

**REVISIÓN CRÍTICA DE LAS GUÍAS DE EVALUACIÓN ECONÓMICA DE
TECNOLOGÍAS SANITARIAS DE AUSTRALIA, CANADÁ, INGLATERRA
Y GALES (NICE), Y SUECIA ^(*)**

Julio 2007

Equipo investigador

Fernando I. Sánchez Martínez

José María Abellán Perpiñán

Jorge E. Martínez Pérez

Grupo de Investigación en Economía de la salud y evaluación económica

Departamento de Economía Aplicada

Universidad de Murcia



^(*) Informe correspondiente a la “Fase 1” del proyecto *Incorporación de la evaluación económica de tecnologías sanitarias al sistema de salud de la Región de Murcia*, desarrollado en el marco del Convenio de Colaboración suscrito con fecha 22 de diciembre de 2006 entre la C.A.R.M.-Consejería de Sanidad y la Universidad de Murcia, en materia de Economía de la Salud, - BORM N° 23, 29/01/2007.



ÍNDICE

Introducción	5
1. Principales cuestiones que deben tratar las guías de evaluación económica	9
1.1. Perspectiva de la evaluación económica.....	9
1.2. Revisión sistemática de la literatura	11
1.3. Población objetivo	16
1.4. Término de comparación de la tecnología evaluada.....	18
1.5. Medida de resultados preferida.....	21
1.6. Costes incluidos en el análisis	23
1.7. Horizonte temporal de los costes y efectos.....	28
1.8. Tasa de descuento	31
1.9. Tratamiento de la incertidumbre.....	33
1.10. Uso de modelos decisionales	36
1.11. Transferibilidad de datos y resultados	41
1.12. Presentación de los resultados	45
2. Revisión de las guías (NICE, Canadá, Australia y Suecia)	49
2.1. Perspectiva de la evaluación económica.....	49
2.2. Revisión sistemática de la literatura	51
2.3. Población objetivo	59
2.4. Término de comparación de la tecnología evaluada.....	64
2.5. Medida de resultados preferida.....	68
2.6. Costes incluidos en el análisis	77
2.7. Horizonte temporal de los costes y efectos.....	82
2.8. Tasa de descuento	87
2.9. Tratamiento de la incertidumbre.....	88
2.10. Uso de modelos decisionales	96
2.11. Transferibilidad de datos y resultados	103
2.12. Presentación de los resultados	111
3. Comentarios sobre las guías y conclusiones.....	119
3.1. Perspectiva de la evaluación económica.....	119
3.2. Revisión sistemática de la literatura	121
3.3. Población objetivo	124
3.4. Término de comparación de la tecnología evaluada.....	126
3.5. Medida de resultados preferida.....	129
3.6. Costes incluidos en el análisis	134
3.7. Horizonte temporal de los costes y efectos.....	136
3.8. Tasa de descuento	138
3.9. Tratamiento de la incertidumbre.....	139
3.10. Uso de modelos decisionales	143
3.11. Transferibilidad de datos y resultados	145
3.12. Presentación de los resultados	149
Referencias bibliográficas.....	153



Introducción

En un entorno en el que la escasez de recursos se revela como una restricción cada vez más perceptible, es esencial elegir entre las distintas alternativas posibles de un modo racional. Esta es la razón por la que la evaluación económica se está convirtiendo en una herramienta imprescindible en la toma de decisiones en el ámbito de la asignación de recursos para el cuidado de la salud.

En la literatura científica relativa a la evaluación económica de tecnologías sanitarias suele utilizarse el término *guidelines* para hacer referencia al conjunto de recomendaciones o directrices de tipo metodológico que, bien según un panel de expertos, bien según el criterio de una agencia evaluadora, deben seguirse para alcanzar un estándar de calidad suficiente.

La revisión que aquí se presenta se ha orientado al segundo caso (guías de agencias evaluadoras), sin que ello haya de interpretarse en el sentido de que el equipo investigador juzgue acertadas todas y cada una de las directrices sugeridas por las agencias. Por esta razón, lo que se ofrece en este informe correspondiente a la “Fase 1” del proyecto no es sólo la descripción en extenso de la citada revisión, sino, lo que es más importante, una valoración crítica de las recomendaciones examinadas, de la que extraer las pautas que inspirarán el desarrollo de la segunda fase del proyecto.



