

Sobre la existencia y los factores de riesgo del sesgo de publicación¹

Kay Dickersin²

El sesgo de publicación es la tendencia de los investigadores, revisores y directores de revistas a proponer o aceptar la publicación de manuscritos en función de la dirección o de la intensidad de los resultados del estudio. Gran parte de lo que se sabe respecto al sesgo de publicación procede de las ciencias sociales y muy poco de la medicina. En este campo, tres estudios proporcionan pruebas directas de la existencia de este sesgo. Es importante prevenir el sesgo de publicación, tanto desde una perspectiva científica (de difusión completa del conocimiento) como desde el punto de vista de quienes combinan resultados de varios estudios similares (metanálisis). Si se pretende que las decisiones terapéuticas se basen en los estudios publicados, las publicaciones científicas deben incluir todos los datos disponibles cuya calidad sea aceptable. Actualmente es difícil o incluso imposible obtener información sobre todos los estudios realizados en torno a un tema determinado. La comunidad científica debe avanzar hacia el registro de los ensayos clínicos y quizá de otros tipos de estudios.

El intelecto humano ... se conmueve y se excita más por las afirmaciones que por las negaciones.

FRANCIS BACON, 1621

Aproximadamente desde 1450, el método habitual de difundir información y adquirir conocimientos ha sido el uso de la palabra escrita. Los méritos de la escritura se reconocían mucho antes de que Gutenberg inventara la imprenta (Moisés no confió a tradición oral la interpretación y transmisión correcta de los Diez Mandamientos). La tradición escrita tiene muchas cualidades valiosas: su conservación, básica para el registro de información histórica sobre familias y comunidades; su existencia como base de comprensión común, fundamental para las leyes; y su utilidad como vehículo para compartir la información que se considera importante, como las noticias de actualidad o los últimos descubrimientos científicos.

¹ Esta traducción del artículo "The existence of publication bias and risk factors for its occurrence" (*Journal of the American Medical Association* 1990;263(10):1385-1389) se publica con autorización de la revista y de la autora. © American Medical Association, 1990. Presentado al Primer Congreso Internacional sobre el Arbitraje Científico en las Publicaciones Biomédicas, Chicago, 10-12 de mayo de 1989.

² Departamento de Epidemiología, Escuela de Higiene y Salud Pública, Universidad Johns Hopkins, Baltimore, Maryland. Las separatas de la versión original en inglés pueden solicitarse a: Dr. K. Dickersin, Department of Ophthalmology, University of Maryland School of Medicine, 22 S. Greene St., Baltimore, MD 21201, Estados Unidos de América.

En la actividad artística es aceptable e incluso muy bien recibida la reinterpretación de la palabra escrita, por ejemplo mediante una traducción. En otros campos las ambigüedades inherentes al lenguaje llevan a una lucha constante por descifrar el significado o la intención de la palabra escrita. La Constitución de los Estados Unidos es un buen ejemplo. En la ciencia es esencial que el trabajo realizado y los resultados obtenidos se describan con una redacción clara, exacta y precisa. Es imperativo que solo haya una interpretación posible de lo escrito.

Además, el avance de la ciencia exige la difusión completa de la investigación, incluidos tanto los experimentos o estudios llevados a cabo como los procedimientos y métodos empleados en dichas investigaciones. En la práctica no es posible ni siquiera deseable que se notifique cada experimento o cada elemento de un experimento. Sin embargo, no parece haber normas aceptadas mediante las cuales un investigador pueda decidir lo que vale la pena divulgar; la decisión de notificar los resultados propios y la forma en que dicha comunicación se lleva a cabo quedan a juicio del investigador.

La cuestión de cómo y cuándo se hacen públicos los resultados de la investigación es de interés por el sesgo de selección potencial: dadas ciertas características de un estudio —diseño, realización y resultado final— ¿puede decirse cuán probable es que se llegue a publicar? Si la respuesta es afirmativa, aquello en lo que se basa nuestro “conocimiento”, los trabajos publicados, es una representación sesgada del conocimiento tomado en su conjunto.

Si las características que determinan la publicación tienen relación con la calidad del estudio, el sesgo de selección que se produce al estudiar solo los trabajos publicados es aceptable o incluso deseable. Si, por el contrario, es la dirección de los resultados del estudio o su significación estadística lo que motiva las diferencias de publicación, la información disponible para la comunidad científica puede tener un sesgo considerable. El sesgo que se genera cuando la publicación de los resultados de estudios se basa en la dirección o en la significación de los resultados se denomina *sesgo de publicación*. Parece que fue Smith (1) quien usó por primera vez esta expresión en una publicación científica, en 1980.

