

15.1 Evidencia científica y gestión del conocimiento



La relación entre conocimiento científico y su aplicación en el ámbito de la salud es compleja pero inobjetable: se trata de identificar los principales problemas y oportunidades para mejorar la fundamentación científica de las decisiones y la reorientación del conocimiento hacia los retos de la atención sanitaria y la salud pública.

Autor: Carlos Campillo-Artero

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

Se recomienda imprimir 2 páginas por hoja

Citación recomendada:

Campillo-Artero, C. Evidencia científica y gestión del conocimiento [Internet]. Madrid: Escuela Nacional de Sanidad; 2014 [consultado día mes año]. Tema 15.1. Disponible en: direccion url del pdf.



TEXTOS DE ADMINISTRACIÓN SANITARIA Y GESTIÓN CLÍNICA
by UNED Y ESCUELA NACIONAL DE SANIDAD
is licensed under a Creative Commons
Reconocimiento- No comercial-Sin obra Derivada
3.0 Unported License.



Resumen:

Es compleja la traslación de información y conocimiento desde los planos teórico y experimental a la cabecera del enfermo o a las dependencias donde se gestan la macro, meso y microgestión.

La traslación es el proceso dual que consiste en aplicar descubrimientos de la investigación básica a problemas de salud en condiciones controladas (investigación clínica) y la aplicación de resultados prometedores de dicha fase a grupos de población

en condiciones controladas o no controladas.

Nuestra avidez por la certeza y reticencia a la incertidumbre nos

Introducción

- 1. Renuencia y temor a la incertidumbre*
- 2. A veces el pensamiento y el comportamiento pueden embaucarnos*
- 3. Cuando lo observado no alcanza lo esperado*
- 4. Deficiencias en la generación de información científica*
- 5. Los problemas de evaluar y de decidir*
- 6. Y la salud pública*
- 7. Algunas soluciones puestas en marcha*
- 8. Ciencia regulatoria y ciencia traslacional: ¿se alcanzará un acuerdo?*

Conclusiones

Referencias bibliográficas

impelen a concluir con fuerza; el problema es que obrando de este modo, soslayamos las debilidades de los supuestos de partida y sacrificamos la credibilidad científica.

En relación con el pensamiento de los decisores, es importante reconocer la existencia de los sistemas de pensamiento automático-intuitivo por ser muy habituales y por ser susceptibles a sesgos: deberían ponerse al servicio de la racionalidad en las políticas y actuaciones.

Con relación al "gap" o abismo entre la eficacia experimental y la efectividad finalmente conseguida, hay muchos factores que contribuyen a dificultar la intensidad y forma en la que los efectos esperados desde el conocimiento no se traducen en la práctica: sobreestimación de la eficacia,

no actuación sobre comportamientos, fallo de gestión del cambio, no abordaje de facetas múltiples del problema en la implantación, intereses, acción prematura o tardía...etc.

Pero, además, se produce una tendencia a manipular los estudios para apoyar conclusiones que favorezcan determinados cursos de acción: así, la substitución de variable primaria por secundarias; el uso de variable intermedia en vez de la medición de impacto, y el uso de variable compuesta (integrando muchas variables) que disuelve el efecto de la variable de impacto.

Junto a lo anterior, los sesgos de notificación (hallazgos comunicados selectivamente) o los de publicación (investigadores, editores o revisores) completan el elenco de barreras a la validez interna y externa.

Se trataría de aplicar en la evaluación análisis de sensibilidad para identificar el efecto de las variables sometidas a incertidumbre, identificar el efecto en subgrupos de población y de proveedores, no aceptar como dogma los niveles de jerarquía de evidencia, y favorecer que los decisores asuman su función sintética, con juicios condicionados y basados en la totalidad de las pruebas científicas existentes, ajustadas al contexto.

En la salud pública hay dificultades metodológicas por las características de las intervenciones, el tiempo dilatado hasta el impacto y otros elementos contextuales.

Analizar la efectividad comparada de dos intervenciones o políticas es una opción prometedora, pero que está afrontando obstáculos importantes (medios, trabas, incentivos, y limitaciones de una actividad a escala artesanal).

El reto de mejorar el complejo, turbulento y cambiante proceso de traslación del conocimiento a la acción puede implicar el fomento de una ciencia regulatoria que facilite la racionalidad de las decisiones colectivas y clínicas.

Nota al pie: *Pasajes de los apartados 5 y 6 de este artículo proceden de la referencia 50.*

Introducción

Puede provocar fatiga hablar de nuevo de traslación de información y conocimiento desde los planos teórico y experimental a la cabecera del enfermo o a las dependencias donde se gestan la macro, meso y microgestión.

Es terreno muy trillado y una actividad ineficiente a un tiempo. **El cociente entre la suma de debates y páginas impresas y los avances reales en traslación es muy alto.** Se han vertido inmensos caudales de tinta y debatido hasta los confines del hartazgo sobre ello y los adelantos siguen discurriendo en

El cociente entre la suma de debates y páginas impresas y los avances reales en traslación es muy alto.

el plano marginal. Traslación y transferencia de conocimiento incluso pueden considerarse haptenos selectivos responsables de reacciones de hipersensibilidad, toda vez que su mención ya provoca en algunos rechazo, irritación o indolencia inmediatos.

Ofrecer una explicación cabal y veraz de la lenta traslación de conocimiento a la práctica o, en uno de sus casos particulares, de la escasa reducción alcanzada hasta la fecha del hiato entre eficacia y efectividad, sigue siendo virtualmente imposible. Cualquier lista que se elabore para enumerar sus causas es muy prolija. Cualquier escrutinio de las razones esgrimidas para dilucidar lo que un ocurrente investigador ha dado en llamar *Perdidos en traslación* (*Lost in translation*) topa con una enmarañada madeja de razones y argumentarios infumables (1).

Más aún, dependiendo del área de conocimiento, el entramado de explicaciones se reorganiza o, en alguna, se transforma por completo. Ocurre, por ejemplo, si se compara el análisis de la traslación de conocimiento desde la investigación oncológica a la cabecera del paciente con dicha traslación desde la teoría de las organizaciones o el análisis de la eficiencia a los efectos de los cambios en la gestión de centros asistenciales o de los recortes presupuestarios en la calidad asistencial. Se observa también al hacerlo desde la teoría económica de los incentivos o de las formas de pago a la mejora palmaria del rendimiento profesional y los resultados asistenciales derivados. Sobra casuística para ejemplificarlo.

Huyendo de la ingenuidad, aquí no se intenta dar la explicación. Rechazando escrutinios plomizos, tampoco se ofrece el enésimo análisis que por erudito no deje de ser estéril o, cuando menos, marginalmente informativo y mayormente reiterativo.

En este *minitema* se pasa revista a unas cuantas de las vertientes y razones —las que considero más importantes por su más vasto alcance— que por amplio consenso científico se acepta que lastran la traslación. Se intenta hacerlo rechazando apriorismos, correcciones políticas que eclipsan verdades y rinden desengaños, y rescatando soluciones propuestas o en curso. Soluciones que por verosímiles no hay que trocarlas ilusoriamente en éxitos, ni por encontrar escollos o quedarse a medio camino, precipitadamente en fracasos.

Se pretende hacerlo asimismo desde un prisma alejado de

Dicen que las crisis espolean la imaginación. De quien la tiene. Se afirma que durante ellas se idean modos de hacer más y mejor con menos. A menudo se queda en eso: en ideaciones.

dogmatismos. Se parte de una actitud que anhela el análisis y el debate sosegados y críticos, y con cierto temor a la apatía, enemigo éste acérrimo del avance del conocimiento y cortapisa no para alcanzar la *arete* —la excelencia de los griegos, que aunque vulgarizada por tanta alusión, nos queda muy lejos—, sino para que en lo más micro se palpen mejoras.

Dicen que las crisis espolean la imaginación. De quien la tiene. Se afirma que durante ellas se idean modos de hacer más y mejor con menos. A menudo se queda en eso: en ideaciones. En ciencia se avanza con esporádicas innovaciones disruptivas separadas largos años por avances marginales, uno de los cuales, con escasa frecuencia y como por arte de birlibirloque asentado en las acumuladas, se convierte en gran salto en conocimiento o tecnológico.

Lo que interesa aquí es pasar revista a algunos obstáculos cruciales a esos avances incrementales en traslación, sus vías de avance y sus incentivos. Se hará a sabiendas de dos limitaciones prácticamente insalvables: por su vastedad, sólo se abordarán algunos de los esenciales, y se ofrecerá información para empezar a esbozar un posible eje vertebrador. Ponga el lector su aportación incremental.

