

Como leer un artículo

Estudios de terapéutica o prevención

El objetivo de esta nueva sección de la revista es ofrecer a nuestros lectores una guía de lectura crítica de los distintos tipos de artículos que uno encuentra en la literatura. EVIDENCIA selecciona para comentar sólo artículos relevantes y de buena calidad metodológica, llegándole al lector información de buena calidad. Pero somos conscientes de que nuestros ávidos lectores están expuestos a todo tipo de literatura médica, y deben tomar decisiones a diario acerca de la validez y relevancia de la información que encuentran.

Las guías de lectura crítica nos ayudan a aprovechar mejor el tiempo. La idea es que siguiendo estos lineamientos, usted pueda decidir rápidamente el destino del artículo en cuestión: incorporarlo a su práctica o descartarlo.

Esperamos que esta nueva sección pueda ser de utilidad para hacer de la lectura crítica una actividad cotidiana para nuestro beneficio y el de nuestros pacientes.

Comenzamos inicialmente con la guía de lectura crítica de artículos de terapéutica, también aplicable a los de prevención, por ser los más frecuentemente hallados en la literatura.

Luego de hacernos una pregunta en base a un paciente, hicimos una búsqueda e identificamos un artículo que nos interesa evaluar para saber si nos puede responder la pregunta inicial (para conocer bien estos pasos previos antes de llegar al artículo, referirse a la sección EOPS).

A continuación iremos viendo paso a paso si podemos aplicar los resultados de ese artículo en nuestro paciente.

La mejor calidad de evidencia para respondernos preguntas de terapéutica o prevención la aportan los ensayos clínicos controlados aleatorizados y los meta-análisis (revisiones sistemáticas de los ensayos aleatorizados)

En esta guía nos referiremos a estudios originales, más adelante publicaremos una guía de lectura crítica de revisiones sistemáticas.

En la figura 1 vemos los pasos necesarios para decidir sobre un artículo:

- | | |
|------|--|
| I. | Validez: Este aspecto se refiere a la metodología del estudio; si fue realizado adecuadamente, sin errores, lo que hace que las conclusiones sean verdaderas (validez interna). |
| II. | Importancia: Una vez que evaluamos la validez de un estudio, nos debemos preguntar si los resultados son relevantes. Puede haber estudios muy válidos cuyos resultados no nos aporten gran cosa. |
| III. | Aplicación: Este es el paso de “bajar al consultorio” y ver si se pueden aplicar sus resultados para ayudar a nuestro paciente. |

En la figura 2 vemos de manera más detallada los pasos a seguir.

I. ¿Es un estudio válido? (validez interna)

Preguntas esenciales (si no se cumplen descartar el estudio)

1. ¿Fue aleatorizada la asignación de los tratamientos? ¿Estaban ocultos los datos de la aleatorización?

La *aleatorización* es la asignación de los tratamientos en base al azar. Es el mejor método existente para crear grupos lo más parecidos posible al comienzo de un estudio. El balance así obtenido incluye tanto factores conocidos (sexo, edad, factores pronósticos, estadio de enfermedad, etc) como desconocidos (otras características no identificadas que puedan estar asociadas con el evento en estudio). La idea es que la única diferencia existente entre los grupos sea el tratamiento, lo que permite atribuirle las eventuales diferencias que puedan encontrarse. Por otra parte, es una manera de garantizar que los investigadores no participen en la asignación de tratamiento.

A pesar de que es fácil identificar un estudio aleatorizado (si no figura en el título, figura en el resumen casi con seguridad), el *ocultamiento* de la secuencia de aleatorización es más difícil de detectar y muchas veces no está reportado. Este ocultamiento hace que los investigadores no sepan a qué grupo se va a asignar el próximo paciente, y por lo tanto no pueden introducir sesgos, conscientes o inconscientes. Lo más seguro en este sentido es la aleatorización a distancia o telefónica. El efecto del no ocultamiento es que se tienden a asignar a los pacientes con mejor pronóstico al grupo experimental, exagerando así los posibles beneficios.

2. ¿Se detalló el seguimiento de todos los pacientes incluidos en el estudio? ¿Se analizó a cada paciente según el grupo al que fue aleatorizado?

Una vez comprobado que el estudio fue aleatorizado o randomizado, hay que observar bien el número de pacientes que entró al estudio y el número al que se aplican los resultados. Idealmente tienen que ser idénticos (aunque rara vez pasa). Los pacientes perdidos en el seguimiento pueden haber tenido eventos que cambien las conclusiones. En general se recomienda como guía que el *seguimiento es adecuado si es $\geq 80\%$* .

