



Revista Americana de Medicina
Respiratoria

ISSN: 1852-1630

revista@ramr.org

Asociación Argentina de Medicina
Respiratoria
Argentina

Sobrino, Edgardo; Fernández Pazoz, Alfonso; Molinari, Luciana; Arce, Santiago

Lectura crítica: Validez interna

Revista Americana de Medicina Respiratoria, vol. 6, núm. 2, septiembre, 2006, pp. 77-83

Asociación Argentina de Medicina Respiratoria

Córdoba, Argentina

Disponible en: <http://www.redalyc.org/articulo.oa?id=382138366005>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica

Red de Revistas Científicas de América Latina, el Caribe, España y Portugal

Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto



Correspondencia
E-mail labneumo@lanari.fmed.uba.ar

Lectura crítica: Validez interna

Autores Edgardo Sobrino, Alfonso Fernández Pazos, Luciana Molinari, Santiago Arce
Instituto de Investigaciones Médicas Alfredo Lanari, Servicio de Neumonología,
Facultad de Medicina, UBA

Una vez realizada la pregunta y la búsqueda, nos encontramos frente a un trabajo, cuya elección depende de varios factores, dentro de los que se hallan: un título adecuado, una revista reconocida, factibilidad de hallarlo en la biblioteca, etc.

Es éste el momento donde se deben generar las tres preguntas básicas que lo ayudarán a decidir si se trata de un artículo digno de ser leído o, por el contrario, un artículo prescindible. Las preguntas son:

1. ¿Es válida la evidencia de este estudio?
2. ¿Son importantes sus resultados?
3. ¿Se pueden aplicar al cuidado de mi paciente?

En esta guía de lectura crítica nos referiremos a las dos primeras preguntas, dejando la tercera para un futuro informe. También podríamos decir que nos ocuparemos de la validez interna (VI) entendida como el grado en el cual el resultado de un estudio se acerca a la realidad mediante una metodología que minimiza el sesgo o error sistemático. Luego nos ocuparemos de la validez externa (VE) o “generalizabilidad”. La VE es la extensión a la cual los resultados del estudio pueden ser aplicados en aquellos más allá de la muestra en estudio.

Dependiendo del tipo de pregunta y del tipo de artículo hallado variará el análisis de la validez interna y variarán los tipos de sesgo a considerar.

Si la pregunta generada está en relación con un tema de **terapéutica**, el tipo de estudio a considerar será un **ensayo clínico**. Si la pregunta está en relación con **pronóstico** el tipo de estudio a considerar será un **estudio de cohorte**, si el tema a considerar es sobre **causalidad** el estudio puede ser una **cohorte** o un **estudio de caso-control** dependiendo de la frecuencia de la entidad, y si es de **diagnóstico** lo mejor será un **corte transversal**^{1,2} (ver tabla 1).

Cuando definimos la VI se mencionó la palabra sesgo. Ahora: **¿Qué es sesgo?** En los estudios epidemiológicos hay dos tipos de errores, aquellos producidos al azar (error aleatorio) y aquellos sis-

temáticos o sesgos. Por lo tanto sesgo es una distorsión de los resultados en una dirección definida.

Las fuentes de variabilidad en general son identificables y predecibles. Señalan fallas de diseño o análisis. Según del tipo epidemiológico dependerá el sesgo a considerar. Hay dos tipos de sesgos y subclasificaciones de estas dos grandes ramas: 1) Sesgo de selección y 2) Sesgo de observación.

– Sesgo de selección: ocurre cuando hay una diferencia sistemática entre las características de las personas seleccionadas para un estudio y las características de aquellas que no lo fueron.

Ej: las personas que responden a una invitación para participar en un estudio sobre los efectos de fumar tiene hábitos de consumo de tabaco distintos a los de las personas que no responden; estos últimos suelen ser grandes fumadores

– Sesgo de observación o de información: errores al obtener la información deseada. Se obtiene información no comparable de los distintos grupos.

Ej: recolección de información diferente por parte del investigador (sesgo de entrevistador). Sujetos comunican eventos o exposiciones de manera no comparable (sesgo de recuerdo).

