

HUMANITAS

HUMANIDADES MEDICAS

TEMA
DEL MES
ON-LINE

¿SE PUEDE CONFIAR EN LOS
RESULTADOS DE LA
INVESTIGACIÓN CLÍNICA?

Salvador Peiró



Director: Prof. Mario Foz

N.º 5, Julio de 2006
ISSN: 1886-1601

HUMANITAS

HUMANIDADES MEDICAS

TEMA
DEL MES
ON-LINE

N.º 5, Julio de 2006

Director

Prof. Mario Foz Sala

Catedrático de Medicina. Profesor Emérito de la Universidad Autónoma de Barcelona

Consejo Asesor

Dr. Francesc Abel i Fabre

Director del Instituto Borja de Bioética (Barcelona)

Prof. Carlos Ballús Pascual

Catedrático de Psiquiatría. Profesor Emérito de la Universidad de Barcelona

Prof. Ramón Bayés Sopena

Catedrático de Psicología. Profesor Emérito de la Universidad Autónoma de Barcelona

Prof. Josep Egozcue Cuixart (†)

Catedrático de Biología Celular. Universidad Autónoma de Barcelona

Prof. Sergio Erill Sáez

Catedrático de Farmacología. Director de la Fundación Dr. Antonio Esteve. Barcelona

Dr. Francisco Ferrer Rusalleda

Médico internista y digestólogo. Jefe del Servicio de Medicina Interna del Hospital de la Cruz Roja de Barcelona. Miembro de la Junta de Govern del Colegio Oficial de Médicos de Barcelona

Dr. Pere Gascón

Director del Servicio de Oncología Médica y Coordinador Científico del Instituto Clínico de Enfermedades Hemato-Oncológicas del Hospital Clínic de Barcelona

Dr. Albert Jovell

Médico. Director General de la Fundación Biblioteca Josep Laporte. Barcelona. Presidente del Foro Español de Pacientes

Prof. Abel Mariné

Catedrático de Nutrición y Bromatología. Facultad de Farmacia. Universidad de Barcelona

Prof. Jaume Puig-Junoy

Catedrático en el Departamento de Economía y Empresa de la Universidad Pompeu i Fabra. Miembro del Centre de Recerca en Ecomía i Salut de la Universitat Pompeu i Fabra de Barcelona

Prof. Ramón Pujol Farriols

Experto en Educación Médica. Servicio de Medicina Interna. Hospital Universitario de Bellvitge. L'Hospitalet de Llobregat (Barcelona)

Prof. Celestino Rey-Joly Barroso

Catedrático de Medicina. Universidad Autónoma de Barcelona. Hospital General Universitario Germans Triás i Pujol. Badalona

Prof. Oriol Romaní Alfonso

Departament d'Antropologia, Filosofia i Treball Social. Universitat Rovira i Virgili. Tarragona

Prof. Carmen Tomás-Valiente Lanuza

Profesora Titular de Derecho Penal. Facultad de Derecho de la Universidad de Valencia

Dra. Anna Veiga Lluch

Directora del Banco de Células Madre. Centro de Medicina Regenerativa de Barcelona

COMENTARIO EDITORIAL

Jaume Puig-Junoy

Departamento de Economía y Empresa y Centro de Investigación en Economía y Salud (CRES) de la Universitat Pompeu Fabra (UPF). Barcelona.

El crecimiento del gasto sanitario –y el cambio tecnológico impulsado por el conocimiento científico es el factor con más responsabilidad en su crecimiento- es un hecho común en la mayoría de países desarrollados que requiere una justificación para seguir políticas de asignación de recursos basadas en el mejor conocimiento disponible (evidencia). Si bien las innovaciones médicas, de forma mayoritaria, conducen a reducciones heterogéneas en la mortalidad y en la morbilidad, su introducción y difusión puede comprometer la financiación de los sistemas públicos de salud. En la mayoría de los casos, la sostenibilidad de las finanzas públicas indica que no será posible continuar pagando el precio de cualquier innovación médica o farmacéutica que pueda ser utilizada para mejorar la salud de la población.

Más allá de los relativamente estériles debates sobre el nivel adecuado de gasto del sistema sanitario, resulta cada vez más claro que un nivel de producción y de gestión adecuada del conocimiento científico (evidencia) es una condición necesaria para la sostenibilidad y eficiencia de los sistemas sanitarios. En este contexto es en el que creo que es útil situar la contribución del artículo de Salvador Peiró sobre la medida en la que se puede confiar en los resultados de la investigación clínica a la hora de tomar decisiones en el ámbito sanitario, tanto a nivel micro, como a nivel meso y macro.

Una de las virtudes de la denominada medicina basada en la “evidencia” (uso de forma consciente, explícita y juiciosa de la mejor

“evidencia” disponible en las decisiones clínicas individuales) es que ha contribuido a poner en el centro de la atención de la política sanitaria los incentivos relacionados con la producción (cantidad y calidad) del conocimiento científico, así como los incentivos relativos a la utilización de este conocimiento en la toma de decisiones en la actividad clínica.

Al intentar fundamentar las decisiones clínicas (a nivel micro) en el mejor conocimiento científico disponible, resulta cada vez más imprescindible para una financiación responsable de los sistemas de salud preguntarse cuestiones como las siguientes: 1) ¿son adecuados los incentivos para que se produzca investigación básica y aplicada en una cantidad/calidad socialmente deseable para fundamentar las decisiones clínicas?; 2) ¿los mecanismos de intervención del mercado y de regulación (desde las patentes a la financiación pública de la investigación sanitaria) garantizan un nivel socialmente eficiente de investigación clínica?; 3) ¿existen incentivos para que se investigue en innovaciones médicas y farmacéuticas que den lugar a tratamientos o diagnósticos con una relación coste-efectividad que resulte socialmente aceptable?; 4) ¿cómo se deberían adaptar las organizaciones y los incentivos de los proveedores de servicios de salud y de inputs -industria farmacéutica y electromédica- para contribuir a reducir la brecha entre eficacia y efectividad?

El artículo de Salvador Peiró pone de relieve las dificultades por conseguir una producción de conocimiento científico de calidad y exenta de

algunos sesgos posiblemente incrementados precisamente por el aumento del valor económico de este propio conocimiento.

A pesar de observar la ausencia de conocimiento sobre la eficacia (y no digamos ya sobre la relación coste-efectividad) de una proporción mayoritaria de los procedimientos diagnósticos y terapéuticos más utilizados, la atención se centra en el margen, en la introducción de innovaciones, especialmente farmacéuticas. Aquí es donde el artículo de este número identifica dificultades importantes que impiden depositar confianza ciega en los resultados de la investigación y que arrojan más preguntas que respuestas.

Las patentes representan una forma de convertir en un bien privado (cuyas rentas son apropiables por parte del innovador) la información resultante de la investigación que, por sí misma, representa un bien público que no daría lugar a rentas "apropiables": el hecho de que alguien utilice, por ejemplo, conocimiento sobre eficacia de intervenciones de estilos de vida, puede ser socialmente rentable, pero no hay incentivos suficientes para que los agentes del sector produzcan este conocimiento en una cantidad/calidad socialmente deseable.

En conjunto, los sesgos que pone de relieve el artículo y que recomiendan precaución en la utilización de los resultados de la investigación se pueden clasificar en diferentes tipos, cada uno de los cuales posiblemente requiere intervenciones públicas distintas, así como la autorregulación de los propios agentes implicados.

Esta última debiera tener un papel destacado en el futuro cercano, especialmente cuando los sesgos de la investigación empiezan a ser demasiado evidentes: la autorregulación debería imponerse de forma paulatina, aunque no florecerá de forma espontánea, entre investigadores, financiadores de la investigación, editores, evaluadores, etc. para mejorar la confianza en su propio "producto", aunque sólo sea para evitar su pérdida de valor (pérdida de confianza en los resultados), que influiría negativamente en las rentas futuras esperadas y en las valoraciones bursátiles de las compañías innovadoras.

Intervención distinta requiere el hecho generalizado, pero escasamente justificable en términos de eficiencia social, de que los aseguradores, especialmente públicos, prácticamente no financian investigación clínica independiente sobre innovaciones (el NICE inglés representa, entre otros, un ejemplo a seguir), abandonando esta financiación en manos de quien tiene incentivos y obligación, por su propia naturaleza y objetivos, de generar conocimiento que produzca rentas apropiables a corto o medio plazo. Los financiadores de los sistemas de salud debieran evolucionar desde la compra de servicios a la compra de resultados en salud (años de vida ajustados por calidad -AVAC-), lo cual les llevará irremediablemente a producir y gestionar mejor el primer recurso que utiliza el sistema sanitario: el conocimiento.



Dr. Salvador Peiró

CURRICULUM VITAE

- Licenciado en Medicina (Universidad de Valencia, 1979), especialista en Medicina Preventiva y Salud Pública, doctor en Medicina (Universidad de Valencia, 1996) y especialista universitario en Economía de la Salud.
- Actualmente es jefe de la Unidad de Investigación en Servicios de Salud de la Escola Valenciana d'Estudis de la Salut de la Conselleria de Sanitat de la Generalitat Valenciana, institución en la que trabaja desde 1992. También preside la Fundación Instituto de Investigación en Servicios de Salud.
- Es editor adjunto de la revista Gaceta Sanitaria (2004-) y de la serie Quaderns de Salut Pública (1995-), y miembro de los consejos científicos de las revistas Medicina Clínica, Revista Española de Salud Pública, Gestión Clínica y Sanitaria y Atlas de Variaciones en la Práctica Médica en el Sistema Nacional de Salud.
- Es investigador asociado del Centre de Recerca en Economia i Salut de la Universitat Pompeu Fabra y del Instituto Aragonés de Ciencias de la Salud, así como colaborador docente de diversas Universidades y Escuelas de Salud Pública.
- Ha sido co-coordinador de la Comisión de Salud Pública e Investigación en Servicios de Salud del Fondo de Investigación Sanitaria, Vicepresidente de la Asociación de Economía de la Salud, miembro del comité de editores de Revista de Calidad Asistencial, patrono de la Fundación Avedis Donabedian, y miembro de diversos grupos de expertos para diferentes instituciones.
- Dedicado fundamentalmente a la investigación en servicios de salud, ha trabajado en líneas de investigación que incluyen el estudio de las variaciones en la práctica médica (codirige el grupo de Variaciones en la Práctica Médica de la Red Temática IRYSS) y otros aspectos de la utilización de servicios sanitarios, utilización inadecuada, calidad asistencial e investigación en resultados.
- En las áreas de investigación comentada es autor de un centenar de publicaciones en revistas nacionales e internacionales con procesos de evaluación por pares. También es autor de otros textos (libros, capítulos, artículos, etc.) en estos campos y ha participado en numerosas ocasiones como conferenciante o ponente de reuniones científicas.

