Primero, mire en **Título** si se menciona el cegamiento de la intervención. Segundo, en **Métodos** se debería describir cómo se evalúa el desenlace y si el evaluador está consciente del tratamiento recibido por el paciente, por ejemplo: placebo o tratamiento convencional con la misma apariencia, sabor y olor que la droga experimental. Además del desconocimiento de los médicos tratantes acerca de la asignación a los grupos.

Este artículo: Si  No  No está claro   
Comentarios:

**2) ¿Fueron los grupos similares al comienzo del ensayo?**

¿Qué es lo mejor?

¿Dónde encuentro esta información?

Si el proceso de aleatorización funcionó (esto es, logró grupos comparables) los grupos deberían ser similares en cuanto a sus características clínicas basales, entre más similares sean es mejor. Un concepto importante es que aquí no importa realizar un análisis estadístico y calcular la p. Dado que el diseño del estudio no está hecho para encontrar diferencias aquí, lo importante es que observe las variables presentadas y determine si le parece que estas diferencias podrían afectar los resultados.

En **Resultados** debería haber una tabla de “Características basales” (“*Baseline Characteristics*”) comparando los grupos aleatorizados en un número de variables que puedan afectar el desenlace de interés (es decir, edad, factores de riesgo, etc.). Si no, debería haber una descripción de las similitudes de los grupos en el primer párrafo de la sección de **Resultados**.

Este artículo: Si  No  No está claro   
Comentarios:

**3) Además del tratamiento asignado ¿Los grupos fueron tratados igualmente?**

¿Qué es lo mejor?

¿Dónde encuentro esta información?

Aparte de la intervención, los diferentes grupos deben ser tratados de la misma manera, por ejemplo, tratamientos adicionales o evaluaciones. Se refiere a las intervenciones recibidas aparte del tratamiento asignado en la aleatorización, que se denominan “cointervenciones”.

Mire en la sección de **Métodos** el horario o esquema de seguimiento y tratamientos adicionales permitidos, etc. y en **Resultados** su uso real. Por ejemplo si detalla la administración de antitérmicos, analgésicos, etc. en forma discrecional entre los grupos.

Este artículo: Si  No  No está claro   
Comentarios:

**Resultados:** ¿Cuál es la magnitud del efecto del tratamiento?

1) ¿Qué tan grande es el efecto del tratamiento?	
<p>Frecuentemente los resultados son presentados como desenlaces dicotómicos (desenlaces que ocurren o no ocurren) como por ejemplo recurrencia del cáncer, curación o muerte. Considere un estudio en el cual a los 2 años de seguimiento murió el 10% (0,10) del grupo tratamiento y el 15% (0,15) del grupo control. Los resultados se pueden expresar de muchas maneras como lo verá más abajo.</p>	
¿Cuál es la medida?	¿Qué significa?
<p><b>Riesgo Relativo (RR)</b> = riesgo del desenlace en el grupo tratamiento / riesgo del desenlace en el grupo control. En nuestro ejemplo, el <math>RR=0,10/0,15=0,67</math>. Es decir que el riesgo de morir es de 0,67 pacientes en el grupo tratado por cada 1 muerto en el grupo control.</p>	<p>El riesgo relativo nos dice <b>cuántas veces más probable</b> es que un evento ocurra en el grupo tratamiento en relación al grupo control. Un <b>RR de 1</b> significa que no hay diferencia entre los dos grupos, así, el tratamiento <b>no tiene efecto</b>. Un <math>RR &lt; 1</math> significa que el tratamiento disminuye el riesgo del desenlace o punto final estudiado. Un <math>RR &gt; 1</math> significa que el tratamiento aumenta el riesgo.</p>
<p><b>Reducción Relativa del Riesgo (RRR)</b>= reducción absoluta del riesgo/ riesgo del desenlace en el grupo control. Una manera alternativa para calcular la RRR es restar el RR a 1 (o sea, <math>RRR=1-RR</math>). En nuestro ejemplo, <math>RRR = 0,05/0,15 = 0,33</math> ó 33% O <math>RRR = 1 - 0,67 = 0,33</math> ó 33%</p>	<p>La reducción relativa del riesgo es el complemento del RR y probablemente es la medida de efecto del tratamiento más comúnmente reportada. Nos dice cuál fue la reducción en la tasa del desenlace en el grupo tratamiento en relación al del grupo control. El tratamiento reduce el riesgo de muerte en un 33% en relación al riesgo de ocurrencia en el grupo control</p>
<p><b>Reducción Absoluta del Riesgo (RAR)</b> = Riesgo del desenlace en grupo control – riesgo del desenlace en el grupo tratamiento. También se conoce como <b>diferencia absoluta de riesgo</b>. En nuestro ejemplo, la <math>RAR = 0,15-0,10 = 0,05</math> ó 5%</p>	<p>La RAR nos dice sobre la diferencia en la tasa de eventos entre los dos grupos, se suele expresar en porcentaje. Tiene en cuenta el riesgo basal del evento de interés sin la intervención. Un <b>RAR de 0</b> significa que no hay diferencia entre los dos grupos, así, el tratamiento <b>no tiene efecto</b>. El beneficio absoluto del tratamiento es un 5% de reducción en la tasa de mortalidad.</p>
<p><b>Número Necesario a Tratar (NNT)</b> = inverso de la RAR y se calcula = <math>1/RAR</math> En nuestro ejemplo, el <math>NNT = 1/ 0,05 = 20</math></p>	<p>El número necesario a tratar representa el número de pacientes que necesitamos tratar con una terapia experimental para prevenir 1 mal desenlace o curar 1 paciente. La significancia clínica puede determinarse en cierta manera mirando los NNT, pero también sopesando los NNT en relación a daños o efectos adversos (NND) de la terapia. En nuestro ejemplo necesitaríamos tratar a 20 personas por 2 años para prevenir 1 muerte.</p>
2) ¿Qué tan precisa es la estimación del efecto del tratamiento?	
<p>El riesgo verdadero del desenlace en la población es desconocido y lo mejor que se puede hacer es estimar el verdadero riesgo en base a la muestra de pacientes del ensayo. Este estimador se denomina <b>estimador puntual</b>. Podemos evaluar qué tan cerca está este estimador del valor real mirando el intervalo de confianza (IC) para cada estimador. Si el intervalo de confianza es bastante estrecho entonces podemos estar seguros que nuestro estimador puntual es un reflejo preciso del valor poblacional. El intervalo de confianza también nos da información sobre la significancia estadística del resultado. Si el valor correspondiente a <b>no hay efecto</b> (1 para el RR) cae fuera del intervalo de confianza del 95% entonces el resultado es estadísticamente significativo a un nivel de 0,05. Si el intervalo de confianza incluye el valor correspondiente a <b>no hay efecto</b> entonces los resultados no son estadísticamente significativos.</p>	



¿Cuál es la medida?	¿Qué significa?
<b>Valor de p</b>	“p” es la probabilidad de tener el resultado obtenido o uno más alejado, dado que la hipótesis nula es verdadera. El proceso de aceptación o rechazo de la hipótesis lleva implícito un riesgo de cometer un error que se cuantifica con el valor de la “p”, que es la probabilidad de aceptar la hipótesis alternativa como cierta, cuando la cierta podría ser la hipótesis nula. El valor de “p” que indica que la asociación es estadísticamente significativa ha sido arbitrariamente seleccionado y por consenso se considera en 0.05 en la mayoría de los casos. Una confiabilidad del 95% lleva implícito una $p < 0.05$ . Cuando rechazamos la $H_0$ (hipótesis nula) y aceptamos la $H_a$ (hipótesis alternativa) como probablemente cierta afirmando que hay una asociación, o que hay diferencia, estamos diciendo en otras palabras que es muy poco probable que el azar fuese responsable de dicha asociación.
<b>Intervalo de confianza 95% (IC 95%)</b>	Cuando precisamos conocer el valor de una variable en una población, debemos estimarla a partir de los datos obtenidos sólo de una muestra de esa población. Los Intervalos de confianza (IC) nos permiten estimar, una vez calculado el valor de la variable en la muestra, entre que <b>rango de valores</b> se encuentra el valor real inaccesible de la población a partir del que podemos obtener de nuestra muestra, con una probabilidad de equivocarnos del 5%. Además el IC nos informa sobre la precisión de nuestra estimación, ya que por ejemplo, no será lo mismo estimar que el IC de la media de glucemia en una población sea de 90 a 110 mg/dl que de 78 a 128 mg/dl. Dicho de otra forma, un IC 95% significa que si repitiéramos 100 veces el estudio, en 95 ocasiones el valor de la media lo encontraremos entre los extremos del intervalo de confianza.
Este artículo: Si <input type="radio"/> No <input type="radio"/> No está claro <input type="radio"/> Comentarios:	