1. Renuencia y temor a la incertidumbre

Un primer impedimento a la traslación parte del hecho de que virtualmente nunca podemos hacer estimaciones exactas y precisas. Una aparente perogrullada que deja de serlo en cuanto comprobamos dos cosas. Primero, lo más frecuente es encontrar estimaciones puntuales, no estimaciones acompañadas de sus (consustanciales) medidas de error e incertidumbre. Segundo —quizá más grave—, en cualquiera de ambos casos las conclusiones extraídas no suelen haber sido ponderadas conforme a la magnitud de la incertidumbre asociada con los datos y la información que las sustentan.

Ello se ha venido en llamar con acierto certeza increíble (*incredible certitude*), partiendo de un precepto delator, la **ley de la credibilidad decreciente**: las suposiciones más fuertes desembocan en conclusiones más poderosas, pero menos creíbles (2). Presenciamos desde hace tiempo una asidua violación de la lógica inferencial: elaboramos hipótesis y formulamos supo-

Nuestra avidez por la certeza y reticencia a la incertidumbre nos impele a concluir con fuerza y, obrando de este modo, soslayamos las debilidades de los supuestos de partida y sacrificamos la credibilidad.

siciones, obtenemos datos para contrastarlas y extraemos conclusiones, pero sin ponderarlas debidamente en función de la incertidumbre. Es incontestable que la veracidad o certeza de las conclusiones depende de la de los supuestos y de la calidad de los datos. Los estudios pueden tener gran coherencia interna pero partir de supuestos indefendibles y operar con datos débiles. Este problema se agrava cuando se implantan medidas o llevan a cabo intervenciones basadas en ese error por omisión.

A menudo, también nos encontramos frente a una tergiversación terca del pensamiento lógico: nuestra avidez por la certeza y reticencia a la incertidumbre nos impelen a concluir con fuerza y, obrando de este modo, soslayamos las debilidades de los supuestos de partida y sacrificamos la credibilidad. Un ejemplo enjundioso: la Congressional Budget Office de los Estados Unidos estima de antemano los costes asociados con la aplicación de una nueva ley al año y transcurridos, cuatro y diez de su aprobación. Los supuestos abundan, no pocos cambios se introducen en esos periodos, pero las estimaciones son puntuales; ni rastro de incertidumbre asociada con ellas. *La incertidumbre del impacto presupuestario no es de naturaleza estadística. Son los analistas, cuando hacen estas predicciones, quienes dudan del realismo de sus supuestos* (2).

Una sencilla contrastación pragmática de lo anterior: ofrezcamos límites de intervalos de confianza en lugar de estimaciones puntuales y comprobaremos la reticencia a aceptar la ambigüedad, a actuar sobre la base de dos posibles valores (los extremos del intervalo) y de todos los comprendidos entre ellos.

Y porfiarnos: exigimos respuestas deterministas, aunque la mecánica de nuestros razonamientos y decisiones sea a menudo de naturaleza probabilística. Como se verá más adelante, la psicología y la economía del comportamiento (*behavioural economics*) arrojan alguna, de momento tenue, luz sobre estas aparentes paradojas (3-5).

Dos grandes peligros que dimanar de todo ello son: las consecuencias derivadas de tomar medidas basadas en supuestos, datos o conclusiones erróneos y el consiguiente desprestigio no acotado de la ciencia.

Por fortuna, la última década muestra el lento, muy lento, pero creciente uso de medidas de error e incertidumbre, así como

el de **análisis de sensibilidad**¹. En lugar de empeñarnos en mantener conclusiones fundamentadas en supuestos o pruebas dudosos, de defender conclusiones respaldadas por pruebas científicas lábiles, provisionales o insuficientes, ¿por qué no mostrar a las claras cómo moldean las conclusiones los distintos supuestos verosímiles, cómo varían las conclusiones cuando partimos de distintos supuestos? ¿Por qué no averiguar cómo cambian las conclusiones al partir de supuestos con distinto grado de certeza? Se realizaría así un acto de honestidad (virtualmente siempre hay incertidumbre) y de rigor científicos (los análisis realizados no la omiten, la incorporan).

Otro dato positivo a destacar es que también se ejercita más que antes, aunque todavía menos de lo deseable, la práctica de, primero, escrutar los supuestos a partir de los cuales, una vez realizados los estudios, se extrajeron las conclusiones y, segundo, aquilatar estas últimas *ex post*, a la luz de los resultados del ejercicio y, sobre todo, antes de convertirlas en medidas, acciones, políticas o intervenciones.

No acaban aquí los peligros. También se ciernen al aplicar medidas que parten de extrapolaciones basadas en conclusiones insostenibles, y al suponer que una situación presente o futura es o será idéntica a aquella pretérita en que se fundamentó el estudio del cual se infirieron las medidas. Es extrapolación osada, ilegítima desde el prisma de la ciencia, la que soslaya las limitaciones de toda investigación y se aleja de los datos observados. La **suposición de la invariancia** (*invariance assumption*) carece de sustento: aunque sea práctica asidua, nada legítima admitir que las asociaciones encontradas en un conjunto de sujetos bajo determinadas circunstancias también se observarán en otros individuos inmersos en parejas y diferentes circunstancias² (2).

Aplicado a numerosas afirmaciones en medicina de inspiración y tono ecuménicos, ¿por qué relajar la vigilancia y creer que no cabe observar efectos secundarios de un fármaco no detectados en un ensayo clínico (con 1000 pacientes) cuando se administren a 10.000 o 100.000? La incidencia es medida de riesgo por antonomasia y la oportunidad de exposición (*exposure*

¹ No confundir el análisis de sensibilidad que aquí se menciona, alterar los parámetros sujetos a incertidumbre para estimar el impacto en las conclusiones, con el de similar nombre en epidemiología clínica, relativo a la capacidad de las pruebas diagnósticas de identificar verdaderos positivos entre los pacientes realmente enfermos.

² Lo que puede ser útil como método de simplificación o reducción de la complejidad para abordar una investigación, entraña el riesgo de convertirse en un atajo para extraer conclusiones triviales que, alimentadas por la pulsión de certidumbre, llevan a desbordar los límites de la sensatez.

opportunity) es *conditio sine qua non* del riesgo. ¿Por qué no examinar al paciente con extremo cuidado y pericia, siempre, a todos, antes de administrar un medicamento y esperar que sea tan eficaz como en el ensayo clínico con que se logró su aprobación? Se olvida o desconoce que los resultados de un ensayo clínico se basan en análisis de efecto en un grupo, no en individuos aislados (estimamos el porcentaje de los que responden, no si lo hará un determinado participante; para esto último han de hacerse otros estudios). En unos la respuesta es intensa, en otros, menor, otros no responden y algunos de los que recibieron placebo mejoran.

Hay quienes defienden a ultranza la primacía de la validez interna de los estudios y se despreocupan o subvaloran la externa³. La credibilidad de los ensayos clínicos, como la de los análisis de políticas o de servicios, depende de ambas. Contravenir este principio y actuar en consecuencia acarrea el riesgo de dilapidar recursos y causar daño por acción u omisión de los fundamentos y la lógica científica (2, 6, 7). Su inevitabilidad obliga.

A todo ello cabe añadir —sólo a título recordatorio por la amplitud del tema— las limitaciones de los ensayos clínicos con controles históricos, junto con la información valiosa que pueden aportar cuando su diseño, realización y análisis cumplen requisitos establecidos de calidad, así como la utilidad de los estudios de casos y controles. Cuando estos últimos se han efectuado observando también los principios de diseño y análisis y se han tomado medidas adecuadas para minimizar la introducción de sesgos y controlar factores de confusión, los resultados que han arrojado han sido extremadamente valioso en la práctica clínica: recuérdense, por ejemplo, su aportación a conocer la asociación entre aspirina e infarto de miocardio, aspirina y síndrome de muerte súbita en recién nacidos o entre folatos y malformaciones del tubo neural (7).

³ Recordamos que la validez interna expresa la medida en que la selección y utilización de los métodos en la muestra de un estudio es adecuada y, por tanto, de la ausencia de sesgos. La validez externa es la medida en que dichos resultados pueden extrapolarse o no a la población de origen, a otras poblaciones o contextos. Un estudio con alta validez interna puede carecer de validez externa. Sin la primera no puede existir la segunda.

Es importante reconocer la existencia de los sistemas de pensamiento automático-intuitivo por ser muy habituales y por ser susceptibles a sesgos: se trataría de ponerlos al servicio de la racionalidad en las decisiones

2. A veces el pensamiento y el comportamiento pueden embaucarnos

Otro origen del freno a la traslación orbita en torno a un postulado de la economía del comportamiento: nuestro comportamiento no se guía por la lógica ni la elección racional (*rational choice*), sino que está determinado por nuestro cerebro falible e intensamente influido por el ambiente o el contexto en que se elige o decide (3,4).

