

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

1. Using evidence to guide practice. MeReC Briefing. 2005;30:1-8.
2. Peiró S. La construcción de la evidencia b(i)ased medicine. *Gestión Clínica y Sanitaria*. 2005;7(4):131-38.
3. 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.
4. Smith R. Medical journals are an extension of the marketing arm of pharmaceutical companies. *PloS Med*. 2005;2(5):e138.
5. Lexchin J, Bero LA, Djulbegovic B, Clark O. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *BMJ*. 2003;326(7400):1167-70.
6. 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.
7. Evaluating clinical evidence - what do I need to know? *WeMeReC Bulletin*. May 2006:1-6.
8. 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.
9. 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.
10. 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.
11. 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.
12. Schulz KF, Grimes DA. Multiplicity in randomised trials II: subgroup and interim analyses. *Lancet*. 2005;365(9471):1657-61.
13. 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.
14. 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