Dando por supuesta la existencia del sesgo de publicación, ¿tiene algún sentido preocuparse por él? Desde el punto de vista científico, desde luego que vale la pena preocuparse. Si se considera que las decisiones sobre los tratamientos médicos deben basarse en todos los datos de buena calidad existentes, hay que insistir en que pueda disponerse de ellos. Sin embargo, las decisiones médicas hasta ahora se han guiado en la práctica por la formación individual del médico y su experiencia personal. Pero ha habido un cambio reciente. Las conferencias de consenso, el análisis de decisiones, los “sistemas expertos”, el uso de los ensayos clínicos como base para la elaboración de estrategias y el metanálisis han hecho que las decisiones relativas al tratamiento médico se tomen con criterios más científicos.

SESGO DE PUBLICACIÓN

Aspectos históricos

No parece que haya criterios científicos formales referentes a cuándo deben o no publicarse los resultados de los estudios. La decisión sobre qué incluir en una publicación y si publicar o no es fundamentalmente personal, aunque dictada en cierta medida por la moda. Cuando en 1680 el químico Robert Boyle publicó sus

experimentos sobre el aire, se le atribuyó haber sido el primero en indicar los detalles de los experimentos y las precauciones necesarias para poderlos repetir. Su publicación anunció un tipo de informe nuevo, en el que se describían las dificultades y los errores. Así, durante los siglos XVII, XVIII y comienzos del XIX los informes científicos habituales daban cuenta no solo de los resultados “positivos” sino también de los resultados “negativos” o “nulos”.

Preocupado por las prácticas de publicación en las ciencias físicas y biológicas, Boyle lamentaba en 1661 que los científicos no notificaran resultados aislados, sintiéndose por el contrario obligados a no publicar hasta tener un “sistema” elaborado que consideraran merecedor de presentación formal: “Pero lo peor aún no se ha dicho y es que, mientras esta forma vanidosa de pensar obliga a describir o un sistema o nada, muchas nociones o experimentos excelentes son suprimidos por hombres serios y modestos...” (2). Parece pues que la idea de ir a la imprenta solo si se dispone de algo “grande” que publicar no es de ninguna manera moderna.

A mediados del siglo XIX el estilo de la escritura científica estaba en proceso de cambio hacia el modelo conciso y tecnificado que hoy nos es familiar. Las limitaciones de tiempo (a medida que la ciencia comenzó a avanzar rápidamente) y de espacio en las revistas, la aparición de grupos de científicos que trabajaban y publicaban juntos, las respuestas a la revisión científica y la dependencia económica respecto a un sistema que recompensaba el éxito rápido fueron factores que llevaron a un cambio en la redacción y en la edición científica. El cambio de estilo que ha tenido lugar a lo largo de los años no es en sí perjudicial. El problema es si la mayor concisión ha llevado a la pérdida de información o si representa una notificación con sesgo.

Pruebas del sesgo de publicación

Quizá como resultado de la dificultad de diseñar estudios sobre el sesgo de publicación, a este respecto se han escrito más quejas que informes de investigaciones llevadas a cabo para evaluarlo. La mayor parte de la investigación sobre el sesgo de publicación se ha hecho en el campo de la psicología y la educación (3–10).

Probablemente fue Sterling (3) el primero que llamó la atención sobre el grave problema que supone la tendencia a publicar resultados positivos y rechazar los negativos. Este autor revisó todos los artículos publicados en cuatro revistas durante un año (1955 o 1956) y comprobó que en 97% de los artículos en los que se usaban pruebas de significación se rechazaba la hipótesis nula. En ciencias sociales y de la conducta otros autores han encontrado pruebas similares de sesgo de publicación (4–9).

Se han hecho dos estudios experimentales al respecto (*New York Times*, 27 de septiembre de 1988) (10); en ambos se encontró que, cuando las demás variables se mantenían constantes, la opinión de los revisores científicos dependía mucho de la dirección y la intensidad de los resultados del estudio. El estudio de Mahoney (10) es especialmente ilustrativo de dónde puede haber sesgos en el proceso de revisión. A 75 revisores científicos de una revista se les asignó al azar uno de cinco manuscritos similares. Las secciones de introducción y de métodos de todos los manuscritos eran idénticas pero variaban las de resultados y discusión. Un grupo de revisores recibió un original que presentaba resultados positivos y otro recibió un original con resultados negativos. A un tercer grupo se les pidió que evaluaran los originales basándose solo en la sección de métodos y en la relevancia del tema y no

se les proporcionaron resultados. Los grupos cuarto y quinto recibieron originales con resultados "ambiguos", con la sección de discusión bien positiva, bien negativa. Los revisores usaron una escala de 0 a 6 para evaluar (de menos a más) los originales respecto a cinco aspectos: relevancia, métodos, presentación de los datos, aporte científico y valor de publicación. Las calificaciones de los revisores se presentan en el cuadro 1. Los estudios con resultados positivos o negativos tenían secciones de métodos idénticas, pero los revisores calificaron los resultados negativos con puntuaciones menores en lo referente a calidad de los métodos y lo mismo ocurrió con los estudios de resultados ambiguos. La presentación de los datos, el aporte científico y el valor de publicación también recibieron calificaciones más bajas cuando los resultados eran negativos o ambiguos. Los estudios negativos recibieron igualmente puntuaciones significativamente menores respecto al valor de publicación.

