

Sumario

INTERPRETANDO LA LITERATURA MÉDICA: ¿QUÉ NECESITO SABER?

PARTE I

■ INTRODUCCIÓN

■ ¿SOBRE QUÉ SE INVESTIGA?

El papel de la industria y de las agencias reguladoras

■ ¿QUÉ SE PUBLICA DE LO QUE SE INVESTIGA? Sesgo de publicación y sesgo en la presentación de resultados

■ UNA VEZ PUBLICADO, ¿ES RELEVANTE EL ENSAYO? ¿PUEDE SER ÚTIL PARA MI PRÁCTICA CLÍNICA?

■ ¿ES EL ENSAYO METODOLÓGICAMENTE CORRECTO?

PARTE II

■ RELEVANCIA CLÍNICA: ¿QUÉ NOS DICEN LOS NÚMEROS?

- ¿Qué es la *p*?
- ¿Qué es el intervalo de confianza?
- Métodos más frecuentes de expresar los resultados

■ ANÁLISIS DE SUBGRUPOS

■ LA PROLIFERACIÓN DE NUEVOS TIPOS DE SESGOS: CONSEJOS PARA SOBREVIVIR EN EL INTENTO

INTERPRETANDO LA LITERATURA MÉDICA: ¿QUÉ NECESITO SABER?

PARTE I

INTRODUCCIÓN

El volumen de la literatura médica que se genera continuamente es inabarcable. Cada año Medline cataloga 560.000 artículos nuevos y Cochrane Central añade aproximadamente 20.000 ensayos aleatorizados nuevos. Estar al día se ha convertido en un verdadero reto para el profesional sanitario. Además, no toda la información disponible es relevante, válida y fiable¹. Aun suponiendo que hemos identificado la información útil, persiste el desafío de interpretarla adecuadamente, de forma que nos permita tomar las decisiones apropiadas en un contexto clínico concreto.

Un caso particular es el del ensayo clínico aleatorizado (ECA) que, junto con las revisiones sistemáticas, se considera el patrón de oro para medir la eficacia de una intervención (farmacológica o no farmacológica).

¿Qué es un ensayo clínico?

Un ensayo clínico aleatorizado es un **estudio experimental** en el que se aplica una intervención (por ejemplo un fármaco para reducir el colesterol o un programa estructurado de ejercicio físico) a un grupo de pacientes (grupo experimental) y el resultado (ej. los niveles de colesterol o la frecuencia de infartos) se compara con el obtenido en un grupo idéntico de pacientes (grupo control) que recibe otra intervención (ej. otro fármaco, un placebo o un consejo no estructurado de ejercicio físico). Los pacientes son **asignados de forma aleatoria y simultánea** al grupo experimental y al grupo control con la misma probabilidad (como lanzar una moneda al aire), de forma que **ambos grupos sean similares en todo** (características de los pacientes, factores pronósticos, otros tratamientos recibidos, etc.) **excepto en la intervención que reciben**. De esta forma, **la diferencia observada podrá atribuirse al tratamiento recibido**, con un margen de error preestablecido y siempre que el ensayo esté bien realizado.

A menudo, la más que razonable actitud de basar las decisiones de manejo de los pacientes en los resultados de la investigación clínica se convierte en una verdadera carrera de obstáculos, al encontrarnos con distintos tipos de limitaciones, a veces referidas a los ya "clásicos" sesgos en el método, y otras, quizás más sutiles, relacionadas con la relevancia de la pregunta de investigación y con la forma en que se presentan los resultados².

A lo largo de este doble boletín intentaremos describir los sesgos más frecuentes que podemos encontrar en relación con los ensayos clínicos. Asimismo propondremos algunas claves para facilitar la interpretación de los resultados de un ECA (¡sin necesidad de ser un experto en estadística!).

¿SOBRE QUÉ SE INVESTIGA? EL PAPEL DE LA INDUSTRIA Y DE LAS AGENCIAS REGULADORAS

La selección de temas y objetivos por el promotor predetermina en buena parte el conocimiento disponible, ya que sólo dispondremos de la “evidencia” cuando alguien haya querido que algo sea “evidente”². En la CAPV alrededor del 80% de los ensayos clínicos multicéntricos que se realizan están patrocinados por la industria farmacéutica. Probablemente esto se debe a que la realización de ECA es un requisito legal para poder registrar y comercializar un nuevo fármaco o para autorizar una nueva indicación.

La consecuencia inmediata es que la investigación a través de ECA está desproporcionadamente enfocada a nuevos

fármacos u otras tecnologías de valor comercial, frente a terapias no farmacológicas (estilos de vida, dieta, ejercicio) o fármacos no rentables (destinados a enfermedades raras o prevalentes en países subdesarrollados, fármacos “viejos”, etc.)². Los ensayos están diseñados para evaluar fármacos, más que para evaluar pacientes o situaciones clínicas o para responder a preguntas relevantes que se presentan en la práctica clínica.