**Validez externa:** ¿Son aplicables los resultados a mis pacientes?

¿Qué tan precisa es la estimación del efecto del tratamiento?
<p>Las preguntas que debe hacerse antes que decida aplicar los resultados de un estudio en su paciente son:</p> <p><b>1) ¿Es mi paciente tan diferente de los del estudio que los resultados no son aplicables a él?</b> Es importante considerar si los pacientes del estudio son comparables a los que se beneficiarán con la aplicación de estos nuevos resultados, es decir mis pacientes. A veces sucede, por ejemplo, que un estudio da resultados alentadores en adultos y luego el tratamiento es utilizado en niños, a pesar de las diferencias fisiológicas y farmacocinéticas.</p> <p><b>2) ¿Se consideraron todas las variables de evaluación del resultado clínicamente importantes?</b> A veces los estudios evalúan variables que no son tan relevantes para los sujetos. Por ejemplo la mejoría de índices de laboratorio, estudios complementarios, etc. En cambio es importante evaluar el impacto clínico que repercute directamente en el niño, la calidad de vida, curación, etc.</p> <p><b>3) ¿Los posibles beneficios del tratamiento sobrepasan los potenciales daños para mi paciente?</b> Aquí cada usuario tendrá que sacar sus propias conclusiones en cuanto la nueva intervención es lo suficientemente buena como para aplicarla a sus pacientes, teniendo en cuenta los costos, efectos adversos, disponibilidad, etc.</p>
Este artículo: Si <input type="radio"/> No <input type="radio"/> No está claro <input type="radio"/> Comentarios:

**PLANILLA PARA LECTURA CRÍTICA: ESTUDIOS SOBRE MÉTODOS DIAGNÓSTICOS**

**Validez interna:** ¿Son válidos los resultados del estudio?

<b>Criterios principales:</b>	
<b>1) ¿Hubo una comparación independiente, ciega entre la prueba diagnóstica y un (“gold”) estándar de referencia apropiado de diagnóstico?</b>	
¿Qué es mejor?	¿Dónde encuentro esta información?
Hay dos temas importantes aquí. Primero que el estándar de referencia debería ser apropiado – lo más cercano a la verdad posible-. Algunas veces puede que no haya un solo test de referencia apropiado y se puede usar una combinación de pruebas para determinar la presencia de enfermedad. Segundo, el estándar de referencia y la prueba diagnóstica que están siendo evaluados deberían aplicarse de manera independiente y ciega. Aquel personal que interpreta los resultados de un test no debería saber de los resultados del otro test.	En la sección de <b>Métodos</b> debería haber una descripción del estándar de referencia usado y si no está seguro respecto a que el estándar de referencia sea el adecuado, puede que sea necesario que realice una búsqueda más profunda en el área. La sección de <b>Métodos</b> también debería describir quien realizó los dos test y si cada uno fue conducido de manera independiente (los 2 test en todos los pacientes) y ciega a los resultados del otro (un resultado no influye al otro).
Este artículo: Si <input type="radio"/> No <input type="radio"/> No está claro <input type="radio"/> Comentarios:	
<b>2) ¿El test diagnóstico fue evaluado en un espectro representativo de pacientes (como aquellos en los cuales se les aplicará en la práctica)?</b>	
¿Qué es mejor?	¿Dónde encuentro esta información?
Es ideal si el test diagnóstico es aplicado al espectro completo de pacientes – es decir, leves, severos, etapa temprana y tardía del desorden de interés. Lo ideal es que el ámbito de estudio sea similar al de aplicación (esto evita el sesgo de espectro). Recuerde que la prevalencia de la enfermedad modifica la performance el test.	La sección de <b>Métodos</b> debería decir cómo fueron reclutados los pacientes y si fueron seleccionados al azar o a medida que fueron admitidos. También debería decir de dónde vienen los pacientes y si es probable que sean representativos de los pacientes en quienes se usará el test.
Este artículo: Si <input type="radio"/> No <input type="radio"/> No está claro <input type="radio"/> Comentarios:	
<b>Criterios secundarios:</b>	
<b>1) ¿Los resultados de la prueba en estudio influyeron en la decisión de realizar el método de referencia?</b>	
¿Qué es lo mejor?	¿Dónde encuentro esta información?
Idealmente tanto la prueba diagnóstica y el estándar de referencia deberían ser aplicados en todos los pacientes del estudio. En algunas situaciones donde el estándar de referencia es invasivo o caro podría haber reparos respecto a someter a la prueba a sujetos con la nueva prueba diagnóstica negativa. Si el gold estándar se aplica solo a los (+) se produce un <b>sesgo de verificación</b> , que sobrestima la capacidad diagnóstica del test. Una alternativa al estándar de referencia es seguir a las personas por un tiempo apropiado (dependiendo de la enfermedad en cuestión) para ver si son verdaderos negativos.	La sección de <b>Métodos</b> debería indicar si el estándar de referencia fue aplicado a todos los pacientes o si un estándar de referencia alternativo (es decir, seguimiento) fue aplicado a aquellos que fueron negativos en la prueba diagnóstica. En <b>Resultados</b> también puede que conste que los pacientes dieron positivo o negativo al test de referencia o en función del seguimiento (personal, telefónico, etc.) se pudo inferir que no presentaron la enfermedad.



Este artículo: Si  No  No está claro   
Comentarios:

**2) ¿Se describieron los métodos para hacer el estudio con suficiente detalle como para ser reproducible?**

¿Qué es lo mejor?

El artículo debe describir con detalle cómo se realizaron las determinaciones diagnósticas. Que estén todos los elementos para que el estudio sea replicable es fundamental. Por ejemplo en un estudio sobre la utilidad de la ecografía en el diagnóstico de una cardiopatía, deben constar el modelo del equipo, el transductor utilizado, la posición del paciente, el entrenamiento del personal que realiza el estudio, etc.

¿Dónde encuentro esta información?

En el apartado de **Materiales y Métodos** del artículo deben estar detallados todos los aspectos técnicos del estudio y el equipamiento.

Este artículo: Si  No  No está claro   
Comentarios:

**Resultados:** ¿Se presentan las características del test?

**¿Son importantes los resultados del estudio?**

Los médicos tenemos una “probabilidad pre-test” acerca de si el paciente está enfermo o no. Los métodos diagnósticos no son infalibles ni perfectos, y eso lo sabemos bien; por lo que normalmente no confirmamos ni descartamos enfermedad tan solo con un resultado, lo que hacemos es cambiar las probabilidades pre-test del paciente a una “probabilidad post-test” de acuerdo a si el resultado es positivo o negativo, lo que simplemente lo acercará o alejará de la posibilidad de estar enfermo. La magnitud de ese cambio dependerá en gran medida de la “capacidad operativa” del test, no es lo mismo una prueba de PPD negativa en un paciente con sospecha de tuberculosis que un cultivo de esputo definitivamente negativo.