Una de las maneras de saber si las pérdidas afectan las conclusiones del estudio es realizar el "análisis del peor escenario". Este responde a la pregunta de qué pasaría si a todos los pacientes perdidos del grupo que le fue bien les asignáramos el peor resultado. Si las diferencias aún persisten en este escenario, uno puede estar tranquilo que las pérdidas no afectaron las conclusiones principales.

Cualquier cosa que suceda luego de la aleatorización puede alterar el resultado (dejar de tomar la medicación, tomar otra, la intensidad del seguimiento, etc). Por eso es fundamental que TODOS los pacientes (incluso los que no cumplen con el tratamiento asignado), sean analizados en los grupos a los que fueron randomizados. Esta forma correcta de análisis se denomina *análisis por intención de tratar*.

Si recordamos que el valor de la randomización es balancear los grupos al azar, si no se analiza a los pacientes en el grupo asignado se está quebrando la aleatorización (ya que las pérdidas, abandonos, etc. no son necesariamente al azar), y corre peligro la validez del estudio. Lo que estoy evaluando es 'la intención de dar un tratamiento', que es lo que uno hace habitualmente en el consultorio. Que nosotros tengamos esa intención no significa que el paciente la cumpla al pie de la letra (compre el medicamento, lo tome correctamente, etc.)

Preguntas 'secundarias' (le agregan validez pero muchas veces no son esenciales)

3. ¿Se mantuvieron ciegos los pacientes y los médicos?

La *ceguera* en los estudios se refiere al desconocimiento del tratamiento asignado. En los estudios *simple ciego* el único que no sabe es el paciente, mientras que en los *doble ciego* desconocen la asignación tanto el médico como el paciente.

Los beneficios que aporta es lograr que ni los síntomas reportados por los pacientes ni la interpretación de los mismos por parte de los médicos esté influenciada por el hecho de conocer el tratamiento al cual fueron asignados.

Es importante notar que no todos los estudios pueden ser ciegos, como por ejemplo un estudio que evalúa una técnica quirúrgica, u otro que evalúe la efectividad de la psicoterapia. En esos casos lo más importante es que quienes evalúen el resultado final sí estén ciegos al tratamiento recibido.

4. ¿Fueron similares las cointervenciones?

Esta pregunta apunta a ver si hubo otras diferencias entre los grupos aparte del tratamiento asignado, como la intensidad del seguimiento, la utilización de otras intervenciones, etc, que puedan influir en los resultados. Una ventaja adicional del doble ciego es que los médicos les ofrezcan a los pacientes seguimientos y cointervenciones similares, lo que evita la introducción de sesgos.

5. ¿Eran los grupos comparables al inicio del estudio?

La manera práctica de chequear la efectividad de la aleatorización es mirando la "Tabla 1" del estudio, y observar si las características de los pacientes (sexo, edad, otros factores de interés) son similares en los distintos grupos. Esto es particularmente importante cuando el número de participantes en el estudio es pequeño.

II. ¿Son los resultados importantes?

Una vez analizado el estudio, y solo si concluye que el mismo fue válido y se puede confiar en sus resultados, debemos evaluar su importancia. En este sentido es fundamental evaluar no solo el efecto de la intervención sino también su *magnitud*.

Para esto vamos a utilizar dos casos clínicos:

Caso1

Nos estamos planteando la posibilidad de administrar aspirina profiláctica a un hombre adulto fumador con el objeto de prevenir un infarto agudo de miocardio (IAM). Utilizaremos los datos del Physician Health Study, en el que 22.071 pacientes tratados con aspirina o placebo fueron seguidos por 60 meses.

- A la proporción de IAM en el grupo tratado a lo largo del seguimiento se la denomina incidencia en el grupo tratado o *Tasa de Eventos en el grupo Tratado* (TET) = 0.012 (1.2%)
- A la proporción de IAM en el grupo placebo, *Tasa de Eventos en el grupo Control* (TEC) = 0.022 (2.2%)

Caso 2

Atendemos a una mujer de 75 años con cifras elevadas de tensión arterial sistólica (TA 170/80). ¿Debe recibir tratamiento antihipertensivo? Encontramos y analizamos el estudio SHEP, que comparaba el cuidado escalonado con el placebo a cinco años de seguimiento y la incidencia de accidentes cerebrovasculares (ACV).