El “sesgo de patrocinio” se refiere a la posibilidad de que el financiador del ensayo pueda influir en su diseño, desarrollo, análisis o publicación, para presentar los resultados del estudio en forma favorable a sus intereses.

¿QUÉ SE PUBLICA DE LO QUE SE INVESTIGA? SESGO DE PUBLICACIÓN Y SESGO EN LA PRESENTACIÓN DE LOS RESULTADOS

No todos los ensayos clínicos se publican, y la no-publicación no se distribuye al azar. Los autores tienden a no enviar los resultados estadísticamente no significativos (también llamados, de forma ciertamente desafortunada, “estudios negativos”) y a los editores no les gustan los trabajos con resultados negativos².

En un estudio realizado para valorar el sesgo de publicación³ los autores hicieron un seguimiento de los 42 ensayos clínicos de 5 diferentes antidepresivos ISRS que habían sido presentados a la Agencia Sueca del Medicamento entre 1983-1994. De los 21 ensayos en que el fármaco fue más eficaz que placebo, 19 fueron publicados como ECA, mientras que de los otros 21 ensayos que no mostraron diferencias, sólo 6 fueron publicados.

El sesgo de publicación también puede darse a la inversa, publicando los trabajos varias veces (en varias revistas, separatas, realizando distintos metaanálisis con los mismos estudios, etc.), o mediante publicaciones fragmentadas (“publicaciones en salami”), provocando la sensación

de que existe más evidencia de la que realmente hay^{1,2,4}.

Por otro lado, las revisiones sistemáticas han demostrado que los ECA financiados por compañías farmacéuticas tienen 4-5 veces mayor probabilidad de obtener resultados favorables a la intervención que los ensayos con otras fuentes de financiación (organismos públicos o entidades sin ánimo de lucro)^{5,6}, siendo excepcionales los ensayos financiados por la industria que favorecen a la terapia tradicional frente a sus productos⁴.

Esta clara tendencia a favorecer al producto investigado por la industria farmacéutica no se debe a una peor calidad de los estudios, al menos en lo que se refiere a aspectos metodológicos como el tamaño de la muestra, la aleatorización o el enmascaramiento⁵. Los motivos están más relacionados con el propio objetivo del estudio (se trata de plantearse la pregunta de investigación de forma “adecuada” para obtener los resultados “deseados”), por ejemplo comparándose frente a placebo, o midiendo múltiples resultados y publicando sólo los favorables^{4,5}.

UNA VEZ PUBLICADO, ¿ES RELEVANTE EL ENSAYO? ¿PUEDE SER ÚTIL PARA MI PRÁCTICA CLÍNICA?

Antes de leer un artículo con detenimiento nos tendremos que preguntar si el tiempo que vamos a dedicarle va a merecer la pena y si podremos aplicarlo a nuestros pacientes⁷. Este ejercicio implica el plantearse hasta qué punto los Pacientes a estudio y las Intervenciones que se van a

Comparar se parecen a nuestros pacientes y a las terapias de que disponemos, y si el Outcome o Resultado que mide el estudio es importante para nuestros pacientes (acrónimo PICO). Ver Tabla 1.

Tabla 1. La pregunta clínica con sus componentes (PICO)

P	Pacientes	¿Son los pacientes a estudio razonablemente similares a los míos?
		Es necesario valorar si son pacientes ambulatorios u hospitalarios, o factores como la edad, gravedad o comorbilidad. Habitualmente los ECA tienden a excluir a niños, pacientes muy ancianos o con comorbilidad importante o que toman varios fármacos simultáneamente ² . Conviene leer con detenimiento los criterios de inclusión y exclusión de los pacientes ¹ .
I	Intervención	¿Cuáles son la intervención a estudio y el comparador?
C	Comparador	<ul style="list-style-type: none">– Si nos encontramos un ECA sobre reflexoterapia en el dolor de espalda, nos preguntaremos si podremos ofrecerlo a nuestros pacientes.– ¿Es importante un ECA de desloratadina frente a placebo en la rinitis?– ¿Es adecuado comparar 20 mg de simvastatina con 20 mg de atorvastatina?
O	Outcome o resultado	¿Qué variables de resultado utiliza?: ¿Intermedias o finales?; ¿variable principal o secundarias?; ¿variable agregada?; ¿durante cuánto tiempo se mide?
		<ul style="list-style-type: none">– ¿Densidad mineral ósea, fractura vertebral o fractura de cadera?– ¿Síntomas, pico flujo o crisis de asma?– 12 semanas puede ser tiempo suficiente para valorar el efecto de un COX-2 sobre el dolor, pero ¿y para valorar el riesgo de infarto agudo de miocardio?