En medicina se ha hecho poco para investigar la posibilidad de sesgo de publicación (11). A pesar de la escasez de datos empíricos, se ha aceptado su existencia como hecho y no como problema hipotético que requiere estudio ulterior. Sin tener en cuenta la falta de datos adecuados, investigadores renombrados han escrito artículos en revistas médicas en los que se menciona el sesgo de publicación como si su existencia fuera conocida y sus causas bien entendidas. "Los investigadores tienen más motivación para enviar a publicación resultados positivos que resultados nulos. Muchos directores de revistas determinan los artículos que publican basándose en ese mismo criterio, y algunos buscan incluso que los valores P no lleguen a 0,05. Los ensayos clínicos publicados constituyen inevitablemente una selección positivamente sesgada" (12).

La información referente a las prácticas de publicación ni se obtiene fácilmente ni es tampoco fácil de encontrar; es probable que esta sea la razón por la que el tema se ha investigado tan poco. Una posibilidad sería hacer una encuesta entre investigadores, preguntándoles por sus hábitos y experiencia. Los problemas que esto plantea fueron puestos de manifiesto en un estudio de Hetherington et al. (13), en el que se envió un cuestionario de una página solicitando información sobre investigaciones perinatales a 42 000 obstetras y pediatras de todo el mundo. De los 395 estudios no publicados notificados a los investigadores, solo 18 habían sido terminados más de 2 años antes de la encuesta. El resto estaban aún en marcha (252) o se trataba de estudios en los que se había dejado de incluir casos durante los 2 años anteriores a la finalización de la encuesta (125), de manera que correspondían al período necesario para que los resultados fueran publicados. Hetherington et al. concluyeron que no es posible estimar el tamaño del sesgo de publicación mediante la identificación retrospectiva de los estudios no publicados.

CUADRO 1. Calificación de un mismo original con distintas secciones de resultados o discusión

Presentación	No. de revisores	Calificaciones medias			
		Métodos	Presentación de los datos	Valor científico	Interés de publicación
Resultados positivos	12	4,2	4,3	4,3	3,2
Resultados negativos	14	2,4	2,6	2,4	1,8
Solo métodos	14	3,4	...	4,5	3,4
Resultados ambiguos, discusión positiva	13	2,5	1,3	1,6	0,5
Resultados ambiguos, discusión negativa	14	2,7	2,0	1,7	1,4

¿Qué datos apoyan en medicina la idea de que la publicación de un estudio tiene relación con la dirección y la intensidad de sus resultados? Varios estudios han proporcionado pruebas indirectas del sesgo de publicación (14-16). Chalmers revisó 23 trabajos que proporcionaban tasas de letalidad de la hepatitis sérica y vio que las tasas notificadas oscilaban de 0,3 a 62%. Las tasas mayores se asociaban a los estudios sobre un pequeño número de pacientes. Los autores sugirieron que estos resultados podrían indicar una tendencia de los investigadores a publicar resultados inusuales.

El cuadro 2 muestra pruebas directas de sesgo de publicación en la literatura médica. Simes (17) comparó los resultados de los estudios publicados y de los estudios registrados en el Banco de Datos Internacional para la Investigación del Cáncer (*International Cancer Research Data Bank*). Se eligieron estudios correspondientes a dos situaciones terapéuticas: a) agentes alquilantes iniciales frente a quimioterapia combinada del cáncer ovárico avanzado; y b) agentes alquilantes o prednisona frente a quimioterapia combinada del mieloma múltiple. Todos los estudios se habían registrado o publicado antes de octubre de 1983. Los resultados combinados de los estudios publicados de tratamientos del cáncer ovárico mostraban un beneficio estadísticamente significativo de la terapéutica combinada, mientras que los resultados combinados de los estudios registrados, que incluían estudios publicados y no publicados, no mostraban una ventaja significativa de supervivencia. De igual manera, se vio una ventaja estadísticamente significativa de supervivencia en la terapéutica combinada en los estudios publicados de tratamiento del mieloma: en los estudios registrados la ventaja era menor, aunque estadísticamente significativa de todas formas.