En este informe, por tanto, se aborda una revisión de las que, a nuestro juicio, son las guías de evaluación económica de mayor calidad, que se corresponden con las de aquellos países en los que la evaluación económica ha alcanzado una mayor relevancia¹. Las guías objeto de revisión son las Australia², Canadá³, el NICE⁴ (para Inglaterra y Gales), y Suecia⁵. El principal objeto de esta revisión no es simplemente, como ya se ha señalado, relatar lo que las guías dicen, sino valorar su contenido a la luz de los postulados teóricos.

Las guías seleccionadas tienen en común el que las agencias responsables de las mismas exigen de los fabricantes de tecnologías sanitarias –sólo medicamentos en los casos de Australia, Canadá y Suecia; cualquier tecnología en el caso británico– la presentación de evidencia sobre el carácter coste-efectivo de sus productos. En consecuencia, las guías revisadas son la norma con la que se va contrastar la calidad de la evidencia aportada por el fabricante. Del resultado de esta comparación dependerá, en gran medida, que la tecnología en cuestión sea aceptada o rechazada

¹ La principal omisión atañe a los Países Bajos, pues es éste un país en el que la evaluación económica ha adquirido notable relevancia y en el que, además, existe una importante comunidad científica. La razón para excluir las guías holandesas de nuestra revisión no es otra que el hecho de que las actuales guías se aprobaron en 2006 y no estaban traducidas al inglés en el momento de redactarse estas páginas.

² *Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (Version 4.1)*. Canberra: Pharmaceutical Benefits Advisory Committee: 2006. Disponible en <http://www.health.gov.au/internet/wcms/publishing.nsf/Content/pbacguidelines-index>.

³ *Guidelines for the economic evaluation of health technologies: Canada* [3rd Edition]. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2006. Disponible en http://www.cadth.ca/media/pdf/186_EconomicGuidelines_e.pdf

⁴ *Guide to the methods of technology appraisal*. London: National Institute for Clinical Excellence; 2004. Disponible en <http://www.nice.org.uk/page.aspx?o=201974>.

⁵ General guidelines for economic evaluations from the Pharmaceutical Benefits Board. Pharmaceutical Benefits Board, 2003. Disponible en http://www.lfn.se/upload/English/ENG_lfnar2003-eng.pdf.



para su utilización en el sistema sanitario público. A este sistema de financiación de las tecnologías sanitarias se le conoce con el término “cuarta barrera” (*fourth hurdle*).

El texto del informe se ha articulado en tres secciones. En primer lugar, se exponen las cuestiones más relevantes que deben ser tomadas en consideración a la hora de realizar una evolución económica. Teniendo en cuenta el objetivo final de esta revisión, que no es otro que el de elaborar una *checklist* que permita su usuario valorar la calidad de los ejercicios de evaluación económica, se ha seleccionado un total de 12 fases o aspectos parciales del proceso de evaluación. Entre ellos no se encuentran –y la omisión es deliberada– ni las cuestiones de equidad, ni el análisis del impacto presupuestario, pues consideramos que no son relevantes para el fin que se persigue. En la sección segunda, se resume lo que al respecto de cada una de estas doce cuestiones dicen las guías que se han revisado. Por último, se recogen algunos comentarios a modo de conclusión para cada cuestión analizada.



1. Principales cuestiones que deben tratar las guías de evaluación económica

1.1. Perspectiva de la evaluación económica

Las guías deberían identificar la perspectiva que se ha de utilizar en la evaluación económica, siendo dos las opciones básicas que caben dentro de los sistemas de salud públicos: adoptar la perspectiva del financiador, esto es, del Sistema Nacional de Salud (en adelante, SNS), o adoptar la perspectiva social (Brouwer y Koopmanschap, 2000). Las diferencias más importantes entre ambos enfoques se reflejarán en la parte de los costes, aunque también podría haber consecuencias del lado de los beneficios. La elección entre una u otra perspectiva depende del tipo de objetivos que se plantee el SNS y en esto las guías han de ser muy claras, en el sentido de justificar nítidamente el porqué de recomendar una alternativa o la contraria.