¿Es adecuado el comparador?

Es ampliamente aceptado que los posibles beneficios y riesgos de cualquier procedimiento nuevo deberían ser evaluados mediante su comparación con los mejores métodos preventivos, diagnósticos y terapéuticos disponibles. Siguiendo este principio, el uso del placebo estaría justificado sólo cuando no hay terapias eficaces con las que comparar (como el caso de las enfermedades “huérfanas” o condiciones clínicas refractarias a tratamientos previos).

Si echamos un vistazo a fármacos comercializados recientemente en base a ensayos frente a placebo (ranelato de estroncio, levocetirizina, buprenorfina transdérmica...), probablemente concluyamos que hubiera sido preferible utilizar un comparador activo (bisfosfonatos, cetirizina, morfina...) de eficacia demostrada.

Otros métodos para diseñar estudios que favorezcan al producto experimental consisten en comparar el fármaco frente a uno que ya de antemano se prevea inferior (losartan frente a atenolol: estudio LIFE), frente a una dosis baja del competidor (atorvastatina 80 mg vs pravastatina 40 mg, estudio PROVE IT_TIME y REVERSAL), o frente a una dosis alta del competidor cuando se trata de demostrar un mejor perfil de efectos adversos (risperidona frente a dosis altas de haloperidol –20 mg–, para valorar efectos extrapiramidales)⁴.

Variable final o “dura” versus intermedia o “subrogada”

Una variable subrogada es un resultado, relativamente sencillo de obtener (densidad mineral ósea, datos endoscópicos o concentraciones de colesterol), que se usa cuando los eventos clínicos importantes, llamados variables finales o “duras” son difíciles de observar, implicando con frecuencia estudios de larga duración (fractura de cadera, sangrado gastrointestinal, infarto de miocardio o muerte)⁷. El valor de una variable subrogada dependerá de la medida en que sea buena predictora de la variable final. Habitualmente las variables orientadas al paciente (VOP) son las duras y las orientadas a la enfermedad (VOE) suelen ser subrogadas. En los siguientes ejemplos puede observarse que los distintos tipos de variables pueden proporcionar diferente información¹:

Doxazosina para la hipertensión

VOE: reduce la presión arterial en pacientes hipertensos de raza negra

VOP: aumenta los eventos cardiovasculares en comparación con clortalidona

THS para la prevención de eventos cardiovasculares

VOE: reduce el LDL-colesterol y aumenta el HDL-colesterol

VOP: no disminuye la mortalidad cardiovascular ni la mortalidad por cualquier causa; aumenta los eventos cardiovasculares en mujeres de más de 60

Variable de resultado (“endpoint”) primaria o principal vs. variable secundaria

Los ensayos clínicos están diseñados para responder a una pregunta concreta (por ejemplo, si un fármaco es más eficaz que otro en prevenir un evento específico), que se denomina variable de resultado principal¹. La variable principal nos indica el objetivo principal del estudio y, por tanto, el tamaño de la muestra se calcula en función de éste.

Sería raro que una sola variable abarcara todos los aspectos que interesan a los investigadores, por lo que habitualmente los ensayos miden varios resultados simultáneamente, siendo importante distinguir la variable principal de las secundarias (son variables de interés, con fines más descriptivos). En la tabla 2 se presentan los aspectos que debemos de tener en cuenta al valorar los resultados de una variable secundaria¹.

Tabla 2. Valoración de las variables secundarias

- Debe estar contemplada en el diseño del estudio, y su inclusión debería estar adecuadamente justificada.
- Es probable que el estudio no tenga una potencia estadística adecuada para las variables secundarias.
- Las comparaciones múltiples: cuantas más variables se midan, mayor es la probabilidad de encontrar algún resultado estadísticamente significativo debido al azar.
- Si el resultado principal no es significativo, la precaución al interpretar los resultados de una variable secundaria ha de ser aún mayor (¡podría ser otra forma de “salvar” el estudio cuando la hipótesis principal no ha conseguido demostrar lo esperado!).

Variable de resultado simple vs. variable “agregada” o compuesta (“composite endpoint”)

Las variables agregadas engloban distintos efectos y requieren menor tamaño de muestra y menor tiempo de seguimiento. La validez de una variable agregada depende de que la intervención afecte de igual forma a todos los componentes, de la frecuencia con la que ocurren y de la importancia clínica que tengan para los pacientes⁷.