Un sesgo actualmente de importancia es el sesgo de inmortalidad en los pacientes con enfermedad pulmonar obstructiva crónica que han sido tratados con corticoides inhalados luego de una internación^{3,4}. Cuando se trata de ensayos clínicos, los cuales generan el mayor grado de evidencia, debemos, para evaluar la validez, hacernos estas preguntas:

¿Fue aleatorizada la asignación de los pacientes al tratamiento?

La aleatorización es la asignación de los tratamientos en base al azar. Es el mejor método para crear grupos lo más parecidos posibles, donde la única diferencia sea el tratamiento instaurado. Permite el balance de las variables conocidas (sexo, edad, etc.) y las variables no conocidas que puedan influir en el

Tabla 1: Diseño de estudios epidemiológicos

Tipo de estudios epidemiológicos	
Estudios analíticos	
- Observacionales	- estudios de casos y controles
	- estudios de cohorte
- Experimentales	- ensayos clínicos
Estudios descriptivos	
- Poblacionales	- ecológicos
	- análisis de situación
- Individuales	- reporte de casos
	- serie de casos
	- estudios de corte transversal

resultado. El trabajo debe tener lo que se conoce como ocultamiento de la secuencia de aleatorización, este dato no siempre está reportado. El ocultamiento hace que los investigadores no sepan a qué grupo se va a asignar el próximo paciente. El efecto del no-ocultamiento es que tienden a asignar a los pacientes con mejor pronóstico al grupo experimental, exagerando así los posibles beneficios.

¿Se detalló el seguimiento de todos los pacientes incluidos en el estudio?

Esta pregunta tiene dos puntos importantes, el primero es si fueron seguidos todos los pacientes luego de ser distribuidos aleatoriamente. Los pacientes que se pierden pueden haber tenido eventos que cambien las conclusiones. Una manera de saber si las pérdidas cambian las conclusiones es realizar un "análisis del peor escenario". Esto responde a la pregunta de qué pasaría si a todos los pacientes perdidos del grupo que les fue bien les asignamos el peor resultado. El segundo punto a considerar es que todos los pacientes, luego de ser aleatorizados, deben ser considerados en el grupo en que fue incluido, cumplan o no cumplan con el tratamiento. Este tipo de análisis se conoce como intención de tratar. Lo que se hace es lo que hacemos habitualmente en nuestra práctica diaria, vale decir que nosotros tengamos la intención de tratar no significa que el paciente lo cumpla al pie de la letra, compre el medicamento, lo tome correctamente, etc.

¿Se mantuvo el ciego en los pacientes y en los médicos tratantes?

Los beneficios que aporta son lograr que ni los síntomas reportados por los pacientes ni la interpreta-

ción de los mismos estén influenciada por el hecho de conocer el tratamiento al cual fueron asignados. En algunos casos se hace imposible mantener el ciego, en una técnica quirúrgica por ejemplo, en estos casos lo más importante es que los que evalúen el resultado final estén ciegos al tratamiento recibido.

Cuando la pregunta es de pronóstico, la validez se evalúa con otras preguntas a considerar:

¿Fue una muestra representativa y bien definida de pacientes en un momento similar en el curso de la enfermedad?

El estudio debe describir bien cómo se definió a los pacientes con la enfermedad y en qué momento de su historia natural. Lo ideal es que la cohorte se encuentre en un estadio inicial de la enfermedad. Por ejemplo en el estudio de sobrevida de la hipertensión pulmonar primaria realizado por el National Heart Lung and Blood Institute se tomó la sobrevida del paciente desde el momento del cateterismo diagnóstico⁵.

Hay que tener cuidado con algunos sesgos que pueden presentarse a la hora de seleccionar a los pacientes, ya que pueden seleccionarse en centros de derivación donde, en general, son pacientes con peor pronóstico (sesgo de referencia). Ej: pronóstico de la fibrosis pulmonar idiopática en un centro de trasplante pulmonar.

¿Fue el seguimiento lo suficientemente prolongado y completo?

Lo ideal es que cada paciente sea seguido hasta su recuperación completa o hasta sufrir el evento adverso de interés. Lo que hay que pensar es que se siguió a los pacientes durante un período de tiempo lo suficientemente largo como para evaluar el evento de interés. Esto depende de la enfermedad de interés. No es lo mismo el seguimiento de los pacientes en terapia intensiva ventilados por trastorno respiratorio agudo, que pacientes con cáncer de pulmón.

