

Lectura crítica de un artículo sobre tratamiento

Autora: Pilar Gayoso Diz. Especialista en Medicina Familiar y Comunitaria. Coordinadora de la Unidad Docente de M. Familia. Unidad de Investigación. Sergas de Ourense.
Grupo MBE Galicia, integrado en la Red Temática de Investigación sobre Medicina Basada en la Evidencia (Expediente FIS: G03/090)
Actualización 24/10/2003.

Escenario, formulación de la pregunta y estrategia de búsqueda hasta seleccionar un artículo.

Escenario

Manuel tiene 68 años, hace 3 años se ha jubilado y desde entonces pasa bastante tiempo en casa ó con los amigos en el bar (su partidita es sagrada). Hace 4 meses acudió a la consulta por cefaleas frecuentes de características tensionales. En la anamnesis y exploración física destacaba un IMC=24, TA sistólica=155 y TA diastólica=100, cifras que se confirmaron en tres ocasiones. La bioquímica solicitada presentaba cifras de LDL-colesterol de 230 mg/dl, glucemia de 85 mg/dl y función renal dentro de la normalidad. En el electrocardiograma se apreciaban signos de hipertrofia ventricular izquierda que la ecografía confirmó.

Tras explicarle que presentaba hipertensión e hipercolesterolemia, estableciste con él una dieta pobre en sodio y grasas, ejercicio diario (30 mn. de paseo) citándolo para un control pasados 3 meses. En ese momento, Manuel quería que le prescribieras unas pastillas; según te contó dos de sus amigos estaban tomando... (un IECA) y tenían su tensión muy bien controlada, “¿por qué no me lo da a mi, doctor?”

Le explicaste las razones por las que convenía empezar cambiando los hábitos de vida y quedáis en que según tenga la tensión en el próximo control valorareis el dar además “pastillas”.

Vendrá dentro de una semana y sabes que probablemente pedirá esas pastillas que tan bien van a sus amigos. Tu piensas que deberías iniciar el tratamiento farmacológico con un diurético, pero hace poco en una sesión del centro se planteó esta cuestión y varios compañeros (que siempre están a “la última”) presentaron algunos artículos que apoyaban la recomendación inicial del tratamiento con IECAs ó ARA II. Sostenían que los diuréticos estaban ya superados.

Ante tus dudas, decides revisar los resultados de trabajos recientes sobre el tema.

Pregunta

Para hacer más eficiente la búsqueda de información, debe estructurarse la duda que se tiene en forma de pregunta. Cuanto mas concreta y específica sea la pregunta, mayor será la probabilidad de hallar una adecuada respuesta en los trabajos identificados. Este es un aspecto crítico, ya que es la base del trabajo posterior.

Una pregunta bien formulada debe constar de las siguientes partes:

- Sujeto (características del paciente que nos preocupa),
- Intervención (cuya eficacia ó efectividad deseamos valorar incluyendo frente a qué se compara),
- Resultado (variable por la que se quiere medir la eficacia ó efectividad de la intervención).

Eficacia es la magnitud en la que una intervención (tratamiento, procedimiento o servicio) mejora los resultados para los pacientes **en condiciones ideales**.

Efectividad es la magnitud en la que una intervención (tratamiento, procedimiento o servicio) mejora los resultados para los pacientes **en la práctica**.

En nuestro caso la pregunta podría ser:

“¿En un paciente varón mayor de 65 años hipertenso (SUJETO), iniciar el tratamiento con diuréticos (INTERVENCIÓN) mejora su riesgo cardiovascular (RESULTADO) en igual ó mayor medida que si se trata con IECAs (COMPARADOR)?”

Estrategia de búsqueda

Cuando nuestra pregunta trata sobre un aspecto de eficacia de una intervención ó tratamiento (de cualquier tipo: educacional, con medicamentos, con terapia física, intervención quirúrgica, etc), disponemos de escalas de calidad de la evidencia que proporcionan diferentes diseños de estudio.

En todas ellas, el **ensayo clínico aleatorizado y controlado (ECA)** proporciona resultados de la mayor calidad, siempre que su diseño y realización haya sido riguroso. Se ha evidenciado que los resultados obtenidos en ECAs realizados con rigor proporcionan resultados menos sesgados en relación a otro tipo de diseños (estudios de cohortes, estudios casos-control...).

Para medir la eficacia de una intervención el ensayo clínico aleatorizado y controlado (ECA), diseñado y realizado de manera adecuada, proporciona los resultados de la mejor calidad y con el menor sesgo.

En un ECA los sujetos son aleatoriamente (cada sujeto tiene la misma probabilidad de ser asignado a uno u otro grupo de comparación, independientemente de los demás sujetos) asignados (al menos) a **dos grupos** uno (grupo experimental) recibe el tratamiento que se está probando y el otro (grupo de comparación o control recibe un tratamiento alternativo). Los dos grupos son seguidos para observar cualquier **diferencia** en los resultados. Así se evalúa la **eficacia** del tratamiento experimental.

La eficacia del tratamiento experimental en los ECA se evalúa a través de las diferencias observadas en los grupos constituidos por los sujetos que integran el estudio distribuidos de manera aleatoria, uno es el grupo intervención y otro es el grupo control.

Para responder a preguntas sobre tratamiento debemos buscar:

- En primer lugar revisiones sistemáticas de ensayos clínicos aleatorios. Presentan mayor calidad las realizadas a partir de los datos originales de los ECAs que aquellas hechas a partir de los datos publicados.
- Si no encontramos ninguna válida, un ECA sobre el tema de interés.