En ocasiones se entiende que el objetivo del SNS es maximizar la salud de la población, sujeto dicho objetivo a la restricción dada por el presupuesto público destinado a financiar la asistencia sanitaria. Partiendo de este objetivo, estaría plenamente justificada la perspectiva del financiador. La consecuencia de optar por esta perspectiva es que no se incluirán todos los costes y beneficios de la tecnología médica que se pretende evaluar. Por el lado de los costes, éstos se limitarán a los



costes sanitarios que recaen sobre el SNS y, a lo sumo, los costes de los servicios sociales (en adelante, SS), tales como los derivados de la atención a pacientes crónicos o con discapacidades. La lógica de incluir los costes de los servicios sociales es que también corren por cuenta del presupuesto público. La perspectiva del financiador obviará, por el contrario, los costes que no recaen sobre el presupuesto, principalmente los que soporta el propio paciente y su entorno y que, en buena medida pueden ser de naturaleza no monetaria (tiempo).

Por el lado de los beneficios, la adopción de la perspectiva del SNS supondrá limitar la medida de los efectos a los beneficios estrictamente sanitarios. Por ejemplo, no se consideraría como beneficio la mejora en el bienestar que se pudiera producir a resultas de utilizar una vía de administración de un medicamento más cómoda para el paciente, aunque de igual eficacia, ya que más comodidad no implica mejor salud.

Si se entiende, por el contrario, que el objetivo del SNS no es únicamente maximizar la salud, sino contribuir, junto con el resto de políticas públicas, a mejorar el bienestar de la población de un país, la perspectiva social será la más adecuada (Johannesson y O'Conor, 1997). En este supuesto todos los costes y efectos (resultados) asociados a las tecnologías evaluadas habrían de tenerse en cuenta, con independencia de quiénes soportasen los primeros y quiénes disfrutasen de los segundos.



En definitiva, cualquier guía metodológica ha de explicitar la perspectiva del análisis y las consecuencias que dicha decisión tiene sobre los métodos de evaluación económica que se presentan en la guía.

1.2. Revisión sistemática de la literatura

Dentro de este apartado, las cuestiones más importantes que debe contemplar una guía se refieren a los siguientes asuntos: la necesidad de una revisión sistemática de la evidencia clínica; las precisiones relativas a la metodología de búsqueda y los criterios de inclusión y exclusión; las características y calidad de los estudios incluidos; los métodos propuestos para sintetizar los resultados de los estudios; y el análisis de la heterogeneidad de los mismos. Comentamos a continuación cada uno de estos aspectos.

En primer lugar, cuando se comparan dos o más alternativas en una evaluación económica puede existir un gran volumen de evidencia clínica acerca de la eficacia y la tolerabilidad de las tecnologías sanitarias analizadas. Los resultados de los estudios clínicos pueden diferir según el tamaño de la muestra y las características de los pacientes, el diseño del estudio, el periodo de seguimiento, los centros y el país en que se ha realizado el estudio, etc. Resulta, en consecuencia, fundamental seleccionar de manera adecuada los estudios clínicos que se utilizarán en el análisis económico, ya que, por ejemplo, la exclusión de algunos estudios podría crear un sesgo a favor de una tecnología sanitaria u otra. Por tanto, es recomendable incluir en



la evaluación económica toda la evidencia clínica disponible sobre las tecnologías comparadas (o al menos explicar las razones para excluir algunos estudios). Una revisión sistemática de la literatura permite identificar toda la evidencia clínica relevante (Sutton, 1999). Interesa conocer si las guías recomiendan la realización de revisiones sistemáticas de la literatura y en qué circunstancias ello se considera indispensable para un correcto análisis económico.

Por otra parte, a la hora de presentar los resultados de una revisión sistemática de la literatura resulta imprescindible ofrecer información completa y transparente sobre los métodos de búsqueda y los criterios de exclusión/inclusión de los estudios. Por ejemplo, el investigador debería referir las bases de datos y otras fuentes alternativas consultadas, las fechas iniciales y finales de la búsqueda, así como los términos (“palabras clave”) utilizados. Aún más relevante es, en nuestra opinión, que se proporcionen detalles precisos sobre las razones para incluir o excluir los estudios identificados. Hay circunstancias en las que el analista podría optar por incluir sólo ensayos clínicos que comparan directamente dos tecnologías, otros casos en los que se excluyen estudios cuando la muestra de pacientes es pequeña; en otras ocasiones se incluyen todos los estudios identificados sin restricciones de ningún tipo sobre su calidad, el contexto o el diseño del estudio (Mandelblatt, 1996). Cualquier criterio de inclusión o exclusión resulta a priori admisible, pero es necesario que dichos criterios sean explícitos para poder juzgar su validez. La guía habrá de indicar el detalle que se requiere en la presentación de los métodos de búsqueda y de selección de los estudios identificados en la revisión de la literatura.



