

# Ronda clínica y epidemiológica

OSCAR OSÍO URIBE<sup>1</sup>, JOHN JAIRO ZULETA TOBÓN<sup>2</sup>

## ERRORES SISTEMÁTICOS EN LA MEDICIÓN DE LA EFICACIA TERAPÉUTICA

ÓSCAR OSÍO URIBE

**L**as mediciones de los efectos de la exposición a un medicamento o de los resultados de una intervención sanitaria pueden ser erróneas por dos motivos: por una fluctuación aleatoria (atribuible al azar) o por desviaciones sistemáticas de la medición. El primero de los errores, la imprecisión, no debe ser confundido con el segundo, el desvío repetido y en un solo sentido de las mediciones. La precisión es una medida de la probabilidad de que se observe una asociación determinada por puro azar, lo que conduce a una conclusión positiva falsa. La precisión de un estudio clínico se determina por la amplitud del intervalo de confianza y en el caso de las revisiones sistemáticas de la literatura (RSL), por la estrategia de dar una mayor valoración a los estudios que muestren los resultados más precisos. La validez de un estudio clínico se refiere a la capacidad que tuvieron sus autores de evitar o disminuir los errores sistemáticos o sesgos durante las fases de su diseño y mientras se ejecutaba el experimento clínico. Los siguientes párrafos estarán dedicados a los errores investigativos que minan la validez de los estudios clínicos, los denominados sesgos en la investigación, fenómeno que algunos autores ven como una inclinación reiterada de la mente del investigador hacia un solo lado de la medición de la asociación y que origina, la mayoría de las veces, una magnificación del efecto terapéutico de una intervención sanitaria, dando la impresión de un beneficio mayor del que en realidad existe. Diseños investigativos pobres hacen que los diferentes tratamientos médicos parezcan mejor de lo que en realidad son, o incluso pueden hacer aparecer como útiles intervenciones dañinas.

.....  
<sup>1</sup> Especialista en medicina interna, MSc Epidemiología Clínica. Director corporación Académica de Patologías Tropicales Universidad de Antioquia.  
cpt\_udea@yahoo.com

<sup>2</sup> Profesor de Ginecología y Obstetricia, MSc en Epidemiología Clínica. Universidad de Antioquia, Medellín, Colombia.  
Dirección electrónica: jjzuleta@une.net.co

## ¿QUÉ ES UN SESGO?

Una disposición sistemática en el diseño de algunos ensayos clínicos que lleva a producir de manera consistente resultados mejores o peores que los producidos por otros diseños investigativos.

## ¿CÓMO SE AFECTAN LAS RSL POR LOS SESGOS?

Se recomienda que las RSL de los efectos de un tratamiento médico incluyan estrategias que ayuden a sus realizadores a disminuir los sesgos que son más comunes, practicando además análisis de sensibilidad que los ayuden a identificar la medida en la cual determinados ensayos clínicos afectan los resultados de la revisión. Desgraciadamente esta recomendación no se sigue siempre por los autores de una RSL, conduciendo de manera frecuente a que las revisiones contengan información y conclusiones incorrectas, llevando a errores y a confusión en los médicos lectores y en las personas que toman decisiones sanitarias para la comunidad.

## ¿CUÁLES SON LOS SESGOS MÁS FRECUENTES QUE AFECTAN LA VALIDEZ DE LA MEDICIÓN DE LA EFICACIA DE UNA INTERVENCIÓN EN SALUD?

**LOS DE SELECCIÓN:** asignación no aleatoria o defectuosa de las personas que se someten a las diferentes exposiciones terapéuticas. Llevan a que las razones de disparidad (OR) se magnifiquen hasta en un 41% cuando la asignación aleatoria es defectuosa, o hasta 30% de aumento del estadígrafo cuando no está claramente descrito el proceso. Una asignación no aleatoria de los pacientes puede llegar al extremo de hacer ver como eficaz un tratamiento que no lo es.

**SESGOS DE OBSERVACIÓN:** la falta de ceguera en los investigadores que observan los resultados de una intervención o el conocimiento que

tengan los pacientes del tipo de tratamiento que recibieron puede llevar a que se sobreestimen de manera sistemática los efectos de un tratamiento particular. La falta de ceguera en una investigación clínica puede llevar a sobreestimaciones de la eficacia de una intervención sanitaria hasta de 17%.”