Aunque sobre el tipo de diseño, está el criterio de calidad metodológica; es más riguroso basar nuestra toma de decisiones con el paciente en un excelente estudio de cohortes que en un defectuoso ensayo clínico aleatorizado.

Niveles de calidad de la evidencia científica (a) Tomado de U.S. Preventive Task Force, 1996.	
I	Evidencia obtenida a partir de al menos un ensayo aleatorizado y controlado diseñado de forma apropiada.
II-1	Evidencia obtenida a partir de ensayos controlados no aleatorizados y bien diseñados.
II-2	Evidencia obtenida a partir de estudios de cohorte o caso-control bien diseñados, realizados preferentemente en más de un centro o por un grupo de investigación.
II-3	Evidencia obtenida a partir de múltiples series comparadas en el tiempo con o sin intervención (b).
III	Opiniones basadas en experiencias clínicas, estudios descriptivos o informes de comités de expertos.
(a) De mayor (I) a menor (III) calidad de diseño y rigor científico. (b) Este tipo de evidencia también incluye resultados «dramáticos» producidos en experimentos incontrolados, como son los derivados de la introducción de la penicilina en los años cuarenta.	

Ese día al acabar la consulta, entras en PubMed para hacer la búsqueda bibliográfica. Como sabes, se trata de la principal base bibliográfica en ciencias de la salud, aunque no es la única (Índice Médico Español, Ovid, Bireme, etc); abarca más de 3000 publicaciones periódicas y es de acceso libre y gratuito a través de internet.

Puedes acceder directamente a PubMed (www.ncbi.nlm.nih.gov/PubMed) ó utilizar alguna de las pantallas de acceso disponibles en castellano como la de la página web de Rafael Bravo Toledo (www.infodoctor.org/rafabravo/busmedline) o www.fisterra.com.

Debes emplear los términos (**descriptores**) en inglés y, si vas a utilizar más de un término de búsqueda, asociarlos mediante los operadores booleanos (AND, OR, etc).

En este caso empleando *hipertensión AND diuretic* obtienes 8492 citas.

Para hacer más accesible la tarea, decides usar un **filtro metodológico**. Se trata de estrategias de búsqueda muy elaboradas diseñadas por expertos documentalistas y seleccionadas por haber demostrado ser las más eficientes para identificar determinado tipo de trabajos.

Sobre el tema de búsqueda nos permite seleccionar:

- si nos interesan aspectos de diagnóstico, tratamiento, pronóstico ó etiología;
- si deseamos que la búsqueda sea más sensible (muy amplia, no se “escapa” ningún trabajo pero habrá mucha “paja”) ó más específica (artículos que se ajustan a los criterios de búsqueda aunque dejando “fuera” otros).

En PubMed los puedes encontrar como Clinical Query; están diseñados por expertos de la Universidad de McMaster bajo la dirección de B. Haynes.

En la página de Rafa Bravo indicada se encuentran en la parte inferior de la pantalla.

Empleando como filtro metodológico aspectos sobre tratamiento y una búsqueda específica, con los mismos descriptores *hipertensión AND diuretic*, recuperas 1148 citas (observa el efecto del filtro utilizado). El filtro seleccionado hace una búsqueda específica de estudios que reúnan las características de ECA, ya que se basan en las consideraciones antes señaladas (observa los descriptores que se han añadido cuando haces la búsqueda con el filtro de tratamiento).

Revisando los resúmenes, identificas un artículo publicado en diciembre de 2002 que te parece puede ayudarte a resolver las dudas: se trata de un ECA que compara la eficacia de los diuréticos (frente a IECAs y calcioantagonistas).

“Major outcomes in high-risk hypertensive patients randomized to Angiotensin-Converting Enzyme Inhibitor or calcium channel blocker versus diuretic” JAMA 2002;288:2981-2997.

Desde el propio resumen en PubMed puedes acceder gratuitamente al artículo desde el enlace que aparece como:

JAMA

Una vez en el mismo, puedes recuperarlo en formato PDF (resulta más clara la disposición del texto y tablas) y sacarlo por la impresora para leerlo en casa con mas calma.

Bases para establecer si un ensayo clínico es válido

Validez se refiere a **validez interna**. Cuando se realiza la lectura crítica de un trabajo, lo primero que debe ocuparnos es establecer su validez interna, definida como la adecuación metodológica de su diseño y desarrollo que garantiza que sus resultados sean no sesgados y por tanto nos permitan una

La lectura crítica de un trabajo sobre tratamiento se inicia evaluando si el diseño es metodológicamente adecuado y el desarrollo garantiza resultados no sesgados, esto es: **validez interna**.

buena estimación de la eficacia real de la intervención de estudio. Cuanto más válido sea un estudio, mejor se aproxima a la verdad en sus resultados.

Para analizar la validez de un ensayo clínico aleatorizado y controlado, se dispone en la literatura de algunas guías que, basándose en aquellos aspectos metodológicos cuyo cumplimiento mejor garantizan la ausencia de sesgos en las estimaciones, permiten al clínico tomar decisiones sobre ello.

Para este ejercicio tomaremos como base la guía elaborada por el grupo CASPe (Critical Appraisal Skill Program España, www.redcaspe.org) para lectura crítica de un artículo sobre tratamiento que puede obtenerse gratuitamente en la citada web.

En general, estos aspectos deben buscarse en el apartado “Material y método” de un artículo original cuya lectura resulta imprescindible para tomar decisiones sobre la validez interna del estudio.

Aspectos esenciales para establecer la validez interna de un ECA.