Otro elemento clave para la transparencia del análisis es la correcta presentación de las características de los estudios elegidos. Debería darse cuenta, al menos, de los siguientes datos acerca de cada estudio: el diseño, las alternativas comparadas, el tamaño de la muestra de pacientes, las características demográficas y el nivel de gravedad de los pacientes, el periodo de seguimiento, el país o los países en que se ha desarrollado y el año en que se ha realizado (o publicado), los resultados de eficacia y/o de tolerabilidad para cada tecnología analizada, y la significación estadística de las diferencias en los resultados. Estos detalles permitirán juzgar la calidad de los estudios incluidos y estimar la robustez de los resultados de efectividad utilizados en el análisis económico. Además, el investigador podría incluir algunos comentarios sobre la calidad de los estudios elegidos, explicando sus criterios de evaluación. A este respecto, por ejemplo, se consideran comúnmente los ensayos clínicos aleatorizados como la fuente de eficacia con mayor validez interna. Así mismo, cuando se analizan dos o más tecnologías, es preferible obtener datos de eficacia de estudios que comparen directamente dichas alternativas en lugar de basar todo el análisis en comparaciones indirectas. Existen guías para juzgar la calidad de los estudios clínicos (por ejemplo, la escala de Jadad) que pueden servir de orientación (Clark, 1999). Veremos más adelante los requerimientos de las diferentes guías en relación con el nivel de detalle en la descripción y el análisis de la calidad de los estudios clínicos utilizados en la evaluación económica.



Una vez identificados y seleccionados los estudios clínicos resulta necesario sintetizar toda la evidencia obtenida. En una evaluación económica (sobre todo cuando ésta se basa en un modelo de decisión) es preciso estimar un valor medio para la eficacia de cada alternativa objeto de comparación (al menos en el caso base). En determinadas situaciones, puede ser suficiente recurrir a los resultados de un único estudio: por ejemplo, podría darse la circunstancia de que sólo existiera un único ensayo clínico que comparara las tecnologías de interés. Sin embargo, lo habitual es que exista más evidencia clínica y que se puedan obtener resultados de eficacia diferentes según el método de síntesis utilizado. El modo más sencillo, aun que no el más adecuado, de sintetizar la evidencia clínica de los estudios elegidos es calcular una media ponderada. Parece más apropiado, por el contrario, realizar un meta-análisis, teniendo en cuenta el tamaño de la muestra de cada estudio y las diferentes características de los estudios (Sutton, 2000). Un caso particular en el que resulta complicado sintetizar los datos de los estudios clínicos es aquél en el que no existen ensayos que comparen directamente dos o más tecnologías. En este supuesto, resulta incorrecto utilizar datos de eficacia absoluta de cada tecnología (aunque estén obtenidos mediante un meta-análisis), porque ello podría inducir un sesgo debido a las diferencias entre los estudios. Sería más correcto identificar un comparador común a cada tecnología (por ejemplo, placebo) y estimar el riesgo relativo de cada tecnología respecto a dicho comparador común (Caldwell, 2005). De este modo se podría valorar si existen diferencias reales entre las tecnologías de interés o si, por el contrario, cualquier diferencia en términos de eficacia hallada en los estudios clínicos se debe exclusivamente a las diferentes características de dichos estudios. La guía



tendrá que indicar qué métodos de síntesis se recomiendan para la evaluación económica y en qué circunstancias.

