

## COLABORACIÓN ESPECIAL

### CUESTIONES CONTROVERTIDAS EN EVALUACIÓN ECONÓMICA (II): MEDIDAS DE RESULTADO EN SALUD DE LAS INTERVENCIONES SANITARIAS

Marta Trapero-Bertran (1,2), Max Brosa Riestra (3), Jaime Espín Balbino (4), Juan Oliva (1) y Key4Value-Grupo II\*.

(1) Universidad de Castilla-La Mancha. Toledo.

(2) Universitat Pompeu Fabra. Barcelona.

(3) Oblikue Consulting. Barcelona.

(4) Escuela Andaluza de Salud Pública. Granada.

Key4Value-Grupo II está formado por las siguientes personas: José María Abellán (Universidad de Murcia), Juan Manuel Cabasés (Universidad Pública de Navarra), Pere Clavé (Hospital de Mataró y Universitat Autònoma de Barcelona), Montserrat Figueras (Novartis Farmacèutica), Anna García-Altés (Agència de Qualitat i Avaluació Sanitàries de Catalunya, AQUAS), Salvador Peiró (Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunitat Valenciana, FISABIO), Aurelio Tobías (Consejo Superior de Investigaciones Científicas) y Joan Ramón Villalbi (Departament de Salut, Generalitat de Catalunya).

Declaración de conflicto de intereses: este trabajo ha sido financiado por Novartis. Las personas que han colaborado en el mismo han recibido honorarios, bien por la coordinación del trabajo bien por la asistencia a las reuniones de trabajo y debate.

#### RESUMEN

En este segundo artículo, de una serie de tres, debatiremos, empleando la técnica Metaplan, sobre cuestiones controvertidas de las medidas de resultado en salud en la evaluación económica de intervenciones sanitarias (EEIS). Las cuatro áreas de debate se centraron en: elección de las medidas de resultado en salud: en abstracto ninguna medida es superior a otra. Extrapolación y transferibilidad de las medidas de resultado en salud: los resultados de una EEIS de un país no deberían ser asumidos en otro sin realizar ciertos ajustes de distinto tipo. Instrumentos adecuados para medir la calidad de vida en España: se apuntó el EQ-5D como conveniente debido a su extendido uso internacional. Y comparaciones indirectas: la combinación de ambas comparaciones, directas e indirectas, sería recomendable si la prueba para las estimaciones indirectas fue consistente y estaba validada. Finalmente, se proponen líneas de investigación para tratar de superar las discrepancias identificadas en cada una de estas áreas, entre las cuáles se encuentran: la realización de estudios de correlación entre puntuaciones de instrumentos específicos y genéricos de medición de calidad de vida, actualizar o crear una base de datos de evaluaciones económicas realizadas en España, la estimación de los conjuntos de utilidades de la población española para instrumentos genéricos y específicos preexistentes o bien, establecer una manera común de mostrar los resultados de un meta-análisis en red.

**Palabras clave:** Economía. Evaluación de resultados. Costes. Costos de la atención en salud. Economía y organización de los cuidados de salud. Bioética. Consenso. Evaluación económica de intervenciones sanitarias. Calidad de vida. Resultados en salud.

#### Correspondencia

Marta Trapero-Bertran  
Universidad de Castilla-La Mancha  
Facultad de Ciencias Jurídicas y Sociales de Toledo  
Análisis Económico y Finanzas  
Coberitzo de San Pedro Mártir s/n  
45071 Toledo  
marta.trapero@uclm.es

DOI:

#### ABSTRACT

### Controversial Issues in Economic Evaluation (II): Health Outcomes of Health Care Interventions

In this second article of a series of three, we will discuss using the Metaplan technique on controversial issues of health outcomes in economic evaluation of health care interventions. The four-discussion areas focus on: choice of health outcomes measures, where any outcome measure is superior to another; extrapolation and transferability of health outcomes measures, which should not be assumed the results of an EEIS of one country to another without making certain adjustments; appropriate instruments to measure quality of life in Spain, where the EQ-5D was indicated as convenient due to its widespread international use; and, indirect comparisons, where the combination of both comparisons, direct and indirect, it would be advisable if the test for indirect estimates is consistent and has been validated. Finally, research lines to try to overcome the identified discrepancies were identified in each of these areas, some of those are: doing studies of correlation between scores of specific and generic instruments measuring quality of life; update or create a database of economic evaluations in Spain; estimating utilities for the Spanish population by existing generic and specific instruments; or, establish a common way to show the results of a meta-analysis network.

**Keywords:** Economics. Outcome Assessment (Health Care). Health care costs. Health Care Economics and Organizations. Bioethics. Methods. Costs. Economic Evaluation of Health Interventions. Quality of life. Health outcomes.

## INTRODUCCIÓN

Las medidas de resultado en salud son un elemento clave en las evaluaciones económicas de intervenciones sanitarias (EEIS). La elección de la medida a emplear depende de la pregunta de investigación y viene determinada por la perspectiva utilizada y por el tipo de evaluación que prevea realizarse: análisis coste-beneficio (metodología que monetiza los resultados), análisis coste-efectividad (que expresa los resultados en unidades clínicas o naturales) o análisis coste-utilidad (que emplea medidas que combinan la calidad y cantidad de vida).

