

## Sesgos de publicación, valor de la información y su efecto en las políticas de salud

### Publication bias, value of information and their effects on health policy decision making

**Dr. C. Carlos Campillo-Artero**

Servicio de Salud de las Islas Baleares. Palma de Mallorca, España.

---

#### RESUMEN

Son notables las deficiencias de la información en que se fundamentan las decisiones en salud pública y servicios de salud. El objetivo de este artículo es sintetizar el conocimiento de algunas de sus causas y consecuencias. Entre las limitaciones que más afectan la toma de decisiones se encuentran la baja calidad y la insuficiencia de la información científica disponible. Esta última consiste en la notificación incorrecta de resultados de los estudios, y abarca manipulaciones de variables primarias y secundarias, abuso de variables sustitutivas y compuestas, omisión *ex profeso* de la notificación de algunos de resultados y sesgos de publicación. Cuanto menor es el valor de la información disponible para la toma de decisiones, tanto mayor es el costo en oportunidad de estas últimas, que se cifra en pérdidas de beneficios en salud por retraso en incorporar intervenciones o tecnologías potencialmente efectivas y eficientes, por falta de pruebas científicas suficientes, o en yatrogenia e ineficiencia por su adopción sin suficiente respaldo científico de su eficacia, efectividad o eficiencia. Medidas incluidas en los modelos de ciencia regulatoria, como son el registro público de intervenciones en salud pública y ensayos clínicos, la sustitución de pruebas de hipótesis por pruebas de estimación y métodos bayesianos, y la declaración de conflictos de interés, son las principales soluciones propuestas. Los políticos, decisores y profesionales de la salud pública y de la clínica no deben escatimar esfuerzos ni demorar más la implantación y consolidación de las soluciones propuestas para reducir las limitaciones descritas. De su efectividad dependerán a la postre, la legitimidad y el sustento científicos de la toma de decisiones.

**Palabras clave:** toma de decisiones, políticas de salud, valor de la información, sesgos.

---

## ABSTRACT

Inadequate information used to support decision making in public health and health services is remarkable. The objective of this paper was to provide a synthesis of some of its causes and effects. Many studies show that two of the main limitations that affect the decision making are low quality and inadequacy of available scientific information. The latter consists of deficient reporting of study results. It encompasses manipulation of primary and secondary variables, abuse of surrogate and composite variables, intentional omission of outcome reporting and publication biases. The lower the value of available information for decision making, the higher the opportunity cost of decisions. This cost can be translated into health benefits lost due to delayed adoption of potentially effective and efficient interventions and medical technologies, or adverse health effects and inefficiencies due to their early diffusion in the absence of sufficient scientific evidence of their efficacy, effectiveness or efficiency. Main solutions put forward are regulatory science measures, public registry of interventions in public health and clinical trials, use of estimation and Bayesian statistics instead of hypothesis testing, and declaration of conflicts of interests. Policy makers, public health professionals and clinicians should neither spare effort nor further delay the implementation and consolidation of these long known measures intended to minimize those limitations. Legitimacy and the level of scientific evidence of decision making will eventually depend on the effectiveness of those solutions.

**Key words:** decision making, health policy, value of information, biases.

---

## INTRODUCCIÓN

Las deficiencias de la información en que se sustenta la toma de decisiones en salud pública y en servicios de salud, así como sus consecuencias, han sido una de las constantes preocupaciones que en las tres últimas décadas se han expresado al diseñar, implantar y evaluar políticas de salud.