TET = 0.052 (5.2%) vs. TEC = 0.082 (8.2%) Placebo

Ahora bien, tenemos los resultados de cada rama del estudio; ¿de qué manera se puede extraer una medida sumaria que resuma el resultado del estudio?

Una de las maneras de expresar los resultados (la más frecuente) es *en términos relativos*: ¿cuánto reduce la incidencia de IAM la aspirina en relación al placebo?

Reducción de Riesgo Relativo (RRR) = (TEC-TET) / TET

Caso1

Aspirina = $(0.022 - 0.012) / 0.022 = 45\%$

Caso2

HTA Sistólica = $(0.082 - 0.052) / 0.082 = 37\%$

O sea que la aspirina reduce un 45% los IAM en relación al placebo y el tratamiento de la HTAS un 37% los ACV.

La RRR es una medida de utilidad para el paciente, ya que es fácil de entender que la aspirina “me reduce el riesgo de IAM un 45%”.

Si solo dispusiéramos de la RRR podríamos creer que ambos tratamientos son similarmente efectivos ya que la RRR es similar. Este es el principal peligro del *RRR* ya que *no nos dice nada acerca del impacto absoluto del tratamiento*, que como veremos es muy distinto en ambos casos.

Otra de las formas de expresar los resultados de un estudio es *en términos absolutos*: ¿cuánto reduce la aspirina la incidencia de IAM comparado con el placebo, o cual es el beneficio atribuible a la misma?

Reducción de Riesgo Absoluto (RRA) o Riesgo Atribuible.

RRA = TET - TEC

Aspirina = $0.022 - 0.012 = 0.01$ o 1%

HTAS = $0.082 - 0.052 = 0.03$ o 3%

Medido con esta vara se observa que el impacto absoluto es muy distinto entre una y otra intervención: mientras la aspirina previene en términos absolutos el 1% de los IAM, el tratamiento de la HTAS previene el 3% de los ACV (su impacto en términos reales es tres veces mayor). ¡Recuerden qué similar era el RRR y cuan distinto es el impacto real!.

Pero busquemos una manera de expresar los resultados más clara e intuitiva y que dé una idea del impacto real del tratamiento. Esta medida es relativamente reciente y muy utilizada en la actualidad (en EVIDENCIA la usamos frecuentemente).

Número Necesario para Tratar

$$\text{NNT} = 1 / \text{RRA} \quad (\text{o } 100 / \text{RRA en } \%)$$

$$\text{Aspirina} = 1 / 0.01 = 100 \quad (\text{o } 100/1)$$

$$\text{HTAS} = 1 / 0.03 = 33 \quad (\text{o } 100/3)$$

El NNT expresa el número de personas que hace falta tratar con el tratamiento activo para prevenir un evento en relación al placebo durante el periodo de seguimiento del estudio.

Necesito tratar a 33 pacientes con HTAS durante cinco años para prevenir un ACV, en cambio debo darle aspirina preventiva a 100 adultos por seis años para prevenir un IAM. Esta medida es de utilidad principalmente para los médicos, para entender rápidamente el impacto del tratamiento en la población.

1. ¿Cuál es la magnitud del efecto?

Ya vimos las distintas maneras de expresar los resultados de un estudio, y la utilidad de cada una.

La magnitud del efecto (que expresa la relevancia clínica) la evaluamos según el tamaño de estas medidas:

- La RRR cercana al 100% tiene más magnitud, y la cercana al 0% implica que el tratamiento es inefectivo
- A mayor RRA mayor es la magnitud del efecto (en términos absolutos), y a la inversa
- Cuanto más pequeño es el NNT (más cercano a uno), mayor es el impacto de la intervención.

2. ¿Cuál es la precisión de los resultados?

Como ocurre con todas las estimaciones en medicina, los resultados de los estudios no son exactos, por lo que es imprescindible no solo conocer el resultado obtenido sino también cuán precisa es esta estimación (o el nivel de confianza que puedo tener).

Para todas las medidas mencionadas se puede calcular (y en general se reporta en los estudios) el *Intervalo de Confianza*. Esta medida expresa un rango de valores dentro del cual uno espera encontrar, con un determinado grado de confianza, el *verdadero* valor que representa la relación entre las variables estudiadas. El intervalo de confianza del 95% (el más utilizado, IC 95%) responde a lo siguiente: si yo repitiera este estudio 100 veces, en 95 de ellas el resultado se encontraría en ese intervalo. Es por ello que cuanto más angosto es ese intervalo, más precisión tiene la estimación.