Psicólogos y neurocientíficos han identificado dos sistemas que actúan en nuestro cerebro. El **sistema 1**, automático, rápido, inconsciente, intuitivo, movido por afectos y emociones. Y el **sistema 2**, reflexivo, lento, consciente y racional. Acostumbramos actuar automáticamente, muchas de nuestras decisiones se basan en automatismos, utilizan reglas generales, empíricas (*rules of thumb*), la heurística, que suelen funcionar bien⁴. Pero a veces introducen sesgos cognitivos, nos hacen cometer errores que distorsionan poco o mucho nuestro proceso "racional de toma de decisiones" (5).

Numerosos errores de pensamiento, falacias y sesgos son fruto del sistema 1. Aquellos científicos han comprobado la existencia de dichas reglas generales y sesgos cognitivos asociados con ellas que actúan de forma automática. Estas claves automáticas tienen sus raíces en emociones, prejuicios, aversiones, tendencias personales por defecto a actuar o a permanecer en la inacción. Y al dirigir algunos de nuestros pensamientos y acciones, nos precipitan a tomar decisiones incorrectas y costosas y a cometer una amplia gama de errores.

A pesar de haberse demostrado la fuerza del sistema tipo 1 y del importante papel rector de desempeña la heurística en nuestras decisiones y elecciones, una mayoría no desdeñable de actuaciones, políticas y medidas que desplegamos se vinculan con el sistema tipo 2.

La economía del comportamiento persigue el objetivo dual de minimizar la toma de decisiones heurísticas, empleando estrategias que intentan alejarla de sesgos cognitivos, y —

⁴ George Pólya enunciaba una serie de recetas "heurísticas" que ilustran bien el concepto:

Si no consigues entender un problema, dibuja un esquema.

Si no encuentras la solución, haz como si ya la tuvieras y mira qué puedes deducir de ella (razonando a la inversa).

Si el problema es abstracto, prueba a examinar un ejemplo concreto.

Intenta abordar primero un problema más general (es la "paradoja del inventor": el propósito más ambicioso es el que tiene más posibilidades de éxito). <http://es.wikipedia.org/wiki/Heur%C3%ADstica>

reconociendo que no podemos desprendernos del sistema 1— convertir dichos sesgos en elementos inherentes de los procesos automáticos y hacerlos trabajar en beneficio (no en detrimento) del que toma decisiones para conseguir resultados congruentes con sus deseos. Se sabe que nuestro comportamiento puede alterarse modificando la arquitectura de elección (*choice architecture*), esto es, el marco en que se nos presentan las opciones que tenemos para elegir. Y este conocimiento ha granjeado buenos resultados al aplicarse en distintos ámbitos (3-5).

En 2010, un grupo de investigadores construyeron el modelo o marco *Mindspace*, que se utiliza con frecuencia para mejorar la efectividad de las políticas encaminadas a modificar comportamientos. *Mindspace*, además de ser su acrónimo, vertebra los efectos (sesgos y automatismos mencionados) que se ha demostrado que afectan el comportamiento actuando, no sólo, pero sobre todo, en el sistema tipo 1. Dichos efectos se describen a vuelapluma en la siguiente tabla (5):

<p>Mensajero. La fuente del mensaje es a veces más influyente que el mensaje. De ahí la importancia de la credibilidad de la fuente.</p>
<p>Incentivos. La heurística y los sesgos cognitivos pueden desviar sustancialmente la dirección y el efecto pretendidos de los incentivos, así como modular las respuestas que con ellos se consiguen a la postre. La gente se esfuerza más en evitar pérdidas que en obtener ganancias de la misma magnitud. Inducirla a sentir la pérdida si fracasan puede motivar más que premiarla en la misma magnitud por conseguir algo.</p>
<p>Normas. En ocasiones percibimos erróneamente actitudes y comportamientos de los demás al pensar que son los de la mayoría. Las percepciones de la realidad tienden a estar muy influidas por anécdotas y experiencias personales.</p>
<p>Actuar por defecto. Tendemos a preferir que las cosas sigan igual. A menudo elegimos la opción por defecto sin ninguna meta <i>in mente</i>. Por ello es útil que la opción que se ofrezca por defecto sea la deseable.</p>
<p>Prominencia. Lo que elegimos está fuertemente influido por lo que se hace llamar la atención. Por ello con frecuencia decidimos aplicando un solo criterio e ignorando lo relevante.</p>
<p>Preparación. Las decisiones pueden alterarse inconscientemente exponiendo (<i>preparando</i>) con antelación a las personas a influencias contextuales.</p>
<p>Afecto. Muchas decisiones contienen cargas afectivas. Elaborar esas afecciones antes de tomarlas puede evitar tomar algunas con consecuencias indeseadas duraderas.</p>
<p>Compromisos. Muchos son conscientes de que carecen de fuerza de voluntad e intentan circunvalarlo. Los llamados <i>dispositivos de compromiso</i> producen respuestas automáticas deseadas que modulan el sistema tipo 1.</p>
<p>Ego. La gente tiene la necesidad psicológica de preservar su autoimagen. No está claro cómo pueden modularse los efectos del ego para favorecer comportamientos deseados</p>

Muchos factores contribuyen a dificultar la intensidad y forma en la cual los efectos esperados desde el conocimiento no se traducen en la práctica: sobreestimación de la eficacia, no actuación sobre comportamientos, fallo de gestión del cambio, no abordaje de facetas múltiples del problema en la implementación, intereses, acción prematura o tardía, etc.

Se desconoce cuán duraderos son estos efectos, si los que funcionan en determinados contextos lo harán en otros distintos y si los que son efectivos con ciertas personas lo serán con otras. Pero, en suma, todos parecen desempeñar un indudable papel que puede ayudar a explicar ciertos comportamientos y decisiones y, por ende, arrojar alguna luz sobre si median en el proceso de traslación información y conocimiento a la práctica.

3. Cuando lo observado no alcanza lo esperado

Más empírico es y cercano nos resulta el hecho habitual de que la efectividad alcanzada en distintas intervenciones sea menor que la esperada a juzgar por estimaciones de su eficacia (y de deseos que centellean *in mente*). Se ha comprobado en muy diversos terrenos; por ejemplo, medicación, pruebas diagnósticas, intervenciones en seguridad, clínica, estrategias para mejorar la adecuación de indicaciones, utilización de herramientas informáticas, implantación de modelos de gestión de servicios o de enfermedades crónicas, formas de pago o de incentivación (8).

La explicación de que el efecto de esta traslación de conocimiento no se traduzca en la práctica con la intensidad y la forma esperadas no estriba en un elemento aislado, sino en un entramado de factores que actúan aisladamente o interactuando entre ellos. Entre los mencionados reiteradamente destacan: sobreestimación de la eficacia; omisión de elementos clave de índole organizativa y estructural de los dispositivos asistenciales o gerenciales donde se despliegan las intervenciones; ausencia de adaptación a las condiciones donde se implantan por no haber realizado estudios previos de viabilidad y factibilidad; carencias en su mantenimiento; mala gestión, y diversos tipos de ineficiencias. Con frecuencia se olvidan también factores culturales del medio, económicos (insuficientes recursos o errores en su asignación), psicológicos y conductuales (8-10).

Respecto a los factores vinculados con el comportamiento que pueden representar obstáculos a la traslación de conocimiento y, por tanto, explicar efectividades menores, sobresale la virtual ausencia de acciones dirigidas a modificar comportamientos, que es necesario realizar antes o durante el despliegue de las intervenciones o de la adopción de la tecnología. Para adoptar adecuadamente algunas tecnologías, como las enumeradas, primero deben modificarse visiones, procedimientos muy consolidados, la organización de determinados sistemas y

servicios, y también percepciones, prejuicios, comportamientos (8-10).

A veces, incluso cuando se utilizan esas estrategias de cambio, no se escogen las más efectivas (se conoce la efectividad de medidas como la retroalimentación, recordatorios, sesiones de formación, diseminación pasiva de información) (8).

Un factor explicativo adicional es que muchas intervenciones, para ser efectivas, han de incidir en distintas facetas del problema que se afronta y, por consiguiente, la efectividad alcanzada dependerá de que su implantación se haya acompañado simultáneamente de estrategias encaminadas a modificar cada una de dichas facetas. Por ello debe insistirse en la necesidad de actuar mediante estrategias multifacéticas. Con éstas, no con las aisladas, se puede conseguir modificar o preparar los factores contextuales necesarios para que las intervenciones o las tecnologías adoptadas sean efectivas (8-10).