### **SESGOS DERIVADOS DE LA FALTA DE UTILIZACIÓN DE INSTRUMENTOS PARA MEDIR LA CALIDAD DE LOS ESTUDIOS CLÍNICOS INDIVIDUALES:**

cuando en una revisión sistemática no se utilizan herramientas que evalúen la calidad de la asignación aleatoria de los pacientes y la observación que hace el investigador en cada uno de los ensayos clínicos incluidos no es ciega, se magnifica de manera importante la eficacia de la intervención. Se han propuesto decenas de estos instrumentos de medición de la calidad en forma de listas y escalas. Uno de los más sencillos y valiosos es la escala de Jadad, propuesta por un anestesiólogo colombiano, doctorado en síntesis del conocimiento. Esta escala asigna un punto por cada respuesta positiva a las siguientes preguntas: ¿En la descripción del estudio se explica que la asignación de las personas a la exposición fue aleatoria? ¿En la descripción del estudio se explica que los observadores estaban ciegos para los resultados de la exposición? ¿Se hace alguna descripción de las pérdidas y de los razones del abandono del tratamiento que tuvieron los individuos dentro de los grupos conformados? Además de los tres puntos de calificación anteriores, si la aleatorización o la ceguera fueron apropiados se suman otros dos puntos, y se restan dos si la aleatorización o la ceguera no fueron adecuadas, permitiendo de esa manera un puntaje global que puede ir de cero a cinco puntos. Las RSL que incluyen estudios individuales con puntajes de dos o menos unidades tienen la tendencia sistemática a magnificar hasta en un 25% la eficacia de los tratamientos médicos, en comparación con aquellas revisiones que incluyen estudios con puntajes  $\geq 3$ .

**SESGOS GEOGRÁFICOS:** pueden explicar las diferencias observadas en la eficacia de un tratamiento cuando sus beneficios son evaluados en un determinado medio geográfico o en ciertos contextos culturales, sociales o religiosos. Un ejemplo de este error sistemático sería la tendencia a presentar mejores resultados de eficacia en los diversos tratamientos con acupuntura hechos en los países orientales que cuando estas técnicas se evalúan en Australia o en Estados Unidos.

**SESGOS DE TAMAÑO MUESTRAL:** resultan cuando en una misma RSL se mezclan estudios clínicos pequeños con otros con gran número de pacientes, aumentando la estimación global de la eficacia hasta en un 30%. Como norma empírica no aceptada por todos los autores, se recomienda no incluir estudios clínicos que tengan menos de diez pacientes para no incurrir en este sesgo.

**SESGOS DE LENGUAJE:** suceden cuando la búsqueda de la literatura se limita a un solo idioma, el inglés por ejemplo. Los autores no ingleses tienen tendencia a enviar artículos que muestran sus éxitos terapéuticos a revistas inglesas que ellos consideran le darán mayor resonancia internacional a sus hallazgos, mientras que aquellos experimentos terapéuticos fallidos tienen la tendencia a publicarse en lenguas propias, diferentes de la inglesa. Los sesgos de lenguaje conducen a sobrestimar la eficacia de los tratamientos de manera poco conocida, posiblemente modesta.

**SESGOS DE DUPLICACIÓN:** ocurren cuando los resultados de algunos ensayos clínicos son informados más de una vez, por una amplia gama de razones. Los estudios clínicos más impactantes tienden a ser duplicados o repetidos en artículos que aparecen en diferentes revistas médicas, muchas veces en cabeza de autores completamente diferentes de los de la investigación original, llevando a sobrestimaciones de la eficacia terapéutica hasta de un 20%.

En una publicación científica también se puede incurrir en sesgos de tipo estadístico cuando se muestran datos como estadísticamente significativos, que en realidad no lo son, cuando se emplean pruebas inapropiadas, cuando se manipulan los datos de la investigación o cuando los investigadores reclaman en sus escritos haber tenido éxito con nuevos tratamientos sin tener el respaldo de los números. Mención especial merecen los resúmenes de los artículos que evalúan eficacia terapéutica, ya que usualmente son optimistas y tienen la tendencia a dar partes de victoria que no se corresponden con la verdad, como seguramente detectaría el lector juicioso y con tiempo que se lea todo el artículo.