En este segundo trabajo, de una serie de tres, se debate sobre las controversias metodológicas en las medidas de resultado en salud, uno de los dos elementos básicos de las EEIS junto a la valoración de los costes, centrándose en cuatro cuestiones principales. La primera cuestión gira en torno a la elección de las medidas de resultado en salud más adecuadas para cada enfermedad y la posibilidad de llegar a un consenso para elaborar una lista de las mismas a nivel español. Existen tres medidas principales de resultados en salud: clínicas y naturales (mortalidad y morbilidad, esperanza de vida), calidad de vida y valores monetarios como la disposición a pagar por las mejoras en la salud de un servicio o de un cambio en el estado de salud<sup>1</sup>. La forma más simple de medir los resultados en un ensayo clínico es utilizando unidades naturales, aunque dificulta las comparaciones entre diferentes condiciones en salud. Esto se soluciona parcialmente cuando incorporamos la mortalidad como medida de resultado (años de vida ganados), si bien no deja de ser un recurso insatisfactorio ante intervenciones que afectan principalmente a la dimensión de la calidad de vida. Por este motivo, existen las medidas de resultado que incorporan la calidad de vida relacionada con la salud (CVRS), que van más allá de la medida clínica y la mortalidad. En concreto, el año de vida ajustado por calidad (AVAC), índice cardinal

que combina los años de vida con la utilidad o bienestar derivado de la calidad con que estos se disfrutan, es la medida de resultado recomendada en buena parte de las guías de EEIS. También existen las medidas de resultado basadas en el paciente<sup>2</sup>, definidas como medidas a corto plazo que comprenden instrumentos como cuestionarios, entrevistas y otros métodos relacionados con la medida de la salud, la enfermedad y los beneficios de las intervenciones en salud desde la perspectiva del paciente.

La segunda cuestión aborda el tema de la extrapolación y transferibilidad de las medidas de resultado en salud de estudios realizados en otros países, un aspecto muy relevante debido al gran número de evaluaciones económicas existentes realizadas fuera de España. Los métodos y resultados de un estudio son transferibles si (a) los potenciales usuarios pueden valorar su aplicabilidad en su contexto y (b) si son aplicables en ese contexto<sup>3</sup>. La mayoría de guías metodológicas para la realización de EEIS abordan el tema de la transferibilidad, si bien lo hacen con diferente grado de profundidad y detalle<sup>4</sup>. Generalmente, las guías reconocen las potenciales diferencias en los parámetros clínicos de un país a otro y destacan la posibilidad de que esto lleve a diferencias en los resultados de las EEIS<sup>4</sup>. También se ha sugerido que la variabilidad en los datos de costes o resultados en salud se deben más a diferencias metodológicas entre los estudios que a diferencias entre países<sup>5</sup>. Un elemento destacado es que los datos de estudios que comprenden varios países infravaloran, de manera importante, la variabilidad a nivel de cada país<sup>5</sup>.

En referencia a los métodos más adecuados de transferibilidad de resultados, algunos autores argumentan que cuando los métodos para derivar las utilidades se replican, no parece que haya demasiadas diferencias entre los diferentes ámbitos geográficos<sup>4</sup>. Cuando hay datos individuales de pacientes pero el ámbito geográfico de interés no par-

ticipa en el ensayo, se sugiere que el mejor método es desarrollar un “modelo basado en eventos”, construido alrededor de las características generalizables de la enfermedad o del pronóstico de los pacientes y utilizar la información del ensayo para estimar la probabilidad de ocurrencia de los eventos de interés que se espera que tengan un impacto en el uso de los recursos y/o en los resultados en salud<sup>6</sup>. Un método común, de utilización creciente, es aplicar la reducción en el riesgo relativo del ensayo clínico en los eventos de interés observados en el mismo (por ejemplo, reducción relativa en el riesgo de muerte, infarto de miocardio, eventos adversos,...) al riesgo basal en el ámbito geográfico de interés, a no ser que haya alguna buena razón para creer que esa reducción en el riesgo relativo de otro ámbito geográfico en particular sea más relevante. De hecho, hay que destacar que la mayoría de guías sugieren que ámbitos geográficos específicos estimen los riesgos basales para utilizar en las evaluaciones económicas y que, de esta manera, el efecto del tratamiento observado en los ensayos clínicos (por ejemplo, la reducción en el riesgo relativo) podría así ser más generalizable entre diferentes contextos<sup>4</sup>.

Existe también la posibilidad de obtener valores agregados para una jurisdicción que abarque a otras, a partir de valores obtenidos en jurisdicciones menores<sup>7</sup>. Otros trabajos también indican que las diferencias se pueden explicar, en gran medida, por las variaciones metodológicas de los estudios de valoración<sup>8</sup>. Asimismo, la literatura reciente se ha concentrado en la aplicación de modelos multinivel para derivar una estimación común de efectos aleatorios para todas las jurisdicciones de interés, equivalente a la estimación obtenida a través de un meta-análisis de efectos aleatorios<sup>6,9</sup>. Por otra parte, algunos autores<sup>3,10</sup> han desarrollado listas de verificación (check-lists) para evaluar el nivel de explicación sobre la transferibilidad de la información o han elaborado mapas o gráficos que ayudan a decidir si una EEIS es

o no transferible<sup>11</sup>. Merece la pena destacar la relevante aportación de investigadores españoles en este campo, con la elaboración de un índice en dos fases para medir el grado de transferibilidad de los resultados de los estudios de EEIS<sup>12</sup>. Finalmente, otros autores sugieren una lista de factores críticos y no críticos para determinar la transferibilidad potencial de las evaluaciones económicas<sup>13</sup>.