Hay factores de otra índole que también obstaculizan la traslación de conocimiento a la práctica: no cambiar el *statu quo* de personas clave en los sistemas; precipitarse en su incorporación (priorizar la acción sobre la evidencia científica); demorarse en exceso (priorizar la inacción), o emular intervenciones externas y aplicarlas en un contexto distinto soslayando factibilidades y la necesidad de introducir medidas de adaptación a la realidad local (8, 10, 11).

Por último, cabe recordar que la traducción de pruebas científicas en cambios de práctica en ocasiones exige realizar esfuerzos a medio plazo, aplicar medidas permanentes, no acciones iniciales o esporádicas, así como mantenerlas y evaluarlas periódicamente e introducir correcciones cuando sus resultados así lo indiquen (10).

4. Deficiencias en la generación de información científica

Trasciende con creces los objetivos y el alcance pretendidos de este documento abordar los innumerables problemas de diseño de investigaciones básicas y aplicadas (y en todas sus parcelas), las variantes de la elección inadecuada de tipos de estudios, o los principales errores en análisis estadístico que explican una fracción no desdeñable del lento avance de la traslación de información y conocimiento.

Esta acotación no es óbice para revisar varias limitaciones importantes del valor de la información científica disponible que merman su traslación a la práctica.

- La primera engloba 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*).
- La segunda 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* o *outcome reporting bias*).
- La tercera corresponde a los **sesgos de publicación** (*publication bias*), que engloban el conjunto de estudios que no se publican porque sus resultados no alcanzan significación estadística o por baja magnitud de los efectos estimados. En un alto porcentaje de casos estos sesgos se explican 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 (12-19).

4.1 Manipulación de variables

Por escasez o ausencia de datos, con fines de manipulación interesada o por desconocimiento, tres son los más frecuentes tipos de manipulación de variables encontrados en estudios y evaluaciones clínicos y de salud pública, y en evaluación de políticas y de servicios de salud: Substitución de variable primaria por secundaria; uso de variable intermedia en vez de la medición de impacto; uso de variable compuesta (integrando muchas variables) que disuelve el efecto de la variable de impacto.

Todos se orientan a la postre hacia la demostración de la superioridad absoluta o relativa —mayor eficacia, efectividad, seguridad o eficiencia absolutas, relativas o incrementales— de una intervención frente a una o más alternativas de comparación escogida.

Primero, 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 (19).

Segundo, otra limitación se concreta en la aportación de pruebas de la superioridad de una intervención por medio de variables de proceso o intermedias (*surrogate variables*) en la cadena causal que une la exposición supuesta 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 número de reuniones anuales celebradas por una comisión, 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, o los resultados que ha logrado tener el trabajo de esa comisión.

El problema aquí estriba en que las variables sustitutivas utilizadas a veces no son predictivas indudables de la variable final o se desconoce si lo son. Es posible que la variable sustitutiva realmente no forme parte de la cadena causal, ocupe un lugar apéndice en ella o su contribución a la causa sea minúscula o sólo lo haga en presencia de terceras variables no siempre presentes. La validación de su capacidad predictiva es una omisión frecuente. 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 (12-14, 20).

Sirvan a guisa de ilustración algunos ejemplos extraídos de 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

Tipos de manipulación de variables:

1 Sustitución de variable primaria por secundaria;
2 Uso de variable intermedia en vez de la medición de impacto;
3 Uso de variable compuesta (integrando muchas variables) que disuelve el efecto de la variable de impacto.

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 (las llamadas Δ "deltas" fijadas a priori) 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) (21). (La tendencia a maximizar el uso de pruebas con alta sensibilidad y baja especificidad aumenta la actividad y los costes evitables.)

Se ha mostrado, además, que la magnitud del efecto estimada en los ensayos clínicos que utilizan básicamente variables sustitutivas suele ser mayor que la de aquellos que utilizan variables primarias relevantes de resultado. Por defecto se incurre en sobreestimación. Se aprecia, asimismo, que este hecho no se explica por diferencias en los riesgos de sesgos u otras características importantes entre ambos grupos de ensayos (22).

Tercero, dicha superioridad también se intenta demostrar por medio de una variable compuesta (*composite variables*), esto es, una variable que resulta de la combinación de n variables individuales (13). Esta manipulación puede ser legítima desde un prisma científico, como las variables compuestas del modelo del buen gobierno (*governance*) de Kaufmann et al. del Banco Mundial o distintos **índices compuestos usados para estimar riesgos de enfermedades, de hospitalización o pronósticos** (23, 24). Pero con el paso del tiempo su uso es más frecuente 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). Afortunadamente, está aumentando el número de revistas que obligan a notificar los resultados de la variable compuesta junto con los de las que la integran (13, 21).

4.2 Notificación selectiva de resultados

Esta limitación responde al denominado sesgo de notificación de resultados, que se introduce cuando los hallazgos de un estudio se notifican selectivamente en función de la magnitud del efecto detectada, la significación estadística alcanzada, los intereses de los financiadores o investigadores del estudio o por mera ignorancia.

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 colige de las revisiones sistemáticas de ensayos clínicos con tratamientos farmacológicos (14, 20, 25, 26).

4.3 Sesgos de publicación

La tercera limitación la encarnan los sesgos de publicación, la forma más extrema del sesgo de notificación de resultados. En este caso, los estudios en los cuales no se supera cierta magnitud del efecto estimado, no se alcanza significación estadística, por la concurrencia de intereses de los financiadores o por ignorancia, no se publican (15, 16).

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 (estudios negativos), o los promotores de la tecnología evaluada cuando no se demuestra que ésta 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 postpublicación, cuando, por las mismas razones, esos estudios se excluyen de las revisiones sistemáticas, los metanálisis u otros informes de síntesis (16, 23, 25).

Hay sesgo de notificación de resultados cuando los hallazgos de un estudio se notifican selectivamente...

Hay tres tipos de sesgos de publicación:
 1 Pre-publicación (de los propios investigadores)
 2- Publicación (de los editores)
 3 Post-publicación (en las revisiones sistemáticas y metanálisis)

El sesgo de publicación se ha estudiado en profundidad en ciencias de la salud. Desde que se acuñó el término en los años 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 (15, 16, 27, 28-30).

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 respaldan la eficacia de sertralina y paroxetina y revelan que el riesgo de suicidio en niños tratados supera con creces sus escasos beneficios. Este resultado fue congruente con los del metanálisis realizados con citalopram y venlafaxina (31-33).

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 (15, 16, 27, 28).

A la difundida reticencia inopinada a valorar y evaluar sin prejuicio los estudios negativos se aúna la contrastada insuficiencia de información, incluso en revistas prominentes, para valorar cabalmente su validez (34).

Las soluciones propuestas para mitigar los efectos de estos sesgos se resumen en las siguientes: algunas de las medidas incluidas en los modelos de ciencia regulatoria (*regulatory science*: 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, la publicación de revistas de estudios negativos, y la declaración de los conflictos de intereses, la denuncia y notificación pública de los que se detecten, y el fomento de la estadística bayesiana (26-28, 30, 31, 35, 36).

Sobre las ventajas que ofrece el enfoque bayesiano frente al frecuentista se ha escrito mucho, pero en un relativamente reciente documento se encuentra una breve y útil síntesis, enmarcada en su debido contexto, para quien desee una excelente introducción (7).

5. Los problemas de evaluar y decidir

Evaluación económica de intervenciones: una limitación adicional menos mencionada

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. La cuarta barrera: se suma a las de eficacia, seguridad y calidad.

Su aplicación se resume en la fijación de umbrales de coste-efectividad (coste/AVAC), como expresión de la predisposición social a pagar y del coste de oportunidad. Salvo excepciones explícitas y *grosso modo*, las tecnologías cuya razón de coste-efectividad es igual o menor que el umbral acordado se financian; aquellas por encima del umbral se excluyen de la financiación (37-40). En algunos países, las instituciones y empresas privadas son con mucho la principal fuente de financiación de las evaluaciones económicas (39).

El primer obstáculo viene dado —de nuevo— por la forma como se lidia con la incertidumbre en las evaluaciones económicas. En ellas, a pesar de la insistencia de muchos investigadores y de las innumerables veces que se insiste en la necesidad de utilizar guías de notificación de resultados, también es asiduo encontrar estimaciones sin medidas de la incertidumbre asociada con ellas. ¿Cambian y, si lo hacen, cómo, los resultados, por ejemplo, de análisis de coste-efectividad, cuando se modifican las

Problemas en los estudios de evaluación coste-efectividad: No se maneja la incertidumbre con los necesarios análisis de sensibilidad Se combina con los sesgos de publicación y para los subgrupos con coste-efectividad sobre el umbral establecido

estimaciones de costes y de efectividad usadas para realizarlo? ¿Cuán sensibles son, si lo son, las razones de coste-utilidad notificadas en una evaluación a las diferentes estimaciones de costes y utilidades incluidas en evaluaciones independientes publicadas? ¿Cuál es el conjunto de valores que pueden tomar dichas razones congruentes con los datos disponibles procedentes de diversos análisis independientes publicados? (41, 42).