La diversidad de la naturaleza y las causas de dichas deficiencias, las numerosas esferas afectadas y el tipo y calado de sus efectos son tan vastos que trascienden con creces el alcance de cualquier estudio o aproximación analítica individual. El conocimiento disponible sobre las limitaciones que afronta la toma de decisiones en los sistemas de salud proviene de un aluvión de heterogéneos debates e investigaciones, la mayoría de los cuales abordan aisladamente alguna de sus muchas facetas.<sup>1-5</sup>

Entre las limitaciones más prominentes -porque atañen a la raíz del problema- destacan la baja calidad y la insuficiencia de la información científica en que se fundamentan las decisiones en políticas de salud. Durante los últimos diez años se ha acumulado un cuerpo de conocimiento cabal de la notable magnitud de este problema -lo que se ha dado en llamar bajo *nivel de evidencia*- y de sus efectos en la toma de

---

decisiones en el ámbito de la salud. Es ingente la cantidad de estudios que han puesto de manifiesto limitaciones graves en la notificación de resultados de estudios dirigidos a estimar la viabilidad, eficacia, efectividad o la eficiencia de intervenciones, proyectos, programas, políticas de salud y tecnologías médicas (medicamentos, aparatos).<sup>1-8</sup>

Las pruebas científicas que se necesitan para informar, orientar y evaluar políticas e intervenciones de salud pública que abordan temas tales como los determinantes de salud, la promoción de la salud y la prevención de enfermedades, las relaciones entre salud y medio ambiente, salud y desigualdades sociales, política fiscal y salud o la efectividad de políticas intersectoriales, son complejas en esencia y provienen de muy diversas fuentes y tipos de estudios.

Lo mismo ocurre con las pruebas científicas que informan las decisiones regulatorias vinculadas, por ejemplo, con los esquemas de aprobación de la comercialización de medicamentos, dispositivos y aparatos médicos, la cobertura y el aseguramiento, la fijación de precios, los mecanismos de control y vigilancia poscomercialización y la reinversión, así como con las que fundamentan las políticas de organización y gestión de sistemas y servicios de salud, tanto en los niveles macro, meso y micro, como en las esferas de la atención primaria y especializada.<sup>9,10</sup>

A la postre, como indican los análisis del valor de la información (*value of information analysis*), cuanto menor es el *valor* de la información disponible para tomar decisiones, tanto mayor es el costo en oportunidad de estas últimas. Su importancia se comprueba de forma palmaria cuando, por ejemplo, dicho costo se expresa como

- Pérdida de beneficios en salud por el retraso en incorporar intervenciones o tecnologías potencialmente efectivas por falta de pruebas científicas suficientes.
- Yatrogenia o reducción de la seguridad clínica (morbilidad y mortalidad evitables).
- Ineficiencia (dilapidación de recursos invertidos en salud, incorporación de intervenciones o tecnologías con bajas razones de costo-efectividad incremental) resultante de su adopción o incorporación prematura, esto es, sin suficiente respaldo científico de su eficacia, efectividad o eficiencia.<sup>2,3</sup>

El objetivo de este artículo se centra en sintetizar el conocimiento de algunas de las causas y consecuencias de las limitaciones en la información científica disponible para la toma de decisiones en salud pública y en los servicios de salud, a escalas macro, meso y micro y enumerar las soluciones propuestas para mitigar sus efectos.

## **ALGUNAS PECULIARIDADES DE LAS EVALUACIONES DE INTERVENCIONES EN SALUD PÚBLICA**

La búsqueda, la selección, el análisis y la síntesis de información científica para diseñar y evaluar políticas y sustentar la toma de decisiones ponen de relieve marcadas diferencias entre dos de sus grandes áreas de aplicación: la de las intervenciones en salud pública y la de las tecnologías médicas en los servicios de salud.