PUBLICACIONES

SOBRE PRESCRIPCIÓN FARMACÉUTICA

- Villanueva P, Peiro S, Librero J, Pereiro I. Accuracy of pharmaceutical advertisements in medical journals. *Lancet* 2003; 361: 27-32.
- Bernal-Delgado E, Galeote-Mayor M, Pradas-Arnal F, Peiró-Moreno S. Evidence based educational outreach visits: effects on prescriptions of non-steroidal anti-inflammatory drugs. *J Epidemiol Community Health* 2002; 56(9): 653-8.
- García Sempere A, Peiró S. Gasto farmacéutico en atención primaria: variables asociadas y asignación de presupuestos de farmacia por zonas de salud. *Gac Sanit* 2001; 15: 32-40.
- Peiró S, Meneu R. Los indicadores de prescripción. Problemática actual, limitaciones y perspectivas. En: Meneu R, Peiró S (eds). Elementos para la gestión de la prescripción y la prestación farmacéutica. Barcelona: Masson, 2004.

INVESTIGACIÓN EN CALIDAD ASISTENCIAL

- Librero J, Marín M, Peiró S, Verdaguer Munujos A. Exploring the Impact of Complications on Length of Stay in Major Surgery Diagnosis Related Groups. *Int J Qual Health Care* 2004; 16: 51-57.
- Verdaguer Munujos A, Peiro S, Librero J. Variabilidad en la utilización de recursos hospitalarios en pacientes con enfermedad pulmonar obstructiva crónica. *Arch Bronconeumol* 2003; 39(10): 442-448.
- Peiró Moreno S. De la gestión de lo complementario a la gestión integral de la atención de salud: gestión de enfermedades e indicadores de actividad. En: Ortun V (ed). Gestión clínica y sanitaria: de la práctica a la academia, ida y vuelta. Barcelona: Masson, 2003; 17-87.
- Librero J, Peiró S, Márquez Calderón S. Inter-hospital variations in caesarean sections. A risk adjusted comparison in the Valencia public hospitals. *J Epidemiol Community Health* 2000; 54(8): 631-636.

- Librero J, Peiró S, Ordiñana R. Chronic comorbidity and outcomes of hospital care: length of stay, mortality and readmission at 30 and 365 days. *J Clin Epidemiol* 1999; 52: 171-179.
- Oterino D, Peiró S, Ridao M, Marchan C. Variations in diagnostic and therapeutic intensity between home and conventional hospitalisation. *Int J Qual Health Care* 1998; 10 (4): 331-338.

EN UTILIZACIÓN INADECUADA DE SERVICIOS SANITARIOS

- Moya-Ruiz C, Peiró S, Meneu R. Effectiveness of feedback to physicians in reducing inappropriate use of hospitalization. *Int J Qual Health Care* 2002; 14: 305-312.
 - Sempere Selva T, Peiró S, Sendra Pina P, Martínez Espín C, López Aguilera I. Inappropriate use of an Accident and Emergency Department: Magnitude, Associated Factors and Reasons. *Ann Emerg Med* 2001; 37: 568-79.
 - Peiro S, Meneu R, Lorenzo S, Restuccia JD. Assessing the necessity of hospital stay by means of the appropriateness evaluation protocol: a different perspective. *Int J Qual Health Care* 2001; 13: 341-3.
 - Peiró S, Meneu R, Roselló-Pérez ML, Portella E, Carbonell-Sanchis R, Fernández C, et al. Validez del protocolo del uso inapropiado de la hospitalización. *Med Clin (Barc)* 1996; 107: 124-129.
- Martínez E, Peiró S, Meneu R, Roselló ML, Portella E. How many days for an appendectomy? *Int J Quality in Health Care* 1994; 6: 347-352.
 - Peiró S, Portella E. Identificación del uso inapropiado hospitalario: la búsqueda de la eficiencia. *Med Clin (Barc)* 1994; 103: 65-71.

SOBRE VARIACIONES EN LA PRÁCTICA MÉDICA

- Peiró S. Variaciones en la práctica médica y utilización inadecuada de tecnologías. En: Gonzalez López Valcercel B (ed). *Difusión de nuevas tecnologías sanitarias y políticas públicas*. Barcelona: Masson, 2005; 101-34.
- Librero J, Peiró S, Bernal-Delgado E et al. Variaciones en intervenciones de cirugía general en el Sistema Nacional de Salud. *Atlas Var Pract Med SNS* 2005; 1(2): 63-81.
- Librero J, Rivas F, Peiró S, Allepuz A, Montes Y, Bernal-Delgado E, Sotoca R, Martínez N por el Grupo VPM-IRYSS. Metodología del Atlas de variaciones en cirugía ortopédica y traumatología en el Sistema Nacional de Salud. *Atlas Var Pract Med Sist Nac Salud* 2005; 1: 43-48.
- Peiró S, Meneu R, Marqués Espí JA, Librero J, Ordiñana Calabuig R. La variabilidad en la práctica médica: relevancia, estrategias de abordaje y política sanitaria. *Papeles de Economía Española* 1998; 76: 165-175.
 - Marión J, Peiró S, Márquez S, Meneu R. Variaciones en la práctica médica. Importancia, causas e implicaciones. *Med Clin (Barc)* 1998; 110 (10): 382-390.
- Marqués JA, Peiró S, Medrano J, Librero J, Pérez-Vázquez MT, Aranaz J et al. Variabilidad en las tasas de intervenciones de cirugía general por áreas de salud. *Cir Esp* 1998; 63: 445-453.

INVESTIGACIÓN CLÍNICA

- Almela P, Benages A, Peiró S, Mínguez Pérez M, Peña A, Pascual I, Mora F. Development of a risk score system to identify patients with Upper GI bleeding suitable for safe outpatient management. *Gastrointest Encoscop* 2004; 59: 772-781.
 - Almela P, Benages A, Peiró S, Mínguez M, Peña A, Pascual I et al. Outpatient Management of Upper Digestive Hemorrhage not Associated with Portal Hypertension. A large prospective cohort. *Am J Gastroenterol* 2001; 96: 2341-8.
- Peiró S. Medicina basada en la evidencia: necesaria, aunque insuficiente. Limitaciones -y autolimitaciones- de la Medicina basada en la evidencia para la práctica clínica y la atención sanitaria. *Med Clin (Barc)* 2002; 118 (supl.3): 49-56.
 - Librero J, Peiró S. Comparaciones abstrusas en los resúmenes de ensayos clínicos en revistas médicas españolas. *Med Clin (Barc)* 2001; 117: 732-733.
 - Navarro Collado MJ, Peiró S, Trénor Gomis C, Ruíz Jareño L, Pérez Igualada A, Guerola Soler N. Factores asociados a resultado funcional y calidad de vida en la rehabilitación tras artroplastia de rodilla. *Med Clin (Barc)* 2000; 114 (7): 250-4.

¿SE PUEDE CONFIAR EN LOS RESULTADOS DE LA INVESTIGACIÓN CLÍNICA?

RESUMEN

En la medicina actual se están produciendo dos fenómenos de forma simultánea. De un lado, ha ido ganando fuerza la idea de que los tratamientos y otros aspectos del manejo de los pacientes deben basarse en resultados de investigación y, en especial, en los ensayos clínicos. De otro, va creciendo la sospecha de que muchos de estos resultados de investigación no son más que falsos hallazgos. La manifestación más palpable de este fenómeno sería la elevada tasa de no reproducibilidad por investigaciones subsiguientes de buena parte de los estudios publicados.

Que un hallazgo de investigación sea “verdadero” (y no un falso positivo) depende de tres factores relacionados con probabilidades (la probabilidad previa a la realización del estudio de que tal resultado fuera verdadero, el poder estadístico del estudio, y el nivel de significación estadística elegido) y de un factor sistemático (la presencia de sesgos). Un estudio sesgado sería aquel en el que, con independencia de la intencionalidad, uno o más factores del diseño, del muestreo, de la obtención de los datos, del análisis, de la presentación de los resultados o de la propia publicación del estudio, tienden a producir conclusiones erróneas sobre la cuestión investigada. En este texto se revisan los sesgos vinculados a la financiación de la investigación clínica por instituciones o empresas cuyos intereses pueden verse afectados por los resultados de estos estudios.

El sesgo de financiación en la investigación clínica ha sido demostrado en numerosos trabajos que, básicamente, reportan que los ensayos patrocinados por la industria farmacéutica tienen, en conjunto, una probabilidad 3-5 veces superior a los no patrocinados de presentar resultados favorables a quien financió el estudio. Este fenómeno se produce en las presentaciones a congresos médicos, en las evaluaciones económicas, es independiente de la calidad del ensayo, aparece en todas las revistas, incluso las de gran prestigio, y se da en los tratamientos médicos, quirúrgicos y con otros dispositivos, y en cualquier área o especialidad clínica.

Los mecanismos que contribuyen a la presencia de este sesgo parecen ser: 1) la desatención al principio de equipoise, uno de los principios centrales de la ética

de la experimentación en humanos; 2) la presencia de cláusulas de control en los contratos de investigación impuestas por la industria y aceptadas por las instituciones sanitarias o académicas y las empresas privadas de investigación; 3) el uso de alternativas de comparación inadecuadas, o en dosis incorrectas; 4) la selección de pacientes y los criterios de inclusión/exclusión empleados; 5) las manipulaciones de los datos ex-post, incluyendo cambios en las medidas de resultados, y los análisis excesivamente creativos; 6) la omisión de los resultados negativos y la exageración de los positivos, o la confusión en el reporte de los resultados; y 7) el sesgo de publicación, que favorece la difusión de los ensayos favorables y limita o impide la de los resultados no deseados.