Si hay incertidumbre, debe notificarse. Si no existe una única razón, por ejemplo, de coste/AVAC, congruente con los datos disponibles, sino que cabe considerar verosímiles diferentes costes y utilidades publicados, hay que incluirlas en los análisis de sensibilidad y notificarlas también, todo el recorrido de sus valores congruentes con los datos observados. Si se han estimado en otros estudios o mediante análisis de sensibilidad determinísticos o probabilísticos, tienen que difundirse. Si no se han llevado a cabo dichos análisis, hay que realizarlos y, antes de hacerlo, llamar a la prevención y disuadir de la precipitación en trasladar resultados inconclusos a la práctica. El recorrido de valores de dichas razones congruentes con los datos observados puede atemorizar a algunos, pero es mucho más informativo que las estimaciones puntuales. A pesar de saberlo, trabajamos con razones de coste/AVAC puntuales sin rastro de planos de coste-efectividad como expresión gráfica de simulaciones con estimaciones tan verosímiles como la única barajada.

Actuando por omisión, se informan decisiones que pasan por alto información crucial, y las decisiones finales que se tomen pueden ser erradas, producir daño o, cuando menos, no contribuir a producir el máximo beneficio neto esperado en salud dada esa información. El coste de oportunidad de dichas decisiones puede ser alto. De hecho, como se deduce de los resultados de los *análisis del valor de la información (value of information analysis)*, se cifra en la diferencia entre el beneficio neto que se obtendría decidiendo sobre la base de la información menos imperfecta (la que incluye la derivada de los resultados de los análisis de sensibilidad) y el que se obtendrá con la información más imperfecta (aquella que excluye esos resultados).

Una actitud inexcusable, sencillamente porque se ha reiterado hasta el hartazgo la necesidad de efectuar análisis de sensibilidad. Y cuando se sospecha que la decisión, por ejemplo, de aplicar una política, solicitar una prueba diagnóstica, implantar un programa o administrar un tratamiento u otro puede cambiar porque el valor esperado de la decisión basada en la información disponible es menor que el de la decisión tomada con información adicional, no

recurrir a análisis del valor de la información e informar la toma de dicha decisión sin obtener información adicional, también es inexcusable.

La importancia de lo anterior puede intuirse fácilmente pensando en las consecuencias de administrar un tratamiento menos efectivo, con mayor riesgo de iatrogenia o menos coste-efectivo que el alternativo (pérdida de efectividad evitable, eventos adversos evitables y gasto evitable, lo cual multiplicado por n pacientes aproxima la magnitud del menoscabo de la eficiencia social provocado). O en adoptar políticas diseñadas *ex ante* con deficiencias o sin haberlas acompañado de análisis de viabilidad y estimaciones adecuadas de sus costes (37-40).

Esta limitación también la agravan los sesgos de publicación de estas evaluaciones: es más probable que aquellas que estiman razones de coste-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 coste-efectividad estimadas superan el umbral (7). También puede producirse por omitir la incertidumbre tratada en los párrafos precedentes.

Con análisis de sensibilidad debidamente justificados y notificados, el valor incluso ilustrativo que tienen los planos de coste-efectividad incremental y las curvas de aceptabilidad y, cuando es menester, de los resultados de los análisis del valor de la información correspondiente, no pocas de las consecuencias comentadas podrían controlarse. (41, 42).

En el área de la evaluación económica se ayudaría a transferir más y mejor conocimiento, a reducir en alguna medida el tamaño de la brecha entre eficacia y efectividad y a mejorar un poco más la práctica clínica, gerencial y política. De paso, también disminuirían las decisiones basadas en dogmatismos por ignorancia o deficiente educación y formación y se reduciría la desconfianza social.

Información para la toma de decisiones: siempre menos de lo necesario...

Las deficiencias de la información que fundamenta la toma de decisiones en clínica, regulación, macro, meso y microgestión, 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. Son otro motivo más que explica el entorpecimiento de la traslación de información y conocimiento a la práctica.

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 innumerables estudios, la mayoría de los cuales abordan aisladamente alguna de sus muchas facetas (27, 28, 43-45). Pocos lo hacen de forma cabal. Todo ello se traduce en obstáculos a la traslación, en impedimentos a la aplicación de conocimiento científico en la mejora clínica, de la gestión de la regulación, del diseño, la aplicación y evaluación de políticas e intervenciones sanitarias.

Entre las limitaciones más prominentes —porque atañen a la raíz del problema— sobresalen 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 integral de la notable magnitud de este problema —de lo que se ha venido 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 (8, 27, 28, 35, 36, 43-45).

No hace mucho, Sir Michael Rawlins puso en jaque a los dogmáticos de los niveles de evidencia: *quienes toman decisiones* —subrayó— *han de ejercitar el juicio; las jerarquías no son ningún sustituto*. Ordenar las pruebas científicas de forma estanca, sin tener en cuenta que las gradaciones de *niveles de evidencia* no pueden dar siempre acomodo a la variedad de estudios disponibles con frecuencia sobre un tema, excluyendo, además,

Las decisiones se toman con un bajo nivel de evidencia. La jerarquización de las evidencias por niveles no debe tomarse como un dogma, ni servir a un modelo lineal o reduccionista de decisión. La toma de decisiones implica un complejo ejercicio de síntesis que supone un juicio condicionado y basado en la totalidad de las pruebas científicas existentes, ajustadas al contexto.

a los cualitativos, *supone adoptar un enfoque extremadamente simplista y seudocuantitativo para valorar todas las pruebas científicas existentes*. El juicio ha de dar cabida a la evaluación de la medida en que cada componente de la *evidencia científica es fiable, generalizable, si los beneficios son mayores que los riesgos,...* Y cita a William Blake: *God forbid that truth should be confined to mathematical demonstration*.

Lo que subyace en el centro de la toma de decisiones es el juicio condicionado por la totalidad de las pruebas científicas de base existentes (7).

Las pruebas científicas necesarias para informar, orientar y evaluar rigurosamente políticas e intervenciones de salud pública que tratan temas tales como los determinantes de salud, la promoción de la salud y la prevención de la enfermedad, 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 postcomercialización 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 (46-49).

En definitiva, como indican los análisis del valor de la información, cuanto menor es el *valor* de la información disponible para tomar decisiones tanto mayor puede ser el coste de oportunidad de estas últimas. Su importancia se comprueba de forma palmaria cuando, por ejemplo, dicho coste se expresa como: a) pérdida de beneficios en salud por el retraso en incorporar intervenciones o tecnologías potencialmente efectivas por falta de pruebas científicas suficientes, b) iatrogenia o reducción de la seguridad clínica (morbimortalidad evitable) o c) ineficiencia (dilapidación de recursos invertidos en salud, incorporación de intervenciones o tecnologías con bajas razones de coste-efectividad incremental) resultante de su adopción o incorporación prematura, esto es, sin suficiente respaldo científico de su eficacia, efectividad o eficiencia (43, 44).

6. Y la salud pública

La búsqueda, la selección, el análisis y la síntesis de información científica para diseñar, implantar y evaluar políticas y apoyar la toma de decisiones ponen de relieve marcadas y notorias 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.

También se conocen las características diferenciales de la información procedente de estudios experimentales y observacionales en investigación clínica. 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 (20, 27, 28).

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 o unidades territoriales (50).

En las políticas, la aleatorización resulta virtualmente 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 y modelos como, por ejemplo, los cuasi experimentales pre-post, con o sin grupo control, series temporales interrumpidas, regresiones truncadas, y un largo etcétera del arsenal metodológico(27, 28).

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 tomar 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 (20, 27, 28, 45).

Por añadidura, 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 (20, 27).

A este ovillo de factores se añade 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 mayor 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, por recortes presupuestarios o pérdida de interés (8, 20, 27).

La lenta traslación de resultados de estas investigaciones en

En la Salud Pública difícilmente caben ensayos clínicos aleatorizados como en las intervenciones asistenciales. Los diseños se diferencian según la diana: individuos, grupos o poblaciones enteras. La complejidad de las intervenciones en Salud Pública, el plazo del impacto, y la importancia de factores comportamentales, dificulta la validez interna; los factores contextuales también limitan la validez externa.

salud pública en atribuible en su justa medida también a la omisión o subestimación de la importancia crucial que reviste la extrapolación de sus resultados a contextos, estructuras y modelos de gestión de otros sistemas de salud. Se sigue concediendo primacía a la validez interna, condición necesaria pero no suficiente para maximizar la probabilidad de éxito en la traslación de resultados a la práctica. También debe atribuirse a la desconsideración de factores contextuales y organizativos en los estudios, cuyo análisis debe ser otra de las piedras angulares de las investigaciones si se pretende traducir resultados en cambios de práctica. En la bibliografía se encuentran enumeraciones enjundiosas de otros elementos cuya omisión o análisis de soslayo explican la lenta traslación (2, 9, 51).