Son notorias las características diferenciales de la información procedente de estudios experimentales y observacionales en investigación clínica. Pero en salud pública son escasas las intervenciones que, como la efectividad de la vacunación, pueden evaluarse mediante ensayos clínicos aleatorizados y controlados, porque en ellas la unidad de intervención no es el individuo sino grupos de personas o poblaciones. Ejemplos clásicos son los programas de educación y promoción de la salud en escuelas, las campañas para prevenir accidentes de tráfico, los proyectos de fluoración del agua o para yodar la sal, la construcción de zonas verdes o la

promoción del uso de la bicicleta en zonas urbanas. También lo ejemplifican las intervenciones que se articulan a través de cambios en las políticas fiscales, como las dirigidas a aumentar la carga impositiva de los alimentos ricos en sal y en grasas o incluso su prohibición, como ha ocurrido recientemente en Dinamarca, o las reguladas por medio de leyes, como la que prohíbe fumar en espacios públicos.<sup>1,4,11</sup>

En la esfera de la salud pública, los diseños de las evaluaciones de la efectividad y eficiencia de las intervenciones están determinados en buena medida por la diana de la intervención: los proyectos, dirigidos a individuos (vacunación, quimioprevención, dejación del hábito tabáquico), los programas, destinados a grupos de población de mayor o menor tamaño (cribado de enfermedades), y las políticas, que engloban a la población general (leyes, regulaciones, estrategias). En los primeros pueden aleatorizarse individuos y en los segundos, grupos de individuos, centros asistenciales, unidades territoriales y otros. En las políticas, la aleatorización resulta imposible porque la intervención abarca a toda la población; además, no suele disponerse de grupo de no expuestos. Por ello -y aunque no abundan las evaluaciones-, si se intenta alcanzar alto grado de rigor, se recomienda evaluar los proyectos con ensayos aleatorizados y controlados o, en su defecto, con estudios observacionales; los programas, mediante ensayos aleatorizados de conglomerados (*cluster randomized trials*) y las políticas, con diseños como, por ejemplo, los *cuasi* experimentales pre-post o series temporales interrumpidas.<sup>1,4</sup>

La marcada complejidad de muchas intervenciones en salud pública y las diferencias consignadas entre los diseños empleados, su amplio abanico temático, las numerosas áreas de intervención afectadas (por ej., los proyectos intersectoriales) y las diferentes disciplinas participantes en las evaluaciones (epidemiología, ciencia política, economía, urbanismo, sociología) explican que la información necesaria para la toma de decisiones se encuentre dispersa en bastantes bases de datos y de revistas electrónicas. Estos factores dan cuenta, además, de que la cifra de revisiones sistemáticas de intervenciones en salud pública sea mucho menor que la de las realizadas para estimar la eficacia o efectividad de medicamentos, aparatos y otras tecnologías médicas y de que sean ocasionales las revisiones elaboradas por agencias de evaluación de tecnologías médicas sobre los efectos de políticas o programas en salud pública.<sup>1,4,5,11</sup>

Además, y a diferencia de las evaluaciones de las intervenciones médicas en individuos, en la evaluación de intervenciones en salud pública es difícil discernir entre los efectos atribuibles a las propias intervenciones de los atribuibles a factores del contexto social, económico y político en que estas se despliegan y a las interacciones entre unos y otros.<sup>1,11</sup>

A este entramado de factores cabe añadir que numerosas intervenciones incluyen modificaciones del comportamiento (actividad física, alimentación), lo cual complica más aún la evaluación, y que la heterogeneidad de las poblaciones o grupos estudiados acostumbra ser mucho mayores que en los estudios clínicos. El riesgo de introducir sesgos y de no controlar adecuadamente numerosos factores de confusión potenciales cobra mayor relevancia en ellas, como mayor dificultad entraña acotar el marco de extrapolación de resultados a otros grupos de población. No pocas intervenciones surten efecto a medio y largo plazo. Tampoco escasean las que terminan antes de su fecha de finalización esperada por motivos de índole política o por recortes presupuestarios.<sup>1,8,11</sup>