Por último, tal y como se ha señalado en los párrafos precedentes, los estudios clínicos elegidos para la evaluación económica presentan, en general, características diferentes. Es necesario tener en cuenta esta heterogeneidad en los métodos de síntesis que se utilizan para obtener los datos de efectividad. Por ejemplo, ante la presencia de heterogeneidad estadística en los resultados de los estudios clínicos, lo más apropiado es realizar un meta-análisis de efectos aleatorios (*random effect*), antes que un meta-análisis de efectos fijos (*fixed effect*) (Ades, 2005). Si la heterogeneidad atañe a las características de los pacientes de los diferentes estudios, se podrían realizar análisis de subgrupos para estimar la variabilidad de la eficacia clínica según el tipo de paciente analizado. Por otro lado, cuando es el nivel de calidad de los estudios el que se muestra heterogéneo puede resultar acertado realizar un análisis de sensibilidad excluyendo algunos de los estudios considerados en el caso base. Finalmente, el caso más complejo emerge cuando no hay comparación directa entre las tecnologías de interés. En este caso se ha sugerido antes la necesidad de estimar el riesgo relativo de las tecnologías respecto a un comparador común. Sin embargo, es posible que las diferencias en las características de los pacientes tengan un impacto también sobre el riesgo relativo respecto a este comparador común. Por ejemplo, pacientes con edad más avanzada o con una enfermedad más grave, podrían tener una eficacia ‘relativa’ mayor (o menor). Por esta razón, se han desarrollado recientemente métodos para incorporar la heterogeneidad entre pacientes en la



comparación indirecta de tecnologías sanitarias (meta-regresiones, *mixed treatment comparison*, etc.) (Ades, 2006; Nixon, 2006). Se analizará qué tipo de indicaciones ofrecen las guías acerca de la cuestión de la heterogeneidad de los estudios.

1.3. Población objetivo

Dentro de este apartado, las cuestiones más importantes que debe contemplar una *guía* son tres: la identificación de la población objetivo y el detalle con el que se ha de definir, la necesidad de analizar subgrupos y, por último, las implicaciones de la población objetivo determinada en términos de impacto presupuestario.

En una evaluación económica ha de identificarse con claridad cuál es la población objetivo del análisis, esto es, resulta necesario especificar las características de los pacientes que se pueden beneficiar del tratamiento estudiado. Es importante proporcionar el nivel de detalle adecuado en la definición de la población objetivo (por ejemplo, edad, sexo, nivel de gravedad, presencia de enfermedades concomitantes etc.). Es fundamental realizar una descripción detallada de la población, ya que tanto los costes como los beneficios de un tratamiento podrían variar al cambiar las características de los pacientes (Nord, 2001).

De otro lado, en determinadas circunstancias puede resultar conveniente establecer subgrupos dentro de la población objetivo. Esto ocurrirá cuando exista una presunción fundada de que los resultados, en términos coste-efectividad, son



diferentes para distintos subconjuntos de la población objetivo (Coyle, 2003), debido a diferencias previsibles en la efectividad o en los costes. Así, por ejemplo, algunos tratamientos podrían tener una efectividad significativamente superior en hombres que en mujeres (o al contrario), o en pacientes jóvenes que en pacientes de edad avanzada (o viceversa). También los efectos secundarios de una tecnología sanitaria podrían depender del nivel de gravedad de la enfermedad de un paciente o del tipo de enfermedades concomitantes. Del mismo modo, la existencia de economías de escala puede motivar que los costes dependan críticamente del tamaño de la población a la que se destina la tecnología evaluada, de suerte que un programa podría ser más costoso cuanto menor fuera el grupo de pacientes o usuarios del mismo. Dado que una tecnología sanitaria puede resultar coste-efectiva para un subgrupo de la población y no para otros, parece relevante identificar dichos subgrupos y adaptar la evaluación económica a los subgrupos de pacientes para quienes los costes y/o los efectos pueden ser significativamente diferentes (Briggs, 1994). Por otro lado, pueden existir problemas para obtener suficientes datos de todos los subgrupos de pacientes y la calidad de dichos datos puede reducirse al aumentar la heterogeneidad. La guía tendrá que indicar cuándo es necesario realizar un análisis de subgrupos y qué métodos se sugieren a tal fin.

Finalmente se ha de advertir de que puede resultar complicado definir de manera apropiada y detallada los candidatos a recibir una nueva tecnología sanitaria. Si bien en los ensayos clínicos la población objetivo resulta en general bien definida, puede ser más difícil en algunas circunstancias transferir los resultados de estos estudios a



la práctica real (Rothwell, 2005). También el tamaño de la población a la que se destina la tecnología evaluada podría resultar importante a la hora de evaluar el coste-efectividad de dicha tecnología y su impacto presupuestario ser un elemento importante para tomar decisiones sobre su reembolso por parte del sistema sanitario.