Tomemos el caso de la HTAS:

Si el estudio reportara un RRR = 37%, IC95% -5 a 90%, ¿qué concluiría Ud?; ¿y si reportara una RRR = 37%, IC95% 30 a 45%?

Desde ya, uno le tiene más confianza al segundo intervalo: no solo es más preciso (angosto) sino que no es compatible con la posibilidad de que el tratamiento no sea efectivo (todo el intervalo está por encima del cero (valor de nulidad), que sería el valor en el que ambos tratamientos son iguales.

En cambio, el primer IC es poco preciso (rango muy amplio) y además incluye al valor de nulidad. Esto significa que puede hasta no existir diferencia entre los grupos y por ende, el resultado obtenido en el estudio no nos da una información confiable.

Entonces la conclusión en base al primer IC es que la RRR de 37% podría ser clínicamente interesante pero no es significativa. No puedo concluir que un tratamiento supera al otro (incluye al cero). La conclusión en base al segundo IC es que la RRR es de 37% y el tratamiento es claramente superior (no incluye al cero). La interpretación es que en base a este estudio podemos decir, con un grado de confianza del 95%, que el tratamiento es *al menos* un 30% y *hasta un* 45% superior al placebo.

III ¿Puedo aplicar estos resultados válidos e importantes a mi paciente?

Una vez que decidió que el estudio es válido y relevante, el próximo paso es decidir si es posible aplicarlo en nuestro paciente.

1. ¿Son aplicables estos resultados a mi/s paciente/s?

Los pacientes de los ensayos clínicos son en general una población muy seleccionada, por lo que es importante evaluar si los resultados se pueden aplicar a mi paciente. Aquí una pregunta útil para hacernos es:

- ¿Es mi paciente tan diferente a los pacientes del estudio que no puedo aplicar los resultados? Fundamentalmente se debe evaluar si los participantes del estudio tienen características particulares que impiden trasladar los resultados a nuestra práctica. Esto define la *validez externa* o la posibilidad de generalizar las conclusiones del estudio. Estas características pueden ser biológicas, socioeconómicas, epidemiológicas, etc.

Si no existiera ninguna razón de peso que comprometa la validez externa, uno puede concluir que los resultados de la intervención serán similares en mis pacientes

- Análisis de subgrupos

Se denomina análisis de subgrupos cuando uno reporta separadamente resultados de distintas subpoblaciones de pacientes (hombres vs. mujeres, jóvenes vs. ancianos, etc.). Salvo que sean análisis previamente especificados (y se haya realizado el diseño del estudio teniéndolos en cuenta) *se debe tomar el resultado principal de toda la población del estudio*. Por una cuestión estadística, aunque también de sentido común, cuanto más subgrupos analizo, mayor es la probabilidad de encontrar algún resultado positivo *solamente debido al azar*. Es por ello que uno debe estar muy atento a este tipo de análisis, que, aunque puede sugerir hipótesis de trabajo interesante, no suele cambiar nuestra conducta.

- ¿Cuan grande sería el beneficio potencial en mi paciente?

Esta es un area que puede ser más imprecisa, ya que hay que aplicar el ojo clínico para estimar el riesgo individual de mi paciente.

Si uno quiere estimar cual podría ser el NNT en pacientes similares al mío, debemos estimar una medida que resuma una idea del riesgo que tiene de padecer el evento de interés *en relación a los pacientes del grupo control* del estudio. Así, si yo considero que mi paciente tiene un perfil de riesgo más alto que los pacientes del estudio, la *fracción de riesgo* será mayor a uno (dos si creo que tiene el doble de riesgo, tres el triple, etc.); y si es un paciente de menor riesgo que el de los pacientes del estudio, tendrá una *fracción de riesgo* menor a uno.

Tomemos un ejemplo: paciente de 75 años, con HTAS, múltiples factores de riesgo cardiovascular, y con antecedentes de un ataque isquémico transitorio hace un año. Si lo comparamos con la población incluida en el estudio SHEP observamos que nuestro paciente tiene un riesgo individual más alto de ACV que el promedio de los pacientes del grupo control (digamos = 3 o el triple del riesgo basal.