### **Notificación deficiente de resultados: tres causas frecuentes**

Una de las limitaciones del valor de la información científica disponible para la toma de decisiones en salud pública y servicios de salud responde a la notificación deficiente de resultados de los estudios, que, con fines didácticos, puede encuadrarse en tres grandes grupos. El primero comprende una amalgama de manipulaciones de

variables: las relativas a las variables primarias y secundarias de los estudios, el uso de variables sustitutivas (*surrogate variables*) y la síntesis de resultados mediante variables compuestas (*composite variables*). El segundo corresponde a la notificación selectiva de resultados de los estudios o informes, es decir, la omisión *ex profeso* de la notificación de algunos de sus resultados (*selective outcome, reporting outcome, reporting bias*). El tercero son los sesgos de publicación (*publication bias*), que engloba el conjunto de estudios que no se publican porque sus resultados no alcanzan significación estadística o por la baja magnitud de los efectos estimados. En un alto porcentaje de casos se explica por los intereses de sus autores, patrocinadores o financiadores o porque los investigadores confunden el significado de las pruebas de hipótesis a resultados de una formación estadística deficiente (*publication bias*).<sup>12-19</sup>

### Manipulación de variables

Tres son los más frecuentes tipos de manipulación de variables encontrados en estudios y evaluaciones clínicas y de salud pública. Todos traducen el intento de demostrar, mediante dicha manipulación, la superioridad absoluta o relativa -mayor eficacia, efectividad o eficiencia absolutas, relativas o incrementales- de una intervención frente a la alternativa de comparación escogida.

La primera consiste en la sustitución de la variable definida como primaria en el proyecto de investigación por alguna de las secundarias. Cuando la magnitud del efecto de la intervención conforme a la variable protocolizada como primaria es insuficiente para demostrar dicha superioridad, en los resultados del estudio se notifica alguna de las secundarias que sí muestra dicha superioridad y por ello se presenta como primaria en el informe final.<sup>19</sup>

La segunda manipulación son las demostraciones de la superioridad de una intervención mediante variables de proceso o intermedias (*surrogate variables*) en la cadena causal que media entre la exposición y el resultado final. Son ejemplos frecuentes el uso de la carga viral, la reducción del tamaño de un tumor, el porcentaje de los que entran en remisión completa tras un ciclo de tratamiento, el incremento de la densidad ósea medido por densitometría, el porcentaje de empresas adheridas a un programa de reducción del contenido calórico de los alimentos, el aumento de la superficie dedicada a zonas verdes y carriles de bicicleta o el descenso en la venta de cajetillas de cigarrillos en sustitución, respectivamente, de la mortalidad por sida, la supervivencia en cáncer, la reducción de la incidencia de fracturas, y la disminución de la incidencia de enfermedades cardiovasculares o de tumores asociados con el sedentarismo o el tabaco. El problema estriba en que a menudo las variables sustitutivas utilizadas no son buenas predictoras de la variable final o se desconoce si lo son. Recuérdese, por ejemplo, que no fue sino tras años después de la comercialización de las estatinas que se demostró la asociación entre hipercolesterolemia y aumento del riesgo de enfermedades cardiovasculares\*.<sup>11-14</sup>

En la tercera, dicha superioridad se intenta demostrar por medio de una variable compuesta (*composite variables*), esto es, una variable que resulta de la combinación de *n* variables individuales.<sup>13</sup> Esta manipulación puede ser legítima desde un prisma científico, como las variables compuestas del modelo del buen gobierno (*governance*) de Kaufmann y otros, del Banco Mundial o numerosos índices compuestos usados para estimar riesgos de enfermedades, de hospitalización o pronósticos.<sup>21,22</sup> Pero, cada vez se utilizan con más asiduidad cuando las variables de resultado principales (por ej., *hazard ratio* de mortalidad, riesgo relativo de infarto en pacientes con coronariopatía estable tras la colocación de un *stent* comparados con los tratados con medicamentos), no muestran la superioridad de la nueva alternativa, que sí señala, por el contrario, la variable compuesta por esa variable principal combinada con otras

(menos relevantes clínicamente, como la reducción del porcentaje de readmisiones o de la incidencia de infecciones nosocomiales o de la estancia media). Por fortuna, aumenta progresivamente el número de revistas que obligan a notificar los resultados de la variable compuesta junto con los de las que la integran.<sup>13,20</sup>