Por su importancia, son tres los aspectos a analizar en primera instancia. Si no se puede responder afirmativamente a ellos, la validez del estudio es cuestionable y resulta más recomendable buscar otro trabajo para resolver la duda clínica que se tiene.

1.- Pregunta claramente definida.

Resulta imposible tomar decisiones sobre la validez interna de un ensayo y también sobre su aplicabilidad (ó validez externa) si no se explicita claramente: tipo de pacientes incluidos; intervención cuya eficacia se estudia (y en comparación con qué); variables resultado mediante las que se evalúa la eficacia.

Pregunta bien definida supone explicitar pacientes incluidos, definir intervención a estudio y variables usadas para medir resultados.

En nuestro ejemplo: JAMA 2002; 288: 2981-2997.

Pacientes.- Se incluyen 33.357 personas de 55 ó más años, de ambos sexos, diagnosticados de HTA grado I/II y al menos 1 factor de riesgo (FR) cardiovascular (IAM ó ACV en los 6 meses anteriores; hipertrofia de ventrículo izquierdo (HVI) documentada por electrocardiograma ó ecocardiografía; D. Mellitus tipo II; hábito tabáquico; HDL-colesterol <35 mg/dl).

Intervención.- Se comparan tres grupos: clortalidona (12.5 a 25 mg/dl) , amlodipino (2.5 a 10 mg/dl) ó lisinopril (10 a 40 mg/dl). Fueron administrados solos en una primera etapa y en combinación con un 2º fármaco en régimen abierto si era precisa la asociación en un 2º y 3º paso. El objetivo era obtener cifras tensionales bajo 140/90 mm Hg.

Resultado.- El resultado primario (aquel por el que se va a evaluar la eficacia) a medir era una variable combinada de ocurrencia de IAM no mortal ó coronariopatía mortal.

Se establecieron antes de iniciar el estudio además cuatro resultados secundarios: mortalidad por todas las causas; ACV (cause o no la muerte del paciente); coronariopatía combinada (ocurrencia del resultado primario, revascularización coronaria o ingreso hospitalario por angina); enfermedad cardiovascular combinada (ocurrencia de la variable anterior, ACV, angina, fallo cardíaco ó enfermedad vascular periférica).

Como resultados de seguridad principales se evaluó: hospitalización por sangrado digestivo y angioedema.

Otras variables clínicas incluyeron: cancer, HVI de nueva aparición, insuficiencia renal crónica terminal y creatinina sérica.

Es interesante señalar que las variables resultado por las que se mide la eficacia de la intervención son variables clínicas, con clara repercusión en la supervivencia y calidad de vida de los pacientes. Aunque se

trata de un estudio sobre efecto de fármacos antihipertensivos, su eficacia no se establece en función de resultados intermedios (como podría ser el descenso de las cifras tensionales), sino de resultados clínicos finales.

La toma de decisiones clínicas para aplicar a pacientes individuales, debe basarse preferentemente en estudios de riguroso diseño que midan variables resultado clínicas. En otro caso, los beneficios potenciales para el paciente pueden ser inferidos (fisiopatología, mecanismo de acción del fármaco) pero no se dispone de evidencia fiable sobre el mismo.

2.- Asignación aleatoria (con ocultación de la secuencia de aleatorización)

Por esta característica, cada uno de los pacientes que acepta participar en el ECA, después de recibir la información completa, tiene la misma probabilidad de recibir cualquiera de los tratamientos a estudio. Con independencia del grupo a que se haya asignado el paciente anterior.

Con ello se busca evitar **sesgos de selección**, procurando que los grupos de pacientes así formados presenten una distribución similar de sus características (tanto aquellas que se sabe pueden condicionar el resultado del tratamiento, como todas las demás). Como consecuencia se puede relacionar las diferencias en las variables resultado con las diferencias en la intervención estudiada (puesto que los grupos eran similares al inicio del estudio).

Se dispone de tablas de números aleatorios y de programas informáticos que generan listados de números pseudoaleatorios y diseñan de forma aleatoria el orden en que se asignan los grupos de tratamiento. Se habla de aleatorización por bloques (de tamaño n) cuando cada n pacientes garantizamos que haya una distribución homogénea en todas las ramas de estudio.

La **ocultación de la secuencia de aleatorización** supone establecer mecanismos para evitar que los clínicos puedan sospechar a qué tratamiento corresponde cada grupo, poniendo en riesgo la ausencia de sesgos de selección. Este aspecto es de especial importancia cuando el periodo de captación de pacientes es largo.

En nuestro ejemplo: JAMA 2002; 288: 2981-2997.

En la página 2982 (2ª columna 2º párrafo) los autores indican que la asignación fue aleatoria (generada informáticamente en el centro de referencia y comunicada telefónicamente para cada paciente reclutado), estratificada por centro (para evitar que por azar el número de pacientes en algún centro sea muy diferente entre ramas de tratamiento ó incluso no haya una rama de tratamiento) y en bloques de tamaño 5 ó 9 (esto supone una estrategia de aleatorización que garantiza que cada 5 ó 9 pacientes, según el tamaño del bloque, haya el número definido asignado a cada brazo).

Estos elementos parecen suficiente garantía sobre la asignación aleatoria y la ocultación de la secuencia de aleatorización.

3.- Seguimiento adecuado hasta el final del estudio. Análisis por intención de tratar.

A pesar de haber logrado grupos similares, mediante una adecuada aleatorización, durante un ECA se producen siempre pérdidas de pacientes (**pérdidas de seguimiento** que hacen que se desconozca la situación del paciente al finalizar el estudio).