Se han obtenido más pruebas a partir de una encuesta a 318 autores de ensayos publicados a los que se preguntó si habían participado en algún estudio no publicado (18) (cuadro 3). Los 156 autores que respondieron notificaron 271 estudios no publicados y 1 041 estudios publicados. Los ensayos clínicos no publicados apoyaban la terapéutica experimental 14% de las veces, frente a 55% de los ensayos clí-

CUADRO 2. Investigaciones del sesgo de publicación en medicina

Fuente, año	Tema	Procedencia de los datos	Método de investigación	Resultados
Simes, 1986 (17)	Ensayos clínicos oncológicos	Registro de ensayos clínicos oncológicos	Publicaciones y registro	Los estudios publicados muestran mayor eficacia del tratamiento combinado
Dickersin et al., 1987 (18)	Ensayos clínicos aleatorizados y controlados	Archivo de ensayos clínicos aleatorizados y controlados	Cuestionario	Los estudios publicados apoyan más frecuentemente el tratamiento experimental
Sommer, 1987 (19)	Investigación del ciclo menstrual	Miembros de una sociedad científica	Cuestionario	Los estudios publicados son más a menudo estadísticamente significativos
Chalmers et al., 1989 (24)	Ensayos clínicos perinatales	<i>Abstracts</i> del RECPO*	Informes completos del RECPO	La intensidad de los resultados divulgados como <i>abstracts</i> no muestra asociación con la publicación final del trabajo

* RECPO indica Base de Datos de Ensayos Clínicos Perinatales de Oxford (*Oxford Database of Perinatal Trials*).

CUADRO 3. Resultados de ensayos clínicos aleatorizados y controlados (ECAC) que fueron publicados comparados a los de ECAC finalizados y no publicados*

Tendencia o significación estadística	ECAC publicados [†]		ECAC finalizados no publicados [†]	
	No.	% del total con la tendencia especificada	No.	% del total con la tendencia especificada
A favor de la terapéutica nueva ($P < 0,05$)	423	55,1	26	14,6
Tendencia favorable a la terapéutica nueva	123	16,0	40	22,5
Sin diferencia entre las terapéuticas	170	22,2	79	44,4
Tendencia a favor de la terapéutica de control o estándar	25	3,3	23	12,9
Favors control or standard therapy ($P < .05$)	26	3,4	10	5,6
Total de ECAC con tendencia especificada	767	100,0	178	100,0
No. de ECAC con tendencia de resultados no especificada	274	...	26	...
Total de ECAC	1041	...	204 [‡]	...

* Reimpreso con autorización de la editorial, © 1987 Elsevier Science Publishing Co, Inc.

[†] χ^2 (con 4 grados de libertad) = 111,09: $P < 0,001$.

[‡] No incluidos 34 ensayos clínicos no publicados realizados por un solo autor.

nicos publicados. Las razones principales que los autores daban para explicar que el estudio no se hubiera publicado fueron que los resultados no apoyaban el tratamiento sometido a prueba y la falta de interés (cuadro 4). De los datos parece deducirse que el hecho de que el estudio no se publicara dependía sobre todo de que los resultados no se redactaran, más que de las decisiones de los revisores científicos o los directores de revistas.

Sommer (19) encuestó a los 140 miembros de la Sociedad para Investigaciones sobre el Ciclo Menstrual y detectó 73 estudios publicados y 218 estudios no publicados (la tasa de respuesta fue de 67%). Treinta por ciento de los 73 estudios publicados, 38% de los 42 estudios en proceso de publicación y 29% de los 28 estudios no publicados tenían resultados estadísticamente significativos. Cuando solo se consideraron los estudios en los que se notificaba la significación estadística del resultado, los porcentajes eran de 61, 76 y 40%, respectivamente. Según Sommer, el mejor factor para predecir si el estudio se publicaría o no era la publicación previa por el investigador interrogado. Los investigadores que solo podían aducir un estudio entre sus méritos (76%) tendían a no publicar, mientras que los que tenían dos o más generalmente habían publicado uno o más (67%). Además, si el primer estudio se publicaba, era más probable que el investigador realizara otros estudios (68%) respecto a los investigadores que no habían publicado el primero (35%).