### **Notificación selectiva de resultados**

El sesgo de notificación de resultados se introduce cuando los hallazgos de un estudio se notifican selectivamente en función de la magnitud del efecto detectado, la significación estadística alcanzada o los intereses de los financiadores o investigadores del estudio. El conocimiento de sus causas y consecuencias se ha puesto de relieve merced a análisis pormenorizados de revisiones sistemáticas de intervenciones en salud pública, estudios clinicoepidemiológicos y evaluaciones de tecnologías, incluidas las evaluaciones económicas. En ellas se ha comprobado que su valoración se omite con frecuencia en revisiones de intervenciones en salud pública, quizás por la amplia variedad de diseños utilizados, la complejidad de dichas intervenciones y la marcada heterogeneidad en la notificación de sus resultados. Lo mismo se deduce de las revisiones sistemáticas de ensayos clínicos con tratamientos farmacológicos.<sup>11,14,23,24</sup>

### **Sesgos de publicación**

La tercera causa de las deficiencias en la notificación de resultados de estudios para informar decisiones son los sesgos de publicación, la forma más extrema del sesgo de notificación de resultados. En este caso, los estudios en los cuales no se sobrepasa cierta magnitud del efecto estimado, no se alcanza significación estadística, por la concurrencia de intereses de los financiadores o por ignorancia, no se publican.<sup>15,16</sup>

Se distinguen tres tipos: sesgo de prepublicación, aquel que introducen los investigadores al decidir no publicar estudios en los cuales no se han encontrado diferencias estadísticamente significativas, o los promotores de la tecnología evaluada cuando no se demuestra que esta es eficaz. (La publicación de ensayos de no inferioridad no excluye este sesgo, porque su objetivo es demostrar equivalencia terapéutica.) El de publicación, cuando la decisión de no publicar esos estudios la toman los editores de revistas médicas. Y el de pospublicación, cuando, por las mismas razones, esos estudios se excluyen de las revisiones sistemáticas, los metanálisis u otros informes de síntesis.<sup>16,23</sup>

El sesgo de publicación se ha estudiado en profundidad en ciencias de la salud. Desde que se acuñó el término en los cincuenta del pasado siglo hasta la actualidad, se han descrito sus consecuencias sobre todo en el contexto de las revisiones sistemáticas en una amplia gama de especialidades médicas. Asciende la cifra de revisiones que incluyen estimaciones de su existencia y magnitud obtenidas con métodos gráficos y estadísticos. La creciente realización de metanálisis ha galvanizado aún más su estudio. Las guías más aceptadas a escala internacional para su realización y revisión han revelado que cuando la magnitud de estos sesgos es notable aumenta el riesgo de que en los metanálisis se acaben notificando intervalos de confianza estrechos de estimaciones puntuales espurias.<sup>1,4,15,16,25,26</sup>

Lamentablemente, en varias ocasiones se han comprobado los efectos deletéreos reales de los sesgos de publicación. Quizá uno de los casos más flagrantes y difundidos haya sido el de los inhibidores selectivos de la recaptación de serotonina para el tratamiento de la depresión infantil: exceptuando la fluoxetina y contraviniendo las conclusiones de los estudios publicados, los metanálisis de estudios publicados y no publicados no repaldan la eficacia de sertralina y paroxetina y revelan



que el riesgo de suicidio en niños tratados sobrepasa con creces sus escasos beneficios. Este resultado es congruente con los del metanálisis realizados con citalopram y venlafaxina.<sup>27-29</sup>

La falta de significación estadística es la causa más frecuente de subnotificación de estudios de intervenciones en salud pública. En ella, tanto la notificación incompleta como el sesgo de publicación reducen el *nivel de evidencia* de las pruebas científicas aportadas, contribuyen a sobrestimar la magnitud de los efectos notificados y menoscaban la difusión y la influencia de los estudios realizados en la toma de decisiones, así como el sustento científico de las políticas.<sup>1,4,15,16</sup>