Los motivos son múltiples: abandono voluntario de participantes sin molestias de salud, cambio de domicilio; aparición de efectos adversos no recogidos por los investigadores; fallecimiento no informado, etc. El significado que cada una de las situaciones anteriores tiene en la valoración de eficacia de un tratamiento es muy diferente: en el primer y segundo casos los pacientes abandonan en buen estado de salud; en el tercero y cuarto debido a posibles efectos adversos del tratamiento.

Además, estas pérdidas pueden ocurrir de forma homogénea en todas las ramas del estudio ó ser más frecuentes en una.

Por otra parte, durante el periodo de seguimiento puede ocurrir que los pacientes no tomen el tratamiento prescrito (olvido, molestias asociadas al mismo, etc); precisen tratamientos adicionales por su situación de salud; incluso que se produzcan cambios de rama de tratamiento.

Estamos ante situaciones de potenciales **sesgos de seguimiento** que cuestionan la validez del estudio.

Para analizar el impacto de las pérdidas de seguimiento producidas, se realizan **análisis de sensibilidad**. Consiste en comparar los resultados obtenidos suponiendo que los pacientes perdidos en cada rama presentaron ocurrencia de un resultado o no lo presentaron. Si la estimación del efecto en ambos casos es muy dispar (pasando de un efecto favorable al tratamiento A, a un efecto favorable al tratamiento B), las pérdidas suponen un sesgo que cuestiona la validez de los resultados. En muchas ocasiones los autores indican en el apartado de material y método si han realizado este análisis y su resultado.

El análisis de sensibilidad compara la hipotética presencia o ausencia de resultados de los pacientes perdidos en cada una de las ramas del estudio. El análisis por intención de tratar analiza a cada paciente perdido en su grupo inicial suponiendo que ha presentado un resultado poco favorable a la intervención que se evalúa. Los ensayos clínicos que muestran éstos análisis contrarrestan sesgos de seguimiento ante pérdidas, suponen un valor añadido a su validez.

El **análisis por intención de tratar** supone que cada paciente es analizado en el grupo al que fue inicialmente asignado, aunque no haya tomado la medicación prescrita. De esta forma, se mantiene la homogeneidad en las características de los grupos y el análisis se realiza en la posición más conservadora. Si se demuestra efecto del tratamiento, este resultado tiene mayor validez que en el análisis por protocolo (incluyendo solo a los pacientes que hayan cumplido el tratamiento en la forma prescrita). Además esta situación se aproxima mejor a la realidad clínica en la que la adherencia del paciente al tratamiento constituye uno de los principales problemas para conseguir los objetivos de mejora de nivel de salud.

En nuestro ejemplo:

Los autores señalan que se realizó el análisis por intención de tratar (página 2983, 2ª columna, métodos estadísticos).

En la figura 1 (página 2985) se detallan los pacientes asignados aleatoriamente a cada grupo:

- Clortalidona , 15255 pacientes iniciales, 419 pérdidas (339 perdidas, 80 abandonos), que supone un 2.7%.
- Lisinopril , 9054 pacientes iniciales, 276 pérdidas (218 perdidas, 58 abandonos) que supone un 3%.
- Amlodipino , 9048 pacientes iniciales, 258 pérdidas (200 perdidas, 58 abandonos) que supone un 2.8%.

El porcentaje de pérdidas es muy bajo y de similar cuantía en los tres grupos de tratamiento.

Aspectos adicionales para establecer la validez de un ECA.

Además de estos tres aspectos, esenciales para valorar la validez interna de un ECA, hay otros aspectos cuyo rigor mejora la calidad del estudio y por tanto la validez de los resultados como estimación de la realidad. Estos son:

4.- Diseño ciego

Cuando el diseño del estudio evita que se identifique durante el estudio cual es tratamiento que recibe cada paciente. Se distinguen situaciones de: simple ciego (paciente), doble ciego (paciente y clínico que

lo sigue), triple ciego (paciente, clínico, persona que analiza los resultados) según quienes desconozcan el tratamiento dado a cada paciente.

En estudio con diseño ciego establece a priori más garantías de que la atención dada a los pacientes durante el seguimiento no sea diferente según grupos. Por otra parte, que se eviten sesgos en la medición de las variables resultado.

Los estudios "ciegos" para pacientes, clínicos y estadísticos implicados en el mismo, dan mayor validez a un ECA.

En nuestro ejemplo:

Los fármacos de estudio presentaban igual aspecto y características, se realizó un doble enmascaramiento (página 2982, 3ª columna 1º párrafo). De esta forma, tanto paciente como clínico desconocían el tratamiento asignado.

Las variables resultado (página 2983, 1ª columna, 1º párrafo) fueron documentadas en un 99% de los casos en las tres ramas de tratamiento, se revisaron las altas hospitalarias ó certificados de defunción de forma centralizada por revisores médicos. En una submuestra aleatoria del 10% de los pacientes con diagnóstico de ACV ó coronariopatía se revisó, de forma independiente, con más detalle la información clínica para validar el proceso diagnóstico. El grado de acuerdo diagnóstico entre clínicos investigadores y el comité de revisores fue del 90% para la variable resultado primaria, 84% para ACV, 85% para hospitalización por fallo cardíaco. Estos datos fueron similares en las tres ramas de tratamiento.

5.- Similitud de las características basales de los grupos

Sobre todo en relación con aquellas características que pudieran tener relación con el resultado (p. ej. Sexo ó raza en el caso de eventos cardiovasculares e hipertensión). Si las ramas de estudio presentan una distribución diferente de las mismas, las diferencias halladas en las variables resultado (p. ej. IAM, ACV) pudieran deberse a esas características distintas y no al efecto del tratamiento en valoración. Estaríamos ante el efecto de variables potencialmente **confusoras**, de las cuales existen un alto número no conocidas y por tanto no controlables.