Adicionalmente, los estudios patrocinados están desproporcionadamente enfocados hacia las tecnologías e intervenciones que tienen valor comercial frente a las de menor valor comercial, hacia el “cómo” intervenir frente al “cuándo” o el “si se debería intervenir o no”, y a los beneficios de las intervenciones antes que hacia sus riesgos, de modo que sólo se dispone de “evidencia” cuando alguien ha querido –y financiado– que algo sea “evidente”. En la última década se han desarrollado numerosas iniciativas para limitar el impacto del sesgo de financiación: registro obligatorio de los ensayos, cambios en los “requisitos de uniformidad” para evitar los trabajos “cautivos”, guías de relaciones con la industria, normas gubernamentales, controles de buena práctica en los ensayos, la publicación de todos los ensayos y alguna otra.

En todo caso, a la hora de decidir si se usarán los resultados de una investigación para tratar a pacientes debería tenerse en cuenta que la confianza en los resultados de un ensayo ha de ser mayor: 1) cuanto más grande sea el tamaño muestral, 2) más grande sea el tamaño del efecto, 3) menor sea el número de asociaciones valoradas, 4) mayor sea la rigidez del estudio (en su diseño, definiciones operativas, medidas de resultado y análisis), 4) menores sean los intereses financieros implicados, 5) más parecidos sean el comparador utilizado y los pacientes del estudio a los tratamientos y pacientes reales.

CAN WE TRUST IN CLINICAL RESEARCH FINDINGS?



SUMMARY

Nowadays, two simultaneous phenomena are taking place in the field of Medicine. On one side, the idea of basing the treatments and other issues involving the patient management on clinical research findings, especially clinical trials, has become stronger. On the other side, a suspicion is growing that a lot of these research findings are nothing but false. The most noticeable manifestation of this phenomenon would be the high rate of non-reproducibility through subsequent researches of most of the published studies.

A research finding that is “true” (and not a false positive) is dependent on three factors involving probabilities (the probability previous to the carrying out of the study that such result was true, the statistical power of the study, and the chosen level of statistical significance) and on a systematical factor (presence of biases). A biased study would be that in which, independently from the intentionality, one or more factors regarding the design, the sampling and data collection, the analysis, the presentation of results or the study’s publication itself, tend to bring about erroneous conclusions regarding the research subject. In this article, the biases related to the financing of the clinical research through institutions or enterprises whose interests can be affected by the results of these studies are reviewed.

The financial bias in clinical research has been shown in many studies that, basically, report that the trials sponsored by the pharmaceutical industry have, as a whole, a 3-5 times higher probability of presenting favourable results to their sponsors, than those without sponsorship. This phenomenon is produced in cases such as presentations within medical congresses, economical assessments, it is independent from the quality of the trial, it appears in every journal, even the most prestigious ones, it has place within the medical, surgical and other devices treatments, and every clinical field or specialty.

The mechanisms that contribute to the presence of this bias seem to be: 1) the disregard to the principle of

equipoise, one of the core principles of ethics regarding human experimentation; 2) the presence of control clauses within the research contracts, inflicted by the industry and accepted by healthcare or scholar institutions and private research enterprises; 3) the use of non-adequate comparative alternatives, or incorrectly dosed ones; 4) the patient selection and the inclusion/exclusion criteria; 5) the manipulations of the ex-post data, including changes in the results measures, and the excessively creative analysis; 6) the omission of the negative results and the exaggeration of those positive, or the confusion regarding the report of the results; and 7) the publication bias, that fosters the diffusion of favourable trials and limits or hinders that of non-desired results.

Besides, the sponsored studies are disproportionately focused on the technologies and treatments that have a commercial value in spite of those having less commercial value, on “when” instead of “how” to intervene or even “if one should intervene or not”, and on the outcomes of interventions more than on their risks, in such a way that there is only “evidence” when somebody wants -and finances- something to be “evident”. In the last decade, several initiatives aimed at limiting the impact of the financial bias have been developed: mandatory trial recording, changes in the “uniformity requirements” to avoid the “caught” works, guides to handle the relations with the industry, governmental rules, good practice controls in the trials, the publication of every trial and some other one.

In any case, at the time of deciding whether the results of a research will be used or not to treat patients, it should be taken into account that the trust in the results of a trial must be greater: 1) as long as the sample size is bigger, 2) the effect range is greater, 3) the number of valued associations lessens, 4) the stiffness of the study (regarding its design, operative descriptions, result measures and analysis) is increased, 5) the involved financial interests decrease, 5) the comparator used and the study patients look more alike the treatments and the real patients.



¿SE PUEDE CONFIAR EN LOS RESULTADOS DE LA INVESTIGACIÓN CLÍNICA?

SALVADOR PEIRÓ

Escola Valenciana d'Estudis de la Salut. Conselleria de Sanitat de la Generalitat Valenciana. Valencia.

LOS “FALSOS POSITIVOS” EN INVESTIGACIÓN BIOMÉDICA

De la mano del movimiento “Medicina Basada en la Evidencia” (MBE) la investigación clínica ha ido ganando predicamento en la medicina moderna, y el ensayo clínico ha sido entronizado en el lugar más alto de una supuesta escala de jerarquía de lo “probado” y, a fin de cuentas, lo “verdadero”.¹ A la par que la medicina intenta adquirir esta imagen más científica, va creciendo la sospecha de que muchas, si no la mayoría de las conclusiones de la investigación actual, podrían no ser más que falsos hallazgos.² No sólo en lo que afecta a la investigación clínica, sino también en la investigación epidemiológica e incluso, y quizás sobre todo, en la investigación más rabiosamente actual, como la genética y molecular.³⁻⁷

La manifestación más palpable de este fenómeno sería la elevada tasa de no reproducibilidad –o, si se prefiere, de falta de confirmación– de muchos hallazgos de investigación por las investigaciones subsiguientes. Ejemplos típicos serían la no confirmación de que los folatos reduzcan la morbilidad cardiovascular, o que la terapia hormonal sustitutiva reduzca este mismo problema en mujeres, o que la vitamina C disminuya la incidencia de un sinnúmero de cánceres o que el magnesio sea eficaz en el infarto de miocardio. Recientemente, Ioannidis siguió 45 artículos citados más de 1.000 veces (artículos publicados entre 1990 y 2003 en *The Journal of the American Medical Association* [JAMA], *The*

New England Journal of Medicine [NEJM] y *The Lancet*) y que habían concluido que una intervención era efectiva.⁸ El 16% de estos trabajos ya habían sido refutados por estudios de mayor tamaño o mejor diseño; en otro 16% el efecto era mucho menor que el relatado en el estudio inicial; un 24% permanecían sin estudios de confirmación posteriores, y sólo el 44% habían sido replicados.⁸ Estos resultados en una selección de “lo mejor” de “las mejores revistas” sugiere que el panorama promedio (artículos promedio, de las revistas promedio) debe de ser muy preocupante.

Buena parte de esta alta tasa de “no reproducibilidad” se debe a la inadecuada pero muy extendida estrategia de basar las conclusiones de investigación en la significación estadística formal (típicamente, en un valor de p menor que 0,05).^{2, 9-11} Una conclusión de investigación quedaría definida, en esta estrategia, por haber identificado (resultado positivo) o no haber identificado (resultado negativo) una relación estadísticamente significativa entre dos variables (intervención y *outcome*). Sin embargo, que la relación detectada sea “verdadera” (y no un falso positivo) depende de tres factores relacionados con probabilidades (la probabilidad previa a la realización del estudio de que tal resultado fuera verdadero, el poder estadístico del estudio, y el nivel de significación estadística elegido)¹¹ y de un factor sistemático (la presencia de sesgos).²

Respecto a los componentes probabilísticos, en los trabajos de Wacholder¹¹ y Ioannidis² se



analiza la probabilidad de obtener un resultado de investigación “falso positivo”, e incluso se aportan modelos, muy similares a los usados para la valoración de pruebas diagnósticas, para estimar el valor predictivo positivo de un resultado de investigación en función de su probabilidad previa de ser verdadero y su poder estadístico. Respecto al componente sistemático, los sesgos, vendría conformado por la combinación de uno o más factores del diseño, del muestreo, de la obtención de los datos, del análisis, de la presentación de los resultados o de la propia publicación del estudio, que tenderían a producir conclusiones erróneas sobre la cuestión investigada. Los sesgos pueden deberse tanto a baja calidad de la investigación como a manipulación intencionada de la misma.

En este trabajo se revisará este componente sistemático y, en concreto, los sesgos vinculados a la financiación de la investigación por instituciones o empresas cuyos intereses se ven afectados por los resultados de la propia investigación que financian.¹² Los trabajos sobre estos sesgos se refieren, en general, a estudios financiados por la industria farmacéutica y, en menor medida, la industria electromédica, pero se producen en otros muchos entornos y los intereses implicados no son sólo económicos. Ejemplos de otros intereses y en otros campos incluyen parte de las investigaciones que demuestran efectos nocivos del alcohol (financiadas por advertiser), trabajos que demuestran beneficios clínicos de la fe (en buena parte financiados por la llamada iglesia de la ciencia), muchos trabajos en el campo de la alimentación financiados por la industria alimentaria o los financiados por la industria del tabaco sobre tabaquismo pasivo. Los sesgos asociados a la financiación de los estudios por entidades que no son neutrales respecto a los resultados del estudio no son el único sesgo y, al menos en los ensayos clínicos, no son el principal elemento a tener en cuenta para otorgar más o menos confianza a los resultados de una investigación. Pero tampoco puede decirse que su importancia sea menor, o que no vaya creciendo –o recibiendo una creciente atención– en los últimos años.