Las soluciones propuestas para mitigar los efectos de estos sesgos se resumen en las siguientes medidas incluidas en los modelos de ciencia regulatoria (*regulatory science*, es la regulación basada en pruebas científicas): la inclusión obligatoria de todos los ensayos clínicos y evaluaciones de intervenciones de salud pública en registros públicos, incluidos sus proyectos de investigación aprobados, la mejora de la educación en estadística, la sustitución de las pruebas de hipótesis por pruebas de estimación, el fomento de la estadística bayesiana, y la declaración de los conflictos de intereses y la denuncia y notificación pública de los que se detecten.<sup>1,4,6,7,24-27</sup>

### **Un sesgo de publicación a menudo olvidado**

En los sistemas de salud de algunos países, la evaluación económica es uno de los criterios decisorios de financiación y cobertura de intervenciones en salud pública y nuevas tecnologías médicas. Su aplicación se resume en la fijación de umbrales de costo efectividad (costo/AVAC), como expresión de la predisposición social a pagar y del costo en oportunidad. Salvo excepciones explícitas, las tecnologías cuya razón de costo efectividad, es igual o menor que el umbral acordado se financian; aquellas por encima del umbral, se excluyen de la financiación.<sup>30-33</sup>

En algunos países, las instituciones y empresas privadas son con mucho la principal fuente de financiación de las evaluaciones económicas.<sup>32</sup> Se ha demostrado la existencia de sesgos de publicación de estas evaluaciones: es más probable que aquellas que estiman razones de costo efectividad, mayores que el umbral fijado, no se publiquen. Se tiende a sobrestimar la eficiencia así expresada de no pocas tecnologías (sobre todo de medicamentos y programas en salud pública), bien sea como consecuencia de dicho sesgo o a resultas de la exclusión selectiva en los análisis de sensibilidad de evaluaciones publicadas de aquellos supuestos o subgrupos de pacientes o de poblaciones en los cuales las razones de costo efectividad estimadas, superan el umbral. Su efecto en las decisiones de financiación y cobertura puede ser directo, el costo de oportunidad de las decisiones, elevado, y las consecuencias en el gasto observado (evitable) y en la eficiencia social alcanzada, notable.

### **El caso de las intervenciones preventivas**

El volumen y la calidad de la información disponible para elaborar políticas y tomar decisiones sobre intervenciones preventivas, siguen siendo menores que los de las terapéuticas. Se estima que en los países de la OCDE el gasto público en salud dedicado a prevención no alcanza el 5 %. Las ventajas competitivas de las intervenciones curativas y el plano relegado que ocupan las preventivas se explican por diversos motivos: la prevención carece de beneficiarios identificables y se caracteriza por costos inmediatos y beneficios a medio y largo plazo, junto con la consecuente percepción de bajos retornos de inversión. Pocos estudios evalúan intervenciones complejas en prevención y promoción de salud de alcance poblacional,

la mayoría de las publicadas tienen como unidad de análisis individuos, no poblaciones. Los diseños utilizados son muy variados, como su calidad científica. Comparados con los de las intervenciones curativas, también es reducida la cifra de estudios que evalúan con rigor políticas e intervenciones preventivas multifacéticas.<sup>1,4,8,9,11,24,25,30,31</sup>

Por estos motivos y por limitaciones metodológicas específicas, la realización y publicación de evaluaciones económicas de las intervenciones preventivas también son menos frecuentes que las de las curativas. A estos elementos deben añadirse el menor número de revisiones sistemáticas sobre su efectividad y eficiencia, junto con los sesgos de publicación, sobre todo a expensas de las modalidades de pre y pospublicación.