La asignación aleatoria correctamente realizada, sobre todo en muestras amplias produce una distribución similar entre los grupos de todas sus características.

En todo artículo que comunique los resultados de un ECA, el **CONSORT** establece que la primera tabla (TABLA 1) detalle las características basales más relevantes en cada una de las ramas de estudio.

El lector debe buscar esta tabla y valorar su similitud en la proporción de cada una (% de mujeres, personas mayores de 65 años, edad media, % de diabéticos tipo II etc) según su magnitud absoluta, no según la significación estadística de pruebas que comparen estas magnitudes. Debe también considerarse si falta en la tabla alguna característica de la que existe evidencia de relación con el resultado a medir.

El lector de un artículo de un ECA debe buscar y revisar la tabla Nº 1 en la que se exponen las características basales de cada una de las ramas de estudio.

En nuestro ejemplo:

La Tabla 1 (página 2984) muestra la distribución, al inicio del estudio, en los grupos de clortalidona, amlodipino y lisinopril de edad, sexo, grupo étnico, nivel educativo, HTA previamente diagnosticada y tto. Antihipertensivo, presencia de otros factores de riesgo cardiovascular, IMC, historia de coronariopatía, uso de ac. Acetilsalicílico ó tto. hormonal sustitutivo. La distribución es muy similar entre ellos.

6.- Igual tratamiento a los grupos (salvo la intervención a estudio)

El objetivo de un ECA es establecer una relación de **causalidad** entre el tratamiento recibido (intervención ó control) y el efecto en términos de ocurrencia de eventos establecidos como variables resultado.

Para ello, además de partir de una situación inicial de grupos similares en todas sus características, es muy importante evitar las **cointervenciones** diferenciales. Si los pacientes de una rama reciben mayor atención, son vistos con mayor frecuencia, se incide más en las modificaciones de hábitos de vida, etc, resulta imposible establecer la relación intervención a estudio – resultado clínico.

En nuestro ejemplo:

Al describir el tratamiento en el apartado de material, los autores indican (página 2982, 3ª columna 1º párrafo) que las medidas no farmacológicas siguieron las guías de práctica clínica vigentes.

En el apartado de resultados se detalla (página 2984) la adherencia a las visitas de seguimiento y la medicación antihipertensiva abierta (conocida por el clínico) que se añadió en cada grupo en los estadios 2y 3.

Como has visto, la información necesaria para evaluar la validez interna de un ECA está en el apartado de material y método en su práctica totalidad.

Si con esta información puedes concluir que dispones de suficientes garantías sobre la validez del estudio (resultados no sesgados), solo en este caso, puedes pasar a analizar la importancia clínica de los resultados.

La revisión del apartado de material y métodos nos proporciona la información necesaria sobre la validez de un ECA y el interés que nos va a aportar conocer los resultados.

Bases para establecer si un ensayo clínico aporta resultados clínicamente importantes

La importancia clínica se analiza a través de la magnitud del efecto hallado, no tiene relación con la significación estadística (valor de p). La significación estadística indica la probabilidad de cometer un error tipo I: es decir afirmar que las diferencias halladas reflejan la realidad, cuando no es así y únicamente se deben al azar. Cuando se analizan muestras de gran tamaño, diferencias de una mínima magnitud resultan muy significativas estadísticamente y viceversa. Por otra parte, cuando en el análisis estadístico se realizan múltiples comparaciones aumenta la probabilidad de hallar significación estadística ($p < 0.05$) en algunas como artefactos del análisis.

Por estas y otras razones que ya se han tratado en capítulos anteriores, cuando se realiza la lectura crítica de un artículo que comunica los resultados de un estudio, la importancia clínica de los mismos deberá siempre valorarse por la magnitud de los mismos y no por su significación estadística.

1.- ¿Cuál es la magnitud del efecto?

Analizar en que forma se comunican los resultados, su adecuación y magnitud del efecto:

- RR, OR, RAR, NNT, en el caso de variables binarias (ACV si/no, etc).
- Diferencia de medias, de medianas etc. en el caso de variables continuas (tiempo de supervivencia libre de enfermedad, puntuación escala calidad de vida, etc.).

En aquellos casos en los que haya habido diferencias durante el seguimiento en los grupos, con posible influencia en las variables resultado, de factores diferentes a la intervención, el hecho de que se comunique que se realizó alguna técnica de análisis multivariante que permite controlar el efecto de las variables potencialmente confusoras, añade rigor y fiabilidad a los resultados.

2.- Precisión con la que se dan los resultados

En todo estudio en el que se trabaja con una muestra de la población, los resultados obtenidos reflejan la realidad de esa muestra únicamente.

Pero tanto al investigador como al lector le interesa aproximarse al conocimiento de lo que ocurre en la población, en otros pacientes (similares a los estudiados pero NO los mismos).

El diseño adecuado del estudio y el análisis estadístico de los resultados de un ECA, se orientan a **inferir** de la forma más rigurosa posible la eficacia ó efectividad del tratamiento ó intervención (fármaco, educación, tipo de cura en pie diabético, etc) en pacientes que atendemos ó atenderemos en el futuro.

Además de la magnitud del efecto es necesario conocer la precisión con la que se estima. Para ello, disponemos de los intervalos de confianza.