EL SESGO DE FINANCIACIÓN EN LA INVESTIGACIÓN CLÍNICA

Hace unos años, Vandenbroucke et al. comentaban en una escueta carta al director del *British Medical Journal* (BMJ) los resultados de los ensayos clínicos que habían evaluado la seguridad, respecto al riesgo de trombosis venosa, de los anticonceptivos orales de tercera generación.¹³ De los 9 estudios existentes sin patrocinio de la industria farmacéutica, 8 mostraban un incremento del riesgo relativo (RR) de trombosis entre 1,5 y 4,0 veces, y en 1 las diferencias no eran significativas. El RR conjunto de trombosis en los 9 ensayos era 2,4 veces mayor para las nuevas formulaciones y estadísticamente significativo. De los 5 estudios bajo patrocinio de los fabricantes, el primero, finalizado 2 años antes y del que se sabía que también mostraba una mayor incidencia de trombosis, no había sido publicado; otro que en origen reportaba un RR de 1,5 había sido publicado en varias ocasiones, pero con diversos reanálisis que reducían el exceso de riesgo en cada publicación, y los 3 restantes no mostraban diferencias significativas. El RR conjunto de los 4 estudios publicados era de 1,1 y no significativo.

Posteriormente, Lexchin et al. publicaron una revisión de 30 artículos (18 comparaciones) que habían evaluado la relación entre financiación y resultados.¹⁴ Los estudios patrocinados mostraban una *odds ratio* (OR) de que los resultados fueran favorables al patrocinador 4,05 veces mayor (IC95%: 2,98-5,51) que los no patrocinados. La calidad metodológica era similar en ambos grupos, aunque la tasa de publicación de los patrocinados era menor. Otro trabajo similar, publicado el mismo año y que revisaba 8 estudios (incluyendo 1.140 originales) que habían evaluado la relación entre resultados y financiación, mostraba una OR de conclusión favorable al patrocinador de 3,60 (IC95%: 2,63-4,91),¹⁵ en concordancia con el de Lexchin.

Hasta donde existen estudios, este fenómeno se produce en las presentaciones a congresos médicos (una revisión de todos los resúmenes de estudios de evaluación de medi-



camentos de un congreso en Estados Unidos puso de manifiesto que todos los estudios con financiación de la industria tenían resultados positivos, por sólo un 67% de los no financiados);¹⁶ en las evaluaciones económicas (un trabajo que revisaba las evaluaciones económicas de fármacos oncológicos halló que sólo el 5% de los estudios patrocinados por la industria favorecían al fármaco tradicional, porcentaje que era del 38% en los estudios sin este patrocinio);¹⁷ en las revistas de prestigio como el BMJ,¹⁸ el NEJM y el JAMA (en estas 2 últimas revistas, y en casi 500 originales publicados en 2001, los autores con relaciones con la industria mostraban una probabilidad entre 10 y 20 veces menor de presentar resultados favorables al tratamiento tradicional);¹⁹ en los ensayos de fármacos, quirúrgicos y de otros tratamientos,²⁰ y en los de cirugía ortopédica²¹ o neurocirugía.²²

La conclusión global de estos trabajos, y otros previos²³ es que existe una relación positiva y estadísticamente significativa entre la fuente de financiación y el resultado de los estudios clínicos, de manera que son excepcionales los ensayos financiados por la industria farmacéutica que favorecen la terapia tradicional frente a las nuevas terapias, o a otros productos respecto al producto de la firma patrocinadora del estudio. Adicionalmente, algunos de estos mismos trabajos ponen de manifiesto que esta asociación no se debe a una peor calidad de los estudios patrocinados que (al menos en aspectos como el tamaño de la muestra, la aleatorización, el enmascaramiento, y otros atributos técnicos) suelen ser iguales o mejores que los estudios patrocinados por los gobiernos o entidades sin ánimo de lucro.^{14, 18, 24, 25}

MECANISMOS QUE EXPLICAN EL SESGO DE FINANCIACIÓN

Desatención al principio de equiponderación

La equiponderación (*equipoise*) es uno de los principios centrales de la ética de la experimentación en humanos. Introducido para proteger a

los pacientes de investigaciones redundantes, implica que sólo en presencia de alternativas aparentemente equivalentes, cuando hay incertidumbre sobre el valor de una sobre otra, o cuando la comunidad médica discrepa en la interpretación de la evidencia disponible, se da la situación en que éticamente es justificable un nuevo estudio para decidir la mejor opción médica (los pacientes no serán asignados a un brazo que de antemano se sabe peor que la alternativa).^{24, 26, 27} En el contexto de los ensayos clínicos, y pese a la importante discusión sobre la aplicación práctica de este concepto,²⁸ la *equipoise* implica que los participantes no sufrirán un incremento relativo de daños derivado de la aleatorización a un brazo particular, que los resultados del estudio no pueden ser predichos consistentemente de antemano y que, sobre un conjunto suficiente, el número de ensayos que hayan aceptado o rechazado una hipótesis debe ser similar.²⁷

Djulbegovic valoró el cumplimiento del principio de *equipoise* en una revisión de 136 ensayos sobre tratamientos en mieloma múltiple. Los resultados revelaban que este principio se cumplía en los trabajos financiados con fondos públicos y organizaciones sin ánimo de lucro (47% frente a 53% favorecían al tratamiento innovador frente al previo, $p=0,61$), pero no en los financiados, totalmente o en parte, por los fabricantes (74% frente a 26% favorecían al tratamiento innovador, $p<0,005$).²⁴ Más recientemente, Fries et al. revisaron los resúmenes presentados al Congreso de 2001 del *American College of Rheumatology* y encontraron que el 100% de los trabajos financiados favorecían al patrocinador.²⁷ Los autores señalaban que el sesgo de publicación no era suficiente para conseguir lo que denominaban una “violación sistemática” del principio de *equipoise*, y señalaban los “sesgos de diseño” como principal fuente de este problema.

La calidad técnica de los ensayos en el trabajo de Djulbegovic, medida por la escala de Jadad, no presentaba diferencias significativas entre estudios patrocinados o no, pero los ensayos patrocinados utilizaban en mucho mayor



medida el placebo que los controles activos (al contrario que los no patrocinados).²⁴ El trabajo sugiere que en los ensayos patrocinados se comparan los nuevos tratamientos frente a placebo y no frente a los tratamientos (previos) más usuales, aspecto que implicaría una ruptura importante del principio de *equipoise* (los pacientes del grupo placebo habrían sido derivados a un brazo de tratamiento que sólo podría ser igual o peor que el brazo alternativo). Este problema afectaría, más que a la industria farmacéutica, a la exigencia de las Agencias de Medicamentos de que se aporten ensayos frente a placebo para la autorización de nuevos medicamentos, exigencia que, aunque pueda tener alguna ventaja para valorar posibles efectos adversos del nuevo fármaco, es muy discutible cuando existen tratamientos previos consolidados. En todo caso, se trata de un problema mayor desde el punto de vista de la ética de investigación y del tratamiento individual de los participantes en los ensayos.

Negocios, contratos e investigación clínica

Los estudios clínicos, más allá (o más acá) de su objetivo final de incrementar el conocimiento clínico, son, en sí mismos, el producto de una industria que se dedica a su realización. En esta industria participan los financiadores, pero también los centros de investigación universitarios, grandes hospitales, empresas de audit, etc. Y mueve un considerable volumen de fondos. Por ejemplo, las firmas farmacéuticas estadounidenses destinaron 10.000 millones de dólares a investigación clínica en 2004.²⁹ Como en otros campos, las relaciones entre agentes están regidas por los contratos entre las partes y, en este terreno, diversos trabajos han señalado la presencia de cláusulas contractuales éticamente inaceptables, entre ellas la reserva de derechos de publicación a favor del patrocinador, la propiedad de los datos (incluso negando el acceso a los investigadores), las reservas de revisión o de reanálisis, y otras.³⁰⁻³⁴ Se ha señalado, refiriéndose a los hospitales estadounidenses de

alto nivel de investigación y con políticas explícitas respecto a la reserva de derechos de publicación, que entre el 30 y el 50% de los contratos de investigación deben ser renegociados por la presencia de cláusulas de control inaceptables a favor del patrocinador.³⁵

En EE.UU. existe un importante movimiento académico para garantizar la libertad de los investigadores y la veracidad de los resultados de las investigaciones. Los centros universitarios han desarrollado guías de relaciones adecuadas con la industria a las que, al menos formalmente, se han acogido la mayor parte de las Universidades,^{36, 37} los editores de revistas médicas, en muchos casos antes de los nuevos Requisitos de Uniformidad,³⁸ han acordado no aceptar los trabajos cuyo contrato no garantice el derecho de los investigadores sobre los datos y la publicación, y se ha incrementado notablemente el control sobre los aspectos éticos y de buena práctica de la investigación clínica. Pese a ello, estudios recientes muestran que los centros de investigación universitarios mantienen una adhesión muy precaria a las guías a las que dicen acogerse³⁹ y que los responsables de los contratos exhiben una gran variabilidad -y mayor tolerancia- a la hora de aceptar cláusulas restrictivas en estos contratos.⁴⁰

Que las universidades y grandes hospitales sean laxos a la hora de aceptar comportamientos que favorecen la investigación dirigida y el control de los datos y la publicación por el patrocinador empieza a ser un problema menor. Los centros académicos cada vez tienen menor importancia en la producción de investigación clínica. La financiación de la industria para investigación clínica pasó, en EE.UU., de poco más de 3.000 millones de dólares en 1994 a 10.000 millones en 2004; pero en 1994 el 63% de esta financiación era canalizada por centros académicos, mientras que en 2004 este porcentaje alcanzó sólo el 26%.²⁹ La investigación no académica en Estados Unidos esta dominada por las *Contract Research Organizations* (organizaciones de investigación por contrato, CRO), auténticas empresas de investigación que obtuvieron 1.600 millones de dólares de la industria



en 1994 y 7.600 en 2004.²⁹ Se ha comentado que las CRO realizan los ensayos de forma más rápida y más eficiente que los centros académicos, pero que también son más sensibles a los intereses de los financiadores,³⁰ aunque no hemos hallado ningún estudio que compare específicamente la investigación realizada por las CRO y los centros académicos. De hecho, la información publicada sobre las CRO es mínima.