Una vez asegurado que el tiempo de seguimiento es el adecuado, el paso siguiente es saber si ha sido completo (sesgo de seguimiento). La pérdida de pacientes puede alterar la validez de los estudios pronósticos. En estos casos hay dos maneras de evaluar los trabajos: la primera es la regla del "5 y 20". Una pérdida menor del 5% probablemente no afecte la validez del estudio, y una pérdida mayor del 20% verá afectada seriamente su validez. La otra manera es con un análisis de sensibilidad o, lo que es lo mismo, preguntarse ¿qué hu-

biera sucedido si...? todos los pacientes perdidos hubieran desarrollado el evento adverso o, lo contrario, también lo podríamos resumir como el análisis del mejor o peor escenario⁶.

¿Se utilizaron criterios objetivos y no sesgados de los resultados?

La medición de resultados como la muerte o la recuperación total no resultan difíciles, pero se complican cuando la medición es, por ejemplo, grado de disnea o intensidad de dolor residual o calidad de vida. Para evitar sesgos de medición, los investigadores deben para definir claramente los objetivos y los resultados de interés.

Sin embargo si los investigadores conocen las características de los pacientes pueden sesgar el estudio, por lo que es importante en estos casos también mantener un ciego por parte de los investigadores de las características clínicas de los pacientes evaluados.

Otras preguntas son si se identificaron subgrupos pronósticos o si se validaron los resultados en otro grupo. Los estudios de pronóstico en general tienden a ser complejos con análisis estratificados y análisis multivariados.

¿Qué es un análisis multivariado?: es una herramienta estadística que determina el peso o contribución de una variable en relación a otras en relación con un único evento. Por ejemplo, numerosos factores contribuyen al desarrollo de enfermedad coronaria, obesidad, tabaquismo, sedentarismo, diabetes, hipertensión arterial, etc. Estas variables son llamadas factores de riesgo o variables independientes. El análisis multivariado determina la contribución independiente de cada factor de riesgo para el desarrollo de enfermedad

coronaria⁷. Si desea ahondar en este tema recomendamos el artículo de Mitchell Katz.

Hemos dado una visión somera y rápida de ensayos clínicos, randomizados y estudios de cohorte, dos de los principales estudios epidemiológicos (ver tabla 2) a la hora de generar evidencia. Ambos tipos de estudio son de fundamental importancia. Los estudios de cohorte pueden ser utilizados para determinar si la eficacia observada en ensayos clínicos puede ser trasladada a efectividad en poblaciones más amplias y con situaciones más realistas⁷ (ver tabla 2).

Quién quiera continuar ahondando este tema sugerimos la serie del British Medical Journal sobre estudios de cohorte^{8, 9, 10} y ensayos clínicos¹²⁻¹⁵.

Finalmente, y antes de iniciar nuestra evaluación sobre importancia, necesitamos evaluar.

¿Qué es un confundidor?

El efecto confundidor aparece cuando un factor de riesgo y un evento son afectados por la relación con una tercera variable que se comporta como un nuevo factor de riesgo para el evento y se asocia con el otro factor de riesgo (ver figura 1). Un ejemplo de efecto confundidor se ha suscitado entre la enfermedad periodontal y la enfermedad coronaria, con un marcado aumento en aquellos que tienen enfermedad periodontal (Hazard ratio 2.66 IC 95% 2.34-3.03) en el análisis univariado. Sin embargo la periodontitis está asociada a numerosos factores entre los que se encuentran edad, tabaquismo, aumento de la masa corporal, pobreza. Cuando se ajustó la periodontitis a la enfermedad coronaria con estas variables, la relación ya no fue más significativa (Hazard ratio 1.21 IC 95% 0.98- 1.5), dejando de ser una asociación estadísticamente significativa¹⁶.

Tabla 2: Comparación entre estudios de cohorte y ensayos clínicos randomizados

	Estudios de Cohorte	Ensayos clínicos randomizados
Población en estudio	Población de pacientes que son observados en un contexto determinado	Población altamente seleccionada en base a detallados criterios de inclusión y exclusión
Asignación de la intervención	Decisiones basadas en prácticas médicas o decisiones de los pacientes	Decisiones basadas por chance y controladas por los investigadores
Resultados	Puede ser definida luego de la intervención y puede incluir eventos raros o inesperados	Son determinados antes que los pacientes entren al estudio y se focalizan en riesgos y beneficios predichos.
Seguimiento	Generalmente seguimientos largos	Generalmente seguimientos más cortos debido a costos y presiones en la generación de evidencia
Análisis	Sofisticadas técnicas requiriendo análisis multivariados para tratamiento de factores confundidores	Técnicas más simples y directas.