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. Aquí la traslación es pobre fundamentalmente por carencia de investigación. 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 costes 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 (8, 20, 26-28, 37, 48).

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 postpublicación.

7. Algunas soluciones puestas en marcha

Efectividad comparada: ¿acabará siendo útil?

Mucho se ha escrito sobre la efectividad comparada (EC, *comparative effectiveness*) y no es objetivo de este documento describirla *in extenso*. Sí, por el contrario —y circunscritos en el marco de análisis acotado aquí—, verter reflexiones que ayuden a apreciar su capacidad para contribuir a su fin primigenio, dados los obstáculos que está arrostrando su despliegue (7, 52-56). La EC se ideó genéricamente como táctica para facilitar la transferencia de información científica y conocimiento experimental a la prestación de servicios de salud. Se trata en este caso concreto de reducir el hiato entre eficacia y efectividad (aportar resultados clínicamente relevantes obtenidos en estudios realizados en condiciones habituales de prestación de servicios para la toma de decisiones de clínicos, pacientes y gestores) (7, 52-56)⁵.

Al mismo tiempo, con la EC se intenta sortear numerosos escollos que siguen obstaculizando la obtención de información clínicamente relevante, práctica, utilizable directamente en la cabecera de la cama del paciente, y mermando la eficacia del modelo de investigación clínica vigente:

- limitada capacidad para llevar a cabo ensayos clínicos con financiación pública en los países más industrializados por insuficiencia de los fondos asignados;
- descensos del retorno de inversión privada (por reducción de la cifra de nuevas entidades moleculares comercializadas y aumento de la razón coste/nueva entidad molecular comercializada);
- extrema complejidad procedimental en la gestión de ensayos (industria artesanal o *cottage industry*);
- incentivos ausentes o de escaso o insuficiente vigor; dificultad creciente para reclutar pacientes y alcanzar cotas altas

Analizar la efectividad comparada de dos intervenciones o políticas es una opción prometedora, pero que está afrontando obstáculos importantes (medios, trabas, incentivos, y limitaciones de una actividad a escala artesanal)

⁵ La investigación en efectividad comparada genera y sintetiza evidencia sobre beneficios y riesgos de métodos alternativos para prevenir, diagnosticar, tratar o monitorizar problemas de salud. Se trata de comparar intervenciones existentes y ver cuál funciona mejor en según qué condiciones. http://www.nap.edu/openbook.php?record_id=12648&page=29

de participación en los estudios de especialistas, profesionales de la atención primaria y la medicina comunitaria, académicos, gestores para ensayos clínicos, y

- porcentajes reducidos de ensayos finalizados; barreras para realizar estudios donde se atiende a los pacientes; intrincadas relaciones entre clínica, industria y universidad; traslado de los estudios a países en desarrollo y un largo etcétera.

Eliminar redundancias, burocracia, en suma, acabar con las profundas ineficiencias de esta *industria artesanal* (7, 52-56).

Trascurridos algunos años desde su "lanzamiento", la EC también afronta duros impedimentos para traducir sus resultados en cambios en la práctica clínica.

- **Primero**, la EC no escapa de las consecuencias de los incentivos mal alineados, los derivados de algunas formas de pago (pago por acto, *fee-for-service*), los alentados por intereses pecuniarios y fines corruptos, y los que se ofrecen con insalvables deficiencias desde su diseño e implantación (7, 52-57).
- **Segundo**, los diseños, análisis, realizaciones y notificaciones de sus estudios padecen las mismas limitaciones que las de esos mismos estudios aplicados en contextos alejados de la EC. Cuando arrojan resultados que contravienen la práctica clínica vigente desencadenan el mismo rechazo que los precedentes de estudios también llevados a cabo en otros ámbitos o con otras finalidades.
- **Tercero**, a semejanza de los sesgos descritos antes desde el prisma de la economía del comportamiento, en EC otros sesgos ejercen asimismo sus efectos y merman a la postre la conversión de información científica en cambios en la práctica clínica. Por ejemplo, el sesgo de confirmación (la tendencia a aceptar pruebas científicas que confirman ideas preconcebidas y rechazar las que las refutan), la fuerte oposición a abandonar prácticas y tecnologías demostradamente no efectivas o con elevadas razones de coste-efectividad por intereses de variada índole o negación a

aceptar errores (*reversals*), el de prointervención (la compulsión a adoptar nuevas tecnologías aunque su beneficio en salud sea marginal y no justifique en absoluto su coste), el de protecnología —conocido también como fenómeno de la fascinación tecnológica— (creer inopinadamente que lo nuevo es mejor que lo existente y adoptarlo en la práctica) (7, 52-59).

- El sesgo de prointervención no sólo es importante en el plano de los cognitivos. Su efecto a corto plazo es la adopción de la nueva tecnología, lo cual, aunado a otras fuerzas y dadas ciertas condiciones, coadyuva a su difusión. Los beneficios en salud y sociales derivados dependerán de su eficacia relativa y de su razón coste-efectividad. En la vertiente de la información y de su traducción en cambios en la práctica, un factor que puede reforzar su adopción es la comprobación de que, a menudo, los resultados del primer ensayo clínico de eficacia de una tecnología son sobreestimaciones de su efecto, hecho que se contrasta en posteriores ensayos o estudios de efectividad (7, 52-59).
- **Cuarto**, la incapacidad de atender las verdaderas necesidades de información de gestores, clínicos, pacientes y todos los agentes interesados, a sabiendas de que éstas y sus expectativas pueden diferir marcadamente entre ellas.
- **Quinto**, insuficiente, bajo, escaso o nulo uso, en dependencia de la geografía, de herramientas de apoyo a la decisión (7, 52-59).
- Y **Sexto**, la insuficiente formación y capacitación de decisores.

A pesar de estas dificultades, numerosos investigadores perseveran en potenciar la EC, y siguen asignándose fondos públicos en distintos países para financiarla. Se ha afirmado, por ejemplo, que *Simplemente, conociendo lo que funciona y lo que no mejorará la eficiencia productiva, despojándonos [pronto o tarde] de prácticas médicas inseguras a cualquier precio*. Es más, se conocen los riesgos y los beneficios que se ciernen sobre los estudios que se inician en el contexto de la EC, incluidos los

Otros problemas de los estudios de efectividad comparada tienen que ver con incentivos no alineados, con limitaciones habituales en todo tipo de estudios, problemas cognitivos para aceptar los resultados preferidos (incluido el sesgo pro-intervención), la desatención a las necesidades de información de los decisores, la falta de uso de instrumentos de apoyo a la decisión, y la escasa de capacitación de los propios decisores

ensayos clínicos (por ejemplo, los asociados con el tratamiento habitual frente al nuevo, con los brazos de los estudios, la aleatorización, el enmascaramiento, el grado de adherencia a los protocolos, el reconocimiento de las consecuencias de la incertidumbre) (52).

La EC ha arrojado datos empíricos sumamente útiles para ajustar medidas regulatorias relativas a efectividad y seguridad relativas, fijación de precios y cobertura y para contener gasto sin menoscabar la calidad de la atención.

Ha contribuido asimismo a informar con datos de efectividad, seguridad y costes reales aprobaciones condicionales de medicamentos, a matizar análisis coste-efectividad, sobre todo mediante la aplicación de estos análisis a subgrupos de pacientes (efectos heterogéneos del tratamiento, *treatment heterogeneity effects*), así como sobre la heterogeneidad real de prestadores (mantenimiento de pericia, curvas de aprendizaje, unidades de referencia, terciarismo). Ha potenciado, además, el uso de análisis del valor de la información y la aplicación de diseños metodológicos poco empleados o inéditos en medicina hasta la fecha (como los adaptativos o los discontinuos) (7, 53).

La EC aporta ventajas al informar mejor la regulación y la decisión, permitir identificar efectividades en subgrupos de pacientes y de proveedores, y para modelos más avanzados de evaluación económica (estudios pragmáticos, de fijación de precios en función del valor...)