Elección, dosis y posología del comparador

La elección de la alternativa a comparar, aspecto en clara relación con el principio de *equipoise*, juega un papel determinante en los resultados de los ensayos. En este terreno se ha señalado que en diversos ensayos patrocinados que comparaban la efectividad de fluconazol en pacientes neutropénicos se utilizaron como tratamientos alternativos la anfotericina-B oral o la nistatina (ambos de ineficacia demostrada en estos pacientes por su nula absorción digestiva),⁴¹ o que en otros ensayos, patrocinados por diferentes firmas, se comparaban los antipsicóticos atípicos a dosis usuales con haloperidol a dosis fijas de 20 mg/día (doble de la máxima usual), aspecto que determinaba que los efectos secundarios (extrapiramidalismo) fueran muy superiores en el grupo de haloperidol,⁴² o que en uno de los ensayos que comparaba paroxetina frente a amitriptilina ésta última se administró 2 veces al día (en lugar de la pauta habitual nocturna) consiguiendo que la “somnolencia diurna” se convirtiera en un importante efecto adverso del comparador.⁴²

Pero probablemente el trabajo que muestra este sesgo en toda su crudeza es una revisión de los ensayos sobre antiinflamatorios no esteroideos (AINE) publicados entre 1987 y 1990. Los autores encontraron 61 artículos que representaban 69 ensayos de los que el 81% tenían financiación de la industria farmacéutica. En el 48% de los ensayos patrocinados la dosis equivalente del AINE de la compañía patrocinadora era mayor que la del producto de la competencia

comparado, y en todos los ensayos aleatorizados que comparaban dos fármacos y concluyeron que un fármaco era superior a otro (16 ensayos, 28%), el fármaco más eficaz fue siempre el de la firma patrocinadora.⁴³ Al extremo, y al menos en algunos grupos farmacológicos como los AINE, que no sería necesario ni hacer ensayos, pues bastaría saber quién los financia para conocer el resultado.

La selección de pacientes y los criterios de inclusión y exclusión

Poco más del 2% de los pacientes incluidos en ensayos de AINE eran mayores de 65 años y representan alrededor del 10% de los pacientes en la consulta hospitalaria;⁴⁴ el 92% de los pacientes con depresión de una consulta externa de psiquiatría serían excluidos de los ensayos clínicos de antidepresivos.⁴⁵ Y las cosas no están mejor en otros campos¹: los ensayos tienden a excluir o infrarrepresentar a ancianos, mujeres y pacientes con comorbilidad,⁴⁶⁻⁴⁹ mientras que hiperrepresentan a los pacientes derivados a especialistas hospitalarios que no son comparables a los atendidos en el ámbito de la atención primaria.⁵⁰

Pero el problema no es sólo de representación. Parece que existe una tendencia a reclutar pacientes de bajo riesgo para los ensayos que buscan medir la seguridad y a reclutar poblaciones de alto riesgo del suceso que se quiere evitar (pero no de otros) para los ensayos cuyo objetivo fundamental es medir beneficios. La consecuencia de esta metódica es que los pacientes reales tienen más efectos adversos que los de los ensayos de seguridad (eran poblaciones sanas, con bajo riesgo de efectos adversos) y que los tratamientos en pacientes reales son menos efectivos que en los ensayos (porque la probabilidad del suceso era mayor en los ensayos). Aunque existen motivos técnicos para alguna de estas estrategias (los ensayos de eficacia en poblaciones de alto riesgo son más eficientes, ya que suelen requerir menor tamaño muestral y menor tiempo de seguimiento),



en la práctica un ensayo que ha evitado expresamente la población diana del fármaco (el ejemplo de los AINE y la exclusión de los mayores de 65 años es obvio) tiene escaso valor clínico. Escaso valor clínico no es contradictorio con elevado valor comercial: un reciente trabajo que valoraba la publicidad farmacéutica que citaba estudios clínicos halló un 46% de afirmaciones publicitarias no soportadas por el estudio que referenciaban, en su mayor parte porque extendían los beneficios del ensayo a grupos de población expresamente excluidos en el estudio.⁵¹

Las manipulaciones *post-hoc* y los análisis creativos

Buena parte de los ensayos clínicos no tienen sus protocolos registrados (por ejemplo, un estudio mostraba que el 25% de los ensayos en cáncer de próstata y el 40% en cáncer de colon no se hallaban incluidos en ningún registro).⁵² La imposibilidad de acceder a muchos protocolos permite la realización de cambios *post-hoc* en el diseño, tanto en la medida principal de resultados (eligiendo una que haya resultado significativa) como re-análisis, análisis de subgrupos no previstos, ocultación de medidas de resultado, análisis múltiples y otros cambios respecto al diseño y la hipótesis inicial que pueden falsear los resultados de un estudio clínico.

Un reciente y conocido ejemplo de este tipo de manipulación fue la publicación del estudio CLASS en *JAMA*, que presentaba un ensayo con 2 brazos (celecoxib frente a ibuprofeno o diclofenaco) y 6 meses de seguimiento, con resultados de seguridad digestiva claramente favorables a celecoxib. Poco tiempo después, la Food and Drug Administration (FDA, organismo que exige los protocolos de los estudios que se usan para registrar o modificar las indicaciones autorizadas de un fármaco, aunque esta información no siempre es pública), divulgaba en Internet los resultados de este mismo ensayo conforme al protocolo original: no había 2 brazos sino 3 (en el artículo se combinaron los bra-

zos ibuprofeno y diclofenaco), el ensayo no era a 6 meses sino a 12-15 (se había seleccionado el momento del seguimiento en que las diferencias eran máximas), y no existían diferencias significativas en seguridad digestiva entre celecoxib e ibuprofeno, al margen de cambios en la definición de la medida principal de resultados que afectaban sensiblemente a las conclusiones finales.⁵³⁻⁵⁵

Otro ejemplo de manejo creativo de los datos puede leerse en la “expresión de preocupación” recientemente publicada por los editores del *NEJM*^{56, 57} tras descubrir -a partir de los documentos aparecidos durante los procesos judiciales en relación con rofecoxib- que en la publicación del ensayo VIGOR⁵⁸ algunos pacientes del brazo rofecoxib que habían sufrido un infarto de miocardio habían “desaparecido”, y que este hecho cambiaba sustancialmente la interpretación de la seguridad cardiovascular del fármaco.

OMISIÓN, EXAGERACIÓN Y CONFUSIÓN EN EL REPORTE DE LOS RESULTADOS

Respecto al reporte de los resultados, parece que en los ensayos patrocinados -y especialmente en los resúmenes- se intenta presentar sólo los resultados favorables y en la forma más favorable posible. Las técnicas incluyen la omisión de resultados adversos o no significativos (fueron notorias en su día las ausencias en el reporte de efectos adversos en la publicación del VIGOR, rofecoxib frente a naproxeno, puestas en evidencia al divulgarse el informe de la FDA sobre este ensayo),⁵⁹ la exageración (empleando medidas de reducción relativa de riesgo, en lugar de la reducción absoluta o el número de pacientes a tratar),⁶⁰ el establecimiento de conclusiones no sustentadas por los resultados del ensayo¹⁷ y el empleo de un lenguaje deliberadamente confuso.

Respecto al establecimiento de conclusiones no sustentadas por el propio ensayo, Friedberg señaló que mientras que la probabilidad de que los resultados de las evaluaciones económicas favorecieran al brazo innovador frente al tradi-



cional era 5,2 veces mayor si el estudio era patrocinado, la probabilidad de que las conclusiones de los autores favorecieran a dicho brazo era 11,4 veces mayor (incluso en 6 de los 14 estudios patrocinados con resultados negativos o desfavorables, las conclusiones de los autores favorecían al tratamiento del patrocinador).¹⁷ En cuanto al lenguaje, puede verse un ejemplo arquetípico en el resumen del artículo que publicaba los resultados del brazo sin fractura vertebral previa del ensayo FIT, en el que sus autores relatan una reducción del riesgo de fractura favorable al fármaco evaluado ¡aunque no significativa!,⁶¹ presentando la ausencia de diferencias entre fármaco activo y placebo como una reducción de riesgo.

Quizás, el trabajo más esclarecedor en este campo (también en el de las manipulaciones *post-hoc*) sea el publicado recientemente por Chan et al., que revisaron los ensayos clínicos aprobados por los Comités Éticos de *Copenhagen* y *Frederiksberg* (Dinamarca) en 1994-95, siguieron las publicaciones (102 ensayos y 122 artículos que emplearon 3.736 medidas de resultado) y las compararon con el protocolo original.⁶² En conjunto, el 50% de los resultados de eficacia y el 65% de los de seguridad fueron incompletamente reportados, desapareciendo especialmente los no significativos. El 62% de los ensayos tenían al menos una medida primaria de resultados que había sido cambiada, introducida u omitida y, al encuestar a los autores, el 86% de los que respondieron negaron que se hubieran producido estos cambios u omisiones pese a la clara evidencia de lo contrario.

El sesgo de publicación

No todos los ensayos clínicos se publican, y la no publicación no se distribuye al azar. Diversos estudios han puesto de manifiesto que los trabajos con financiación de parte interesada sufren demoras,³⁵ se publican menos,^{63, 64} o se publican parcialmente.⁶⁵ Los autores tienden a no enviar los resultados negativos (sobre todo

en trabajos patrocinados) y a los editores no les gustan los trabajos con resultados negativos.⁶⁶⁻⁶⁹ Y las compañías farmacéuticas apenas respondieron a la “amnistía” de los editores para publicar los miles de ensayos nunca publicados (apenas dos centenares de ensayos se acogieron a la misma).

Un trabajo de Melander et al. ejemplifica este sesgo. Estos autores siguieron 42 ensayos clínicos de 5 diferentes inhibidores selectivos de la recaptación de serotonina (ISRS) que habían sido presentados a la Agencia Sueca del Medicamento entre 1983 y 1999.⁷⁰ En 21 ensayos el fármaco fue más efectivo que el placebo, y 19 fueron publicados en artículos específicos. En otros 21 ensayos no se observaron diferencias entre fármaco y placebo y sólo 6 fueron publicados como artículos específicos. Cuatro ensayos nunca fueron publicados (ni como original ni como parte de revisiones o metaanálisis) y en los 4 no había diferencias entre fármaco y placebo. Los *flow-chart* de este trabajo mostraban un descarado sesgo de publicación que favorecía notoriamente a los respectivos fármacos. Los autores concluían que las decisiones basadas en los datos publicados no serían soportadas por el conjunto de la información.