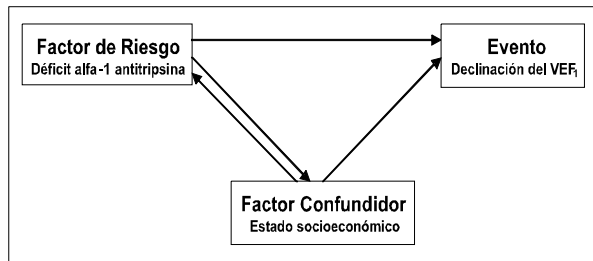


Figura 1: Relación entre factor de riesgo evento y confundidor.

Otro efecto confundidor en Neumonología, y no tan bien estudiado como el ejemplo previo, es el que relaciona la administración de alfa-1 antitripsina (tratamiento supletorio en pacientes con déficit) y caída anual del VEF_1 . El estado socioeconómico del paciente es un confundidor en esta relación¹⁷ (ver figura 1).

Una vez que determinamos su validez y se puede confiar en sus resultados, debemos saber si es importante.

¿Son importantes los resultados?

En este sentido no sólo es importante el efecto sino también su magnitud y precisión.

¿Qué son las medidas de asociación?

Las medidas de asociación describen la fuerza de asociación entre la exposición (intervención) y el resultado en los estudios clínicos (ensayos clínicos, estudios de cohorte, estudios caso-control).

Hay dos tipos de medida de asociación: las relativas y las absolutas. Dentro de las relativas esta el riesgo relativo (RR), el odds ratio (OR) y la reducción del riesgo relativo (RRR) como principales.

Con respecto a las absolutas se encuentran la reducción de riesgo absoluto (RRA) y número necesario para tratar (NNT).

Además, los intervalos de confianza (IC) deben darse para cada medida de modo de cuantificar la incertidumbre. Normalmente se utilizan IC del 95% (lo que significa que dentro de ese intervalo está la chance de tener el resultado en el 95% de los casos).

¿Cómo se calculan?

Se puede realizar con una tabla de contingencia o tabla de dos por dos. ¿Qué es una tabla de contingencia? Se trata de una tabla de distribución de frecuencias utilizada para calcular riesgos. Se cruzan dos variables, ej: presencia o no de una enfermedad con presencia o no de un factor de exposición.

		Enfermedad		
		Sí	No	
Exposición	Sí	Enfermos expuestos a	No enfermos expuestos b	Riesgo en expuestos a/a + b
	No	Enfermos no expuestos c	No enfermos no expuestos d	Riesgo en no expuestos c/c + d
		a + c	d + b	n

Tomemos como ejemplo el primer ensayo clínico realizado, que casualmente está relacionado con nuestra especialidad¹⁸.

Estreptomycin vs. placebo.

Pacientes Randomizados	Muertos a los 6 meses	Vivos a los 6 meses
Estreptomycin (55)	4	51
Placebo (52)	14	38

En el grupo tratado con estreptomycin murieron 4 de 55 (7.3%) denominado riesgo del evento en el grupo experimental (REE). En el grupo placebo murieron 14 de 52 (26.9%) denominado riesgo del evento en el grupo control (REC).

¿Cuál es la diferencia de riesgo de muerte entre ambos grupos? (los tratados con estreptomycin vs. los no tratados con estreptomycin)

La diferencia de riesgo se puede calcular con la reducción de riesgo absoluto (RRA).

$$\begin{aligned} \text{REC} - \text{REE} &= \text{RRA (en este caso)} \\ 26.9\% - 7.3\% &= 19.6\% \end{aligned}$$

En grupo placebo (19.6%) más personas mueren en relación al grupo tratamiento. El RRA es una medida de asociación o de efecto. También recibe el nombre de diferencia de riesgos, ya que representa la diferencia de la incidencia de la enfermedad en la población expuesta al factor de riesgo y la incidencia en la población no expuesta a dicho factor. O sea, representa la diferencia entre la tasa de incidencia de la población expuesta y no expuesta

Siempre que se informe la RRA es importante agregar la incidencia de la patología en las poblaciones que se está comparando, dado que de lo contrario la interpretación se hace muy difícil.