Se han notificado también varios estudios sobre la publicación completa de trabajos inicialmente publicados como resúmenes (*abstracts*) (20–24). Los datos de estos estudios han sido llamativamente homogéneos, mostrando que solo de 30 a 60% de los resultados publicados como resúmenes se publican finalmente como trabajos completos. McCormick y Holmes (23) vieron que los resúmenes pediátricos que habían sido propuestos para presentación y rechazados fueron finalmente pu-

CUADRO 4. Situación de los ensayos clínicos aleatorizados y controlados (ECAC) y razones por las que no se publicaron los ECAC finalizados (18)*

Situación del ECAC	ECAC interrumpido	ECAC finalizado		Sin respuesta	Total (%)
		y artículo no enviado a publicación	y artículo enviado a publicación		
Artículo en proyecto, en elaboración o en revisión científica	0	15	10	...	25 (12)
Resultados negativos	16	35	7	...	58 (28)
Falta de interés	6	16	2	...	24 (12)
Problemas de tamaño muestral	20	3	0	...	23 (11)
Métodos inadecuados	6	2	1	...	9 (4)
Efectos colaterales	12	1	0	...	13 (6)
Problemas del grupo externo	9	1	0	...	10 (5)
Controversia	0	3	2	...	5 (2)
Desconocida o sin contestación	5	26	1	5	37 (18)
Total (%) de ECAC	74 (36)	102 (50)	23 (11)	5 (2)	204†

* Reimpreso con autorización de la editorial, © 1987 Elsevier Science Publishing Co, Inc.

† No incluidos 34 ensayos clínicos finalizados llevados a cabo por un solo autor.

blicados como trabajos completos solo de 13 a 22% de las veces, mientras que de 49 a 54% de los resúmenes aceptados se publicaron finalmente como trabajos completos. Esto puede indicar que los artículos seleccionados representan estudios de mejor calidad, pero también podría indicar mayor interés editorial en los resultados presentados en los artículos seleccionados. Chalmers et al. (24) investigaron sistemáticamente los informes sumarios contenidos en el Registro de Ensayos Clínicos Perinatales de Oxford (*Oxford Database of Perinatal Trials*) y que fueron publicados entre 1940 y 1984. La investigación del registro mediante nombre de los autores reveló que aproximadamente 37% fueron publicados posteriormente como trabajos completos. Ni la calidad del estudio, juzgada por el informe sumario, ni tampoco sus resultados mostraron asociación con que el trabajo se publicara o no finalmente. Como este diseño no proporciona información alguna sobre el proceso que media entre el análisis de los datos y la decisión de publicación, los resultados no indican necesariamente la ausencia de sesgo de publicación.

Con la información de que se dispone es difícil estimar la magnitud del sesgo de publicación, incluso *grosso modo*. Según los resultados de las investigaciones sobre el tema, los estudios publicados se hallan respecto a los no publicados en una razón de entre 128:1 (13) y 1:1 (17); la mayor parte de los estudios dan una razón entre 10:1 y 1:1 (1, 9, 13, 17-19). Nosotros estamos llevando a cabo en la Universidad Johns Hopkins de Baltimore, Maryland, un estudio prospectivo que proporcionará una mejor estimación de la magnitud del problema. El plan del proyecto es investigar los estudios aprobados en 1980 por los consejos de revisión institucional de nuestra institución y los ensayos clínicos financiados por los Institutos Nacionales de Salud de Bethesda, Maryland, en 1979, para ver si hay sesgo de publicación, y, si lo hay, cuáles son los factores de riesgo. Los factores de riesgo potenciales son las características de diseño del estudio como el tamaño de la muestra, tipo de grupo control y número de centros colaboradores; las características del investigador; la fuente de financiación; y la intensidad de los resultados del estudio.

Los directores de revistas y el sesgo de publicación

¿De dónde salió el sesgo de publicación? En la literatura médica parece que se da por supuesto que el sesgo hacia la publicación de resultados llamativos tiene su origen en los directores de revistas. Esta creencia tiene cierta base. En 1980 la *British Medical Journal* indicó que su artículo ideal describiría "resultados que afecten la práctica clínica ... y hallazgos que mejoren el pronóstico o simplifiquen el tratamiento de enfermedades comunes..." (25). Un suelto publicado en 1983 en la sección *Views* [Perspectivas] de la misma revista aconsejaba "a quienes buscan la publicación rápida de un trabajo (especialmente si son resultados negativos)" que enviaran el original a una de las revistas en las que se paga por publicar. El texto acababa indicando "en cuanto al número de personas que lo verán entonces ... bien, los resultados negativos nunca han llamado mucho la atención..." (26).

Esta situación ha llevado a discutir si tendría sentido una revista para publicar resultados negativos. "La necesidad científica de una publicación de resultados negativos o diferencias nulas es evidente y la 'Revista de Resultados Negativos' se ha disputado durante años, casi como broma morbosa. Esa revista no solo sería decididamente aburrida sino catastrófica desde el punto de vista financiero" (12).

Maxwell (12) llegó a proponer que, al menos, los directores de revistas deberían formalizar un registro de resultados negativos accesible mediante el *Index Medicus*, en el que se citase el tema, los nombres y direcciones de los autores y el título del estudio. Esta solución del problema sería tan poco satisfactoria como la otra.