1.4. Término de comparación de la tecnología evaluada

Por lo que respecta a la elección de la tecnología que se utilizará como referencia o “comparador”, las cuestiones más importantes que debe contemplar una guía son: los términos de comparación que se recomiendan y los factores relevantes para su elección, la definición de la “práctica vigente” y cómo determinarla y, por último, la inclusión de las alternativas “no hacer nada” o “intervención mínima” como comparadores.

En primer lugar, resulta de interés conocer si la guía recomienda uno o más comparadores específicos para la tecnología analizada en la evaluación económica. La elección del comparador es un aspecto clave en el estudio económico, ya que se evalúa si una nueva tecnología es coste-efectiva, no en términos absolutos, sino respecto a la tecnología alternativa, esto es, en términos incrementales (Drummond, 1996). Por esta razón, la elección de un término de comparación inadecuado puede dar lugar a conclusiones erróneas. Es importante que la guía indique cuál o cuáles son los comparadores ideales para el tratamiento de estudio y sobre la base de qué criterios se pueden elegir dichos comparadores en diferentes circunstancias.



A este respecto, existen muchos factores relevantes a la hora de elegir uno o más comparadores para la intervención estudiada. En algunas circunstancias, puede existir un elevado número de tecnologías aprobadas para una indicación terapéutica, o bien los tratamientos pueden variar según el tipo de paciente. En teoría, la solución ideal a este problema sería comparar la nueva tecnología con todas las alternativas existentes. Es evidente que resulta muy difícil poder realizar una evaluación económica teniendo en cuenta todas estas alternativas por razones de costes y de disponibilidad de datos, lo que exige seleccionar algunos comparadores. Como norma general, la nueva tecnología tendría que ser comparada con la práctica vigente o habitual, aunque ésta pueda tener diferentes definiciones y aunque puedan existir casos en los que éste no sea el único comparador. Sin embargo, hay otras alternativas susceptibles de ser relevantes y que podrían no coincidir con la práctica habitual, como pueden ser los casos del tratamiento más efectivo o del tratamiento más barato. De ahí la importancia de tener en cuenta muchos elementos a la hora de elegir el comparador (o comparadores), como la población objetivo, la disponibilidad y calidad de los datos, o el tipo de intervención (Drummond, 2005). Por ejemplo, hay ocasiones en las que el nuevo tratamiento tiene que ser comparado a lo largo de una secuencia de tratamientos (en el caso de fármacos, sobre todo) y se ha de comparar diferentes intensidades de un programa (como en el supuesto de un programa de cribaje).



En determinadas circunstancias puede ser difícil identificar la práctica vigente en un país, ya que con frecuencia la práctica habitual no consiste en un único tratamiento, sino en la combinación de varios. En tal caso, existen diferentes opciones cuyas implicaciones metodológicas conviene sopesar. Por ejemplo, se podría adoptar como término de comparación la práctica clínica más comúnmente usada (si existe una práctica dominante), o bien emplear cada una de las tecnologías alternativas en uso como términos de comparación múltiples, o, finalmente, utilizar una combinación lineal de las intervenciones existentes ponderadas según su frecuencia de uso (Garber, 1996). Se analizará si la guía propone una definición específica de práctica habitual y si indica la forma de identificarla (bases de datos, estudios publicados, opiniones de expertos, etc).

Como se ha advertido al comienzo de este apartado, la elección de un comparador inapropiado puede dar lugar a la obtención de resultados erróneos sobre la relación coste-efectividad de una nueva tecnología. Aunque la comparación con la práctica habitual se considera en general la elección ideal, cabe la posibilidad de que esta alternativa arroje resultados equívocos. Ello puede ocurrir en el caso de que el *statu quo* no haya sido, a su vez, objeto de evaluación y, de hecho, no sea coste-efectivo. Una solución a este problema podría ser la incorporación en el análisis de la opción “no hacer nada”, o alternativamente, una tecnología clasificable como “intervención mínima” (el tratamiento de menor coste que tenga una efectividad mayor que un placebo).