Se conocen las causas, la magnitud y las consecuencias de las deficiencias de las pruebas científicas que informan la elaboración de políticas y la toma de decisiones en los planos de la macro, meso y microgestión, tanto en el ámbito de la salud pública como en el de los servicios de salud. El reto actual estriba en no escatimar esfuerzos ni demorar más la implantación y consolidación de las medidas propuestas hace tiempo para reducir las limitaciones descritas. De su efectividad dependerá a la postre la legitimidad y el sustento científicos de las políticas de salud.

## Agradecimiento

Agradezco a *Gaspar Tamborero* sus valiosos comentarios sobre el borrador de este artículo y su incondicional apoyo.

## REFERENCIAS BIBLIOGRÁFICAS

1. Faggiano F, Vigna-Taglianti F. Systematic reviews of effectiveness of Public Health practice. *Ital J Public Health*. 2006;3:29-33.
2. Eckermann S, Willan AR. Expected value of information and decision making in HTA. *Health Econ*. 2007;16:195-209.
3. Griffin SC, Claxton KP, Palmer SJ, Sculpher MJ. Dangerous omissions: the consequences of ignoring decision uncertainty. *Health Econ*. 2011;20:212-24.
4. Pearson M, Peters J. Outcome reporting bias in evaluations of public health interventions: evidence of impact and the potential role of a study register. *J Epidemiol Community Health*. 2012;66:286-9.
5. Institute of Medicine. For the Public's Health: Revitalizing law and policy to meet new challenges. Washington, D. C.: The National Academies Press; 2011.
6. Institute of Medicine. Building a national framework for the establishment of regulatory science for drug development: Workshop summary. Washington, D. C.: The National Academies Press; 2010.



7. Institute of Medicine. Strengthening a workforce for innovative regulatory science in therapeutics development: Workshop summary. Washington, D. C.: the National Academies Press; 2011.
8. Campillo C. When health technologies do not reach their effectiveness potential: A health services research perspective. *Health Policy*. 2012;104:92-8.
9. Centre for health economics. Value-based pricing for pharmaceuticals: Its role, specification and prospects in a newly devolved NHS. York: University of York; 2011.
10. Claxton K, Longo R, Longworth L, McCabe C, Wailoo A. The value of innovation. University of York (Centre for Health Economics, , Leeds University, School of Health and Related Research, University of Sheffield). York: University of York; 2009.
11. Ordúñez P, Campillo C, editores. Organización Panamericana de la Salud. Consulta regional: prioridades para la salud cardiovascular en las Américas. Mensajes claves para los decisores. Washington, D. C.: OPS; 2011.
12. Berger VW. Does the Prentice criterion validate surrogate endpoints? *Stat Med*. 2004;23:1571-8.
13. Kerche la Cour J, Brok J, Gotzsche P. Inconsistent reporting of surrogate outcomes in randomized clinical trials: cohort study. *BMJ*. 2010;341:c3653. doi: 10.1136/bmj.c3653.
14. Montori VM, Permyer-Miranda G, Ferreira-González I. Validity of composite end points in clinical trials. *BMJ*. 2005;330:594-6.
15. Dickersin K. The existence of publication bias and risk factors for its occurrence. *JAMA*. 1990;263:1385-9.
16. Chalmers TC, Frank CS, Reitman D. Minimizing the three stages of publication bias. *JAMA*. 1990;263:1392-5.
17. Ivers NM, Taljaard M, Dixon S, et al. Impact of CONSORT extension for cluster randomized trials on quality of reporting and study methodology: review of random sample of 300 trials, 2000-8. *BMJ*. 2011;343:d5886. doi: 10.1136/bmj.d5886.
18. Murray DM, Varnell SP, Blitstein JL. Design and analysis of group-randomized trials: a review of recent methodological developments. *Am J Public Health*. 2004;94:423-32.
19. Mathieu S, Boutron I, Moher, D. Comparison of registered and published primary outcomes in randomized controlled trials. *JAMA*. 2009;302:977-84.
20. Ocana A, Tannock IF. When are "positive" clinical trials in oncology truly positive? *J Natl Cancer Inst*. 2011;103:16-20.
21. Kaufmann D, Kraay A, Mastruzzi M. Governance Matters VIII. Aggregate and individual governance indicators, 1996-2008. Policy Working Paper 4978. Washington, D.C.: The World Bank; 2009.