La tercera cuestión se centra en los instrumentos adecuados para medir la calidad de vida relacionada con la salud en España y su incorporación a las EEIS. Varias guías metodológicas internacionales especifican el método por el cual las estimaciones de las utilidades deberían obtenerse, sugiriendo que es preferible un instrumento genérico<sup>14</sup>. Ya hace años se trató de identificar aquellos instrumentos más destacados en las diferentes áreas y especialidades médicas<sup>2</sup>. Se ha afirmado también que el SF-36 es uno de los instrumentos de CVRS más conocidos y utilizados a escala internacional<sup>15</sup> y que ilustra la madurez que ha alcanzado en nuestro entorno la medición y estudio de la CVRS. Grupos de investigación muy activos en España han realizado contribuciones importantes al desarrollo general de este instrumento. Sin embargo, uno de los posibles puntos débiles del uso del SF-36 en España, así como de muchos otros instrumentos de CVRS, es que se trata de una herramienta cuyo contenido ha sido desarrollado íntegramente en el contexto cultural de Estados Unidos y que no puede utilizarse en evaluaciones económicas, por tratarse de un perfil de salud cuyas dimensiones no pueden combinarse en un único índice si no es mediado por el instrumento SF-6D<sup>16</sup>. Por otro lado, el EQ-5D es un instrumento genérico de CVRS creado para obtener un índice de salud útil para la evaluación económica. Su uso está muy extendido en toda Europa<sup>17-18</sup> y, recientemente, se ha desarrollado en España un conjunto de valores de la nueva versión EQ-5D-5L que además ha sido incorporada a la Encuesta Nacional de Salud (ENS).

Por último, la cuarta y última cuestión afronta el papel de las comparaciones indirectas<sup>19-20</sup> en las medidas de resultado para su incorporación a las EEIS. Quienes toman las decisiones en salud se enfrentan a brechas significativas entre las necesidades de la investigación comparativa de efectividad y seguridad y la limitada disponibilidad de datos comparativos<sup>21</sup>. La brecha es particularmente pronunciada para los nuevos tratamientos, que a menudo se introducen sin la información aportada por ensayos aleatorios contra todas las alternativas clínica y económicamente relevantes. La toma de decisiones basada en la evidencia requiere la comparación de todas las intervenciones alternativas pertinentes. A falta de ensayos controlados aleatorios que incluyan una comparación directa de todos los tratamientos de interés, las comparaciones indirectas de tratamientos y los meta-análisis en red proporcionan evidencia útil para seleccionar con criterio el/los mejor/es tratamiento(s)<sup>19</sup>. En las comparaciones mixtas de tratamientos se combinan los resultados procedentes de comparaciones directas e indirectas, sintetizando así una mayor integración de la evidencia disponible que el meta-análisis tradicional, basado únicamente en comparaciones directas.

Los resultados de una prueba indirecta combinada con la evidencia directa pueden reforzar la evaluación entre los tratamientos comparados<sup>20</sup>. Pero es fundamental cuestionar el uso de los estudios comparativos observacionales si no existen ensayos clínicos aleatorizados disponibles<sup>20</sup>. En este sentido, es importante recordar que en los meta-análisis en red de ensayos clínicos aleatorizados, el valor de la asignación al azar no se sostiene entre los ensayos. Por lo tanto, las comparaciones indirectas y los meta-análisis en red son de hecho una fuente de evidencia observacional, sujeta también a posibles sesgos de confusión. Existen varias guías de verificación de buenas prácticas en investigación para reportar resultados de meta-análisis en red<sup>19-20</sup>. La

estandarización de los métodos aumentaría la credibilidad general y aplicabilidad de las comparaciones de tratamientos indirectos y meta-análisis en red<sup>19</sup>. Las comparaciones indirectas y el meta-análisis en red presentan problemas metodológicos<sup>22</sup> y algunos autores sostienen que la estimación de comparaciones indirectas debería realizarse a partir de meta-análisis de efectos aleatorios<sup>23</sup> y otros afirman lo contrario<sup>24</sup>. También es discutible cuándo usar la evidencia indirecta y cómo evaluar su validez interna y externa, el posible desequilibrio en la distribución de los modificadores de efecto entre estudios y comparaciones, la inconsistencia estadística entre las comparaciones indirectas y la existencia de sesgos en los estudios individuales y en los de revisión.

El objetivo de este segundo estudio de la serie fue plantear estas cuatro cuestiones controvertidas en materia de resultados en salud en las EEIS a un grupo de expertos con el fin de identificar puntos de acuerdo y desacuerdo y proponer líneas de investigación futuras que ayuden a resolver estas cuestiones.