		ENFERMO	SANO	
Resultado de la prueba diagnóstica que evaluamos	POSITIVO	VP a	FP b	Total positivos a+b
	NEGATIVO	FN c	VN d	Total negativos c+d
		Total de enfermos a+c	Total de sanos b+d	TOTAL INDIVIDUOS a+b+c+d

¿Cuál es la medida?

**Sensibilidad (S)** o tasa de verdaderos positivos (TVP):  
**Sensibilidad = Verdaderos Positivos / Enfermos = VP / (VP + FN)**

¿Qué significa?

La sensibilidad de un método diagnóstico es la capacidad de identificar a los pacientes enfermos. Indica la proporción de pacientes enfermos a los que el test les dio positivo, y se expresa como porcentaje (0-100%) o fracción (de 0 a 1). Una prueba con alta sensibilidad tiene pocos falsos negativos (pacientes enfermos a los que la prueba no detectó).



<p><b>Especificidad (E)</b> o Tasa de Verdaderos Negativos (TVN):  <b>Especificidad= Verdaderos Negativos/ Sanos= VN/ (VN + FP)</b></p>	<p>La especificidad es la capacidad de un método diagnóstico de reconocer a las personas sanas.          Una especificidad alta evita falsos positivos (pacientes sanos a los que el test cataloga como positivos).</p>
<p><b>Poder Predictivo Positivo (PPP)</b> o Valor Predictivo Positivo (VPP):  <b>PPP= Verdaderos positivos/ Positivos= VP/ (VP + FP)</b></p>	<p>El poder predictivo positivo mide la probabilidad de estar enfermo si la prueba es positiva. Es la proporción de pruebas positivas que son verdaderas o corresponden a pacientes enfermos.          Una prueba con alto poder predictivo positivo tiene pocos falsos positivos (resultados positivos equivocados).</p>
<p><b>Poder Predictivo Negativo (PPN)</b> o Valor Predictivo Negativo (VPN):  <b>PPN: Verdaderos Negativos/ Negativos= VN/ (VN + FN)</b></p>	<p>El poder predictivo negativo mide la probabilidad que tiene un paciente de estar sano si la prueba le dio negativa. Es la proporción de resultados negativos verdaderos o correspondientes a pacientes sanos. Una prueba con alto poder predictivo negativo tiene pocos falsos negativos (resultados negativos erróneos).</p>
<p>Existe otra medida de capacidad operativa que combina toda la información de la tabla y que es aplicable al caso particular: el <b>cociente de probabilidades</b> o razón de probabilidades.  <b>LR+ : tasa de Verdaderos Positivos/ tasa de Falsos Positivos = tVP/ tFP = S/ 1-E</b>  <b>LR - : tasa de Falsos Negativos/ tasa de Verdaderos Negativos = tFN/ tVN = 1-S/ E</b></p>	<p>Razón de probabilidades o coeficiente de verosimilitudes (o "Likelihood Ratio": LR): Expresa la chance de estar enfermo cuando la prueba es positiva (LR positivo) o sano cuando es negativa (LR negativo). Compara cuántas veces la prueba acierta contra las veces que se equivoca. Un LR= 1 significa que las chances de un paciente de estar enfermo o sano son iguales (la prueba no ayuda al diagnóstico para la patología en estudio). Para que un método diagnóstico tenga utilidad se espera que su LR+ sea alto (idealmente &gt; 10) y su LR- sea bajo (&lt; 0,1).</p>

**Resultados no dicotómicos - Curvas ROC**

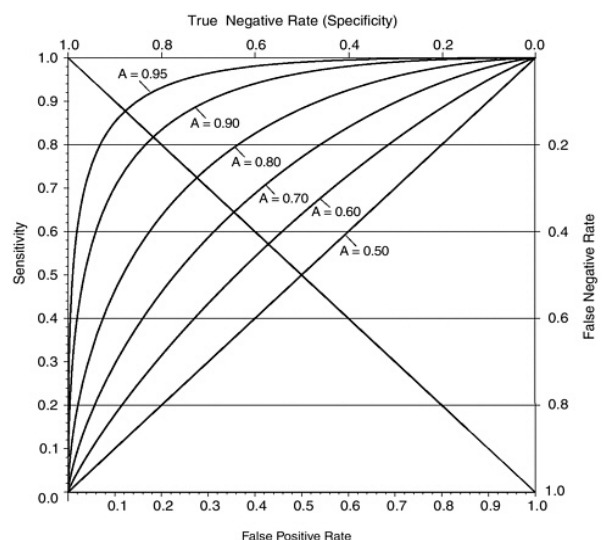
Las curvas ROC o Receiver Operator Characteristics se utilizan para los estudios donde el resultado no es dicotómico (positivo o negativo), sino medidas numéricas continuas.

Se grafican los diferentes puntos de corte de un método relacionando la Tasa de verdaderos positivos o sensibilidad (en las ordenadas) y la tasa de falsos positivos o 1 – Especificidad (en las abscisas). Esto permite no sólo comparar puntos de corte de un método determinado sino también comparar dos métodos diferentes.

- **Cuanto mayor es la sensibilidad menor es el número de falsos negativos.**
- **Cuanto mayor es la especificidad menor es el número de falsos positivos.**

El mejor punto de corte es el del ángulo superior izquierdo, porque balancea la sensibilidad y especificidad.

Tener en cuenta que si el test es de screening conviene máxima sensibilidad, y si es de confirmación conviene máxima especificidad.



Este artículo: Si  No  No está claro   
 Comentarios:



**Validez externa: Aplicabilidad.**

**¿Son aplicables los resultados a mis pacientes?**

**1) ¿Son aplicables los resultados a mis pacientes?**

En este caso se debe verificar si las poblaciones donde se realizó la prueba y donde se utilizará son similares como para presumir que ésta se comportará de manera similar. Para esto sirve controlar los criterios de inclusión, de exclusión y el ámbito de realización del estudio (atención primaria, internación, etc.).

**2) ¿Está la prueba disponible, es accesible, reproducible y fácilmente interpretable?**

Aquí se debe evaluar la viabilidad de realizar la prueba en su medio de trabajo, es decir si no es excesivamente costosa, riesgosa e inclusive objetable por los pacientes.

**3) ¿Los resultados modificarán mi tratamiento?**

Los resultados de las pruebas diagnósticas serán de utilidad especialmente si modifican conductas médicas. Es decir cuando hay una alta presunción clínica de enfermedad (alta probabilidad pre-test) deberá tener un muy alto LR(-) para descartar enfermedad y viceversa. Por lo tanto es fundamental conocer los cocientes de probabilidad de los test diagnósticos antes de decidir realizarlos, en especial en los casos que estos tengan un alto costo económico y/o emocional para los pacientes y el sistema de salud.

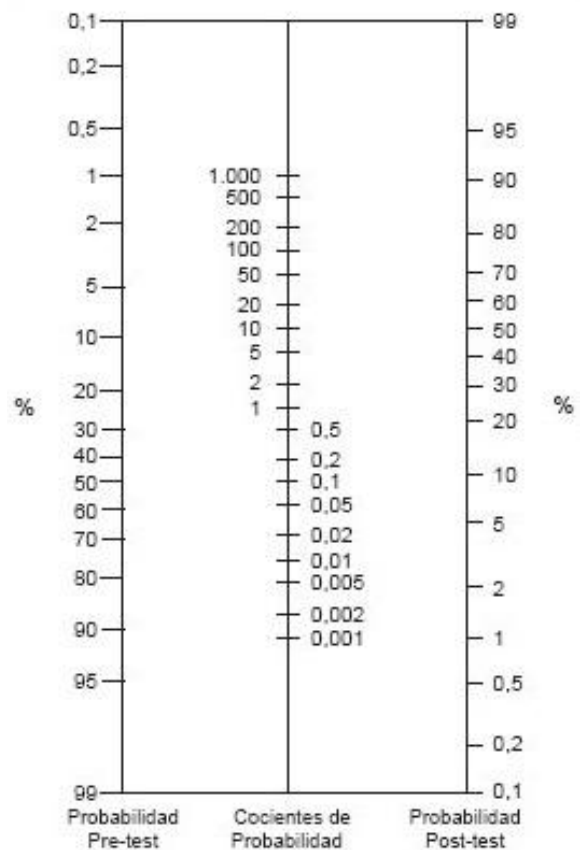
**4) ¿Obtendrán beneficio los pacientes como consecuencia del examen?**

Para contestar este punto, será necesario medir el impacto de la realización de una prueba diagnóstica como si esta fuera una intervención, es decir medir el impacto de su uso sobre la salud de los pacientes.