Sumamente útiles está siendo también los resultados que la EC brinda a los estudios de coste—efectividad relativos a costes, precios fijados y precios endógenos y sus efectos en las estimaciones que brinda la evaluación económica. Como lo son los de los estudios pragmáticos de organización eficiencia y prestación de servicios, tanto a escala sistémica como de mesogestión, o los dirigidos a establecer reglas de fijación de precios en función del valor (*value-based pricing*), de aseguramiento basado en valor (*value-based insurance*), de copagos o de fijación dinámica de precios (*dynamic pricing*), que se aplica, por ejemplo, Medicare, pagando más a proveedores que prestar servicio más coste-efectivos, el mismo a tratamientos equivalentes y menos a los no efectivos o con razones de coste-efectividad muy elevadas) (53).

La EC, aunada a los estudios de variaciones de la práctica médica, también aporta información para tomar decisiones sobre desinversión. Un hecho reiteradamente comprobado es que la magnitud de las fuerzas que impelen la adopción de nuevos medicamentos, aparatos, procedimientos o modelos de gestión, con independencia de su valor incremental, son mucho mayores que la de las que se oponen a retirar los no efectivos o no coste-

efectivos consolidados de la cartera de servicios (57-59).

Ensayos pragmáticos: una entre otras soluciones

Así se denominan los ensayos realizados cuyo diseño y cuya hipótesis se escogen y formulan para obtener información experimental necesaria para tomar decisiones. Se distinguen de los explicativos en que mientras en estos últimos la finalidad principal es estimar eficacia en condiciones experimentales, en los pragmáticos interesa conocer primordialmente los efectos, riesgos, beneficios, y costes de tecnologías o intervenciones tal como se observan en la práctica habitual, no en condiciones experimentales. En ellos se seleccionan alternativas concretas, las que están siendo utilizadas. Se incluyen unidades de estudio (por ej., pacientes, centros asistenciales, intervenciones), escenarios y lugares mucho más diversos que en los explicativos, los que remedan más fielmente las condiciones habituales de práctica, sin los criterios de selección estrictos característicos de los explicativos. Se realizan por añadidura durante periodos más largos (para detectar efectos a medio plazo no detectados por los explicativos) e incorporan más variables de resultado que estos últimos (6).

Se ha avanzado con ellos, aunque hoy día su profusión es menor que la necesaria para obtener la información que se necesita. Su futuro dependerá en gran medida de los fondos que reciban, la prioridad que se les otorgue, el compromiso de los decisores, el apoyo político *de facto* o el grado de participación activa en ellos de todos los agentes. Y sin olvidar la imaginación y versatilidad que tengamos para idear nuevos diseños y modelos o acomodar los disponibles a necesidades de información práctica.

7. Ciencia regulatoria y ciencia traslacional: ¿se alcanzará un acuerdo?

Bien mirado, y desde una perspectiva internacional, el conjunto de mecanismos de regulación de la investigación, de la aprobación de nuevos medicamentos, aparatos, cirugía y pruebas diagnósticas, de patentes, de vigilancia y evaluación postcomercialización, y de la cobertura y fijación de precios, entre otras acciones, conforman algo rayano en un desbarajuste.

Por ejemplo, y en aras de la brevedad, cada país aplica sus propios criterios para aprobar o no la comercialización de nuevas

La traslación es el proceso dual que consiste en aplicar descubrimientos de la investigación básica a problemas de salud en condiciones controladas (investigación clínica) y la aplicación de resultados prometedores de dicha fase a grupos de población en condiciones controladas o no controladas

entidades moleculares (heterogéneos en conjunto); también varían las funciones, el carácter de las recomendaciones (vinculante o no) y la eficiencia de la labor de las agencias de evaluación de tecnologías vistas desde el prisma internacional. Todos los reguladores evalúan la eficacia, seguridad y calidad de los nuevos fármacos (aunque no todos con los mismos criterios); sólo algunos incluyen en dicha evaluación su coste-efectividad incremental. Los financiadores fundamentan las decisiones de cobertura y fijación de precios, entre otras, sobre la base de maximizar beneficios en salud con presupuestos restringidos. En dichos criterios, así como en los factores de gobiernan las decisiones, también se aprecian marcadas variaciones entre países y en un mismo país en el transcurso de tiempo (53, 60). En otros países se desconocen, por qué no se explican abiertamente.

De esta variabilidad, y de la incertidumbre consiguiente —que se suma a la procedente de la información que alimenta las decisiones— resultan externalidades que se interponen en el proceso de traslación de conocimiento.

La conciencia despertada sobre este entramado de problemas ha dado origen hace algunos años a la ciencia regulatoria (*regulatory science*). Se entiende por ella la aplicación del método científico al desarrollo de nuevos enfoques y herramientas dirigidas a mejorar, por ejemplo, el desarrollo, la revisión y la vigilancia de nuevos medicamentos (incluidos los biológicos) y dispositivos médicos, que requieren aprobación previa a su adopción y difusión. En ella confluyen numerosas disciplinas. Incluye como funciones básicas la evaluación preclínica de su eficacia y seguridad, la realización de ensayos clínicos, la vigilancia postcomercialización de su efectividad, seguridad y uso adecuado, y la elaboración y aplicación de políticas regulatorias pertinentes (35, 45).

De acuerdo con otra definición, es el estudio sistemático de la estructura y el comportamiento de la regulación mediante la observación y la experimentación, para conocer el impacto de normas, principios y leyes que regulan la investigación de esas actividades (35, 45).

Por traslación se entiende el proceso dual que consiste en aplicar descubrimientos de la investigación básica a problemas de salud en condiciones controladas (investigación clínica) y la aplicación de resultados prometedores de dicha fase a grupos de población en condiciones controladas o no controladas (1, 61, 63).

Partiendo de estas definiciones genéricas, es fácil comprender tres hechos y los retos que plantean. Primero, los innegables vínculos entre ciencia traslacional y ciencia regulatoria y la necesidad de armonizarlas. Segundo, la ciencia regulatoria debería desempeñar un papel determinante en la regulación de la investigación básica y la traslación de sus resultados en las dos fases señaladas. Y, tercero, el potencial que atesora su adecuada armonización para mejorar la traslación de información y conocimiento, así como para reducir, en definitiva, la brecha entre eficacia y efectividad.

Si no somos capaces de actuar partiendo de una fraguada visión integral e integradora de las partes que se han intentado hilvanar aquí, continuaremos afrontando altos costes de oportunidad a escala social y pagando el precio inflado de la ineficiencia, también social.

Conclusiones

Aunque ímprobos, los esfuerzos dirigidos a mejorar e intensificar la traslación de información y conocimiento a la clínica siguen siendo insuficientes. Aquí se ha pasado revista a algunos de los principales factores que, formando un complejo entramado de relaciones entre ellos, contribuyen a explicar su lento avance. No hay una explicación simple y lineal a un problema poliédrico, no sólo por sus múltiples facetas y aristas, sino, además, por su distinta naturaleza y sus dispersos orígenes.

Contamos con teorías y modelos científicos. Aunque menos de lo deseado, se ejercita el rigor metodológico. Lo cuantitativo parece prevalecer en valor otorgado sobre lo cualitativo. No es suficiente. La traslación y la toma de decisiones efectivas en la esfera de la salud y lo sanitario claman por ello, así como por idear nuevos enfoques, diseños y métodos, asumir el carácter consustancial de la subjetividad en la ciencia, hacer uso del juicio clínico, canalizar los sesgos cognitivos en el sentido deseable, convencernos y convencer a los demás de la primacía de la eficiencia y el bienestar sociales como uno de los propósitos postreros del avance científico y que sin políticas ni incentivos consonantes con todo ello se seguirá progresando a ritmo lento.

Si no somos capaces de actuar partiendo de una fraguada visión integral e integradora de las partes que se han intentado hilvanar aquí, continuaremos afrontando altos costes de oportunidad a escala social y pagando el precio inflado de la ineficiencia, también social.

Referencias bibliográficas

1. Simon R. *Lost in translation: problems and pitfalls in translating laboratory observations to clinical utility.* *Eur J Cancer* 2008;44:2707-13.
2. Mansky CF. *Public Policy in an uncertain world. Analysis and decisions.* Cambridge, MA: Harvard University Press; 2013.
3. King D, Greaves F, Vlaev I, et al. *Approaches based on behavioral economics could help nudge patients and providers toward lower health spending growth.* *Health Aff* 2013;32:661-8.
4. Croskerry P. *From mindless to mindful practice — Cognitive bias and clinical decision making.* *N Engl J Med* 2013;368:2445-8.
5. Kahneman D. *Thinking, fast and slow.* New York: Farrar, Straus and Giroux; 2011.
6. Tunis SR, Stryer DB, Clancy CM. *Practical clinical trials. Increasing the value of clinical research for decision making in clinical and health policy.* *JAMA* 2003;290:1624-32.
7. Rawlings M. *What constitutes credible evidence of effectiveness? 19th Annual Lecture.* London: Office of Health Economics; 2012. <http://www.ohe.org/object/display.cfm?serv=2&id=131#131>
8. Campillo C. *When health technologies do not reach their effectiveness potential: A health services research perspective.* *Health Policy* 2012;104:92-8.
9. Wahr JA, Prager RL, Abernathy JH, et al. *Patient safety in the cardiac operating room: human factors and teamwork. A scientific statement from the American Heart Association.* *Circulation* 2013;128:1-31.
10. Campillo-Artero C. *Proyectos que fracasan en los servicios de salud: una aproximación etiológica.* *Gest Clín Sanit* 2007;9:43-7.