El sesgo de publicación también puede darse a la inversa, y hay pruebas de trabajos publicados varias veces que llevan a sobreestimar el efecto de un fármaco en los metaanálisis. Por ejemplo, Tramer et al. encontraron 84 ensayos que valoraban la eficacia antiemética del ondansetrón (incluyendo 20.181 pacientes, de los que 11.980 recibieron ondansetrón). El 17% de los ensayos y el 28% de los pacientes (3.335) estaban duplicados en diferentes publicaciones sin referencias cruzadas entre las mismas, y su inclusión mejoraba un 23% la eficacia antiemética del ondansetrón.⁷¹ En este apartado del sesgo de publicación cabe citar también la realización de ensayos en países con menor control sobre los aspectos éticos y de buena práctica de la investigación y la publicación, destacando China o Rusia, países que no han reportado todavía ni un solo ensayo negativo.⁷²



La elección del tema a estudio

Los sesgos, en sí, son sólo una parte del problema. Los estudios patrocinados están desproporcionadamente enfocados hacia las tecnologías e intervenciones que tienen valor comercial frente a las de menor valor comercial, hacia el “cómo” intervenir frente al “cuándo” o el “si se debería intervenir o no” y a los beneficios de las intervenciones antes que hacia sus riesgos.⁷³⁻⁷⁵

La selección de temas y objetivos por el patrocinador predetermina en buena parte el conocimiento disponible, ya que sólo dispondremos de “evidencia” cuando alguien haya querido –y financiado– que algo sea “evidente”. El resultado de este fenómeno es que tendremos “evidencia de grado A” (apoyada por ensayos clínicos aleatorizados, según las escalas de jerarquía de la evidencia al uso) en lo que respecta a fármacos, dispositivos quirúrgicos y otras tecnologías propietarias, frente a ninguna evidencia o evidencias de “baja” calidad en las tecnologías “no apropiables” (estilos de vida, dieta, ejercicio...) y fármacos no rentables (enfermedades huérfanas, típicas de países subdesarrollados...). Bajo este esquema de disponibilidad de información, la regla, en principio razonable, de favorecer el uso de unas u otras tecnologías en función de la “evidencia” disponible, favorece a los tratamientos con fármacos frente a las también razonables intervenciones de estilos de vida o, incluso, otros fármacos o tecnologías menos rentables.¹

El impacto de este fenómeno en la práctica clínica es frecuente e importante. Como ilustración vale la pena citar una reciente guía de hipertensión, que señala que “las medidas de estilos de vida no han demostrado prevenir las complicaciones cardiovasculares en pacientes hipertensos y nunca deberían retrasar innecesariamente el tratamiento farmacológico”,⁷⁶ frase técnicamente impecable pero que obvia el hecho de que nadie ha financiado los ensayos que podrían demostrar el valor de estas medidas. Una segunda ilustración la ofreció Dieppe, editor de una importante revista de reumatología, al señalar que el problema importante en el manejo farmacológico del dolor artrósico es si

los AINE aportan alguna ventaja sobre los simples analgésicos, mientras que las “comparaciones entre AINE están siendo una distracción dirigida por la industria de la pregunta de investigación adecuada”.⁷⁷

PROBLEMAS COMPLEJOS Y SOLUCIONES INSUFICIENTES

En un interesante artículo publicado en 1989, Gotzsche revisó la calidad de los ensayos de AINE. En sus palabras: “93 ensayos mostraron resultados estadísticamente significativos que en 73 ensayos favorecían al nuevo fármaco y en 8 al control activo. Todos los ensayos (n=39) que mostraron diferencias significativas en efectos secundarios favorecieron al nuevo fármaco. La elección de dosis, comparaciones múltiples, cálculos erróneos, el análisis intra-grupos y de subgrupos, las unidades de muestreo erróneas [...], los cambios en las escalas de medida previos al análisis, las diferencias basales y el reporte selectivo de los resultados significativos fueron algunas de las causas demostradas o posibles del gran porcentaje de resultados que favorecían a los nuevos fármacos. El 76% de las conclusiones o resúmenes contenían afirmaciones dudosas o inválidas. Los sesgos favorecieron consistentemente al nuevo fármaco en 81 ensayos y al control en solo 1. No es obvio cómo podría realizarse un metaanálisis fiable a partir de estos ensayos”.⁷⁸ En 2004, 15 años después del trabajo de Gotzsche, Whittington et al. publicaban una revisión sistemática sobre el efecto de diversos inhibidores selectivos de la recaptación de la serotonina en niños y adolescentes, comparando los resultados de los trabajos publicados y no publicados.⁷⁹ La incorporación de los resultados de los trabajos no publicados invertía las conclusiones de los publicados: los ISRS, en este grupo de edad y con alguna excepción, no mostraban mayor eficacia que el placebo y sí muchos más efectos adversos.⁸⁰

La conclusión obvia es que el “sesgo de financiación” persiste en la actualidad. Incluso



es probable que haya aumentado, dado el mayor volumen de ensayos patrocinados y la mayor importancia de los mismos en la promoción de los productos, creando un entorno caracterizado por: sólo se dispone de “evidencia” cuando alguien ha tenido interés en que algo sea “evidente”; existen importantes “evidencias” de comportamientos inaceptables en la investigación clínica patrocinada; el reporte de los resultados de los ensayos no sólo es incompleto, sino sesgado e inconsistente con los protocolos; y, sobre estos problemas, el sesgo de publicación produce una importante sobreestimación de la efectividad de los tratamientos y una no menos importante infraestimación de sus efectos adversos.

Estos problemas se pueden trasladar a los metaanálisis (incluso de instituciones muy respetadas),⁸¹ a las guías de práctica (a través de la incorporación de los resultados de los estudios clínicos), y pueden llegar a los clínicos a través de las revistas o, sobre todo, de la propia promoción farmacéutica. Su impacto sobre el manejo de los pacientes permanece inexplorado, aunque cabe presumir que en algunos casos - los AINE no selectivos, los COXIB, los ISRS- puede ser importante. Pero también se producen otros problemas. Los sesgos de interés en la investigación clínica minan la credibilidad de todos los ensayos y ponen en peligro la confianza en las profesiones sanitarias, en las instituciones académicas y en la propia industria farmacéutica y electromédica. Se trata de un problema grave porque la atención sanitaria se mantiene, sobre todo, sobre la confianza en que todos los agentes implicados tienen en cuenta, como interés primario, la mejor atención a los pacientes; y los intereses económicos (o los idearios o las creencias religiosas) están supeditados a este interés primario. Nadie espera (y probablemente no sería muy beneficioso) que las industrias relacionadas con la atención sanitaria sean organizaciones no lucrativas, pero sí que sus beneficios deriven de innovaciones reales demostradas con investigaciones válidas y no de una investigación sesgada, ni que la investigación clínica y las revistas médicas

sean, en palabras del exeditor del BMJ, una extensión de la división de marketing de las compañías farmacéuticas.⁸²

Por ello, y sobre todo en la última década, han aparecido importantes iniciativas para limitar el impacto del sesgo de financiación: registro obligatorio de los ensayos, cambios en los “requisitos de uniformidad” para evitar la publicación de los trabajos “cautivos”, guías de relaciones con la industria, normas gubernamentales, controles de la buena práctica de los ensayos, la publicación de todos los ensayos (recientemente asumida por buena parte de las grandes firmas farmacéuticas) y alguna otra. Muchas de estas iniciativas nacen en la propia industria farmacéutica que en los últimos años ha realizado notables esfuerzos por restaurar la confianza en los resultados de la investigación clínica.

En general, todas estas medidas -aun con la contrapartida de aumentar la complejidad de los ensayos, alargarlos y encarecerlos- son, teóricamente, de utilidad para reducir los sesgos asociados a la financiación. Pero el problema parece persistir, como muestra una reciente revisión de los grandes ensayos cardiovasculares publicados desde 2000 a 2005 (67% de los financiados por la industria favorecen al nuevo fármaco frente a 39% de los financiados por fuentes sin ánimo de lucro, sobre 324 ensayos cardiovasculares publicados en JAMA, NEJM y Lancet).⁸³ Nadie debería extrañarse de esta resistencia de un problema complejo a soluciones simples y en la que muchos agentes tienen intereses legítimos contrapuestos (la industria quiere que los ensayos favorezcan sus productos pero también que no generen desconfianza; los gobiernos quieren fármacos eficaces y seguros, pero también que sus industrias generen riqueza y sus centros de investigación sean potentes, etc.).

Colectivamente tiene interés explorar las políticas que puedan favorecer la producción de la información clínica que la sociedad necesita, en las áreas en que la necesita y con la calidad y veracidad que necesita. No es un tema fácil, pero la respuesta seguro que incluye una parte de mayor inversión pública en investigación, otra de incremento de la transparencia, algo de



mayor educación en ética de la investigación clínica, comportamientos más estrictos de los editores y, probablemente, cambios en los sistemas de patentes, regulación de los ensayos y en el papel de los comités de ensayos.¹² La investigación clínica tiene dimensión universal y, desde un país relativamente periférico en producción de conocimiento, estas grandes políticas son difícilmente abordables. Pero probablemente son las líneas por las que debe transcurrir la búsqueda de un mayor equilibrio entre la validez de los estudios y los intereses de los diversos agentes sanitarios.

CONVIVIENDO CON LOS FALSOS Y LOS VERDADEROS POSITIVOS DE LA INVESTIGACIÓN CLÍNICA

Individualmente, a nivel de los clínicos que deben decidir si usarán los resultados de una investigación para tratar a sus pacientes, el problema es más complejo y resulta difícil fijar normas rígidas que separarían los ensayos cuyos resultados positivos tienen una mayor probabilidad de ser “verdaderos positivos”. Obviamente hay que leer los ensayos, pero no parece factible –ni deseable– que los clínicos se conviertan en expertos cazadores de sesgos. Tampoco sirve, obviamente, desconfiar de todos los ensayos o de, lo que hoy sería casi lo mismo, los ensayos financiados por la industria. Esta práctica, al margen de ofrecer poca guía sobre cómo y a quién tratar, llevaría precisamente a que los pacientes no se beneficiaran de los resultados “verdaderos positivos”.