Nomograma de Fagan:

Fagan en 1975 describió un nomograma para el teorema de Bayes, basado en la capacidad de convertir el teorema de Bayes en una función sumatoria lineal simple. El nomograma de Fagan tiene 3 columnas: la primera es la probabilidad de tener la enfermedad antes de aplicar la prueba (prevalencia), la segunda es el cociente de probabilidad (LR) y la tercera la probabilidad post-test.

Con una regla se traza una línea entre la probabilidad pre-test y el cociente de probabilidad. La prolongación de esta línea corta en la tercera columna la probabilidad de tener la enfermedad en función del resultado de la prueba.



**PLANILLA PARA LECTURA CRÍTICA: ESTUDIOS SOBRE FACTORES DE RIESGO**

**Validez interna:** ¿Son válidos los resultados del estudio?

<b>Criterios principales:</b>	
<b>1) ¿Se han utilizado grupos de comparación similares, en cuanto a factores importantes determinantes del resultado, aparte del que se investiga?</b>	
¿Qué es mejor?	¿Dónde encuentro esta información?
El ensayo clínico aleatorizado sería el estudio ideal para que los grupos sean similares. Pero para estudiar factores de riesgo (FR) suele ser imposible este diseño, especialmente por razones éticas. Por lo tanto el seguimiento de cohortes y los estudios de casos y controles son los más utilizados. O sea que todo diseño debería contemplar que los grupos de comparación o la cohorte sean de sujetos similares en todo menos en los factores de estudio. En una cohorte se determina la aparición de un evento específico en un grupo de individuos inicialmente libres de la enfermedad en estudio, y se analiza la exposición o no a ciertos factores de riesgo. En el estudio de casos y controles en cambio, se parte de seleccionar individuos que hayan presentado el evento y se revisa hacia el pasado si éstos fueron expuestos o no al FR, y se compara con un grupo control.	En <b>Métodos</b> debería decir cómo fueron asignados a los grupos los pacientes o si fue una cohorte. En general no son estudios experimentales sino observacionales. También en el apartado sobre <b>Diseño</b> puede que explique directamente si es un estudio de cohorte (prospectiva o retrospectiva) o si es del tipo Casos y Controles.
Este artículo: Si <input type="radio"/> No <input type="radio"/> No está claro <input type="radio"/> Comentarios:	
<b>2) ¿Se han evaluado las exposiciones y los resultados de la misma forma en los grupos que se comparan?</b>	
¿Qué es lo mejor?	¿Dónde encuentro esta información?
Aquí es importante prever la aparición de sesgos en el estudio de la exposición a los factores de riesgo en los grupos comparados. Por ejemplo, si el entrevistador sabe que está frente a un enfermo, puede preguntar con más énfasis acerca de la exposición al FR que si está frente a una persona sana (sesgo del entrevistador); lo mismo que una persona enferma tiende a recordar mejor el pasado relacionado a la enfermedad que padece (sesgo del recuerdo). Tampoco se pueden controlar más de cerca a los portadores de los FR ya que podríamos verificar la presencia de enfermedad con mucha más anticipación que en el otro grupo (sesgo de vigilancia), o tender a hacer un diagnóstico positivo con menos elementos (sesgo de diagnóstico).	La sección de <b>Métodos</b> debería decir que todos los participantes son controlados y entrevistados de la misma forma. El diseño de un estudio ciego, tanto para los sujetos de estudio como para los investigadores, ayudará a disminuir estos sesgos.
Este artículo: Si <input type="radio"/> No <input type="radio"/> No está claro <input type="radio"/> Comentarios:	

3) ¿Ha sido el seguimiento lo suficientemente largo y completo?	
¿Qué es lo mejor?	¿Dónde encuentro esta información?
El seguimiento largo y completo parece una obviedad, pero es fundamental asegurarse que si la enfermedad puede aparecer luego de muchos años, todos los pacientes sean seguidos hasta ese momento y no sacar conclusiones prematuras. Claro que esto aumenta la probabilidad de perder pacientes, por lo que el seguimiento debe ser muy bien planeado en las cohortes, cosa que se evita en los estudios de casos y controles.	Primero, mire en la sección de <b>Métodos</b> si se menciona hasta cuando se planeó el seguimiento, y si éste es razonable para la aparición del evento que se quiere estudiar. Si no aparece en esta sección, vea en <b>Resultados</b> cuál fue el resultado del seguimiento, hasta cuando duró y cuantos pacientes fueron perdidos durante este tiempo.
Este artículo: Si <input type="radio"/> No <input type="radio"/> No está claro <input type="radio"/> Comentarios:	
Criterios secundarios:	
1) ¿Es correcta la relación temporal?	
¿Qué es lo mejor?	¿Dónde encuentro esta información?
¿Precede la exposición al agente nocivo a la presentación del resultado adverso? A veces este punto puede no ser tan obvio como parece. Imagínese si está estudiando factores de riesgo de bajo peso al nacer y el probable factor nocivo es la presencia de infección del tracto urinario materno durante el embarazo, la única forma de asegurar si ésta tiene relación con el bajo peso al nacer es conocer en qué momento del mismo ocurre la infección, ya que si esta aparece un par de semanas previas al parto ya no tendría lógica su relación con el bajo peso del recién nacido.	En <b>Métodos</b> debería describir en detalle la temporalidad de la exposición al factor de riesgo con respecto a la aparición del evento. A veces esto no está explicitado pero se puede deducir de la información que se brinda junto a un razonamiento fisiopatológico.
Este artículo: Si <input type="radio"/> No <input type="radio"/> No está claro <input type="radio"/> Comentarios:	
2) ¿Existe un gradiente en la relación dosis-respuesta?	
¿Qué es lo mejor?	¿Dónde encuentro esta información?
Se puede atribuir con mayor seguridad un efecto adverso a una determinada noxa si, a medida que aumenta la exposición en cantidad o duración se incrementa también el riesgo de un resultado adverso. Supongamos que investigamos con un estudio de casos y controles con el cigarrillo como FR para la aparición de cáncer de pulmón; el hallar que a mayor consumo de tabaco existe mayor riesgo de aparición de cáncer se refuerza la hipótesis planteada. Otras veces esto no es evaluable, ya que el FR estará presente o ausente, sin posibilidad de dosificación (ej: mutación de un gen)	Primero en la sección <b>Materiales y Métodos</b> fíjese si los autores se plantean realizar un análisis que tenga en cuenta el grado de exposición al factor de riesgo. Luego en la sección <b>Resultados</b> debería estar explicitado este aspecto, aclarando si hubo un aumento o no de los eventos a medida que aumentó la exposición.
Este artículo: Si <input type="radio"/> No <input type="radio"/> No está claro <input type="radio"/> Comentarios:	

**Resultados:** ¿Cuál es la magnitud del efecto del tratamiento?

**1) ¿Cuál es la fuerza de la asociación entre la exposición y el resultado?**

Como comentamos antes el ensayo clínico aleatorizado sería el estudio ideal, pero para estudiar factores de riesgo suele ser imposible este diseño. Por lo tanto el seguimiento de cohortes y los estudios de casos y controles son los más utilizados. El análisis de los resultados en el caso de las cohortes se realiza utilizando el Riesgo Relativo siempre que el evento estudiado sea dicotómico (se curó o no se curó, enfermó o no enfermó) dividiendo las incidencias del evento entre los grupos. Para el caso de estudios retrospectivos (tipo caso-control) donde no se puede hablar de incidencias, la medida de magnitud de resultados utilizada es el Odds Ratio o Razón cruzada.