11. *Pronovost PJ, Goeschel CA, Olsen KL, et al. Reducing health care hazards: lessons from the commercial aviation safety team. Health Aff 2009;28:w479-89.*
12. *Berger VW. Does the Prentice criterion validate surrogate endpoints? Stat Med 2004;23:1571-8.*
13. *Kerche la Cour J, Brok J, Gotsche P. Inconsistent reporting of surrogate outcomes in randomized clinical trials: cohort study. BMJ 2010;341:c3653 doi: 10.1136/bmj.c3653.*
14. *Montori VM, Permanyer-Miranda, G, Ferreira-González I, et al. 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. *Mathew S, Bouton I, Moher D, et al. Comparison of registered and published primary outcomes in randomized controlled trials. JAMA 2009;302:977-84.*
20. *Ordúñez P, Campillo C, eds. Organización Panamericana de la Salud. Consulta regional: prioridades para la salud cardiovascular en las Américas. Mensajes claves para los decisores. Washington, DC: OPS; 2011.*
21. *Ocana A, Tannock IF. When are "positive" clinical trials in oncology truly positive? J Natl Cancer Inst 2011;103:16-20.*

22. Ciani O, Buyse M, Garside R, et al. Comparison of treatment effect sizes associated with surrogate and final patient relevant outcomes in randomised controlled trials; meta-epidemiological study. *BMJ* 2013;346:f457 doi: 10.1136/bmj.f457.
23. Kaufmann D, Kraay A, Mastruzzi M. Governance Matters VIII. Aggregate and individual governance indicators, 1996-2008. Policy Working Paper 4978. Washington, DC: The World Bank; 2009.
24. Meneu R. Evaluación del buen gobierno sanitario. Algunas recomendaciones para su mejora. En: Ortún V, Dir. El buen gobierno sanitario. Madrid: Springer Healthcare Communications; 2009.
25. Higgins JPT, Green S, eds. *Cochrane Handbook for Systematic Reviews of Interventions*. Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Disponible en: www.cochrane-handbook.org.
26. Waters E, Priest N, Armstrong R, et al. 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.
27. Faggiano F, Vigna-Taglianti F. Systematic reviews of effectiveness of Public Health practice. *Ital J Public Health* 2006;3:29-33.
28. 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.
29. 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.
30. Bero LA, Glantz SA, Rennie D. Publication bias and public health policy on environmental tobacco smoke. *JAMA* 1994;272:133-6.

31. *Herxheimer A. Antidepressants and adverse effects in young patients: uncovering the evidence. CMAJ 2004;17:487-9.*
32. *The Lancet. Depressing research. [Editorial] Lancet 2004;363:1335.*
33. *Whittington CJ, Kendall T, Fonagy P, et al. Selective serotonin reuptake inhibitors in childhood depression: systematic review of public versus unpublished data. Lancet 2004;363:1341-5.*
34. *Hebert RS, Wright SM, Dittus RS, et al. Prominent medical journals often provide insufficient information to assess the validity of studies with negative results. J Negative Res Biomed 2002;1:1.*
35. *Institute of Medicine. Building a national framework for the establishment of regulatory science for drug development: Workshop summary. Washington, DC: The National Academies Press; 2010.*
36. *Institute of Medicine. Strengthening a workforce for innovative regulatory science in therapeutics development: Workshop summary. Washington, DC: the National Academies Press; 2011.*
37. *Schwappach DLB, Boluarte TA, Suhrcke M. The economics of primary prevention of cardiovascular disease — a systematic review of economic evaluations. Cost effectiveness and resource allocation 2007;5:5 Doi: 10.1186/1478-7547-5-5.*
38. *Bell CM, Urbach DR, Ray JG, et al. Bias in published cost-effectiveness studies: systematic review. BMJ 2006;332:699-701.*
39. *Catalá-Lopez F, Alonso-Arroyo A, Aleixandre-Benavent R, et al. Coauthorship and institutional collaborations on cost-effectiveness analyses: A systematic network analysis. PLoS ONE 7(5): e38012. doi:10.1371/journal.pone.0038012.*
40. *Hillman AL, Eisenberg JM, Pauly MV, et al. Avoiding bias in*

the conduct and reporting of cost-effectiveness research sponsored by pharmaceutical companies. NEJM 1991;324:1362-5.

41. *Fenwick E, O'Brien BJ, Briggs A. Cost-effectiveness acceptability curves — facts, fallacies and frequently asked questions. Health Econ 2004;13:405-15.*
42. *Fenwick E, Marshall DA, Levy AR, et al. Using and interpreting acceptability curves: an example using data from a trial of management strategies for atrial fibrillation. BMC Health Serv Res 2006;6:52.*
43. *Eckermann S, Willan AR. Expected value of information and decision making in HTA. Health Econ 2007;16:195-209.*
44. *Griffin SC, Claxton KP, Palmer SJ, et al. Dangerous omissions: the consequences of ignoring decision uncertainty. Health Econ 2011;20:212-24.*
45. *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.*
46. *Campillo-Artero C, Bernal-Delgado E. Reinversión en sanidad: fundamentos, aclaraciones, experiencias y perspectivas. Gac Sanit 2013;27:175-9.*
47. *García-Armesto S, Campillo-Artero C, Bernal-Delgado E. Disinvestment in the age of cost-cutting sound and fury. Tools for the Spanish National Health System. Health Aff 2013;110:180-15.*
48. *Centre for health economics. Value-based pricing for pharmaceuticals: Its role, specification and prospects in a newly devolved NHS. York: University of York; 2011.*
49. *Claxton K, Longo R, Longworth L, et al. The value of innovation. Centre for Health Economics, University of York, Leeds University, School of Health and Related Research, University of Sheffield; 2009.*

50. Campillo-Artero C. Sesgos de publicación, valor de la información y su efecto en las políticas de salud. *Rev Cub Salud Pública* 2012;38(Supl 5):714-24.
51. Glasgow RE, Lichtenstein E, Marcus AC. Why don't we see more translation of health promotion research to practice? Rethinking the efficacy-to-effectiveness transition. *Am J Public Health* 2003;93:1261-7.
52. Feudtner Ch, Schreiner M, Lantos JD. Risks (and benefits) in comparative effectiveness research trials. *N Engl J Med* 2013;
53. Chandra A, Jena AB, Skinner JS. The pragmatist's guide to comparative effectiveness research. *J Econ Persp* 2011;25:27-46.
54. Timbie JW, Schneider EC, Van Busum K, et al. Five reasons that many comparative effectiveness studies fail to change patient care and clinical practice. *Health Aff* 2012;31:2168-75.
55. Giffin RB, Woodcock J. Comparative effectiveness research: who will do the studies. *Health Aff* 2010;29:2075-81.
56. Iglehart JK. The political fight over comparative effectiveness research. *Health Aff*. Doi: 10.1377/hlthaff.2010.0901.
57. Schroeder SA, Frist W. Phasing out fee-for-service payment. *N Engl J Med* 2013,368:2029-32.
58. Prasad V, Cifu A, Ioannidis JPA. Reversals of established medical practices. Evidence to abandon ship. *JAMA* 2012;307:37-8.
59. Prasad V, Gall V, Cifu A. The frequency of medical reversal. *Arch Intern Med* 2011;171:1675-6.
60. Asociación de Economía de la Salud. Bernal E, Campillo C, González López-Valcárcel B, Meneu R, Puig-Junoy J, Repullo JR, y Urbanos R (2011). *La sanidad pública ante la crisis. Recomendaciones para una actuación pública sensata y respon-*

sable. Barcelona: Asociación de Economía de la Salud. http://aes.es/Publicaciones/DOCUMENTO_DEBATE_SNS_AES.pdf

61. Hauser AS, Johnston C. Translational research for a new administration: what sort of change to believe in? *Ann Neurol* 2008;64:A5-6.
62. Roman J. Creating a culture of discovery through clinical trials and translational research. *Am J Med Sci* 2009;337:155.
63. Lean MEJ, Mann JI, Hoek JA, et al. Translational research. From evidence based medicine to sustainable solutions for public health problems. *BMJ* 2008;337:705-6.