¿Cuántas personas con tuberculosis debo tratar con estreptomicina para prevenir 1 muerte?

Esta es una medida más intuitiva y sirve para mejorar la interpretación de las medidas de efectos. Se calcula como la recíproca de la RRA.

$$\text{NNT} = 1 \div (19.6\% \div 100) = 5.1 = 6$$

Quiero decir que debo tratar a 6 con estreptomicina para salvar 1 vida.

¿Cuál es el riesgo de morir en el grupo de estreptomicina en relación al grupo placebo?

La respuesta es que por cada 4 que mueren en el grupo placebo, muere 1 en el grupo tratamiento.

¿Cómo llegamos a esa conclusión?

Riesgo relativo =

riesgo en expuestos / riesgos en no expuestos

$$\text{RR} = \text{REE} / \text{REC} =$$

$$7.3 / 26.9 = 0.27 \text{ (IC 95\% 0.10 - 0.77)}$$

Esto puede interpretarse diciendo que por cada 0.27 pacientes que mueren el grupo tratamiento muere 1 en el grupo control. Esto es lo mismo a decir 1 de cada 4 expuestos.

El IC del 95% nos dice que si realizamos 100 veces el mismo estudio, el resultado en el 95% de los casos va a estar entre 0.10 y 0.77. Esto le da mucho valor, porque si el IC pasara la unidad no tendría el efecto protector y probablemente podríamos decir que no se trata de un dato significativo.

Ej: (RR IC 95%: 0.10-1.22) quiere decir que podría disminuir el riesgo en 1 de cada 10 o podría aumentar la mortalidad en 1 de cada 5.

Como conclusión:

- Si el RR es > 1 la enfermedad es más frecuente en la población expuesta
- Si el RR es < 1 la enfermedad es menos frecuente en la población expuesta. Factor protector:
- Si el RR = 1 la incidencia es igual en los expuestos que en los no expuestos, por lo que no se podría demostrar asociación entre la enfermedad y la exposición.
- La cuantificación de riesgo sirve para poder comparar los riesgos de ocurrencia de un mismo evento en poblaciones diferentes o de distintos eventos de un mismo grupo.
- El concepto de riesgo no es diferente al de incidencia.

- RR responde a la pregunta ¿Cuántas veces más probable es que las personas expuestas desarrollen la enfermedad en comparación con las no expuestas?

¿Cuánto menos frecuente es morir por tuberculosis en el grupo tratado vs. el grupo no tratado?

$$\text{RR} = \text{RRA} / \text{REC} = 19.6\% \div 26.9\% = 0.73 = 73\%$$

El tratamiento con estreptomicina disminuye un 73% el riesgo de morir en lugar del placebo. En este punto hay que ser cuidadoso, porque si la RRA fuera del 1.96% y el REC del 2.69%, la reducción también sería del 73%. Por esta razón, la reducción de riesgo relativo (RRR) es frecuentemente usada para promocionar la literatura de las empresas farmacéuticas.

Odds ratio: cuando no se dispone de la información necesaria para calcular las tasas de incidencia en expuestos y en no expuestos, se debe utilizar otra medida de asociación denominada odds ratio. El ejemplo típico es en los diseños de casos y controles. En este caso el investigador selecciona los casos de la enfermedad, que habitualmente son los prevalentes. Como no pueden conocerse los casos nuevos, no puede determinarse la incidencia y es imposible calcular el RR.

Odds = número de personas que experimentan el evento dividido el número de personas que no lo experimentan.

En nuestro ejemplo, sería 4 de 51 (odds 0.078) y 14 de 38 (odds 0.36%). Por lo tanto, el odds ratio será de $0.078 / 0.36 = 0.21$. O sea que el odds ratio de morir en los tratados es de 1 a 5 con respecto a los no tratados.

Como conclusión:

- Para calcular el odds ratio es indispensable conocer el odds.
- El odds se obtiene dividiendo el número de personas que experimentan el evento sobre el número de personas que no lo experimentan.

	Sí	No	Odds
Sí	Enfermos expuestos a	No enfermos expuestos b	a / b
No	Enfermos no expuestos c	No enfermos no expuestos d	c / d
	a + c	d + b	n

¿Cómo medimos la precisión o cómo medimos la incertidumbre?