No existe una guía de *cribado* que permita seleccionar qué resultados de investigación son, aun en primera instancia, merecedores de confianza. Pero si hubiera que hacerla probablemente se tendrían en cuenta algunos de los elementos que siguen:^{2, 12, 84}

- Cuanto más grande sea el tamaño muestral de los estudios realizados en un determinado campo, mayor confianza ofrece el estudio (mayor probabilidad de que un estudio posi-

tivo sea un verdadero positivo, y un negativo un verdadero negativo).

- Cuanto más grande sea el tamaño del efecto (que no tiene nada que ver con los ceros que lleve p tras la coma), mayor la probabilidad de que las diferencias entre las intervenciones comparadas sean reales. Valore este efecto en frecuencias absolutas, como reducción absoluta del riesgo o como número de pacientes a tratar para evitar un resultado adverso, y no como reducción del riesgo relativo.
- Cuanto menor sea el número de asociaciones valoradas en un campo y más lógica –lógica clínica o fisiopatológica– la selección de las hipótesis, mayor confianza ofrecen los resultados del estudio. La viceversa también es sugestiva: cuantos más análisis exploratorios se hayan realizado, mayor probabilidad de que los resultados positivos sean falsos positivos. Además de en las llamadas *fishing-expeditions*, esta situación también se produce en los campos de investigación muy novedosos, con muchos equipos de investigación valorando la misma o similares hipótesis.
- Cuanto mayor sea la pulcritud del estudio en diseño, definiciones operativas, medidas de resultado (*endpoints* combinados, secundarios), análisis (subgrupos diversos), mayor la confianza que genera. Un ensayo clínico bien planteado, cuyo protocolo de investigación es público y se registró antes de iniciar el trabajo de campo, centrado en un único end-point clínicamente relevante y que no explota causalmente el análisis de subgrupos es un arquetipo de solidez. Esto no obsta para que el ensayo incluya otros *endpoints* o presente análisis de subgrupos, pero los resultados de estos subanálisis deben verse con mucha prudencia (más como generadores de hipótesis para otros ensayos que como conclusiones de la investigación realizada).
- Cuanto menores sean los intereses financieros en el estudio (o los prejuicios en el campo de investigación), mayor confianza. Esto no se refiere sólo a que el estudio esté finan-



ciado o no por parte interesada, sino al impacto financiero de un resultado negativo (mucho mayor en un *blockbuster* bajo patente que en medicamentos con escasa cuota de mercado o patente caducada). En este sentido, un resultado negativo en un estudio patrocinado por el fabricante y en un fármaco bajo patente ofrece gran confianza porque, previsiblemente, ha superado un diseño contrario, muchos reanálisis y muchas barreras.

- Finalmente, cuanto más similar sea el comparador utilizado (fármaco, dosis, vía de administración, etc.) al tratamiento estándar actual, y más parecidos sean los pacientes incluidos en el estudio a los que se ven normalmente por las consultas, mayor confianza.

Una guía para la confianza incluiría otras cosas (cómo se realizó la aleatorización, el enmascaramiento, si el análisis fue por intención de tratar y similares). Pero si el ensayo se ha publicado en una revista - con un mínimo de calidad seguro que en el proceso de *peer review* se han valorado estos aspectos, y la mayor parte de los elementos considerados en los párrafos previos se pueden mirar muy rápidamente (la mayoría de ellos en el resumen). No corresponden a ninguna métrica exacta, pero pueden ayudar a conformarse una opinión, obviamente subjetiva, sobre las posibilidades de que el resultado positivo de un ensayo sea algo a aplicar ya a nuestros pacientes o si valdría la pena esperar a que nuevos estudios confirmen o rechacen la utilidad del tratamiento propuesto.

Agradecimientos y advertencias. Parte de este manuscrito está basado en trabajos previos del autor y, en especial, la referencia 12 de la bibliografía. **Financiación:** Ninguna. **Conflicto de intereses:** ninguno en relación con este manuscrito.

Bibliografía

1. Peiró S. Medicina basada en la evidencia: necesaria, aunque insuficiente. Limitaciones -y autolimitaciones- de la Medicina basada en la evidencia

- para la práctica clínica y la atención sanitaria. *Med Clin (Barc)* 2002; 118(supl 3): 49-56.
2. Ioannidis JP. Why most published research findings are false. *PLoS Medicine* 2005; 2: e124. Epub 2005 Aug 30.
3. Ioannidis JP. Molecular bias. *Eur J Epidemiol* 2005; 20: 739-45.
4. Ioannidis JP. Microarrays and molecular research: noise discovery? *Lancet* 2005; 365: 454-5.
5. Ioannidis JP. Genetic associations: false or true? *Trends Mol Med* 2003; 9: 135-8.
6. Davey Smith G, Ebrahim S, Lewis S, Hansell AL, Palmer LJ, Burton PR. Genetic epidemiology and public health: hope, hype, and future prospects. *Lancet* 2005; 366: 1484-98.
7. Colhoun HM, McKeigue PM, Davey Smith G. Problems of reporting genetic associations with complex outcomes. *Lancet* 2003; 361: 865-72.
8. Ioannidis JP. Contradicted and initially stronger effects in highly cited clinical research. *JAMA* 2005; 294: 218-28.
9. Sterne JA, Davey Smith G. Sifting the evidence- what's wrong with significance tests? *BMJ* 2001; 322: 226-31.
10. Risch NJ. Searching for genetic determinants in the new millennium. *Nature* 2000; 405: 847-56.
11. Wacholder S, Chanock S, Garcia-Closas M, El Ghormli L, Rothman N. Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. *J Natl Cancer Inst* 2004; 96: 434-42.
12. Peiró S. La construcción de la evidencia basada en Medicina. *Gestión Clínica y Sanitaria* 2005; 7: 131-8.
13. Vandembroucke JP, Helmerhorst FM, Rosendaal FR. Competing interests and controversy about third generation oral contraceptives. *BMJ* readers should know whose words they read. *BMJ* 2000; 320: 381-2.
14. Lexchin J, Bero LA, Djulbegovic B, Clark O. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *BMJ* 2003; 326: 1167-70.
15. Bekelman JE, Li Y, Gross CP. Scope and impact of financial conflicts of interest in biomedical research: a systematic review. *JAMA* 2003; 289: 454-65.
16. Finucane TE, Boulton CE. Association of funding and findings of pharmaceutical research at a meeting of a medical professional society. *Am J Med* 2004; 117: 842-5.
17. Friedberg M, Saffran B, Stinson TJ, Nelson W, Bennett CL. Evaluation of conflict of interest in economic analyses of new drugs used in oncology. *JAMA* 1999; 282: 1453-7.



SALVADOR PEIRÓ - ¿SE PUEDE CONFIAR EN LOS RESULTADOS DE LA INVESTIGACIÓN CLÍNICA?

18. Kjaergard LL, Als-Nielsen B. Association between competing interests and authors' conclusions: epidemiological study of randomised clinical trials published in the BMJ. *BMJ* 2002; 325: 249.
19. Friedman LS, Richter ED. Relationship between conflicts of interest and research results. *J Gen Intern Med* 2004; 19: 51-6.
20. Bhandari M, Busse JW, Jackowski D, Montori VM, Schunemann H, Sprague S, et al. Association between industry funding and statistically significant pro-industry findings in medical and surgical randomized trials. *Can Med Assoc J* 2004; 170: 477-80.
21. Leopold SS, Warme WJ, Fritz Braunlich E, Shott S. Association between funding source and study outcome in orthopaedic research. *Clin Orthopaed Relat Res* 2003; 293-301.
22. Shah RV, Albert TJ, Bruegel-Sanchez V, Vaccaro AR, Hilibrand AS, Grauer JN. Industry support and correlation to study outcome for papers published in Spine. *Spine* 2005; 30: 1099-104.
23. Yaphe J, Edman R, Knishkowsky B, Herman J. The association between funding by commercial interests and study outcome in randomized controlled drug trials. *Fam Pract* 2000; 18: 565-8.
24. Djulbegovic B, Lacey M, Cantor A, Fields KK, Bennett CL, Adams JR, Kuderer NM, Lyman GH. The uncertainty principle and industry-sponsored research. *Lancet* 2000; 356: 635-8.
25. Kjaergard LL, Gluud C. Funding, disease area, and internal validity of hepatobiliary randomized clinical trials. *Am J Gastroenterol* 2002; 97: 2708-13.
26. Freedman B. Equipoise and the ethics of clinical research. *N Engl J Med* 1987; 317: 141-5.
27. Fries JF, Krishnan E. Equipoise, design bias, and randomized controlled trials: the elusive ethics of new drug development. *Arthritis Research and Therapeutics* 2004; 6: R250-5.
28. Weijer C, Shapiro SH, Cranley Glass K. For and against: clinical equipoise and not the uncertainty principle is the moral underpinning of the randomized controlled trial. *BMJ* 2000; 321: 756-8.
29. Steinbrook R. Gag clauses in clinical-trial agreements. *N Engl J Med* 2005; 352: 2160-2.
30. Bodenheimer T. Uneasy Alliance - Clinical investigators and the pharmaceutical industry. *N Engl J Med* 2000; 342: 1539-44.
31. Rosenberg SA. Secrecy in medical research. *N Engl J Med* 1996; 334: 392-4.
32. Larkin M. Whose article is it anyway? *Lancet* 1999; 354: 136.
33. Elstein AS. MDM policy regarding financial support of authors. *Med Decis Making* 1997; 17: 497-8.
34. Peiró S, García-Altés A, Meneu R, Librero J, Bernal E. La declaración del conflicto de intereses en las publicaciones científicas. ¿Tiempo para las luces y los taquígrafos en la trastienda de la investigación financiada por la industria? *Gac Sanit* 2000; 14: 472-81.
35. Blumenthal D, Causino N, Campbell E, Louis KS. Relationships between academic institutions and industry in the life sciences -an industry survey. *N Engl J Med* 1996; 334: 368-73.
36. AAMC Task Force on Financial Conflicts of Interest in Clinical Research. Protecting subjects, preserving trust, promoting progress I: policy and guidelines for the oversight of individual financial interests in human subjects research. *Acad Med* 2003; 78: 225-36.
37. AAMC Task Force on Financial Conflicts of Interest in Clinical Research. Protecting subjects, preserving trust, promoting progress II: principles and recommendations for oversight of an institution's financial interests in human subjects research. *Acad Med* 2003; 78: 237-45.
38. García AM. Nueva versión de los «Requisitos de uniformidad para los manuscritos enviados a revistas biomédicas y de ciencias de la salud». *Gac Sanit* 2004; 18: 163-5.
39. Schulman KA, Seils DM, Timbie JW, Sugarman J, Dame LA, Weinfurt KP, et al. A national survey of provisions in clinical-trial agreements between medical schools and industry sponsors. *N Engl J Med* 2002; 347: 1335-41.
40. Mello MM, Clarridge BR, Studdert DM. Academic medical centers' standards for clinical-trial agreements with industry. *N Engl J Med* 2005; 352: 2202-10.
41. Johansen HK, Gotzsche PC. Problems in the design and reporting of trials of antifungal agents encountered during meta-analysis. *JAMA* 1999; 282: 1752-9.
42. Safer DJ. Design and reporting modifications in industry-sponsored comparative psychopharmacology trials. *J Nerv Ment Dis* 2002; 190: 583-92.
43. Rochon PA, Gurwitz JH, Simms RW, Fortin PR, Felson DT, Minaker KL, et al. A study of manufacturer-supported trials of nonsteroidal anti-inflammatory drugs in the treatment of arthritis. *Arch Intern Med* 1994; 154: 157-63.
44. Rochon PA, Berger PB, Gordon M. The evolution of clinical trials: inclusion and representation. *Can Med Assoc J* 1998; 1373-4.
45. Zimmerman M, Mattia JI, Posternak MA. Are subjects in pharmacological treatment trials of depression representative of patients in routine clinical practice? *Am J Psychiatry* 2002; 159: 469-73.
46. Gurtwitz JH, Col NF, Avorn J. The exclusion of the elderly and women from clinical trials in acute myocardial infarction. *JAMA* 1992; 268: 1417-22.