Intervalo de confianza es el intervalo dentro del que se encuentra la verdadera magnitud del efecto (nunca conocida exactamente) con un grado prefijado de seguridad. A menudo se habla de "intervalo de confianza al 95%" (o "límites de confianza al 95%"). Quiere decir que dentro de ese intervalo se encontraría el verdadero valor del efecto en el 95% los casos. O también podemos expresarlo como que tenemos una seguridad del 95% de que el verdadero valor estará entre los límites de ese intervalo. Cuanto más estrecho el intervalo de confianza, mayor es la **precisión** con la que se estima el efecto en la población.

Ejemplo: Se mide el efecto de una intervención de educación sanitaria en el abandono del tabaco como proporción de fumadores que dejan el tabaco después de 6 meses. El estudio consigue una proporción de abandono del 60%, si el IC95% es 10%, 80% estamos ante una intervención de resultado incierto ya que el verdadero porcentaje de abandono puede ser tan bajo como 10% ó tan alto como 70%; ¿la incorporarías a tu práctica?, ¿la recomendarías a tus pacientes?. Si el IC95% es 54%, 71%, ¿qué te parece ahora la intervención?.

En el caso de RR ó OR en los que 1 indica igual efecto del tratamiento que del control, cuando el IC incluye al 1 (IC - 0.4, 2.4) nos indica que no se ha podido demostrar diferencia en el efecto entre tratamiento y control.

Actualmente, la mayoría de las publicaciones exigen que los autores de los artículos comuniquen los intervalos de confianza de los resultados; sobre todo en los resultados principales.

En nuestro ejemplo:

Los autores comparan los resultados del grupo que recibió clortalidona respecto a los grupo lisinopril y amlodipino respectivamente. En la tabla 5 encontramos los resultados principales detallando el número de eventos de cada variable, que han ocurrido en cada grupo (columnas 2 a 7) y la medida del efecto estimada como riesgo relativo (RR). Para calcular otras medidas del efecto únicamente necesitamos el número de pacientes analizados (en este caso igual al número de pacientes que iniciaron el ECA al hacerse un análisis por intención de tratar, (figura 1) y el número de eventos ocurridos (columnas 2,4 y 6 de la tabla 5).

El riesgo relativo (RR) es una medida de la **eficacia** de un tratamiento. Si es igual a 1, el efecto del tratamiento no es distinto del efecto del **control**. Si el RR es mayor (o menor) que 1, el efecto del tratamiento es mayor (o menor) que el del control. Nótese que el efecto que se está midiendo puede ser adverso (ej. muerte, IAM, discapacidad) o deseable (ej. Curar, dejar de fumar).

Un RR = 2.4 indica que la probabilidad de presentar el evento (P. ejemplo IAM) es 2.4 veces mayor para los pacientes que reciben el tratamiento respecto a los que reciben el control → el tratamiento tiene un efecto negativo, causa daño (evento dañino).

Un RR = 0.72 indica que el tratamiento produce una disminución de la probabilidad de que ocurra el evento (p. ejemplo IAM) en un 28% [(1-0.72) x100] , en comparación con los que reciben el control. El tratamiento tiene un efecto positivo, causa beneficio.

Para interpretarlo (en términos de efecto beneficioso ó dañino del tratamiento) hay que considerar la magnitud (cantidad absoluta) en relación al tipo de efecto (deseable ó adverso) que se está midiendo.

Bases para establecer si un ensayo clínico aporta resultados clínicamente importantes: Se valora la magnitud del efecto con las medidas de RR, OR, RAR, NNT, en el caso de las variables binarias y a diferencia de medias, de medianas etc en el caso de variables continuas

Podríamos resumirlo como sigue:

- Evento deseable (curación, alivio del dolor, mejora calidad de vida...)
 - RR > 1 → efecto beneficioso del tratamiento.
 - RR < 1 → efecto perjudicial del tratamiento.
- Evento adverso (muerte, IAM, ACV, fractura osteroporótica...)
 - RR > 1 → Efecto perjudicial del tratamiento.
 - RR < 1 → Efecto beneficioso del tratamiento.

En la tabla siguiente se resumen algunas medidas de efecto y sus intervalos de confianza para la variable principal de resultado y algunos eventos de variables secundarias, entre los grupos.

Grupos que se comparan	Variables resultado	Medidas de efecto
AMLODIPINO VS CLORTALIDONA	Variable principal de resultado: IAM no mortal ó muerte por coronariopatía	Tasa grupo control: 1362/15255 (8.92%) Tasa grupo experimental: 798/9048 (8.81%) RR = 0.98 IC 95% 0.90, 1.07 RAR = 0.1% IC 95% -0.6% , 0.8% RRR = 1% IC 95% -7% , 9% NNH = 921 IC 95% 118 a infinito
	Componentes de variables secundarias: a) ICC	Tasa grupo control: 870/15255 (5.7%) Tasa grupo experimental: 706/9048 (7.8%) RR = 1.38 IC 95% 1.25 , 1.52 RAR = -2.1% IC 95% -2.8% , -1.4% RRR = -37% IC 95% -48% , -25% NNH = 48 IC 95% 37 , 70
	b) Muerte u hospitalización por ICC	Tasa grupo control: 724/15255 (4.7 %) Tasa grupo experimental: 578/9048 (6.4%) RR = 1.38 IC 95% 0.90, 1.07 RAR = -1.7% IC 95% -2.3% , -1.1% RRR = -36% IC 95% -49% , -23% NNH = 59 IC 95% 44 , 92
LISINOPRIL VS CLORTALIDONA	Variable principal de resultado: IAM no mortal ó muerte por coronariopatía	Tasa grupo control: 1362/15255 (8.92%) Tasa grupo experimental: 796/9054 (8.81%) RR = 0.99 IC 95% 0.91, 1.08 RAR = 0.1% IC 95% -0.6% , 0.8% RRR = 1% IC 95% -7 % , 9% NNH = 732 IC 95% 114 a infinito
	Componentes de variables secundarias: a) ICC	Tasa grupo control: 870/15255 (5.7%) Tasa grupo experimental: 612/9054 (6.8%) RR = 1.19 IC 95% 1.07, 1.31 RAR = -2.1% IC 95% -2.8% , -1.4%