1.5. Medida de resultados preferida

Dentro de este apartado, también son varias las cuestiones que debe contemplar una guía. En primer lugar, las guías tienen que pronunciarse acerca de si, a los efectos de realizar una evaluación económica, se exigirá que los resultados se presenten en forma de variable final o, por el contrario, se aceptarán evaluaciones basadas en variables intermedias (Brazier et al 1999). Por ejemplo, en el caso de medicamentos para la hipertensión, ¿se admitirá la presión arterial como medida de resultado en una evaluación económica o se exigirá la traducción de dicha variable intermedia a años de vida ganados o reducción de la mortalidad?

En segundo lugar, si la guía exige (o recomienda) el uso de variables finales, deberá adoptar una posición al respecto del uso de medidas de Calidad de Vida Relacionadas con la Salud (CVRS). Dado que los perfiles específicos tienen una amplia tradición y se utilizan de forma regular en la evaluación clínica de tecnologías médicas, sería conveniente que las guías trataran el posible uso de perfiles genéricos y específicos en la evaluación económica de tecnologías médicas (Kind P. 2001).

Además, las guías han de aclarar si en las evaluaciones económicas será obligado, recomendable o, simplemente, posible emplear medidas basadas en la psicometría y/o en preferencias (Revicki y Kaplan, 1993). Aunque existe cierto debate acerca del término más adecuado para denominar las medidas de CVRS basadas en



preferencias, nos referiremos a ellas, de modo genérico, como “utilidades” (Torrance 1986).

En el supuesto de que las guías exijan o recomienden el uso de utilidades, habrán de precisar además si estas utilidades deben provenir de pacientes o de la población general (Ubel et al, 2000). Adicionalmente, las guías han de pronunciarse acerca de los métodos más adecuados para estimar estas utilidades (Green et al 2000). Asimismo, han de decidir si se van a utilizar medidas basadas en sistemas de clasificación de salud estándar tales como el *Euroqol* (EQ), el *Health Utility Index* (HUI), el *Quality of Well Being Scale* (QWB) o similares (Tsuchiya et al 2006).

Es importante señalar que la utilización de métodos basados en utilidades para incluir los efectos de los programas sanitarios en la calidad de vida, no exige necesariamente el uso de los Años de Vida Ajustados por Calidad (AVAC) como medida de resultado. Aunque ambos aspectos suelen ir unidos, el uso de utilidades no implica la utilización de AVAC. En consecuencia, las guías tienen que decidir el papel que consideran que han de jugar los AVAC en las evaluaciones económicas de programas sanitarios (Rasanen et al 2006).

Finalmente, un método que está siendo cada vez más utilizado en la valoración económica de tratamientos sanitarios es el de la disposición a pagar. Las guías deberían hacer explícita su postura respecto a otra forma de evaluación económica como es el Análisis Coste-Beneficio (ACB), donde el valor de los cambios en la



salud se mide en dinero (Gyrd-Hansen 2003). Se revisará, pues, el papel que las guías atribuyen a este tipo de métodos de valoración de la salud.

1.6. Costes incluidos en el análisis

Las cuestiones que han de abordar unas guías de evaluación económica en lo que respecta a los costes tienen que ver con la identificación de los recursos, su medida y, finalmente, su valoración. Procedemos a exponerlas seguidamente, por este orden.

Cuando hablamos de identificación de los recursos hacemos referencia a la decisión relativa a qué conceptos de coste se van a incluir en el análisis y cuáles van a ser excluidos (Jacobs et al 2005). Las razones de índole normativa por las que un concepto no ha de incluirse en un análisis de costes, son dos: o bien no se trata de costes reales en ningún caso; o bien no son costes reales bajo una cierta perspectiva que se considera la correcta desde un punto de vista normativo.

En el primer supuesto, se encontrarían costes como los debidos al protocolo utilizado para realizar un ensayo clínico en el que se ha basado el estudio. Otro ejemplo es el de las transferencias (p.ej. prestaciones por desempleo o por incapacidad temporal), ya que no son costes reales para el conjunto de la sociedad, en el sentido de no representar costes de oportunidad.