SALVADOR PEIRÓ - ¿SE PUEDE CONFIAR EN LOS RESULTADOS DE LA INVESTIGACIÓN CLÍNICA?

47. Hutchins LF, Unger JM, Crowley JJ, Coltman CAJ, Albain KS. Underrepresentation of patients 65 years of age or older in cancer-treatment trials. *N Engl J Med* 1999; 341: 2061-7.
48. Lee PY, Alexander KP, Hammill BG, Pasquali SK, Peterson ED. Representation of elderly persons and women in published randomized trials of acute coronary syndromes. *JAMA* 2001; 286: 708-13.
49. Schellevis FG, Velden J vd, Lisdonk E vd, Eijk J ThM van, Weel C van. Comorbidity of chronic diseases in general practice. *J Clin Epidemiol* 1993; 46: 469-73.
50. McKee M, Britton A, Black N, McPherson K, Sanderson C, Bain C. Methods in health services research: interpreting the evidence: choosing between randomised and non-randomised studies. *BMJ* 1999; 319: 312-5.
51. Villanueva P, Peiro S, Librero J, Pereiro I. Accuracy of pharmaceutical advertisements in medical journals. *Lancet* 2003; 361: 27-32.
52. Manheimer E, Anderson D. Survey of public information about ongoing clinical trials funded by industry: evaluation of completeness and accessibility. *BMJ* 2002; 325: 528-31.
53. Hrachovec JB. Reporting of 6 month vs 12 month data in a clinical trial of Celecoxib. *JAMA* 2001; 286: 2398.
54. McCormack JP, Rangno R. Digging for data from the COX-2 trials. *Can Med Assoc J* 2002; 166: 1649-50.
55. Juni P, Rutjes AW, Dieppe PA. Are selective COX 2 inhibitors superior to traditional non steroidal anti-inflammatory drugs? *BMJ* 2002; 324: 1287-8. Erratum in: *BMJ* 2002; 324: 1538.
56. Curfman GD, Morrissey S, Drazen JM. Expression of concern: Bombardier et al., "Comparison of upper gastrointestinal toxicity of rofecoxib and naproxen in patients with rheumatoid arthritis". *N Engl J Med* 2000;343:1520-8. *N Engl J Med* 2005; 353: 2813-4.
57. Curfman GD, Morrissey S, Drazen JM. Expression of Concern Reaffirmed. *N Engl J Med* 2006 Feb 22; [Epub ahead of print].
58. Bombardier C, Laine L, Reicin A, Shapiro D, Burgos-Vargas R, Davis B, et al. Comparison of upper gastrointestinal toxicity of rofecoxib and naproxen in patients with rheumatoid arthritis. VIGOR Study Group. *N Engl J Med* 2000; 343: 1520-8.
59. Juni P, Rutjes AW, Dieppe PA. Are selective COX 2 inhibitors superior to traditional non steroidal anti-inflammatory drugs? *BMJ* 2002; 324: 1287-8. Erratum in: *BMJ* 2002; 324: 1538.
60. Meneu R, Peiró S, Márquez-Calderón S. Efecto de la presentación de los resultados de los ensayos clínicos en la intención de prescribir: relativizando el riesgo relativo. *Aten Primaria* 1998; 21: 446-50.
61. Cummings SR, Black DM, Thompson DE, Applegate WB, Barrett-Connor E, Musliner TA, et al. Effect of alendronate on risk of fracture in women with low bone density but without vertebral fractures. Results from the Fracture Intervention Trial. *JAMA* 1998; 280: 2077-82.
62. Chan AW, Hrobjartsson A, Haahr MT, Gotzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA* 2004; 291: 2457-65.
63. Egger M, Davey Smith G. Misleading in meta-analysis. *BMJ* 1995; 310: 752-4.
64. Wise P, Drury M. Pharmaceutical trials in general practice: the first 100 protocols. An audit by the clinical research ethics committee of the Royal College of General Practitioners. *BMJ* 1996; 313: 1245-8.
65. Chalmers I. Underreporting research is scientific misconduct. *JAMA* 1990; 263: 1405-8.
66. Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. *Lancet* 1991; 337: 867-72.
67. Dickersin K, Min YI, Meinert CL. Factors influencing publication of research results. Follow-up of applications submitted to two institutional review boards. *JAMA* 1992; 267: 374-8.
68. Rennie D, Flanagan A. Publication bias. The triumph of hope over experience. *JAMA* 1992; 267: 411-2.
69. Liebeskind DS, Kidwell CS, Saver JL. Empiric evidence of publication bias affecting acute stroke clinical trials. *Stroke* 1999; 30: 268.
70. Melander H, Ahlqvist-Rastad J, Meijer G, Beer-mann B. Evidence based medicine -selective reporting from studies sponsored by pharmaceutical industry: review of studies in new drug applications. *BMJ* 2003; 326: 1171-3.
71. Tramer MR, Reynolds DJ, Moore RA, McQuay HJ. Impact of covert duplicate publication on meta-analysis: a case study. *BMJ* 1997; 315: 635-40.
72. Vickers A, Goyal N, Harland R, Rees R. Do certain countries produce only positive results. A systematic review of controlled trials. *Control Clin Trials* 1998; 19: 159-66.
73. Dieppe PA, Frankel SJ, Toth B. Is research into the treatment of osteoarthritis with nonsteroidal anti-inflammatory drugs misdirected? *Lancet* 1993; 341: 867-72.
74. Warlow C, Sandercock P, Dennis M, Wardlaw J. Research Funding. *Lancet* 1999; 353: 2250.
75. Rochon PA, Mashari A, Cohen A, et al. Relation between randomized controlled trials published in lea-



SALVADOR PEIRÓ - ¿SE PUEDE CONFIAR EN LOS RESULTADOS DE LA INVESTIGACIÓN CLÍNICA?

- ding general medical journals and the global burden of disease. *Can Med Assoc J* 2004; 170: 1673-7.
76. European Society of Hypertension-European Society of Cardiology Guidelines Committee. 2003 European Society of Hypertension-European Society of Cardiology guidelines for the management of arterial hypertension. *J Hypertens* 2003; 21: 1011-53. Erratum in: *Journal of Hypertension* 2003; 21: 2203-4. *J Hypertens* 2004; 22: 435.
 77. Dieppe P. Evidence-based medicine or medicines-based evidence? *Ann Rheum Dis* 1998; 57: 385-6.
 78. Gotzsche PC. Methodology and overt and hidden bias in reports of 196 double-blind trials of nonsteroidal antiinflammatory drugs in rheumatoid arthritis. *Control Clin Trials* 1989; 10: 31-56.
 79. Whittington CJ, Kendall T, Fonagy P, Cottrell D, Cotgrove A, Boddington E. Selective serotonin reuptake inhibitors in childhood depression: systematic review of published versus unpublished data. *Lancet* 2004; 363: 1341-5.
 80. Peiró S, Cervera Casino P, Bernal Delgado E. Los inhibidores selectivos de la recaptación de serotonina en la depresión infantil: un culebrón que refleja importantes problemas de seguridad de los medicamentos. *Gac Sanit* 2005; en prensa.
 81. 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: 921-8.
 82. Smith R. Medical Journals Are an Extension of the Marketing Arm of Pharmaceutical Companies. *PLoS Medicine* 2005; 2: e138.
 83. Ridker PM, Torres J. Reported outcomes in major cardiovascular clinical trials funded by for-profit and not-for-profit organizations: 2000-2005. *JAMA* 2006; 295: 2270-4.
 84. Montori VM, Jaeschke R, Schunemann HJ, Bhandari M, Brozek JL, Devereaux PJ, et al. Users' guide to detecting misleading claims in clinical research reports. *BMJ* 2004; 329: 1093-6.