Necesitamos tomar una muestra de una población que sea representativa ya que se hace imposible medir a toda la población. Tanto el test de hipótesis (con el cálculo de la p) o los intervalos de confianza (IC) pueden cuantificar esa incertidumbre. Esto significa que a las medidas de efecto recién descriptas debemos agregarle información adicional.

Intervención	Estreptomycin (55)	Control (52)
Sobrevida	51	38
Muerte	4	14
Riesgo	4/55=7.3%	14/52 = 26.9 %
RRA	26.9 - 7.3 = 19.6%	
IC 95%	5.7% a 33.6%	
Valor de p	0.006	

El valor de p de 0.006 nos dice que el RRA 19.6% o más se puede dar sólo en 6 casos de cada 1000 si acaso la estreptomycin fuera igual de efectiva que el control. El IC del 95% significa que realizando 100 veces el mismo estudio los valores de la RRA estarán entre 5.7% y 33.6% en el 95% de los casos.

¿Porque es la $p < 0.05$ indicadora de significación estadística?

Básicamente por convención, en el test de hipótesis, la hipótesis nula dice en este caso que el tratamiento con estreptomycin es igual al control con una p de 0.05. Esto significa que en un 5 % de los casos estaremos diciendo que la hipótesis nula es falsa cuando en realidad es verdadera, o lo que es lo mismo, puedo equivocarme en 1 de cada 20 veces. En caso de equivocarme se trata de un falso positivo o error tipo I.

El riesgo de un falso negativo también existe, como sería en este caso. Decir que la hipótesis nula es cierta cuando en realidad es falsa o, lo que es lo mismo, decir que la estreptomycin es igual al tratamiento placebo, cuando no lo es. La probabilidad de no encontrar una diferencia significativa, cuando la hay, se llama error tipo II y disminuye cuando aumenta el tamaño de la muestra. A mayor tamaño de la muestra menor el riesgo de error tipo II.

	La verdad	
Resultado del estudio	Hipótesis nula verdadera (Tratamiento no resulta)	Hipótesis nula falsa (tratamiento resulta)
Rechazo la hipótesis nula	Error tipo I	
Acepto la hipótesis nula		Error tipo II

Múltiples test

Supongamos que en su hospital se estudiaron 200 pacientes con EPOC, de los cuales 100 fueron vistos por el servicio de clínica médica y 100 por el servicio de neumonología. Usted evalúa 6 variables para ver si hay alguna diferencia en las poblaciones. Las seis variables fueron: hipoxemia, hipercapnia, severidad del VEF₁, internaciones, distribución de sexo, presencia de ojos azules.

Para 5 de estas variables encontró distribuciones similares en ambos grupos con $p > 0.10$. Sin embargo, para la variable presencia de ojos azules se halló en el grupo de clínica 25 de 100 y 38 de 100 en el de neumonología.

¿Qué probabilidad hay de que en el testeo de estas seis hipótesis independientes alguna haya sido por chance < 0.05 ? La probabilidad de rechazar la hipótesis nula y que ésta sea verdadera la habíamos denominado error tipo I y era como mínimo de 1 en 20. Ahora, cuando se evalúan las seis hipótesis en forma independiente, la probabilidad de que una dé significativa es de 0.95 a la sexta, lo que es igual a 0.74. Por lo tanto la probabilidad de que una dé significativa es 26% o 1 de cada 4. Si nosotros queremos mantener la misma probabilidad de 1 en 20 debemos dividir la $p = 0.05$ por 6 = 0.008.

Como conclusión podemos decir que 1) hallazgos raros ocurren por azar. 2) Hay que tener cuidado con el testeo de múltiples hipótesis que puede llegar a resultados erróneos.

Intervalos de confianza: ¿Qué nos informa?

Nos informa un rango de valores dentro del cual se encuentra el verdadero parámetro de la población a partir de los estimadores obtenidos en una muestra.

Informa acerca de la magnitud del efecto a diferencia del valor de p que no solo no informa sobre la magnitud sino que además tampoco nos infor-

ma sobre la dirección del efecto, recordando que los test estadísticos son a dos colas.