Por ejemplo, una variable agregada cardiovascular puede ser definida como muerte, infarto de miocardio no fatal y hospitalización por síndrome coronario agudo. Para valorar la idoneidad de esta variable agregada nos podemos guiar por la Tabla 3. La muerte y el infarto tendrán mayor importancia clínica que las hospitalizaciones por síndrome coronario, aunque éstas últimas sean las que ocurran con mayor frecuencia.

Tabla 3. Guía para interpretar las variables agregadas^a

1. ¿Los distintos componentes tienen la misma importancia clínica?
2. ¿Ocurren los resultados con similar frecuencia?
3. ¿El efecto del tratamiento sobre los distintos componentes es similar?

¿ES EL ENSAYO METODOLÓGICAMENTE CORRECTO?^{7,9}

Los aspectos más “técnicos” de los ECA son probablemente los más difundidos en los libros sobre metodología de la investigación; además, en los últimos años ha mejorado considerablemente la calidad metodológica de los ensayos.

Aleatorización

La fortaleza principal de un ensayo clínico para poder demostrar una relación de causa-efecto es la adecuada aleatorización, ya que garantiza, siempre que el tamaño de muestra sea suficiente, que los grupos sean comparables antes de iniciar el tratamiento.

Es importante que haya Ocultación de la Secuencia de Aleatorización (“OSA”), es decir, que el investigador no pueda saber a qué rama va a ser asignado un paciente ni pueda alterar el orden de aleatorización.

Enmascaramiento

Idealmente, en un ECA ni el paciente ni el investigador deberían saber qué tratamiento están recibiendo el grupo experimental o el grupo control, lo que se conoce como “doble ciego”. Técnicamente implica que las dos opciones a comparar sean idénticas en apariencia y se administren de la misma forma.

El enmascaramiento es especialmente importante cuando las variables que se miden son subjetivas (como calidad de vida, dolor...) o cuando es posible que, al margen de la intervención a estudio, puedan añadirse otros tratamientos concomitantes (el investigador podría añadirlos de forma desigual al grupo experimental y al control).

Si por razones éticas o técnicas el ensayo no puede ser doble ciego (ej. intervenciones quirúrgicas, educativas, tratamientos parenterales, programa de ejercicio físico...), al menos el evaluador de los resultados debería desconocer la intervención asignada (“evaluación ciega por terceros”), y utilizar medidas de resultado objetivas (como los resultados de laboratorio, mortalidad, infarto...).

El seguimiento de los pacientes

Siempre se deben cuantificar y describir las causas de la pérdida de pacientes. Es importante que los motivos de las pér-

didias afecten de la misma forma al grupo experimental y al control. Si en una de las ramas hay más abandonos debido a efectos adversos o a falta de eficacia, los resultados estarán sesgados.

También es importante que, al analizar los resultados, se siga el principio del “análisis por intención de tratar” (ITT), de forma que todos los pacientes se contabilicen en la rama a la que fueron aleatorizados (aunque no completen el estudio). Se considera que el ITT ofrece unos resultados más parecidos a los de la realidad (donde no todos los pacientes cumplen con sus tratamientos) y es más conservador. La excepción a esta regla son los estudios de equivalencia y de no inferioridad, donde se deben presentar tanto el ITT como el análisis por protocolo.

¿Tiene el estudio un número suficiente de pacientes? Potencia estadística

La potencia estadística es la probabilidad de detectar una diferencia significativa real, no debida al azar. Arbitrariamente se considera adecuado un poder estadístico del 80% para la mayoría de los estudios. Antes de realizar un ensayo, se debe realizar el cálculo del tamaño de muestra, teniendo en cuenta la magnitud del efecto que se pretende observar y las pérdidas de seguimiento que se esperan. Es posible pecar por defecto (si el número de pacientes es demasiado pequeño es muy posible que no encontremos diferencias significativas, aunque éstas realmente existan) o por exceso (aumentando desproporcionadamente el tamaño de muestra casi siempre será posible demostrar una diferencia estadísticamente significativa, aunque clínicamente sea irrelevante).

Continúa en Volumen 14 nº 8, PARTE II

Galdera, iradokizun edo parte-hartze lanak nori zuzendu / Para consultas, sugerencias y aportaciones dirigirse a: zure komarkako farmazialaria / el farmacéutico de su comarca o CEVIME/MIEZ - tel. 945 01 92 66 - E-mail: cevime-san@ej-gv.es

Idazkuntza Batzordea / Consejo de Redacción: José Ramón Agirrezabala, Inigo Aizpuru, Miren Albizuri, Iciar Alfonso, María Armendáriz, Arrate Bengoa, Arribu Etxeberria, Julia Fernández, Susana Fernández, Anabel Giménez, Juan José Iglesias, Josune Iribar, Jesús Iturralde, Nekane Jaio, Itxasne Lekue, M^o José López, Javier Martínez, Carmela Mozo, Elena Olloquegi, Elena Ruiz de Velasco, Rita Sainz de Rozas, Elena Valverde.