Según Edward Huth, director de *Annals of Internal Medicine*, podría publicarse una revista médica en soporte informático a menor costo que las revistas impresas. Ese tipo de revista estaría así en mejor posición para publicar resultados negativos o nulos. Sin embargo, prevenía, "en el título o en el resumen debe quedar bien claro que el trabajo llega a una conclusión negativa, con el fin de que los autores o investigadores no crean que se trata de resultados positivos" (*New York Times*, 29 de abril de 1986: C1, C7).

The Journal of the American Medical Association tuvo una vez una sección titulada "Resultados negativos". Un vistazo a los volúmenes de la revista sugiere que fue una sección más o menos regular publicada mensualmente entre 1962 y 1968. Los artículos tenían entre página y media y dos páginas de extensión. Desgraciadamente, no se dispone de información sobre las razones por las que se inició y se continuó dicha sección, a pesar de que se ha intentado averiguarlas (Dr. E. Knoll, comunicación personal, marzo de 1988).

Función de la calidad del estudio

Hasta ahora se ha hecho hincapié en la función potencial de la dirección de los resultados del estudio sobre su publicación, pero hay otros factores de riesgo potencial para que el trabajo no se publique. Los más sobresalientes son la calidad, el tamaño muestral y la fuente de financiación de la investigación.

En 1980, Hemminki (27) describió la calidad de la información enviada al servicio de homologación de fármacos y la relacionó con la situación de publicación o no publicación de los trabajos (cuadro 5). Este autor revisó las peticiones de autorización de fármacos psicotrópicos en Finlandia y Suecia en 1965, 1970, 1974 y 1975, y las de una muestra aleatoria de fármacos no psicotrópicos en Finlandia. Para todos los años en conjunto, 39 y 44% de los estudios incluidos en la solicitud en Fin-

CUADRO 5. Características de los estudios de fármacos según su situación de publicación (27)

Características de los estudios	Fármacos no psicotrópicos (n = 69)		Fármacos psicotrópicos (n = 234)	
	Publicados	No publicados	Publicados	No publicados
% controlados	47	26	47	52
% "buena" calidad	23	46	35	37
% que contenían información sobre efectos adversos	43	83	56	77
Tamaño muestral medio	62	48	83	76

landia o Suecia, respectivamente, no se habían publicado. Los informes no publicados de estudios de fármacos psicotrópicos contenían más a menudo información concerniente a la selección de pacientes y criterios de exclusión que los informes publicados, aunque no ocurría lo mismo respecto a los estudios no publicados de los fármacos no psicotrópicos. En conjunto, la calidad de los informes publicados y no publicados parecía más o menos igual. Según este estudio, la calidad de un ensayo clínico financiado por una compañía farmacéutica no es un factor influyente en las decisiones de publicación.

Influencia del tamaño muestral

El tamaño de una investigación puede tener un efecto directo o indirecto en las decisiones de publicación. En el estudio realizado por mí y mis colegas anteriormente mencionado (18) la mediana del tamaño muestral de los estudios no publicados llevados a cabo por un grupo determinado de investigadores fue 24, mientras que la de los ensayos aleatorizados y controlados "índice" (una muestra de estudios aleatorizados y controlados utilizada para generar la lista de autores encuestados) fue 68. Chalmers et al. (24) vieron que de 76 *abstracts* que describían estudios perinatales, los que tenían un tamaño muestral mayor que la mediana del conjunto eran más susceptibles de publicación completa que los de tamaño muestral inferior a la mediana. (Sin embargo, cuando se estratificó el tamaño muestral en más categorías, la asociación entre tamaño muestral y publicación no fue significativa.) De alguna manera, que los estudios más pequeños no se publiquen tan a menudo puede ser alentador, porque el tamaño muestral puede ser indicativo de la calidad del estudio. Sin embargo, también puede ocurrir que el tamaño muestral pequeño lleve a que el estudio no tenga potencia suficiente para rechazar la hipótesis nula. Si el sesgo de publicación interviene, el estudio no se publicará por sus resultados nulos. Los datos apoyan ese *continuum* hipotético.

En una revisión de 246 estudios publicados de tratamientos oncológicos, Berlin y Begg (28, 29) encontraron una asociación intensa entre el tamaño muestral y el efecto del tratamiento: los estudios con menor tamaño muestral mostraban mayor efecto del tratamiento. La tendencia es más intensa en los estudios aleatorizados comparados con los no aleatorizados. Esto implica que tienden a publicarse preferentemente los estudios de pequeño tamaño muestral con efecto grande, mientras que las investigaciones a gran escala tienden a publicarse independientemente del resultado.