		RRR = -37% IC 95% - 48% , - 25% NNH = 48 IC 95% 37 , 70
	b) Tto. y hospitalización por angor	Tasa grupo control: 1567/15255 (10.3%) Tasa grupo experimental: 1019/9054 (11.3%) RR = 1.11 IC 95% 1.03, 1.20 RAR = -1.0% IC 95% - 1.8% , - 0.2% RRR = -9.6% IC 95% - 17.5% , - 1.7% NNH = 102 IC 95% 56 , 579

Observamos que tanto para la variable principal de resultado como para las variables secundarias el riesgo es similar tanto para los tratados con lisinopril como con amlodipino respecto a los tratados con diurético (clortalidona). El trabajo demuestra la ausencia de mayor efecto de los dos primeros en relación al diurético al mostrar intervalos de confianza que incluyen la unidad en todos los casos.

Por otra parte, los intervalos de confianza son muy estrechos. Esto indica una alta precisión en las estimaciones del efecto.

Se han calculado el número de pacientes que es necesario tratar para producir un evento. Es una medida de la eficacia de un tratamiento. Es el número de personas que se necesitaría tratar con un tratamiento específico (ej. aspirina a quienes han sufrido un ataque cardíaco) para producir, o evitar, **una** ocurrencia **adicional** de un evento determinado (ej. prevención de muerte). Del mismo modo se define **número necesario para perjudicar (NNP) (number needed to harm (NNH))** para evaluar efectos indeseables.

Como ya se ha explicado en capítulos anteriores, se calcula como:

$$NNT = \frac{1}{RAR}$$

Las medidas de magnitud de efecto: riesgo relativo, reducción absoluta del riesgo y número necesario a tratar; y las de precisión: intervalo de confianza aportan las bases para establecer si los resultados son clínicamente importantes.

donde RAR es la **Reducción absoluta del riesgo calculada como la diferencia entre la tasa de eventos en el grupo control (R_c) y en el grupo experimental (R_e): RAR = R_c - R_e**

Cuando se analizan algunos eventos como ICC, muerte u hospitalización por ICC, angina, observamos que aparecen con más frecuencia en los grupos tratados con lisinopril ó amlodipino que en el grupo de clortalidona. Por ejemplo:

Los pacientes que reciben lisinopril tienen una probabilidad de presentar ICC que es un 19% mayor respecto a los que reciben clortalidona, en la muestra de estudio. En la población, con un 95% de seguridad esta probabilidad será entre un 7% y un 31% mayor.

Los pacientes que reciben amlodipino tienen en la muestra de estudio una probabilidad de ICC 38% mayor que los que reciben clortalidona. En la población, con una seguridad del 95% esta probabilidad estará entre un 25% y un 52%.

En este ultimo caso, la probabilidad de muerte ó ingreso hospitalario por la ICC es un 35% mayor en el grupo tratado con amlodipino entre los pacientes del ensayo. Con un 95% de seguridad, en la población esta probabilidad esta entre un 21% y un 50%.

Para el evento “tener ICC” encontramos un NNH de 48 (IC95% 37 a 70). Es decir cada 48 pacientes tratados con amlodipino uno más presentó ICC que si hubiera recibido clortalidona. En pacientes similares, ésto ocurrirá (con una seguridad del 95%) en uno de cada 37 a 70 pacientes tratados con amlodipino.

Bases para establecer si un artículo aporta resultados aplicables en mis pacientes

Una vez establecida la **validez interna** del trabajo a partir de los datos del apartado de material y método y si la **importancia clínica** de los resultados hace que consideremos relevante la eficacia ó efectividad del tratamiento estudiado, debe analizarse la **validez externa** de los resultados. Es decir su **generabilidad** a otras poblaciones de pacientes similares.

Cuando la lectura se hace para responder a una pregunta que surge en la atención a nuestros pacientes, es importante analizar la **aplicabilidad** de los resultados, válidos e importantes, a estos pacientes.

Excelentes ensayos desde un punto de vista de su diseño y desarrollo, pueden no ser aplicables a la toma de decisiones en la práctica clínica, si alguno de los siguientes aspectos presenta limitaciones importantes.

1.- Reproducibilidad de los resultados en el medio del lector

Debemos intentar responder a preguntas como las que siguen:

- ¿Son los pacientes del estudio suficientemente similares a los que yo atiendo?, ó dicho de otra forma si mi paciente estuviera durante el periodo de captación en alguno de los centros de estudio, ¿hubiera sido invitado a participar?
 - Para responder tenemos en material y métodos los criterios de inclusión / exclusión y en la Tabla 1 las características basales de los pacientes finalmente incluidos en el ECA.
- ¿Son las pautas de tratamiento estudiadas y la forma de seguimiento adecuadas para realizarlas en mi centro con los medios de que dispongo?. Esto es especialmente importante cuando se evalúan intervenciones que requieren recursos ó entrenamiento específico (p. ej. una nueva técnica de crioterapia)

Si los pacientes incluidos en el estudio son similares a los de nuestro medio y tenemos posibilidad de realizar la misma intervención, decimos que el estudio es aplicable o reproducible.