## MATERIAL Y MÉTODO

En este trabajo se siguió la metodología ya descrita en el primero de los artículos de la serie<sup>25</sup>. Se identificaron y agruparon en bloques temáticos los elementos controvertidos de mayor interés para los debates, teniendo cada grupo un coordinador, y se identificaron los tres grupos de expertos que formarían parte de las reuniones. En esta segunda sesión hubo 10 participantes. Se llevó a cabo una revisión narrativa sobre los principales aspectos relacionados con los resultados en salud y se utilizó la técnica Metaplan para debatir los principales temas identificados previamente en la revisión de la literatura. Finalmente, se acordaron las líneas de investigación a desarrollar que se describen en el apartado de líneas de avance.

## COMENTARIOS

Frente a la cuestión de la elección de las medidas de resultado en salud y la elaboración de una lista consensuada de las mismas, los expertos constataron la relevancia potencial de crear un listado de medidas asociadas a estados de salud propios de diferentes enfermedades, aunque su viabilidad no fue claramente percibida de manera unánime. En general, la elección de la medida a utilizar se asoció al objetivo del estudio en cuestión (por ejemplo, si el estudio trata de un análisis de coste-utilidad, la medida del resultado debería basarse en preferencias), sin descartar la validez de medidas de resultado específicas e incluso el uso de variables clínicas en los estudios de coste-efectividad. De hecho, se consensuó que, en abstracto, ninguna medida de resultado es superior a otra. La elección depende del objetivo de la evaluación económica y de los condicionantes de factibilidad del estudio.

A partir de un trabajo relevante previo<sup>26</sup>, se llegó a la clasificación de medidas de resultado en salud para las EEIS de la **tabla 1**. Incluso, se suscribieron las reglas que ese estudio estableció en su momento para elegir el tipo de medida de resultado más adecuada. Por ejemplo, si la mortalidad, la morbilidad o los sucesos adversos son el resultado relevante, se deberían usar las medidas clásicas (adecuadas a la intervención) para realizar el análisis coste-efectividad. En cambio, se deberían usar los instrumentos de CVRS validados para realizar un análisis coste-utilidad si hay otros resultados de interés, además de la mortalidad o morbilidad, o si la mortalidad no es lo relevante o se trata de pacientes con dependencia importante o mal estado de salud permanente y la sensibilidad a los cambios es importante. Para obtener un índice de utilidad se tiene que haber definido antes un conjunto de estados de salud a partir de un instrumento de CVRS, sea genérico o específico<sup>27-29</sup>. Posteriormente, se pueden utilizar diferentes técnicas para medir las preferencias, solas o en

combinación (en su acepción anglosajona *time trade-off*, *standard gamble*, *rating scales*, etcétera.). La escala entre los valores 0 y 1 que toman las utilidades es lo que hace comparables las distintas medidas.

Sobre la segunda cuestión, extrapolación y transferibilidad de las medidas de resultado en salud, los participantes constataron que los ajustes a realizar en una EEIS realizada en otro país para su aplicación a España serían de la misma naturaleza que los necesarios para transferir los resultados de estudios realizados en nuestro país a terceros. Se apuntó que la transferibilidad de los resultados de las EEIS depende del contexto del análisis y las características de los pacientes y la atención que reciben en distintos ámbitos y, por ello, no deberían ser asumidos los resultados de una EEIS de un país en otro sin realizar ajustes de distinto tipo. Sin embargo, los participantes reconocieron que en muchas ocasiones estos ajustes no se realizan, asumiendo los resultados originales en nuestro medio. En cuanto a la identificación de los ajustes a realizar, además de la utilización de costes y valores de utilidad locales, los participantes constataron la necesidad de ajustar los modelos de EEIS con datos de riesgos basales, características epidemiológicas y sociodemográficas específicas del ámbito de decisión de la EEIS, aunque se apuntó que en el caso de la adaptación de resultados de EEIS de otros países a España existen diversas dificultades: (a) existencia de variaciones entre distintas áreas geográficas dentro de España; (b) ausencia de datos y/o dispersión y dificultad de acceso a las fuentes de los mismos; y (c) poco conocimiento de métodos claramente establecidos para realizar estos ajustes<sup>4</sup>. En cualquier caso, es importante hacer uso de técnicas estadísticas para explorar la heterogeneidad y, una vez explorada, utilizar modelos de efectos fijos o modelos multinivel para incorporarla o tenerla en cuenta en las EEIS<sup>4</sup>. Los expertos reconocieron que uno de los retos es localizar datos de riesgo basal para un ámbito geográfico específico. En

**Tabla 1**  
**Clasificación de medidas de resultado en salud para las EEIS**

**Medidas clínicas clásicas (análisis coste-efectividad)**

- 1) Supervivencia/mortalidad. Intrahospitalaria, en UCI, a los “n” días del episodio o ingreso o alta o intervención, supervivencia a los “n” años del diagnóstico, por causas específicas, etcétera.
- 2) Mortalidad/Supervivencia (*life-time*): años de vida ganados, años de vida sin discapacidad, mortalidad evitable.
- 3) Morbilidad evitada: Morbilidad específica que la intervención intenta reducir (por ejemplo eventos coronarios).
- 4) Sucesos adversos: reingresos, infecciones nosocomiales, patología yatrogénica; reacciones adversas a fármacos, complicaciones post-quirúrgicas u otras, reintervenciones, recidivas, etcétera.
- 5) Estabilidad parámetros fisiológicos y *surrogate endpoints*: Normalización/mejora de signos clínicos (pulso, temperatura, presión arterial, etc.) o parámetros bioquímicos, etcétera.