Veamos un ejemplo donde se estudia la variable *introducción del gluten mientras dura la lactancia materna (LM)*, comparada a la introducción del gluten en lactantes que ya no reciben leche materna. Estudio de casos y controles:

	CELÍACOS	SANOS
Gluten con LM	Expuestos enfermos (a) N = 46	Expuestos sanos (b) N = 79
Gluten sin LM	No expuestos enfermos (c) N = 408	No expuestos sanos (d) N = 334
	454	413

¿Cuál es la medida?

¿Qué significa?

**Riesgo Relativo (RR)** = riesgo del desenlace en el grupo expuesto al FR / riesgo del desenlace en el grupo control.

El riesgo relativo nos dice **cuántas veces más probable** es que un evento ocurra en un grupo en relación al grupo control. Un **RR de 1** significa que no hay diferencia entre los dos grupos. Un **RR < 1** significa que el factor analizado disminuye el riesgo del desenlace o punto final estudiado (factor protector). Un **RR > 1** significa que la exposición al factor de riesgo aumenta la probabilidad de que ocurra el evento.

**Odds Ratio (OR, razón cruzada)** = Se define como el cociente entre la probabilidad que un evento ocurra y la de que no ocurra.  
**OR** = expuestos enfermos x no expuestos sanos / no expuestos enfermos x expuestos sanos.  
**OR = (a x d) / (c x b)**

La interpretación del **OR** es similar a la del **RR** y como vimos en otras oportunidades siempre está en relación a la unidad. Un **OR = 1** indica que no hay asociación entre el factor de riesgo y el evento en estudio un **OR > 1** indica una asociación positiva (mayor riesgo en los individuos expuestos) y un **OR < 1** sugiere una asociación negativa o efecto protector.

**Incremento Absoluto del Riesgo (IAR)** = riesgo del desenlace en el grupo control - riesgo del desenlace en el grupo con FR.

Diferencia entre el riesgo en el grupo control y el riesgo en el grupo con el factor. El rango es de -1 a 1. El **RAR=0** significa que no hay asociación entre la presencia del factor y el evento. **RAR<0** si la asociación es positiva, es decir la presencia del FR se asocia a mayor ocurrencia del evento y **RAR>0** si la asociación es negativa.

**2) ¿Qué precisión tiene la estimación del riesgo?**

Como comentamos en artículos anteriores, para que las diferencias sean consideradas estadísticamente significativas, el valor de la p debe ser < a 0,05, lo que equivale a que el intervalo de confianza (IC) 95% no debe abarcar el **valor de nulidad** (uno). El valor de p debe interpretarse en el contexto del resto de los resultados obtenidos ya que, aún cuando las diferencias sean mínimas o irrelevantes, si la muestra es lo suficientemente grande la p será < a 0,05.

Continuando con nuestro ejemplo, con la variable introducción del gluten mientras dura la lactancia materna (LM), de la tabla de 2 x 2 se logra un OR = 0,47 (IC 95% 0,32 – 0,72). Podemos decir que el riesgo de padecer enfermedad celíaca en los individuos con introducción del gluten mientras dura la LM varía entre 0,32 a 1 (como máximo) y una relación 0,72 a 1 (como mínimo) comparado con los lactantes con introducción del gluten sin estar todavía con LM.

Este artículo: Si  No  No está claro   
Comentarios:

**Validez externa:** ¿Los resultados me ayudarán en la atención de mis pacientes?

#### Utilidad de los resultados para la asistencia de mis pacientes

Las preguntas que debe hacerse antes que decida aplicar los resultados de un estudio en su paciente son:

##### 1. ¿Son aplicables los resultados a mi práctica clínica?

Para que usted pueda aplicar los resultados obtenidos en el estudio a los pacientes que le consultan diariamente deberá comparar las características raciales, etarias, o de comorbilidad de sus pacientes con respecto a la muestra descrita en el estudio y las diferencias clínicamente importantes en las exposiciones o factores de riesgo entre ambos.

##### 2. ¿Cuál es la magnitud del riesgo?

El OR no informa la frecuencia con la que se producen los efectos no deseados, sólo compara la frecuencia del efecto observado entre un grupo expuesto y otro no expuesto, por lo que de ninguna manera reemplaza el criterio clínico con respecto a la importancia del resultado.

Una buena medida para expresar el riesgo es el número necesario a dañar (NND), que representa el número de pacientes que deben estar expuestos al FR para que 1 de ellos presente la enfermedad, y que la aparición de ésta sea atribuible al FR. En los estudios de casos y controles, al no poder hablar de incidencia, no se puede calcular en forma directa el NND, tendrá que calcularlo con una fórmula para convertir los OR a NND (o NNT). En cambio en estudios de cohortes el NND se calcula 1/IAR.

$$NND (OR > 1) = \frac{[1 + (PEER \times (OR - 1))]}{[(1 - PEER) \times (PEER) \times (OR - 1)]}$$

$$NND (OR < 1) = \frac{[1 - (PEER \times (1 - OR))]}{[(1 - PEER) \times (PEER) \times (1 - OR)]}$$

PEER es la "Tasa de eventos esperada en los pacientes" (del inglés "Patients expected event rate"), es decir, la tasa de eventos esperada en la población general; dato que se obtiene de la literatura médica o incluso en la introducción del propio artículo que estamos analizando. También hay links en internet donde se puede calcular: <http://ktclearinghouse.ca/cebm/practise/ca/calculators/ortonnt>.

##### 3. ¿Debo intentar detener la exposición?

Para poder interpretar los resultados como indicativos de efecto es importante tener en cuenta no sólo los valores descriptos, sino también el diseño del estudio. No todos los diseños tienen el mismo nivel de evidencia ya que depende de la calidad metodológica del estudio. Las revisiones sistemáticas, meta-análisis y ensayos clínicos aleatorizados, tienen mayor nivel de evidencia que los diseños de cohorte o caso control. En los casos en que el diseño del estudio es débil, podría ser necesario un aumento de riesgo muy superior ya que es más probable que las observaciones sutiles se deban a leves dificultades del diseño. Otra pregunta importante es cuál es la consecuencia de reducir o eliminar el agente nocivo, aunque la evidencia sea relativamente débil, la decisión es más sencilla cuando la exposición es fácilmente evitable.

Este artículo: Si  No  No está claro   
Comentarios:

**PLANILLA PARA LECTURA CRÍTICA: ESTUDIOS SOBRE FACTORES PRONÓSTICOS**

**Validez interna:** ¿Son válidos los resultados del estudio?

<b>Criterios principales:</b>	
<b>1) ¿Fue una muestra representativa y bien definida de pacientes en un momento similar en el curso de la enfermedad?</b>	
<b>¿Qué es mejor?</b>	<b>¿Dónde encuentro esta información?</b>
En este tipo de estudios, los pacientes deben encontrarse en un momento similar de la enfermedad (tiempo cero), idealmente en un estadio inicial, salvo que se quiera evaluar el pronóstico de un estadio tardío de ésta. Por ejemplo, si se quiere establecer la importancia como valor pronóstico de una metástasis en un cierta neoplasia, es relevante que todos los pacientes se encuentren por ejemplo, al momento del debut o del diagnóstico inicial del tumor; y que no sean recaídas de una antigua neoplasia porque esto incidiría de por sí negativamente en el pronóstico.	En la sección de <b>Métodos</b> debería describir el estadio en el cual los pacientes ingresaron al estudio (por ejemplo, día de la punción de médula ósea al diagnóstico, fecha de inicio de la sintomatología, etc.). La sección de <b>Métodos</b> también debería proveer información sobre el reclutamiento de los pacientes, si es que estos fueron reclutados de la atención primaria o centros de referencia terciarios (en este último es probable que se encuentren los pacientes con peor evolución o factores de mal pronóstico: sesgo de referencia). Los sesgos o errores sistemáticos, relacionados con el reclutamiento o selección de los pacientes, implican una inclusión diferencial de sujetos en el estudio que genera grupos no comparables.
Este artículo: Si <input type="radio"/> No <input type="radio"/> No está claro <input type="radio"/> Comentarios:	
<b>2) ¿Fue el seguimiento lo suficientemente prolongado y completo?</b>	
<b>¿Qué es lo mejor?</b>	<b>¿Dónde encuentro esta información?</b>
El período de seguimiento de los pacientes debe ser suficiente para permitir la aparición del evento de interés y este varía de acuerdo a la enfermedad en estudio. Por ejemplo si hablamos de riesgo de recaída en un cáncer, si el seguimiento es de pocos años puede que no lleguemos a detectar a todos los que sufren recaídas. Pero a su vez, existe mayor riesgo de pérdida de pacientes cuanto mayor es el tiempo requerido para la aparición del evento, y esto trae un riesgo de sesgo. Existen dos maneras de evaluar el sesgo por pérdida de seguimiento, la regla de 5 y 20, y el análisis del peor/mejor escenario o análisis de sensibilidad. Una pérdida menor al 5% probablemente no afecte los resultados, en cambio si esta es mayor al 20%, invalida las conclusiones.	La sección de <b>Resultados</b> debería decir la mediana o media de seguimiento. La sección de <b>Resultados</b> también debería proveer el número y las razones de los pacientes que no estuvieron disponibles durante el seguimiento. En una tabla se puede presentar la comparación de los dos grupos (los que completaron el seguimiento y los que no) o los autores deberían declarar en el texto si hubo diferencias entre ellos o no.
Este artículo: Si <input type="radio"/> No <input type="radio"/> No está claro <input type="radio"/> Comentarios:	



**Criterios secundarios:**

**1) ¿Se utilizaron criterios objetivos y no sesgados de resultados?**

¿Qué es lo mejor?

Para evitar sesgos de medición diferencial, se deben determinar previamente criterios objetivos que definen cada resultado de interés, un diagnóstico incorrecto, omisiones o un instrumento de medida no adecuado, podrían conducir a una estimación equivocada de los resultados. Por ejemplo, la muerte es un hecho tan objetivo que sería indiscutible... pero el momento en que se diagnostica un recaída es algo propenso al error. ¿qué criterios se utilizaron?. Lo ideal es que los desenlaces menos objetivos sean evaluados de manera ciega, es decir el individuo que evalúa el desenlace no sabe si el paciente tiene un potencial factor pronóstico.

¿Dónde encuentro esta información?

La sección de **Métodos** debería proveer una definición clara o criterios específicos para cada desenlace, y si la determinación es ciega en relación a los factores pronósticos.

Este artículo: Si  No  No está claro   
Comentarios:

**2) ¿Se llevó a cabo un ajuste para los factores pronósticos importantes?**

¿Qué es lo mejor?

Un factor pronóstico es una característica de un paciente (por ejemplo, edad al diagnóstico, etapa de la enfermedad, etc.) que predice el eventual desenlace. El estudio debería ajustar en el análisis según factores pronósticos conocidos para que los resultados no se distorsionen. Los ajustes pueden ser sencillos, como el análisis estratificado, o más complejos y poderosos, como el análisis multivariado.

¿Dónde encuentro esta información?

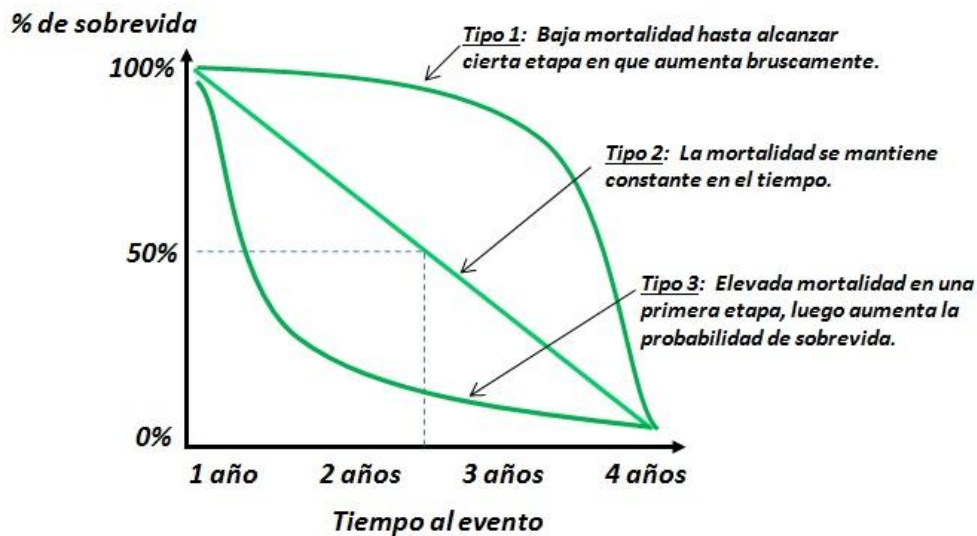
La sección de **Resultados** debería identificar cualquier factor pronóstico y si fueron ajustados por ellos en el análisis o no.  
Mire en las tablas y figuras en búsqueda de evidencia de ajuste (por ejemplo, podrían haber curvas de supervivencia separadas para pacientes en diferentes etapas de la enfermedad o para grupos de edad diferentes).

Este artículo: Si  No  No está claro   
Comentarios:

**Resultados:** ¿Cuáles fueron los resultados?

**1) ¿Cuán amplia es la probabilidad del acontecimiento a lo largo del tiempo de seguimiento?**

Para evaluar la importancia de los resultados obtenidos en el estudio, los resultados deben cuantificarse en términos de probabilidad. La supervivencia o supervivencia se puede definir como el tiempo de espera hasta la ocurrencia de un evento (sea éste bueno o malo), este puede ser muerte, curación, recaída, etc. Luego, la probabilidad de que algo ocurra a través del tiempo puede ser descrita por medio de porcentajes, mediana de supervivencia y las más complejas curvas de supervivencia de Kaplan-Meier. La supervivencia no significa sólo el concepto de defunción o no defunción, porque puede ser también la duración o permeabilidad de un catéter, etc. Entonces las curvas grafican la ocurrencia de eventos a través del tiempo.



¿Cuál es la medida?	¿Qué significa?
<p><b>Incidencia</b> = Es el número de casos nuevos de una enfermedad en una población determinada y en un periodo determinado.</p> $\frac{\text{N}^\circ \text{ de casos nuevos} \times 100}{\text{N}^\circ \text{ susceptibles al inicio del período}}$	<p>Las medidas de incidencia más utilizadas son la incidencia acumulada y la tasa de incidencia.</p> <p><b>Incidenca acumulada:</b> Proporción de nuevos casos de la enfermedad en un período de tiempo. Se expresa en %.</p> <p>Por ejemplo, si estudiáramos un grupo de niños con osteosarcoma de fémur durante un período de 5 años y quisiésemos detectar la presencia de metástasis pulmonares. Se analizan 300 niños de los cuales 80 desarrollan metástasis pulmonares al final del período (<math>80/300 = 0,26</math>), por lo que se concluye que el 26% de los pacientes tienen la probabilidad de desarrollar metástasis pulmonares a los 5 años.</p>
<p><b>Mediana de supervivencia</b> = Es el tiempo en el cual la mitad de los pacientes presentaron el desenlace de interés.</p>	<p>Se calcula como cualquier mediana, ordenando los tiempos al evento de menor a mayor e identificando el valor central o se puede visualizar en la curva de supervivencia.</p> <p>Siguiendo con el ejemplo anterior, deberíamos ordenar los tiempos en los cuales desarrollaron la metástasis pulmonar de menor a mayor, por ejemplo a los 6 meses, al año, etc; e identificar cuando la mitad de los pacientes presentaban metástasis pulmonares y la otra mitad no.</p>
<p><b>Curva de supervivencia</b> = El eje de las abscisas (línea horizontal), corresponde al tiempo de supervivencia y el eje de las ordenadas (línea vertical) a la probabilidad de supervivencia (expresado en porcentaje).</p>	<p>Se realiza una representación en forma de curva en donde se grafica la evolución de los pacientes desde el momento del diagnóstico (tiempo 0) hasta el desarrollo del evento de interés. Para comparar las curvas de supervivencia de dos o más grupos de pacientes, la prueba más usada es la denominada prueba del "log-rank", con la cual se calcula el valor p correspondiente a la comparación de las curvas de supervivencia de los grupos.</p>