ISSN: 1575054-X
D.L.: BI-587-99



Osakidetza
Servicio vasco de salud



Eusko Jaurlaritzaren Argitalpen Zerbitzu Nagusia
Servicio Central de Publicaciones del Gobierno Vasco

Sumario

INTERPRETANDO LA LITERATURA MÉDICA: ¿QUÉ NECESITO SABER?

PARTE I

- INTRODUCCIÓN
- ¿SOBRE QUÉ SE INVESTIGA?
El papel de la industria y de las agencias reguladoras
- ¿QUÉ SE PUBLICA DE LO QUE SE INVESTIGA? Sesgo de publicación y sesgo en la presentación de resultados
- UNA VEZ PUBLICADO, ¿ES RELEVANTE EL ENSAYO? ¿PUEDE SER ÚTIL PARA MI PRÁCTICA CLÍNICA?
- ¿ES EL ENSAYO METODOLÓGICAMENTE CORRECTO?

PARTE II

- RELEVANCIA CLÍNICA: ¿QUÉ NOS DICEN LOS NÚMEROS?
 - ¿Qué es la p ?
 - ¿Qué es el intervalo de confianza?
 - Métodos más frecuentes de expresar los resultados
- ANÁLISIS DE SUBGRUPOS
- LA PROLIFERACIÓN DE NUEVOS TIPOS DE SESGOS: CONSEJOS PARA SOBREVIVIR EN EL INTENTO

INTERPRETANDO LA LITERATURA MÉDICA: ¿QUÉ NECESITO SABER?

PARTE II

RELEVANCIA CLÍNICA:

¿QUÉ NOS DICEN LOS NÚMEROS?

A menudo los ensayos clínicos están repletos de estadística y expresiones numéricas, indescifrables para el lector poco familiarizado. Una vez más, es necesario resaltar que **la estadística no sustituye al sentido común** y, en ocasiones, bajo la apariencia de unos análisis complejos, puede esconderse la ausencia de unos resultados clínicamente relevantes.

A continuación describiremos el significado de la p y de los intervalos de confianza, así como los métodos más frecuentes de expresar los resultados, utilizando como ejemplo el estudio CURE¹⁰.

Resumen de resultados del ensayo CURE, con fines ilustrativos

El estudio CURE comparó la combinación de clopidogrel + aspirina frente a aspirina sola en el síndrome coronario agudo. Tras 9 meses de tratamiento, el resultado principal (variable agregada de muerte cardiovascular, infarto e ictus), fue menor en los pacientes tratados con la combinación que en los tratados con aspirina (9,3% vs. 11,4%, riesgo relativo 0,80 (IC95%: 0,72-0,90)). El riesgo de sangrado fue mayor con la combinación (3,7% vs. 2,7%, $p=0,001$).

¿Qué es la p ? Significación estadística vs. relevancia clínica¹

La p sólo nos mide la probabilidad de que los resultados sean explicados por el azar y de manera arbitraria se considera que una $p < 0,05$ es lo suficientemente pequeña como para aceptar que las diferencias se deben a la intervención y no al azar. Tradicionalmente se le ha dado mucha importancia, pero el valor de p no dice nada acerca de la magnitud del efecto. Podemos tener una p muy significativa pero una diferencia clínicamente irrelevante.

En el estudio CURE el valor de p para el riesgo de sangrado era de 0,001, lo que significa que la probabilidad de que no exista diferencia entre los dos tratamientos es de 1 entre 1000 (muy poco probable, p muy significativa).

¿Qué es el intervalo de confianza (IC)?¹

El IC indica el rango de valores entre los que se encontrará el verdadero valor de la variable en la población, con un nivel de confianza que se establece de antemano (generalmente el 95%). Es importante ver la amplitud del intervalo: un IC muy amplio nos indica que las estimaciones son poco precisas.

En nuestro ejemplo, podemos afirmar con una confianza del 95% que el verdadero valor del beneficio de la combinación estará comprendido entre un riesgo relativo de 0,72 y 0,90.

MÉTODOS MÁS FRECUENTES DE EXPRESAR LOS RESULTADOS¹

RIESGO ABSOLUTO (RA) Y RIESGO RELATIVO (RR)

El riesgo absoluto se refiere a la frecuencia de eventos en un determinado grupo, mientras que el riesgo relativo estima el efecto de una intervención dividido por el efecto de su comparador.