Con respecto a su interpretación en general se dice: asuma usted cual es el menor efecto que consideraría importante para tratar a su paciente, luego observe el valor inferior del IC del 95%, si este es mayor al valor que usted había considerado entonces el trabajo tiene un adecuado poder o tamaño y es concluyente.

Si el resultado es negativo o tiene un efecto protector asuma cual es el mayor efecto que usted consideraría para tratar a su paciente. Luego fíjese en el límite superior del IC del 95% y si este es mayor al que usted pensaba, puede considerar al trabajo concluyente.

Conclusión

Los pasos a seguir en la lectura crítica de un artículo son:

1. Una adecuada pregunta.
2. Búsqueda de acuerdo a esta pregunta.
3. Hallazgo de un artículo y determinación de su validez interna.
 - La validez interna se orientara a la búsqueda de sesgos y confundidores, que dependerán del tipo de trabajo (ensayo clínico, estudio de cohorte, casos y controles, estudios de corte transversal etc).
4. Determinación de la importancia del artículo
 - Magnitud del efecto (medidas absolutas o relativas) RR, OR, RRA, NNT.
 - Precisión de los resultados (valor de p y/o IC).
5. Aplicabilidad al paciente (Validez externa).

Bibliografía

1. Hernández-Avila M, Garrido F, López Moreno S. Diseño de estudios epidemiológicos. *Salud Pública de México* 2000; Vol 42 N°2.
2. Hennekens CH, *Epidemiology in Medicine*, Ed Little Brown and company 1987.
3. Suissa S, Effectiveness of inhaled corticosteroids in chronic obstructive pulmonary disease *Am J Respir Crit Care Med* 2003; 168: 49-53.
4. Suissa S, Observational studies of inhaled corticosteroids in chronic obstructive pulmonary disease: misconstrued immortal Time Bias *Am J Respir Crit Care Med* 2006; 173: 464.
5. D'Alonzo GE, Barst R, Ayres SM, Bergofsky EH, Brundage BH, Detre KM, Fishman AP, Goldring RM, Groves BM, Kernis JT, et al. Survival in patients with primary pulmonary hypertension. Results from a national prospective registry. *Ann Intern Med*: 1991; 115: 343-349.
6. Augustovsky F. Como leer un artículo de terapéutica; *Evidencia en atención Primaria* 2000; Vol 3 N° 1: 28-31.
7. Mitchell H. Katz Multivariable analysis: A primer for readers of medical research *Ann Internal medicine* 2003; 138: 644-650.
8. Black N, Why we need observational studies to evaluate the effectiveness of health care *BMJ* 1996; 312: 1215-8.
9. Gurwitz JH, Sykora K, Mamdami M, Streiner D, Garfinkel S, Sharon-Lise T, Normand, Anderson G, Rochon P. Reader's guide to critical appraisal of cohort studies: 1 Role of design *BMJ* 2005; Vol 330: 895-7.
10. Mamdami M, Sykora K, Streiner D, Garfinkel S, Normand SL, Anderson G, Rochon P. Reader's guide to critical appraisal of cohort studies: 2 Assessing potential for confounding *BMJ* 2005; Vol 330: 960-2.
11. Sykora K, Ping Li, Mamdami M, Rochon P, Anderson G. Reader's guide to critical appraisal of cohort studies: 3 Analytical strategies to reduce confounding *BMJ* 2005; Vol 330: 1021-3.
12. Schultz K, Grimes D. Generation of allocation sequences in randomized trials ; chance, not choice *Lancet* 2002; Vol 359: 515-519.
13. Schultz K, Grimes D. Allocation concealment in randomized trials: defending against deciphering. *Lancet* 2002; 359: 614-618.
14. Schultz K, Grimes D. Blinding in randomized trials: hiding who what *Lancet* 2002; 359: 614-619.
15. Schultz K, Grimes D. Sample site slippages in randomized trials: exclusion and the lost end wayward *Lancet* 2002; 359: 781-785.
16. Hujoel PP, Drangsholt M, Spiekerman C, Deroguen TA. Periodontal disease and coronary heart disease risk *JAMA* 2000; 284: 1406-10.
17. Abboud RT, Ford GT, Chapman KR. Emphysema in alpha-1 antitrypsin deficiency. *Treat Respir Med* 2005; 4:1-8.
18. Medical Research Council Streptomycin treatment for pulmonary Tuberculosis. *Br Med J* 1948; ii: 769-82