## 2) ¿Qué precisión tiene la estimación del riesgo?

Para determinar la precisión de los estimadores necesitamos mirar en los intervalos de confianza (IC) del 95% alrededor del estimador. Entre más estrecho sea el IC, más útil es el estimador. La precisión de los estimadores depende del número de observaciones en las que se basa el estimador. Dado que los periodos más tempranos del seguimiento usualmente incluyen resultados de más pacientes que los periodos más tardíos, es que los estimadores en el lado izquierdo de la curva son comúnmente más precisos. Las observaciones en la derecha o en la cola de la curva usualmente se basan en un número muy pequeño de personas debido a las muertes, abandonos e ingresos tardíos al estudio. Como consecuencia, los estimadores de sobrevivencia al final del periodo de seguimiento son relativamente imprecisos y pueden estar afectados por lo que les ocurra a unas pocas personas.

Este artículo: Si  No  No está claro   
Comentarios:

**Validez externa:** ¿Los resultados me ayudarán en la atención de mis pacientes?

### Utilidad de los resultados para la asistencia de mis pacientes

Las preguntas que debe hacerse antes que decida aplicar los resultados de un estudio en su paciente son:

#### 4. ¿Fueron los pacientes del estudio parecidos a los míos?

Existen características poblacionales que pueden alterar los resultados de un estudio, pacientes pediátricos o adultos, comorbilidades, etc., por lo que es importante evaluar si existe semejanza entre la muestra analizada en el estudio y los pacientes a los cuales uno quiere aplicar los resultados, o sea... mis pacientes. Para esto los autores del estudio deberán describir a los sujetos con el suficiente detalle como para permitir una comparación.

#### 5. ¿Producirá esta evidencia un efecto clínicamente importante en las conclusiones de que decirle u ofrecerle a nuestro paciente? ¿Son útiles los resultados para tranquilizar o aconsejar a los pacientes o su familia?

Creemos que el conocimiento siempre es una herramienta que mejora nuestra práctica diaria, independientemente de que la conclusión sobre el pronóstico de nuestro paciente sea buena o mala. La información sobre el pronóstico de la enfermedad podrá ser útil tanto para la toma de decisiones del médico como del paciente y/o su familia.

Este artículo: Si  No  No está claro   
Comentarios:

**PLANILLA PARA LECTURA CRÍTICA: ESTUDIOS DE META-ANÁLISIS**

**Validez interna:** ¿Son válidos los resultados de esta revisión?

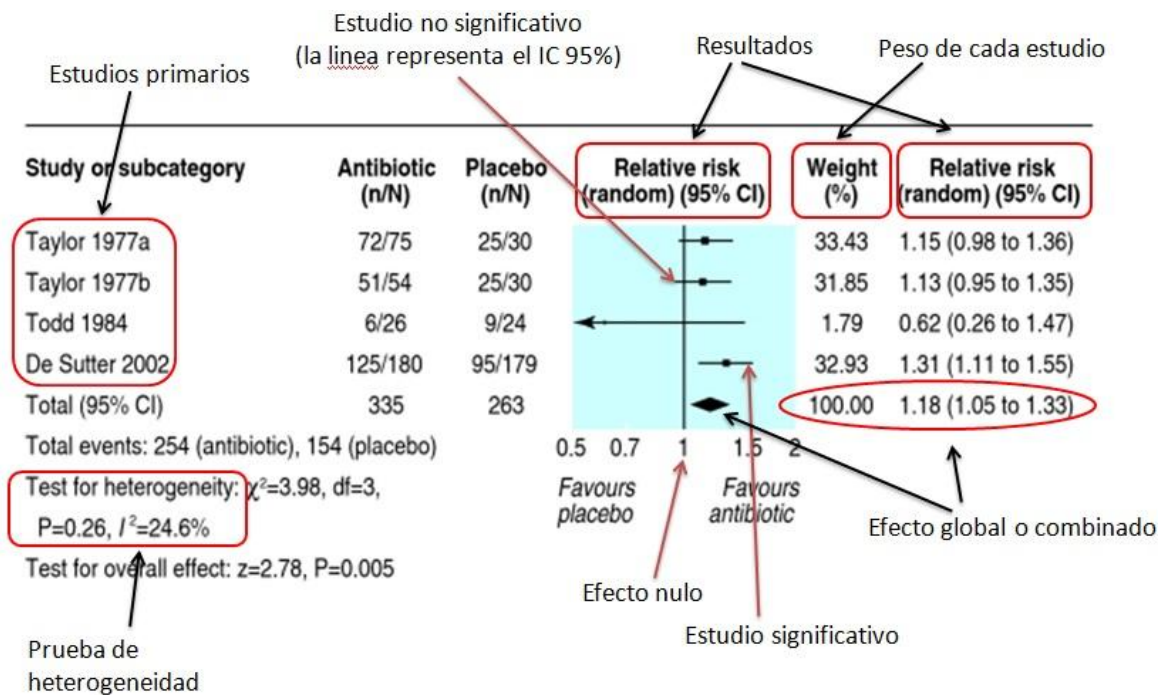
<b>Criterios principales:</b>	
<b>1) ¿Qué pregunta (PICO) abordó la revisión sistemática? ¿Es un problema clínico bien focalizado?</b>	
<b>¿Qué es mejor?</b>	<b>¿Dónde encuentro esta información?</b>
La mayor parte de las preguntas clínicas pueden ser formuladas en términos de una simple relación entre un paciente, una exposición, un test diagnóstico, un potencial agente dañino y uno o más puntos finales de interés. Si la pregunta principal de una revisión no queda clara, probablemente no esté bien focalizada. Se sugiere que identifique los componentes PICO (Población, Intervención, Comparación y Outcome).	En el <b>Título</b> , <b>Resumen</b> o el párrafo final de la <b>Introducción</b> debería definirse la pregunta claramente. Si aún después de leer estas secciones no puede estar seguro cuál es la pregunta en la que se enfoca, ¡busque otro artículo!
Este artículo: Si <input type="radio"/> No <input type="radio"/> No está claro <input type="radio"/>	
Comentarios:	
<b>2) ¿Los criterios de selección usados para incluir a los artículos fueron apropiados?</b>	
<b>¿Qué es lo mejor?</b>	<b>¿Dónde encuentro esta información?</b>
La inclusión o exclusión de estudios en la revisión sistemática debería estar claramente definido desde antes. Los criterios de elegibilidad usados deben especificar los pacientes, intervenciones o exposiciones y desenlaces de interés. En muchos casos el tipo de diseño de estudio también será un componente clave de los criterios de elegibilidad.	En la sección de <b>Métodos</b> se debería describir en detalle los criterios de inclusión y exclusión. Normalmente esto incluye el diseño de estudio.
Este artículo: Si <input type="radio"/> No <input type="radio"/> No está claro <input type="radio"/>	
Comentarios:	
<b>Criterios secundarios:</b>	
<b>1) ¿Es probable que estudios importantes y relevantes se hayan pasado por alto?</b>	
<b>¿Qué es lo mejor?</b>	<b>¿Dónde encuentro esta información?</b>
El punto de partida para una búsqueda exhaustiva de todos los estudios relevantes son las bases de datos bibliográficas más importantes (es decir, Medline, Cochrane, EMBASE, etc.), pero también debería incluir una búsqueda de las referencias de los estudios relevantes y contactar expertos para averiguar sobre estudios no publicados. La búsqueda no se debería limitar al idioma inglés. La estrategia de búsqueda debería incluir tanto términos MeSH como palabras clave.	La sección de <b>Métodos</b> debería describir cual fue la estrategia de búsqueda, incluyendo los términos usados con algún detalle. La sección de <b>Resultados</b> describirá el número de títulos y resúmenes revisados, el número de estudios rescatados en texto completo y el número de estudios excluidos junto con las razones para hacerlo. Esta información puede aparecer en una figura o diagrama de flujo.
Este artículo: Si <input type="radio"/> No <input type="radio"/> No está claro <input type="radio"/>	
Comentarios:	