En el estudio CURE el riesgo absoluto de padecer el resultado principal en el grupo tratado con la asociación fue del 9,3% y del 11,4% en el grupo de la aspirina. El RR es $9,3/11,4=0,82^*$.

REDUCCIÓN ABSOLUTA DEL RIESGO (RAR) Y REDUCCIÓN RELATIVA DEL RIESGO (RRR)

La RAR es la diferencia de la frecuencia de eventos entre los dos grupos.

La RRR expresa la reducción de riesgo en relación al comparador, y se calcula mediante la fórmula $1-RR$.

La industria tiende a dar los resultados en términos de riesgo relativo, más que de riesgo absoluto.

En nuestro ejemplo la RAR sería $11,4\%-9,3\%=2,1\%$ y la RRR sería $1-0,80^*=0,20$, es decir, una reducción del 20%. Para el lector este 20% resulta mucho más impactante que la RAR del 2,1%.

NÚMERO NECESARIO DE PACIENTES A TRATAR (NNT) Y NÚMERO NECESARIO DE PACIENTES PARA PERJUDICAR O “NUMBER NEEDED TO HARM” (NNH)

NNT

Es el número de pacientes que se necesitaría tratar con una intervención específica para producir o evitar un evento adicional. Se calcula como $1/RAR$ (el RAR lo expresaremos entre 0 y 1, no en porcentaje). Cuanto menor sea el NNT mayor será el efecto de la intervención.

El NNT depende del riesgo basal, ya que las intervenciones dirigidas a personas de alto riesgo tienen mayor beneficio absoluto que las que se dirigen a poblaciones de bajo riesgo. El NNT se debe referir siempre a un tiempo determinado y a una población determinada.

En nuestro ejemplo, el NNT sería $1/0,021=48$.

Por cada 48 pacientes tratados con la combinación, en lugar de aspirina sola, durante 9 meses, evitaremos un evento.

NNH

Se define del mismo modo que el NNT para medir efectos indeseables.

Para el sangrado el NNH sería $(1/(0,037-0,027))= 100$.

Por cada 100 pacientes tratados con la combinación, en lugar de aspirina sola, durante 9 meses, uno sufrirá un sangrado adicional.

El NNT y el NNH son medidas intuitivas para el clínico, y pueden facilitar la toma de decisiones.

En nuestro ejemplo, podemos concluir que por cada 100 pacientes tratados con aspirina+clopidogrel en lugar de aspirina sola en el síndrome coronario agudo durante 9 meses, evitaremos 2 eventos (muerte, infarto o ictus) y produciremos una hemorragia. Visto así, el balance entre beneficios y riesgos es favorable, aunque seguramente menos espectacular que pensar que evitaremos el 20% de los infartos, ictus o muertes, tal y como se podía haber interpretado al tener en cuenta solamente el valor del RRR.

* En el estudio, el RR = 0,80 fue calculado según curvas de Kaplan-Meier

ANÁLISIS DE SUBGRUPOS

(¡atención!, sobre todo si son retrospectivos o “a posteriori”)

El análisis de subgrupos se realiza cuando interesa evaluar los resultados de algún grupo especial de sujetos del ECA, habitualmente en función de su edad, sexo, gravedad, comorbilidad, etc. **En los análisis de subgrupos, la probabilidad de encontrar un resultado significativo debido al azar aumenta con el número de análisis que realicemos.** La revista Lancet publicó un ejemplo muy ilustrativo. En el estudio ISIS-2¹¹, aspirina fue muy eficaz en prevenir la muerte y el infarto. Los editores pidieron a los investigadores que incluyeran 40 análisis de subgrupos. Los investigadores, reacios a este tipo de análisis, “demostraron” que para los pacientes nacidos bajo los signos Géminis y Libra del zodiaco el tratamiento con aspirina era perjudicial, y lo publicaron como muestra de los sesgos en la interpretación en el análisis de subgrupos¹².

Cuando el tratamiento no demuestra los efectos esperados, el análisis de subgrupos puede ser aún más controvertido, ya que podría tratarse de “encontrar a toda costa” algún grupo de pacientes a los que el tratamiento beneficie. Obviamente, la transparencia en las publicaciones es fundamental, y en todo momento los autores deberían especificar claramente si los subgrupos estaban definidos en el diseño del estudio y si muestran todos los análisis que se han realizado (¡y no sólo los favorables!).

Cuando el registro único de ensayos clínicos se haga efectivo y tengamos acceso a los protocolos, se solventarán problemas como el análisis de subgrupos, ya que podremos cotejar lo previsto en el protocolo y lo finalmente publicado.

Tabla 4. Criterios de validez en el análisis de subgrupos. Adaptado de 13.