2.- ¿Se consideraron todos los resultados clínicamente relevantes en el estudio?

- ¿Se comunican en el artículo resultados de todos los aspectos que consideras importante conocer en tus pacientes, incluyendo los efectos secundarios adversos de los tratamientos?.
- En caso negativo, ¿puede su desconocimiento afectar tu toma de decisiones?.
- Cada vez es más frecuente encontrar al final de un ECA, tal como establece el acuerdo CONSORT, una declaración de conflicto de intereses. Los autores deben indicar quién patrocina el ensayo, las ayudas recibidas y el papel del promotor. Esto ayuda al lector a descartar posibles sesgos en los resultados como consecuencia del interés del promotor en una nueva molécula.

3.- Beneficios respecto a perjuicios y costes de la intervención a estudio

Este punto plantea una valoración global de los beneficios del tratamiento (ú otro tipo de intervención) que se demuestran en relación a los potenciales perjuicios que su aplicación a tus pacientes pueda suponer (p. ejemplo, en principios activos nuevos se desconocen muchos efectos adversos que son identificados con su uso).

Valorarlos también en relación al coste que supone su incorporación en términos económicos, de adherencia del paciente, entrenamiento profesional, etc.

Esta es una reflexión personal en la que la experiencia personal y el conocimiento del medio profesional propio es muy

La aplicabilidad de los resultados, la exposición detallada de los datos clínicos relevantes y los beneficios y costes de la intervención definen la validez externa de los resultados de un estudio.

importante.

Solución al escenario

El estudio demuestra que los beneficios del tratamiento de la HTA, en pacientes con riesgo cardiovascular, de un diurético son similares a los obtenidos con un IECA ó un calcioantagonista, medidos como eventos cardiovasculares mayores y menores a los 5 años de seguimiento. Por otra parte, los controles de las cifras tensionales eran adecuados.

Se indican (página 2995 1ª columna) las ayudas recibidas y el papel del esponsor (National Heart, Lung and Blood Institute).

Con estos resultados disponemos de razones que apoyan el iniciar tratamiento con un diurético en nuestro paciente. En caso de no controlar su tensión por debajo de 140/90 solo con él, podría asociarse otro fármaco, pero el diurético parece aconsejable que se mantenga tras estos resultados.

Decides proponer una sesión bibliográfica para comentar este estudio con tus compañeros.

Este ejercicio requiere un tiempo para recuperar y leer el artículo siguiendo alguna de las guías de lectura crítica de ECA disponibles.

Una buena forma de iniciarse en esta metodología es leer revistas secundarias en las que podrás encontrar de cada trabajo seleccionado:

- el resumen del trabajo que hacen los autores.
- un comentario sobre la validez e importancia clínica de los resultados del trabajo realizado por un experto siguiendo los mismos principios que hemos trabajado aquí.

Esto te permitirá tener mas elementos de juicio sobre la calidad de los estudios que se publican en el tema que te interesa y sobre el interés y aplicabilidad de sus resultados. En la bibliografía se indican algunas de las más útiles.

Bibliografía

Baker M, Maskrey N, Kirk S. Clinical effectiveness and primary care. Radcliffe Medical Press; 1997. [[Reseña](#)]

Evidence-based Medicine Working Group. Evidence-based medicine. A new approach to teaching the practice of medicine. JAMA 1992; 268: 2420-2425. [[Medline](#)]

Fowkes FG, Fulton PM, Critical appraisal of published research: introductory guidelines. BMJ 1991; 1:136-1.140. [[Medline](#)]

Gómez de la Cámara A. Investigación con medicamentos. Es necesario atender a las normas que la regulan. Aten Primaria 2003; 31: 3 - 5. [[Medline](#)] [[Texto completo](#)]

Gómez de la Cámara A. La investigación en atención primaria. El ensayo clínico y los estudios observacionales de productos farmacéuticos. Aten Primaria 1999; 24: 431-435. [[Medline](#)] [[Texto completo](#)]

Haynes RB, McKibbon KA, Fitzgerald D, Guyatt GH, Walker CJ, Sackett DL. How to keep up with the medical literature: I. why try to keep up and how to get started. Ann Intern Med 1986; 105: 149-153. [[Medline](#)]

Haynes RB, Sánchez RG, Jadad AR, Browman GP, Gómez de la Cámara A. Herramientas para la práctica de la medicina basada en la evidencia (I). Actualización en recursos de información

basados en la evidencia para la práctica clínica. Med Clin (Barc) 2000; 115: 258-260. [[Medline](#)]
[[Texto completo](#)]

Moer D, Schulz K, Altman D for the Consort Group. The CONSORT Statement: Revised Recommendations for Improving the Quality of Reports of Parallel-Group Randomized Trials JAMA. 2001; 285: 1987-1991.

Traducción al castellano publicada en Rev Sanid Milit Mex 2002; 56: 23-28. “La declaración CONSORT: Recomendaciones revisadas realizadas para mejorar la calidad de los informes de ensayos aleatorizados de grupos paralelos”. [[Texto completo](#)]

Sackett D, Richardson WS, Rosenberg W, Haynes RB. Evidence-based medicine. How to practice and teach EBM. Londres: Churchill Livingstone; 1997.

Sacristán JA, Soto J, Galende I. Evaluación crítica de ensayos clínicos. Med Clin (Barc); 1993; 100: 780-787. [[Medline](#)]