2) ¿Los estudios incluidos fueron suficientemente válidos para el tipo de pregunta?	
¿Qué es lo mejor?	¿Dónde encuentro esta información?
El artículo debe describir como se evaluó la calidad de cada estudio usando criterios de calidad predefinidos apropiados al tipo de pregunta clínica (por ejemplo, aleatorización, ciego y seguimiento exhaustivo). Diferentes herramientas permiten valorar críticamente los estudios primarios; para ensayos clínicos por ejemplo se puede utilizar la escala de Jadad, etc.	La sección de <b>Métodos</b> debe describir la evaluación de la calidad y los criterios usados. La sección de <b>Resultados</b> debería proveer información sobre la calidad individual de los estudios.
Este artículo: Si <input type="radio"/> No <input type="radio"/> No está claro <input type="radio"/> Comentarios:	
3) ¿Fueron reproducibles las evaluaciones de los estudios?	
¿Qué es lo mejor?	¿Dónde encuentro esta información?
Cada estudio debería ser evaluado por al menos 2 expertos en lectura crítica, y su opinión debería coincidir. En caso de que esto último no ocurra, un tercer evaluador puede desempatar acerca de la validez del estudio.	La sección <b>Métodos</b> debe describir quienes intervinieron en la evaluación y el método de desempate en caso de desacuerdo.
Este artículo: Si <input type="radio"/> No <input type="radio"/> No está claro <input type="radio"/> Comentarios:	
4) ¿Los resultados fueron similares entre los estudios?	
¿Qué es lo mejor?	¿Dónde encuentro esta información?
Cuando hablamos de heterogeneidad podemos distinguir dos aspectos: por un lado el relativo a las diferencias existentes entre los estudios en cuanto a características de los pacientes incluidos, el tiempo de seguimiento, las dosis empleadas, etc. y por otro lado el concepto de heterogeneidad estadística, que únicamente cuantifica la variabilidad entre los resultados de los estudios, y que puede ser debida a las diferencias reales de planteamiento y ejecución entre los estudios incluidos, o a otras causas. Es decir que es deseable que no exista heterogeneidad, porque si esto pasara no podremos combinar los resultados de los estudios para alcanzar una medida final única.	La sección de <b>Resultados</b> debería plantear si los resultados son heterogéneos y discutir posibles razones. El gráfico de bosque debería mostrar los resultados del test de Chi cuadrado para heterogeneidad y discutir razones de su presencia, si existe. Las pruebas de heterogeneidad pueden ser el test Q de Cochran; se recomienda utilizar un nivel de rechazo en esta prueba al menos de 0.1 (menor de 0,1 revela heterogeneidad). La otra es la prueba del $I^2$ que indica la proporción de la variación entre estudios respecto de la variación total. Esta prueba revela heterogeneidad importante cuando es mayor a 50%.
Este artículo: Si <input type="radio"/> No <input type="radio"/> No está claro <input type="radio"/> Comentarios:	

**Resultados:** ¿Cuáles han sido los resultados?

### ¿Cómo se presentan los resultados?

Una revisión sistemática provee un resumen de datos de los resultados de un número de estudios individuales. Si los resultados de los estudios son similares, un método estadístico (llamado meta análisis) se usa para combinar los resultados de los estudios individuales y se calcula un estimador global de resumen. El meta análisis le da peso a los valores de cada estudio dependiendo de su tamaño. Los resultados individuales de los estudios necesitan ser expresados de una forma estándar, como en riesgo relativo, *odds ratio* o diferencia promedio entre los grupos. Los resultados se presentan tradicionalmente en una figura, como la de más abajo, llamada **Forest plots (gráfico de bosque)**



El forest plot ilustrado arriba representa un meta análisis que evaluó el efecto de los antibióticos en la rinitis purulenta (BMJ 2006). Los estudios individuales están representados por un cuadrado (media) atravesado por una línea horizontal (IC 95%), en este caso del Riesgo Relativo. El tamaño del cuadrado refleja el peso del estudio en el meta análisis. La línea sólida vertical corresponde a "efecto nulo" del tratamiento - que en el caso de diferencias de riesgo es 0 y para el resto es de 1-.

Entonces, cuando el intervalo de confianza incluye al 0 ó al 1 (como en este ejemplo) indica que el resultado no es significativo ( $p>0,05$ ). El diamante en la parte inferior representa el RR combinado o conjunto de los ensayos con su intervalo de confianza de 95% representado por sus extremos. En este caso, se presenta un metaanálisis de 4 estudios primarios con resultado combinado a favor del uso de los antibióticos.

Este artículo: Si  No  No está claro   
 Comentarios:



**Validez externa:** ¿Los resultados me ayudarán en la atención de mis pacientes?

**¿Qué tan precisa es la estimación del efecto del tratamiento?**

Las preguntas que debe hacerse antes que decida aplicar los resultados de un estudio en su paciente son:

- 4) ¿Es mi paciente tan diferente de los del estudio que los resultados no son aplicables a él?
- 5) ¿El tratamiento es posible en mi contexto clínico?
- 6) ¿Los posibles beneficios del tratamiento superan los potenciales daños para mi paciente?

Este artículo: Si  No  No está claro   
Comentarios:

**Bibliografía:**

- Guyatt G. et al. ¿Cómo utilizar un artículo sobre tratamiento o prevención? JAMA. 1993;270:2598-2601.
- Guyatt G. et al. ¿Cómo utilizar un artículo sobre tratamiento o prevención II? JAMA. 1994;271:59-63.
- Jaeschke R. et al. ¿Cómo utilizar un artículo sobre un examen diagnóstico? JAMA. 1994;271:389-391 y 703-707.
- Levine M. et al. ¿Cómo utilizar un artículo sobre efectos nocivos? JAMA. 1994;271:1615-1619.
- Laupacis A. et al. ¿Cómo utilizar un estudio sobre pronóstico? JAMA. 1994;272:234-237.
- Oxman A. et al. ¿Cómo utilizar una revisión de conjunto? JAMA. 1994;272:1367-1371.
- Rowensztein H; Berlin V. "Medicina Basada en la Evidencia (1ra parte). Estudios de intervención". Revista de la Facultad de Ciencias Médicas de Guayaquil. 2009; 12(4):80-85.
- Rowensztein H; Berlin V. "Medicina Basada en la Evidencia (2da parte). Estudios sobre diagnóstico". Revista de la Facultad de Ciencias Médicas de Guayaquil. 2010; 13(1):55-62.
- Rowensztein H; Berlin V. "Medicina Basada en la Evidencia (3ra parte). Estudios sobre factores de riesgo". Revista de la Facultad de Ciencias Médicas de Guayaquil. 2010; 13(2):50-54.
- Rowensztein H; Berlin V. "Medicina Basada en la Evidencia (4ta parte). Estudios sobre factores pronósticos". Revista de la Facultad de Ciencias Médicas de Guayaquil. 2011; 14(1):53-58.