*Medidas de estado funcional*

- 6) Estado funcional general: capacidad para el esfuerzo, capacidad para las actividades de la vida diaria (básicas, instrumentales, etcétera)
- 7) Estado funcional capacidades/órganos/sistemas específicos: visuales (agudeza visual), movilidad articulaciones o miembros (escalas de cadera, rodilla, brazo, etcétera), respiratorias (pruebas funcionales respiratorias), audición (audiometría), estado nutricional, etcétera.
- 8) Función psicosocial: estado cognitivo, ansiedad, depresión, estrés, etcétera.

**Medidas de preferencias - Capacidad de vida relacionada con la salud (análisis coste-utilidad)**

- 9) Medidas genéricas de capacidad de vida relacionada con la salud
  - a. Perfiles de salud: *Medical Outcomes Study Short Form 36* (MOS-SF36), SF-12, *Nottingham Health Profile* (NHP), *Sickness Impact Profile* (SIP).
  - b. Índices de Salud: *Index of Well Being*, EQ-5D-3L, EQ-5D-5L, SF6D, HUI I, HUI II, HUI III, EHS y otros.
  - c. Medidas específicas de capacidad de vida relacionada con la salud
    - c.1. De enfermedad: *Quality of Life after Myocardial Infarction*, *Pediatric Asthma Quality of Life Questionnaire*.
    - c.2. De subpoblación: *Pediatric Quality of Life Inventory*, *Hospice Quality of Life Index*.
    - c.3. De una función: *Psychosocial Adjustment Illnes Scale*, *Urinary Incontinence-Specific QL Instrument*
    - c.4. De un problema o síntoma: *Brief Pain Inventory*, *Hospital Anxiety Depression Scale (HADS)*, *Overall Nausea Index*
- 10) Medidas directas de utilidad: utilidades obtenidas de forma directa de la población.
- 11) Medidas de CVRS específicas con índices de utilidad: mapping a una escala de preferencias genérica o estimación de la tarifa de utilidades para esas medidas.

Fuente: elaborado a partir de Peiró<sup>26</sup>

cualquier caso, los expertos estuvieron de acuerdo en que las bases de datos de evaluaciones económicas (por ejemplo, EURON-HEED) necesitarían aportar información para valorar la potencial transferibilidad de los estudios publicados y, así, poder utilizar-se por parte de los decisores sanitarios.

En la tercera cuestión, instrumentos adecuados para medir la calidad de vida en España, la principal controversia se centró en la conveniencia o no de recomendar un sólo instrumento de forma prioritaria, o dejar más de una opción a elección por parte de los analistas. Así, por un lado se constató la adecuación de utilizar instrumentos específicos<sup>30-33</sup> para usos distintos al análisis de coste-utilidad (mencionando como ejemplo el EORTC/QLQ-C30 en la valoración de los efectos de las intervenciones para el cáncer) o la necesidad de considerar los casos espe-

ciales de intervenciones en pacientes con patologías mentales, y por otro, la utilidad de instrumentos genéricos como el SF-6D (SF-36) o el EQ-5D para la realización de EEIS. En cualquier caso, se apuntó el EQ-5D como conveniente debido a su extendido uso internacional, especialmente ahora que se han estimado para España los valores poblacionales de la versión EQ-5D-5L, los cuales pueden ser presumiblemente más sensibles a cambios menores en el estado de salud respecto a su versión antecesora. Finalmente, cabe destacar que se apuntó la posibilidad de ponderar el AVAC en función de posibles características especiales de los pacientes<sup>34</sup>.

Finalmente, en el caso de las comparaciones indirectas, los participantes constataron que, aunque las comparaciones indirectas son una herramienta útil en la EEIS (inclu-

**Tabla 2**  
**Problemas metodológicos y recomendaciones de las comparaciones indirectas y meta-análisis en red según los miembros de Key4value-Grupo II**

Problemas metodológicos	Recomendaciones metodológicas
1. La geometría de la red y el número de comparaciones (entre pares) 2. Uso de distintas medidas de resultado en los estudios disponibles 3. Ausencia de transitividad en los estudios disponibles 4. En intervenciones de salud pública, ignorar proceso de la intervención y/o la información contextual	1. Evaluar de forma explícita el supuesto de transitividad antes de generar la evidencia indirecta y combinarla con la evidencia directa (especialmente en comparaciones de intervenciones de salud pública) 2. Describir de forma apropiada y precisa el método estadístico utilizado. Las técnicas más empleadas son el método de Bucher*, meta-regresión†, metanálisis multivariante‡, y, los modelos jerárquicos bayesianos§. 3. Establecer un ranking de tratamientos basado en la probabilidad de ser la mejor intervención entre las opciones comparadas 4. Justificar apropiadamente el uso del modelo de efectos fijos 5. Identificar la posibilidad de sesgo de publicación 6. Definir un formato estandarizado para la presentación de resultado.