Luego, para estimar el NNT, divido el NNT del estudio por la fracción de riesgo de mi paciente: $NNT_i = NNT/F$
--

Si recordamos el NNT del estudio (33), el NNT “de mi paciente” será $33/3 = 11$. En este caso, debería tratar a 11 (en vez de a 33) para prevenir un ACV a los cinco años, lo que sugiere un mayor impacto en los pacientes de mayor riesgo.

2. ¿Respetar el tratamiento y sus consecuencias las preferencias y valores de mi paciente?

Estas serían las últimas dos preguntas a considerar antes de decidir aplicar el tratamiento en cuestión:

- ¿Tiene Ud y el paciente una evaluación clara de valores y preferencias?
- ¿Son estos valores cumplidos por el régimen terapéutico y sus consecuencias?

Es muy importante evaluar si el potencial beneficio de la intervención supera el riesgo de efectos adversos. Este es un terreno donde es preciso conocer y charlar con nuestro paciente acerca de sus valores y preferencias, así como también las del propio médico con respecto al tratamiento en consideración. Esto es especialmente difícil cuando se consideran intervenciones invasivas, o con efectos adversos que alteren la calidad de vida de nuestros pacientes..

Federico Augustovski
Unidad de Medicina Familiar y Preventiva
Hospital Italiano de Buenos Aires

Referencias

- Sackett DL y cols. Evidence-based Medicine. How to practice & teach EBM. Churchill Livingstone 1997.
- Evidence Based Medicine Working Group. Evidence Based Medicine. A new approach to teaching the practice of medicine. *JAMA* 1992;268:2420-2425
- Guyatt GH. Users' guides to the medical literature. *JAMA* 1993; 270 (17) : 2096-2097.
- Oxman A, Sackett, DL & Guyatt GH. Users' guides to the medical literature. I. How to get started. *JAMA* 1993 Nov 3; 270 (17) : 2093-2095. En Internet: http://HIRU.MCMASTER.CA/ebm/userguid/1_intro.htm
- Guyatt GH, Sackett DL Cook DJ for the Evidence-Based Medicine Working Group. Users' Guides to The Medical Literature: II – How to use an article about therapy or prevention A. Are the results of this study valid? *JAMA* 1993;270:2598-601
- Guyatt GH, Sackett DL Cook DJ for the Evidence-Based Medicine Working Group. Users' Guides to The Medical Literature: II – How to use an article about therapy or prevention B. What were the results and will they help caring for my patients? *JAMA* 1994;271:59-63
- Dans AL, Dans LF, Guyatt GH, Richardson S. Users' guides to the Medical Literature. XIV. How to decide on the applicability of clinical trial results to your patient. Evidence Based Medicine Working Group. *JAMA* 1998; 279 (7): 545-549.

Figura 2. Pasos a seguir en la lectura crítica de un estudio de prevención o terapéutica

<p>¿Es un estudio válido? (validez interna)</p> <p>Preguntas esenciales (si no se cumplen descartar el estudio)</p> <ol style="list-style-type: none">1. ¿Fue aleatorizada la asignación de los tratamientos? ¿Estaban ocultos los datos de la aleatorización?2. ¿Se detalló el seguimiento de todos los pacientes incluidos en el estudio? ¿Se analizó a cada paciente según el grupo al que fue aleatorizado? <p>Preguntas 'secundarias' (le agregan validez pero muchas veces no son esenciales)</p> <ol style="list-style-type: none">3. ¿Se mantuvieron ciegos los pacientes y los médicos?4. ¿Fueron similares las cointervenciones?5. ¿Eran los grupos comparables al inicio del estudio? <p>II. ¿Son los resultados importantes?</p> <ol style="list-style-type: none">1. ¿Cuál es la magnitud del efecto? (RRR, RRA, NNT)2. ¿Cuál es la precisión de los resultados? (IC) <p>III ¿Puedo aplicar estos resultados válidos e importantes a mi paciente?</p> <ol style="list-style-type: none">1. ¿Son aplicables estos resultados a mi/s paciente/s?<ul style="list-style-type: none">▪ ¿Es mi paciente tan diferente a los pacientes del estudio que no puedo aplicar los resultados?• ¿Cuan grande sería el beneficio potencial en mi paciente?2. ¿Respeto el tratamiento y sus consecuencias las preferencias y valores de mi paciente?<ul style="list-style-type: none">• ¿Tiene Ud y el paciente una evaluación clara de valores y preferencias?• ¿Son estos valores cumplidos por el regimen terapéutico y sus consecuencias?