Función de la fuente de financiación

Davidson (30) revisó 107 estudios publicados en 1984 y los clasificó según la dirección de sus resultados (a favor de la terapéutica nueva o a favor de la terapéutica estándar) y la fuente de financiación (compañías farmacéuticas o financiación "general"). Setenta y seis (71%) de los 107 estudios apoyaban la terapéutica nueva y 31 (29%) favorecían la terapéutica estándar. De los que favorecían la terapéutica nueva, 33 (43%) tenían financiación farmacéutica, mientras que de los que favorecían la terapéutica tradicional solo 4 (13%) tenían apoyo farmacéutico. Esto implica que la terapéutica nueva fue apoyada por un 89% (33/37) de los estudios financiados por la industria farmacéutica y por 61% de los estudios con financiación general. La proporción de artículos que favorecían la terapéutica nueva frente a la tradicional fue considerablemente distinta en las diferentes revistas evaluadas (cuadro 6). Davidson concluyó que "aunque parece improbable que haya conspiraciones para ocultar los resultados desfavorables de los ensayos clínicos, lo que sí parece ocurrir es una exclusión *de facto* de los resultados negativos". La posibilidad de una supresión conspirativa de resultados también ha sido sugerida (31) y refutada (32).

La posibilidad de que la financiación afecte la difusión de los resultados del estudio rebasa el campo de los ensayos clínicos (29). La preocupación por este tema y por otros aspectos relativos a posibles conflictos de intereses ha hecho que muchas revistas, incluida *JAMA*, exijan que se indiquen todos los apoyos financieros.

COMENTARIO

Quizá no sea muy conveniente desarrollar métodos para "curar" el problema del sesgo de publicación. La recopilación de datos de estudios no publicados, por ejemplo, exige un gran esfuerzo y puede que también tenga sesgos. Quizá ese u otros remedios (29) puedan proporcionar información adicional útil para quienes evalúan los trabajos científicos publicados, pero no son tratamientos muy buenos del sesgo de publicación.

Algo intermedio entre el tratamiento y la prevención es insistir en que la comunidad científica cambie sus prácticas. Los investigadores deben notificar los resultados de todos los estudios llevados a cabo. Los directores de revistas deben formalizar normas editoriales que establezcan que la decisión de publicar ha de basarse en la calidad y el razonamiento lógico de los autores y no en la dirección o magnitud

CUADRO 6. Dirección de los resultados de los ensayos clínicos publicados durante 1984 en varias revistas

Revista	No. (%)		Total (%)
	A favor del tratamiento nuevo	A favor del tratamiento estándar	
<i>Annals of Internal Medicine</i>	12 (86)	2 (14)	14 (100)
<i>Archives of Internal Medicine</i>	11 (79)	3 (21)	14 (100)
<i>Lancet</i>	30 (75)	10 (25)	40 (100)
<i>New England Journal of Medicine</i>	18 (67)	9 (33)	27 (100)
<i>American Journal of Medicine</i>	5 (42)	7 (58)	12 (100)

de los resultados del estudio (33). Aunque este es el abordaje más sencillo del sesgo de publicación, no es fácil que se ponga en práctica en todas partes.

La medida más eficaz para prevenir el sesgo de publicación es el registro de todos los ensayos clínicos, o quizá de todas las investigaciones que se emprenden. Ya existen registros en varios campos de estudio (34), principalmente la investigación perinatal (35), oncológica (36) y del SIDA (37). La Base de Datos de Ensayos Clínicos Perinatales de Oxford (35) es uno de los mejores y se ha utilizado como base de investigación metodológica y para cientos de metanálisis (38). El registro previo de todos los estudios es una tarea considerable, pero es el imperativo lógico de la era electrónica en la que vivimos. La prevención, no la cura, es la meta hacia la que hay que avanzar.

AGRADECIMIENTO

Este trabajo fue financiado en parte por el fondo R01 HS 05523 del Centro Nacional de Investigación en Servicios de Salud y Evaluación de la Tecnología Sanitaria, Rockville, Maryland.

Agradezco a Iain Chalmers, miembro del *Royal College of Obstetricians and Gynaecologists*, su lectura y sus muchos comentarios útiles a varios borradores de este trabajo.