Fuente: elaborado a partir del grupo Key4value-Grupo II

\*1: Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in metaanalysis of randomized controlled trials. J Clin Epidemiology. 1997; 50(6): 683-91.

†: Berkey CS, Hoaglin DC, Mollerster F, Colditz GA. A random-effects regression model for meta-analysis. Stat Med.1995;14(4):395-411.

‡: Hasselblad V. Meta-analysis of multi-treatment studies. Med Decision Making.1998;18(1):37-43.

§: Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. Stat Med. 2004;23(20):3105-3124.

yendo desde tratamientos farmacológicos hasta intervenciones de salud pública), no se utilizan suficientemente, relegando su utilidad al ámbito académico<sup>35</sup>. El debate en relación al papel de las comparaciones indirectas en la EEIS en España se estructuró a partir de un artículo de revisión<sup>36</sup> con el objetivo de identificar otros problemas metodológicos y recomendaciones por parte de los expertos participantes. La tabla 2 muestra los cuatro problemas metodológicos que se identificaron y las ocho recomendaciones metodológicas que se acordaron. En conclusión, las comparaciones indirectas y su extensión en los meta-análisis en red pueden ser especialmente útiles en situaciones en las que existen varios tratamientos que se han comparado frente a un comparador común, especialmente cuando la información de comparaciones directas es escasa o inexistente<sup>37</sup>. La combinación de ambas comparaciones, directas e indirectas, sería recomendable si la prueba para las estimaciones indirectas es consistente y ha sido validada<sup>38</sup>.

### LÍNEAS DE AVANCE

Durante este taller de expertos se identificaron varias líneas de investigación en las que se podría avanzar en el ámbito de la medida de resultados en salud y, por lo tanto, en el rigor de las evaluaciones económicas y la generación de información relevante para la toma de decisiones. Una parte de ellas refleja cuestiones presentes en la literatura internacional, y otras son propias de particularidades del medio sanitario español.

Sobre las medidas de resultado más adecuadas para medir la CVRS asociada a diferentes enfermedades se identificaron cuatro posibles líneas de investigación.

La primera es la creación de un catálogo de buenas prácticas en el uso de las medidas de resultado, incluyendo la necesidad de describir la validez de los instrumentos y de los métodos de obtención de utilidades, la recomendación de utilizar instrumentos

genéricos en lugar de medidas *ad-hoc*, y la realización de estudios de correlación entre puntuaciones de instrumentos específicos y genéricos.

La segunda está relacionada con la aplicación de instrumentos específicos como herramientas de valoración del resultado de intervenciones tanto en gestión clínica como en las EEIS (mediante la definición de algoritmos para la obtención de utilidades). Respecto al uso de este tipo de instrumentos, se discutió sobre los potenciales problemas de un exceso de especificidad, que podría llevar a la ‘medición de lo irrelevante’, por un lado, y de la conveniencia de evitar la doble contabilización de beneficios, por otro.

La tercera es considerar la equidad en pacientes pluripatológicos mediante, por ejemplo, la identificación de medidas que permitan ponderar los efectos de distintas patologías en un mismo paciente.

Y la cuarta, definir lo que es valor en salud, como la consecución de unos resultados en salud para el paciente a través de la definición de un proceso integral diseñado y dirigido a este fin<sup>39</sup>.

Sobre la extrapolación y transferibilidad de las medidas de resultado de estudios realizados en otros países, con el propósito de soslayar parte de estas dificultades, se identificó como posible línea de investigación futura la creación en nuestro país de una base de datos de riesgos basales, ya que se constató que a pesar de que los datos existen su disponibilidad puede estar restringida o ser poco conocida. En cuanto a la utilidad de identificar el grado de transferibilidad (o los ajustes necesarios para garantizar la misma) de los estudios de EEIS realizados en España, los participantes opinaron que lo primero sería crear (o actualizar) una base de datos de este tipo de estudios realizados en España, para después intentar estimar factores de ponderación, según las características de nuestra población, para ajustar las medidas de resultado en nuestro país.

Sobre la cuestión de qué instrumentos son los más adecuados para medir la CVRS en las EEIS en España, parece que es conveniente utilizar el EQ-5D por su uso extendido (que en su versión EQ-5D-5L cuenta además con una norma poblacional para España (ENSE 2011-12) y algunas Comunidades Autónomas (ESCA en Cataluña, Encuesta Canaria de Salud o Encuesta de Salud de Navarra y un conjunto de valores de los estados de salud), si bien dicha utilización es compatible con la aplicación de otros instrumentos (SF-6D) para los que también se dispone de un conjunto de valores estimado en España<sup>40</sup>. Las líneas de investigación más relevantes identificadas fueron la estimación de los conjuntos de utilidades de la población española para instrumentos genéricos y específicos preexistentes, así como el desarrollo en España de nuevos instrumentos para la medida de la CVRS en el marco de las EEIS.