Finalmente se debe decir que los médicos hemos venido siendo estimulados durante los últimos años a mejorar nuestras habilidades para usar mejor la información proveniente de la investigación clínica, poniéndola al servicio de nuestros pacientes, dando pie a que como lectores caigamos en sesgos en el momento de interpretar esa información a la luz de nuestros prejuicios o de nuestra propia experiencia. Son ejemplos de estos sesgos de lectura, los de confirmación (tendencia que tiene el lector de un artículo médico a evaluar la evidencia que confirma sus prejuicios de manera diferente a como evalúa la información que los contradice), los sesgos de rescate (intentos deliberados de evadir la evidencia que contradice una expectativa personal o profesional) y los sesgos de quien solicita siempre esperar más evidencias (o sea esperar indefinidamente otra investigación clínica que nos suministre la evidencia que termine confirmando nuestras creencias).

Un ejemplo reciente de sesgo de publicación fue el documentado por Vaitkus y Brar con respecto al uso de la N-acetilcisteína para prevenir la nefropatía inducida por medios de contraste, que ha sido sujeto de numerosos ensayos clínicos y metanálisis. Mediante la comparación de las estimaciones de los efectos del

tratamiento en las investigaciones y las revisiones sistemáticas publicadas y las estimaciones obtenidas a partir de las investigaciones que nunca fueron publicadas, se demostró que las primeras eran más optimistas de la eficacia de la N-acetilcisteína que las últimas. También demostraron que hubo un periodo de publicación de estimaciones que mostraban un mayor beneficio durante el periodo de tiempo que siguió a las primeras publicaciones, concluyendo que ocurrió un sesgo importante de publicación con las primeras investigaciones clínicas individuales, error que luego se amplificó con la publicación de las revisiones sistemáticas de la literatura acerca del tema.

## BIBLIOGRAFÍA

1. www.ebandolier.com Evidence-based health care. Bandolier 80, Diciembre 2001
2. Sterne J, Egger M, Davey G. Investigating and dealing with publication and other biases in meta-analysis. BMJ 2001; 323: 101-105
3. Kaptchuk T. Effect of interpretive bias on research evidence. BMJ 2003; 326:1453-1455
4. Vaitkus P., Brar C. N-Acetylcysteine in the prevention of contrast-induced nephropathy: Publication bias perpetuated by meta-analyses. AM Heart J 2007; 153: 275-280

## ¿SESGOS O MANIPULACIONES?

**JOHN JAIRO ZULETA TOBÓN**

Todas las guías de lectura crítica de literatura médica hacen énfasis en la detección de potenciales sesgos como causantes o modificadores de los resultados encontrados en un estudio. Estos sesgos se buscan

en los elementos metodológicos de la implementación de la investigación, sin embargo existe otra serie de factores que también pueden desviar falsamente los resultados y no se han resaltado en la dimensión que ellos se merecen. Estos factores no se pueden catalogar como sesgos y en algunos casos tocan aspectos éticos de la investigación. Existe un buen número de publicaciones que llaman la atención hacia ellos y a continuación se presentan algunos.

Johansen y Gotzsche<sup>1</sup>, del Centro Nórdico de la Colaboración Cochrane realizaron un metanálisis para evaluar la efectividad del fluconazol en pacientes con neutropenia. El fluconazol tiene una buena absorción por vía oral, contrario a la anfotericina B, que no se recomienda por tal vía. A pesar de ello algunas investigaciones que evaluaron la eficacia del fluconazol lo hicieron comparándolo contra la anfotericina B por vía oral, y no intravenosa, que es la vía con la cual ha demostrado buen efecto. La nistatina se considera un medicamento poco efectivo en pacientes inmunosuprimidos, y a pesar de ello fue el seleccionado como comparación con el fluconazol en otros estudios. Obviamente que esta manipulación sesga los resultados en favor de éste último medicamento. Un segundo elemento denunciado por los autores fue la multiplicidad de publicaciones de los mismos pacientes disimulados como estudios diferentes, incluso con autores diferentes. Esta estrategia llevaría a aceptar que los resultados son reproducibles porque no es una sola investigación sino que son varias las que corroboran un hallazgo. Un tercer llamado de atención es la precaución que se debe tener cuando en un metanálisis se incluyen los resultados de estudios no publicados en los cuales el único que tiene acceso a los datos es el patrocinador del producto en estudio.