REFERENCIAS

1. Smith ML. Publication bias and meta-analysis. *Evaluation Educ* 1980;4:22–24
2. Hall MB. In defense of experimental essay. En: *Robert Boyle on natural philosophy*. Bloomington: Indiana University Press; 1965:119–131.
3. Sterling TD. Publication decisions and their possible effects on inferences drawn from tests of significance- or vice versa. *J Am Stat Assoc* 1959;54:30–34.
4. Smart RG. The importance of negative results in psychological research. *Can Psychol* 1964;5: 225–232.
5. Greenwald AG. Consequences of prejudice against the null hypothesis. *Psychol Bull* 1975;82: 1–20.
6. White KR. The relation between socioeconomic status and academic achievement. *Psychol Bull* 1982;91:461–481.
7. Coursol A, Wagner EE. Effect of positive findings on submission and acceptance rates: a note on meta-analysis bias. *Professional Psychol Res Proc* 1986;17:136–137.
8. Smith ML. Sex bias in counseling and psychotherapy. *Psychol Bull* 1980;87:392–407.
9. Shadish WR, Doherty M, Montgomery LM. How many studies in the file drawer? an estimate from the family/marital psychotherapy literature. *Clin Psychol Rev* 1989;9:589–603.
10. Mahoney MJ. Publication prejudices: an experimental study of confirmatory bias in the peer review system. *Cog Ther Res* 1977;1:161–175
11. Begg CB, Berlin JA. Publication bias: a problem in interpreting medical data *J R Stat Soc Series A* 1988;151(pt3):418–463.
12. Maxwell C. Clinical trials, reviews, and the journal of negative results. *Br J Clin Pharmacol* 1981; 1:15–18.
13. Hetherington J, Dickersin K, Chalmers I, Meinert C. Retrospective and prospective identification of unpublished controlled trials: lessons from a survey of obstetricians and pediatricians. *Pediatrics* 1989;84:374–380.

14. Chalmers TC, Koff RS, Grady GF. A note on fatality in serum hepatitis. *Gastroenterology* 1965;49: 22–26.
15. Juhl E, Christensen E, Tygstrup N. The epidemiology of the gastrointestinal randomized clinical trial. *N Engl J Med* 1977;296:20–22.
16. Vandembroucke JP. Passive smoking and lung cancer: a publication bias? *Br J Med* 1988;296: 391–392.
17. Simes RJ. The case for an international registry of clinical trials. *J Clin Oncol* 1986;4:1529–1541.
18. Dickersin K, Chan S, Chalmers TC, Sacks HS, Smith H, Jr. Publication bias and clinical trials. *Controlled Clin Trials* 1987;8:343–353.
19. Sommer B. The file drawer effect and publication rates in menstrual cycle research. *Psychol Women Q* 1987;11:233–242.
20. Dudley HAF. Surgical research: master or servant. *Am J Surg* 1978;135:458–460.
21. Goldman L, Loscalzo A. Fate of cardiology research originally published in abstract form. *N Engl J Med* 1980;303:255–259.
22. Meranze J, Ellison N, Greenhow DE. Publications resulting from anesthesia meeting abstracts. *Anesth Analg* 1982;61:44–448.
23. McCormick MC, Holmes JH. Publications of research presented at the pediatric meetings. *AJDC* 1985;139:122–126.
24. Chalmers I, Adams M, Dickersin K, et al. A cohort study of summary reports of controlled trials. *JAMA* 1990;263:1401–1404.
25. "The editor regrets ..." [Editorial]. *Br Med J* 1980;280:508.
26. Minerva. [Views]. *Br Med J* 1983;287:1886.
27. Hemminki E. Study of information submitted by drug companies to licensing authorities. *Br Med J* 1980;280:833–836.
28. Berlin JA, Begg CB, Louis TA. An assessment of publication bias using a sample of published clinical trials. *J Am Stat Assoc* 1989;84:381–392.
29. Begg CB, Berlin JA. Publication bias and dissemination of clinical research. *JNCI* 1989;81: 107–115.
30. Davidson PA. Source of funding and outcome of clinical trials. *J Gen Intern Med* 1986;1:155–158.
31. Lauritsen K, Havelund T, Laursen LS, Rask-Madsen J. Withholding unfavourable results in drug company sponsored clinical trials. *Lancet* 1987;1:1091.
32. Nicholson PA. Information for drug trial participants. *Lancet* 1987;2:396.
33. Angell M. Negative studies. *N Engl J Med* 1989;321:464–466.
34. Dickersin K. Report from the panel on the case for registers of clinical trials. *Controlled Clin Trials* 1988;9:76–81.
35. Chalmers I, Hetherington J, Newdick M, et al. The Oxford Database of Perinatal Trials: developing a register of published reports of controlled trials. *Controlled Clin Trials* 1986;7:306–324.
36. Hubbard SM, Henney JE, DeVita, VT Jr. A computer data base for information on cancer treatment. *N Engl J Med* 1987;316:315–318.
37. Dutcher GA. *NLM Technical Bull.* Washington, DC: US Dept of Health and Human Services; 1989;243:17–27.
38. Chalmers I, Enkin M, Kierse M, eds. *Effective care in pregnancy and childbirth.* Oxford, England: Oxford University Press; 1989.