Para concluir, las recomendaciones de líneas de investigación futuras en relación al papel de las comparaciones indirectas en las medidas de resultado incluyen adaptar una guía para la elaboración y evaluación de los meta-análisis en red, desarrollada a partir de guías previas publicadas<sup>37</sup>. Por otro lado, sería necesario acordar una manera común de mostrar los resultados de un meta-análisis en red. Se podría hacer uso de una tabla de posiciones (*league table*) o un diagrama de bosque (*forest plot*)<sup>41</sup>. También es necesario investigar la cuestión de cómo en las comparaciones indirectas se combinan datos de estudios aleatorizados y estudios que no lo son<sup>22</sup>. Finalmente, otra de las líneas de investigación futura<sup>42</sup> es promover y apoyar que en los procedimientos regulatorios y de evaluación de los nuevos medicamentos se incorpore la información de las comparaciones indirectas de los nuevos fármacos comparados con la mejor práctica actual, cuestión que hasta el momento no se ha llevado a la práctica.

## BIBLIOGRAFÍA

1. Torgerson D, Raftery J. Measuring outcomes in economic evaluations. *BMJ*. 1999; 318:1413.
2. Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. *Health Technol Assess*. 1998; 2:1-74.
3. Boulenger S, Nixon J, Drummond M, Ulmann P, Rice S, de Pouvoirille G. Can economic evaluations be made transferable? *Eur J Health Econ*. 2005; 6:334-346.
4. Drummond M, Barbieri M, Cook J, Glick HA, Lis J, Malik F, et al. Transferability of economic evaluations across jurisdictions: ISPOR good research practices task force report. *Value Health*. 2009; 12:409-418.
5. Boehler CEH. Mind the gap! Geographic transferability of economic evaluation in health. PhD Thesis. London: Brunel University; 2013. Disponible en: <http://bura.brunel.ac.uk/handle/2438/7170>.
6. Manca A, Willian AR. Lost in translation: accounting for between-country differences in the analysis of multinational cost-effectiveness data. *Pharmacoeconomics*. 2006; 24:1101-1109.
7. Greiner W, Weijnen T, Nieuwenhuizen M, Oppe S, Badia X, Busschbach J et al. A single European currency for EQ-5D health states: Results from a six-country study. *Eur J Health Econ*. 2003; 4:222-231.
8. Knies S, Evers SM, Candel MJ, Severens JL, Ament AJ. Utilities of the EQ-5D: transferable or not? *Pharmacoeconomics*. 2009; 27:767-79.
9. Pinto EM, Willan AR, O'Brien BJ. Cost-effectiveness analysis for multinational clinical trials. *Stat Med*. 2005; 24:1965-1982.
10. Turner S, Chase DL, Milne R, Cook A, Hicks NJ, Rosten C, et al. The health technology assessment adaptation toolkit: description and use. *Int J Technol Assessment Health Care*. 2009; 25(suppl2):37-41.
11. Welte R, Feenstra T, Jager H, Leidl R. A decision chart for assessing and improving the transferability of economic evaluation results between countries. *Pharmacoeconomics*. 2004; 22:857-876.
12. Antoñanzas F, Rodríguez-Ibeas R, Juárez C, Hutter F, Lorente R, Pinillos M. Transferability indices for health economic evaluations: methods and applications. *Health Econ*. 2009; 18:629-643. Disponible en: <http://www.alde.es/encuentros/antiores/xeea/trabajos/a/pdf/089.pdf>
13. Goeree R, He J, O'Reilly D, Tarride J-E, Xie F, Lim M, Burke N. Transferability of health technology assessments and economic evaluations: a systematic review of approaches for assessment and application. *Clinicoecon Outcomes Res*. 2011; 3:89-104. Disponible en: <http://www.biomedsearch.com/attachments/00/21/93/53/21935337/ceor-3-089.pdf>.