22. Meneu R. Evaluación del buen gobierno sanitario. Algunas recomendaciones para su mejora. En: Ortún V, Director. El buen gobierno sanitario. Madrid: Springer Healthcare Communications; 2009: 11-25.
23. Higgins JPT, Green S, eds. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0 [Internet]. The Cochrane Collaboration; 2011[updated March 2011]. Available from: <http://www.cochrane-handbook.org>
24. Waters E, Priest N, Armstrong R. The role of a prospective public health intervention study register in building public health evidence: proposal for content and use. J Public Health. 2007;29:322-7.
25. Sterling TD. Publication decisions and their possible effects on inferences drawn from tests of significance and viceversa. J Am Stat Assoc. 1959;54:30-4.
26. Bero LA, Glantz SA, Rennie D. Publication bias and public health policy on environmental tobacco smoke. JAMA. 1994;272:133-6.
27. Herxheimer A. Antidepressants and adverse effects in young patients: uncovering the evidence. CMAJ. 2004;17:487-9.
28. Depressing research [editorial]. Lancet. 2004;363:1335.
29. Whittington CJ, Kendall T, Fonagy P, Cottrell D, Cotgrove A, Boddington E. Selective serotonin reuptake inhibitors in childhood depression: systematic review of public versus unpublished data. Lancet. 2004;363:1341-5.
30. Schwappach DLB, Boluarte TA, Suhrcke M. The economics of primary prevention of cardiovascular disease - a systematic review of economic evaluations. Cost Effectiveness Resource Allocation. 2007;5:5. doi: 10.1186/1478-7547-5-5.
31. Bell CM, Urbach DR, Ray JG. Bias in published cost-effectiveness studies: systematic review. BMJ. 2006;332:699-701.
32. Catalá-Lopez F, Alonso-Arroyo A, Aleixandre-Benavent R. Coauthorship and institutional collaborations on cost-effectiveness analyses: a systematic network analysis. PLoS ONE. 2012;7(5):e38012. doi:10.1371/journal.pone.0038012.
33. Hillman AL, Eisenberg JM, Pauly MV. Avoiding bias in the conduct and reporting of cost-effectiveness research sponsored by pharmaceutical companies. NEJM. 1991;324:1362-5.

\* Muy ilustrativos son los siguientes hechos comprobados en estudios en oncología: la medida de eficacia frecuentemente no refleja un beneficio neto (como la supervivencia global) ni incremental (comparado con el de la mejor alternativa disponible), ni una mejora de la calidad de vida durante el periodo de aumento de la supervivencia; por el contrario, suelen usarse variables sustitutivas (como la reducción del tamaño tumoral, la progresión libre de enfermedad, el porcentaje de pacientes que entran en remisión completa) cuando no siempre son buenos predictores de la variable de resultado principal; las magnitudes de las mínimas diferencias clínicamente importantes (deltas) detectadas y publicadas son menores que las especificadas en los protocolos (las fijadas por la industria a veces no son las

clínicamente importantes sino las que es más factible detectar); decisiones de adopción de tratamientos sobre la exclusiva base de valores  $p$ , y el uso de biomarcadores de respuesta al tratamiento y de pronóstico insuficientemente validados (cuya profusión actual es preocupante).<sup>20</sup> (La tendencia a maximizar el uso de pruebas con alta sensibilidad y baja especificidad aumenta la actividad y los costes evitables).

*Carlos Campillo-Artero*. Servicio de Salud de las Islas Baleares. Palma de Mallorca, España. Correo electrónico: [carlos.campillo@ibsalut.es](mailto:carlos.campillo@ibsalut.es)