Safer<sup>2</sup> dedicó un artículo a denunciar una serie de modificaciones en el diseño y en el reporte de estudios clínicos de algunos psicofármacos. Con un buen sustento con base en artículos publicados presentó

trece estrategias dirigidas a favorecer al medicamento que se quiere promover. Se resaltan acá las que tienen que ver con el proceso investigativo, no con la publicación de los resultados. Una primera es sobredosificar el medicamento seleccionado como control para lograr que el grupo que lo reciba manifieste mayor cantidad y mayor intensidad de efectos secundarios, y específicamente lo han hecho usando haloperidol a dosis de 20 mg/día a pesar de que se sabe que la dosis terapéutica está entre 4 y 10 mg/día. Una segunda es realizar incrementos en el medicamento de control a intervalos más cortos o con esquemas diferentes de los rutinariamente recomendados, o también subdosificando el medicamento patrocinado con el fin de demostrar menores efectos secundarios. Una tercera es emplear escalas de medición de desenlaces no validadas que favorecen la intervención. La cuarta es presentar únicamente los resultados positivos de múltiples análisis que no estaban planeados en el protocolo, sino que son productos de una búsqueda en la base de datos ("pesca"). La quinta es enmascarar o disimular los resultados negativos; aunque no está citado por Safer en su artículo, un excelente ejemplo de esta situación fue la vivida con los inhibidores de COX-2 que inicialmente no se presentaron como dañinos desde el punto de vista cardiovascular, sino que se propuso que los medicamentos con los que los compararon eran protectores. La sexta es seleccionar los sujetos de investigación y modificar la duración de los estudios para obtener un desenlace favorable; como ejemplo de la selección, con el fin de demostrar que el tratamiento de prueba es mejor, se incluyen como pacientes resistentes o intolerantes a otros tratamientos a algunos individuos que realmente no cumplen estrictamente los criterios para calificarlos como tal.

Posiblemente la situación más aberrante en este aspecto es la que se presenta en un artículo reciente de la revista *British Medical Journal*,<sup>3</sup> aclarando que es una más entre muchas otras periódicamente denunciadas: el doctor Julio Cruz

publicó entre el año 2001 y 2004 en dos diferentes revistas que se pueden catalogar como serias, tres ensayos clínicos del uso de dosis altas de manitol en pacientes con hematomas subdurales, hemorragia intraparenquimatosa temporal y en coma. En las tres publicaciones aparece acompañado por los doctores Minoja y Okuchi y en una de ellas adicionalmente por el doctor Facco. Estos estudios los reunió el grupo del doctor Roberts en un metanálisis publicado por la colaboración Cochrane en el año 2005. En mayo de 2006 el doctor Jorge Mejía, investigador colombiano, apoyado en versiones de algunos colegas Brasileños, cuestionó la veracidad de las investigaciones del doctor Cruz. El doctor Cruz se había suicidado en el año 2005 y cuando se contactó a los coautores de los artículos, todos negaron que hubieran participado en alguno de los estudios y reconocieron que únicamente habían tenido conversaciones casuales del tema con el autor principal, a pesar de lo cual habían aceptado aparecer como coinvestigadores. La institución a la que el Dr. Cruz decía pertenecer en sus artículos negó cualquier nexo con el investigador. Un comité de ética evaluó la situación y no encontró el origen de los supuestos sujetos de la investigación. La colaboración Cochrane decidió entonces retirar la revisión de su biblioteca en el año 2006.

Estos ejemplos de situaciones antiéticas deben llamar la atención de todo lector al momento de evaluar la literatura médica e invitan a una reflexión con respecto al trasfondo de la investigación y a los intereses que ella mueve. Aunque es indispensable evaluar la validez interna, el impacto de los resultados y la validez externa de estas investigaciones, igualmente es necesario contar con conocimiento del tema con el fin de evaluar aspectos que no se incluyen en las guías, aspectos que un lector desprevenido y de buena fe posiblemente piense que nunca suceden en la investigación y en la publicación.

## BIBLIOGRAFÍA

1. Johansen H, Gotzcheb P. Problems in the design and reporting of trials of antifungal agents encountered during meta-analysis. JAMA 1999; 282: 1752-1759
2. Safer D. Design and reporting modifications in Industry-sponsored comparative psychopharmacology trials. J Nervous and Mental diseases 2002 190: 583-592
3. Roberts I, Smith R, Evans S. Doubts over head injury studies. BMJ 2007; 334: 392-394.



La Revista Iatreia agradece su apoyo a las siguientes instituciones:

**Corporación Académica Ciencias  
Básicas Biomédicas -CCBB-**

**SUSALUD**

Compañía Suramericana de Servicios de Salud S.A.

**COMEDAL**  
Cooperativa Médica de Antioquia



**Fundación  
Fernando Zambrano Ulloa**

 **Coomeva**  
EPS

**EUROETIKA**

  
**SURATEP**  
Administradora de Riesgos Profesionales  
SURAMERICANA