14. Barbieri M, Drummond M, Willke R, Cahncellor J, Jolain B, Towse A. Variability in cost-effectiveness estimates for pharmaceuticals in Western Europe: Lessons for inferring generalizability. *Value Health*. 2005; 8:10-23.
15. Herdman MJ. Reflexiones sobre la medición de la calidad de vida relacionada con la salud en España. *Gac Sanit*. 2005; 19:91-92.
16. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *J Health Econ*. 2002; 21:271-92.
17. Devlin NJ, Krabbe PFM. The development of new research methods for the valuation of EQ-5D-5L. *Eur J Health Econ*. 2013; 14(Suppl1):S1-S3.
18. Ramos-Goñi JM, Rivero-Arias O, Serrano-Aguilar P, López Bastida J, Pinto Prades JL, Cabasés Hita JM et al. Estimación de la tarifa española para evaluar los estados de salud del EQ-5D-5L basadas en las preferencias de los individuos. Madrid: Ministerio de Sanidad y Consumo y Fundación Canaria de Investigación y Salud "FUNCIS" (SESCS N° 2010/05) (en prensa).
19. Hoaglin DC, Hawkins N, Jansen JP, Scott DA, Itzler R, Cappelleri JC et al. Conducting indirect-treatment-comparison and network-meta-analysis studies: Report of the ISPOR Task Force on indirect treatment comparisons good research practices-Part 2. *Value Health*. 2011; 14:429-437.
20. Jansen JP, Fleurence R, Devine B, Itzler R, Barrett A, Hawkins N, et al. Interpreting indirect treatment comparisons and networks meta-analysis for healthcare decision making: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: Part 1. *Value Health*. 2011; 14:417-428.
21. Signotovitch JE, Sikirica V, Erder H, Xie J, Lu M, Hodgkins PS, et al. Matching-adjusted indirect comparison: a new tool for timely comparative effectiveness research. *Value Health*. 2012; 15:940-947.
22. Cipriani A, Higgins J, Geddes JR, Salanti G. Conceptual and technical challenges in network meta-analysis. *Ann Intern Med*. 2013; 159:130-137.
23. Glenny AM, Altman DG, Song F, Sakarovitch C, Deeks JJ, D'Amico R, et al. Indirect comparisons of competing interventions. *Health Technol Assess*. 2005; 9:1-152.
24. ISPOR. ISPOR states its position on network meta-analysis. *Value Health* 2011; 14:414-416.
25. Oliva J, Brosa M, Espín J, Figueras M, Trapero-Bertrán M y Key4Value-Grupo I. Cuestiones controvertidas en evaluación económica (I): perspectiva y costes de intervenciones sanitarias. *Rev Esp Salud Pública*. 2015;89:5-14.
26. Peiró Moreno S. Medición y comparación de resultados en la práctica asistencial. En: Oteo LA, ed. *Gestión Clínica: Desarrollo e instrumentos*. Madrid: Diaz de Santos; 2006.p. 95-144.
27. Brazier JE, Klotkin RL, Crosby RD, Williams GR. Estimating a preference-based index from the Impact of Weight on Quality of Life Instrument (IWQOL-Lite) from the SF-6D. *Value Health*. 2004; 7:490-498.
28. Tsuchiya A, Brazier J, McColl E, Parkin D. Deriving preferencebased condition-specific instruments: converting AQLQ into EQ-5D indices. Sheffield Health Economics Group Discussion Paper 02/01, School of Health and Related Research. Sheffield: University of Sheffield; 2002.
29. Dixon S, McEwan P, Currie CJ. Estimating the health utility of treatment in adults with growth hormone deficiency. *J Outcomes Res*. 2003; 7:1-12.
30. Revicki DA, Leidy NK, Brennan-Diemer F, Thompson C, Togias A. Development and preliminary validation of the multiattribute Rhinitis Symptom Utility Index. *Qual Life Res*. 1998; 7:693-702.
31. Stolk EA, Busschbach JJ. Validity and feasibility of the use of condition-specific outcome measures in economic evaluation. *Qual Life Res*. 2003; 12:363-371.
32. Revicki DA, Leidy NK, Brennan-Diemer F, Sorensen S, Togias A. Integrating patient preferences into health outcomes assessment: the multiattribute Asthma Symptom Utility Index. *Chest*. 1998; 114:998-1007.
33. Kok ET, McDonnell J, Stolk EA, Stoevelaar HJ, Busschbach JJ, Triumph Research Group; Pan-European Expert Panel. The valuation of the International Prostate Symptom Score (IPSS) for the use in economic evaluations. *Eur Urol*. 2002; 42:491-497.
34. Espín J, Brosa M, Oliva J, Trapero-Bertrán M y Key4Value-Grupo III. Cuestiones controvertidas en evaluación económica (III): la evaluación económica de intervenciones sanitarias en tres situaciones especiales: enfermedades raras, los tratamientos al final de la vida y las externalidades en las evaluaciones. *Rev Esp Salud Pública*. 2015;89 (3) (en prensa).
35. Abdelhamid AS, Loke YK, Parekh-Bhurke S, Chen Y, Sutton A, Eastwood A et al. Use of indirect comparison methods in systematic reviews: a survey of Cochrane review authors. *Res Synth Method*. 2012; 3:71-79.
36. Song F, Loke YK, Walsh T, Glenny AM, Eastwood AJ, Altman DG. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. *BMJ*. 2009; 338:b1147. doi: 10.1136/bmj.b1147.
37. Catalá-López F, Tobías A. Síntesis de la evidencia clínica y meta-análisis en red con comparaciones indirectas. *Med Clin (Barc)*. 2013; 140:182-187.

38. Song F, Xiong T, Parekh-Bhurke S, Loke YK, Sutton AJ, Holland R, et al. Inconsistency between direct and indirect comparisons of competing interventions: meta-epidemiological study. *BMJ*. 2011; 343:1-11.
39. Porter ME. What Is Value in Health Care? *NEJM*. 2010; 363:2477-2481.
40. Abellan-Perpiñan JM, Sanchez-Martinez FI, Martinez-Perez JE, Mendez I. Lowering the 'floor' of the SF-6D scoring algorithm using a lottery equivalent method. *Health Economics* 2012; 21:1271-1285.
41. Cipriani A, Barbui C, Salanti G, Rendell J, Brown R, Stockton S, et al. Comparative efficacy and acceptability of antimanic drugs in acute mania: a multiple.treatments meta-analysis. *Lancet*. 2011; 378:1306-1315.
42. Rawlins MD. Comparative effectiveness research: a view from the other side of the pond. *J Comp Eff Res*. 2012; 1:293-295.