• La hipótesis específica se había planteado en el diseño del estudio y es una de las pocas hipótesis que se quiere contrastar
• La magnitud del efecto es grande
• El efecto del tratamiento es altamente significativo (cuanto más bajo es el valor de la p más creíble será la diferencia)
• Los resultados son consistentes con los de otros estudios
• Existe evidencia indirecta a favor de los resultados encontrados (plausibilidad biológica)
• Más creíbles si se observa un beneficio en la población global del estudio

LA PROLIFERACIÓN DE NUEVOS TIPOS DE SESGOS: CONSEJOS PARA SOBREVIVIR EN EL INTENTO

Como ya se ha comentado a lo largo de este boletín, el clínico que desea aplicar los nuevos conocimientos para mejorar la asistencia de sus pacientes se enfrenta a dos grandes retos: la sobrecarga de información y los sesgos en el diseño, desarrollo, análisis y publicación de los estudios.

La solución a esta situación implica medidas efectivas a diferentes niveles²:

- Mayor inversión pública en investigación, dirigida a contestar preguntas socialmente relevantes.
- Transparencia en la investigación: registro de los protocolos de los ensayos clínicos, publicación completa de los resultados.
- Implementación de los principios éticos y normativos en los ámbitos de las administraciones públicas, los centros académicos, la industria farmacéutica y los editores de revistas médicas.
- Cambios en la regulación de los ensayos y en el papel de los comités de ensayos, con un mayor esfuerzo para evi-

tar el desarrollo de ensayos que no aporten conocimiento y para el seguimiento de su publicación.

- Cambios en los requerimientos de las agencias reguladoras para el registro de medicamentos, enfatizando el concepto de “avance terapéutico” (beneficio para el paciente y la sociedad) frente al actual concepto comercial de “innovación”.

Como sugiere Salvador Peiró en su artículo “La construcción de la *evidence based medicine*”², mientras se implementan algunas de estas estrategias y dado que la alternativa de desconfiar sistemáticamente de los ensayos patrocinados es peor que cualquier otra, puede utilizarse la guía de Montori y cols¹⁴ para evitar interpretaciones erróneas de los datos (Tabla 5).

Esta guía está más centrada en la relevancia de los estudios que en los aspectos más técnicos; para estos últimos existen excelentes guías en castellano que nos ayudan a reconocer las limitaciones metodológicas y que se encuentran disponibles en: (<http://www.redcaspe.org/homecasp.asp>).

Tabla 5. Guía para evitar la interpretación sesgada de los datos¹⁴

• Leer únicamente las secciones de “método” y “resultados”, obviando la sección de discusión
• Fijarse en los comparadores inadecuados: ¿es el placebo el comparador adecuado?
• Fijarse en la elección de las variables de resultado: ¿agregadas?
• Cuidado cuando el efecto del tratamiento (la magnitud) es pequeño (ej. estudios con gran número de pacientes con diferencias significativas próximas al efecto nulo). Valorar los resultados en términos absolutos (NNT, reducción absoluta del riesgo), más que en medidas relativas como el riesgo relativo o la reducción relativa del riesgo. Fijarse en los límites de los intervalos de confianza
• Leer los resúmenes evaluados críticamente en revistas secundarias (como ACP Journal Club o Evidence-based Medicine) realizados por un equipo de clínicos y metodólogos sin conflictos de interés
• Atención a los análisis de subgrupos

Finalmente, retomando lo ya dicho en este boletín, quisiéramos añadir un primer punto a esta guía de Montori, la “**pregunta del millón**” que nos deberíamos hacer antes de decidir si merece la pena proseguir con la lectura de un estudio: **¿Es relevante el estudio? ¿Puede ser útil para mi práctica clínica?**

LECTURAS Y FUENTES SECUNDARIAS RECOMENDADAS

A continuación se detalla una relación no exhaustiva de fuentes secundarias fácilmente accesibles en las que se pueden obtener revisiones de temas clínicos y artículos comentados basados en la evidencia, que pueden resultar útiles para la clínica diaria.

Fuentes secundarias (bases de datos, revistas, CATs o temas valorados críticamente...)	
UpToDate <i>Web Osakidetza: en Colecciones Especiales</i>	Revisiones de temas clínicos basados en la evidencia.
Clinical Evidence <i>Web Osakidetza: en Colecciones Especiales</i>	Revisiones sobre eficacia de las intervenciones, basado en la evidencia.
ACP Journal Club <i>Web Osakidetza: en Bases de datos- EMB Reviews</i>	Resúmenes de estudios seleccionados por su validez y relevancia, evaluados críticamente.
Evidence Based Medicine <i>Inglés: Web Osakidetza: en Revistas- Sumarios OVID</i> <i>Castellano: http://ebm.isciii.es</i>	
Bandolera (versión en español de Bandolier) http://www.infodoctor.org/bandolera/	Resúmenes de la literatura médica para una práctica basada en la evidencia.
CATs del grupo MBE del Hospital Donostia <i>Disponible en: www.donostiaospitalea.org/MBE</i>	Temas valorados críticamente por dicho grupo.
Metodología general en lectura crítica	
Argimon Pallás JM, Jiménez Villa J. Métodos de investigación clínica y epidemiológica (3ª edición). 2004. Madrid. Editorial Elsevier	
CASPe: Programa de habilidades en lectura crítica. Contiene herramientas para facilitar la lectura crítica de la literatura. <i>Disponible en: http://www.redcaspe.org/homecaspe.asp</i>	

Accesos desde la página web de Osakidetza: Biblioteca virtual - Recursos electrónicos.

BIBLIOGRAFÍA

- Using evidence to guide practice. *MeReC Briefing*. 2005;30:1-8.
- Peiró S. La construcción de la evidencia b(i)ased medicine. *Gestión Clínica y Sanitaria*. 2005;7(4):131-38.
- Melander H, Ahlqvist-Rastad J, Meijer G, Beermann B. Evidence b(i)ased medicine - selective reporting from studies sponsored by pharmaceutical industry: review of studies in new drug applications. *BMJ*. 2003;326(7400):1171-3.
- Smith R. Medical journals are an extension of the marketing arm of pharmaceutical companies. *PloS Med*. 2005;2(5):e138.
- Lexchin J, Bero LA, Djulbegovic B, Clark O. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *BMJ*. 2003;326(7400):1167-70.
- Als-Nielsen B, Chen W, Gluud C, Kjaergard LL. Association of funding and conclusions in randomized drug trials: a reflection of treatment effect or adverse events? *JAMA*. 2003;290(7):921-8.
- Evaluating clinical evidence - what do I need to know? *WeMeReC Bulletin*. May 2006:1-6.
- Montori VM, Permyer-Miralda G, Ferreira-Gonzalez I, Busse JW, Pacheco-Huergo V, Bryant D, et al. Validity of composite end points in clinical trials. *BMJ*. 2005;330(7491):594-6.
- Guyatt GH, Sackett DL, Cook DJ. Users' Guides to the Medical Literature: II. How to Use an Article About Therapy or Prevention: A. Are the Results of the Study Valid? *JAMA*. 1993;270(21):2598-2601.
- Angina to Prevent Recurrent Events Trial Investigators. Effects of clopidogrel in addition to aspirin in patients with acute coronary syndromes without ST-segment elevation. *N Engl J Med*. 2001;345(7):494-502.
- Randomised trial of intravenous streptokinase, oral aspirin, both, or neither among 17,187 cases of suspected acute myocardial infarction: ISIS-2 (Second International Study of Infarct Survival) Collaborative Group. *Lancet*. 1988;2(8607):349-60.
- Schulz KF, Grimes DA. Multiplicity in randomised trials II: subgroup and interim analyses. *Lancet*. 2005;365(9471):1657-61.
- Freemantle N. Interpreting the results of secondary end points and subgroup analyses in clinical trials: should we lock the crazy aunt in the attic? *BMJ*. 2001;322(7292):989-91.
- Montori VM, Jaeschke R, Schönemann HJ, Bhandari M, Brozek JL, Devereaux PJ et al. Users' guide to detecting misleading claims in clinical research reports. *BMJ*. 2004; 329:1093-6.

Fecha de revisión bibliográfica: agosto 2006

Galdera, iradokizun edo parte-hartze lanak nori zuzendu / Para consultas, sugerencias y aportaciones dirigirse a: zure komarkako farmazialaria / el farmacéutico de su comarca o CEVIME/MIEZ - tel. 945 01 92 66 - E-mail: cevime-san@ej-gv.es

Idazkuntza Batzordea / Consejo de Redacción: José Ramón Agirrezabala, Iñigo Aizpura, Miren Albizuri, Iciar Alfonso, María Armendáriz, Arrate Bengoa, Aritxu Etxeberria, Julia Fernández, Susana Fernández, Anabel Giménez, Juan José Iglesias, Josune Iribar, Jesús Iturralde, Nekane Jaio, Itxasne Lekue, M^a José López, Javier Martínez, Carmela Mozo, Elena Olloquegi, Elena Ruiz de Velasco, Rita Sainz de Rozas, Elena Valverde.

ISSN: 1575054-X
D.L.: BI-587-99



Eusko Jaurlaritzaren Argitalpen Zerbitzu Nagusia
Servicio Central de Publicaciones del Gobierno Vasco