En el segundo caso se sitúan aquellas categorías de costes que suponen un consumo de recursos bajo una cierta perspectiva pero no cuando la perspectiva cambia. Así, cuando bajo la perspectiva normativamente correcta dichos costes se excluyen, éstos no deben considerarse. Por ejemplo, si unas guías consideran que lo normativamente correcto es adoptar la perspectiva del Sistema Nacional de Salud al hacer el estudio de costes, no habrán de identificarse como recursos consumidos los costes que recaigan sobre los particulares (típicamente, el paciente y sus cuidadores informales). A la hora de juzgar los recursos que deben identificarse o no, puede ser útil la siguiente clasificación de costes:

1. Costes sanitarios	1a. Costes directos 1b. Costes sanitarios futuros
2. Costes no sanitarios	2a. A cargo del paciente (gastos monetarios y tiempo) 2b. A cargo de terceros (otras entidades públicas y cuidadores informales) 2c. Costes de productividad

Las guías deben pronunciarse acerca de si es necesario identificar todos los recursos que caen dentro de los anteriores epígrafes. En el caso de que se incluyan costes no sanitarios, como el tiempo utilizado por el paciente y/o sus cuidadores informales, se tiene que decidir, además, si serán todos los posibles usos del tiempo los que se consideren como costes (Olsen y Richardson, 1999). Por ejemplo, en el caso del paciente, la enfermedad le exige dedicar tiempo al tratamiento (visitas al médico, realización de pruebas) y/o le impide realizar actividades que podía realizar antes de estar enfermo. Es decir, la enfermedad tiene para el paciente un coste de oportunidad en términos de tiempo perdido, y este tiempo que pierde puede conllevar un menor



rendimiento (productividad) en el trabajo, debido a las ausencias del trabajo, sean éstas de mayor o menor duración. En último caso, la enfermedad puede dar lugar a una pérdida de productividad laboral del paciente debido a su muerte prematura (Jacobs et al 2003). En el caso de los cuidadores informales, lo relevante es valorar el tiempo que ocupan en cuidar al enfermo o en otras actividades directamente asociadas a la enfermedad del paciente. En general, la cuestión está en decidir si se ha de incluir en la evaluación económica el valor del tiempo que se detrae del trabajo remunerado, pero también el valor del tiempo que se detrae del trabajo no remunerado e incluso del tiempo que se resta a actividades de ocio.

Finalmente, otro concepto sobre el que existe cierto debate en la literatura es el de los costes que, no estando relacionados con la tecnología que se evalúa, se producen en los años de vida ganados como consecuencia de la misma (*survivor costs*). Las guías deberían indicar si tales costes han de incluirse o no.

Una vez que las guías hayan precisado qué recursos se han de identificar, deben orientar acerca del modo en que deben medirse. En este caso, las guías pueden indicar qué grado de precisión se requiere en los estudios de evaluación económica (Smith y Barnett, 2003). Más concretamente:

- Las guías deben explicitar si van a exigir estimaciones propias de los investigadores o, por el contrario, van a permitir el uso de fuentes externas (literatura, opiniones de expertos, guías de práctica clínica...).



- En el supuesto de que se acepte el recurso a fuentes externas, se ha de concretar en qué casos se admitirán y en cuáles no. Por ejemplo, ¿se va a hacer una distinción entre los conceptos de coste que son más o menos importantes para la precisión de las estimaciones finales? Esto es, ¿se exigirán estimaciones propias para los tipos de costes que son más relevantes en la evaluación o se aceptará que todo el estudio de costes se construya a partir de datos procedentes de la literatura?

- En el supuesto de que se exija (o recomiende) la realización de estimaciones propias, las guías pueden preferir (o no) que el método se ajuste a una de las dos opciones metodológicas posibles: *top down* o *bottom up* (“microcoste”). En el primer caso se empieza por grandes partidas de coste y se van asignando, mediante ciertos procedimientos, partes de tales costes a diversos conceptos más concretos. En el segundo caso, se miden de forma pormenorizada los recursos consumidos por cada paciente.

- En el caso de que las guías aconsejen la realización de estudios de microcoste, puede que ofrezcan indicaciones metodológicas sobre la forma de tratar algunas de las cuestiones más problemáticas, como el reparto de los costes generales (*overhead costs*) o las amortizaciones.



Finalmente, por lo que respecta a la valoración monetaria de los recursos consumidos, son varias las cuestiones a las que las guías deberían prestar su atención. En primer lugar está la distinción entre costes financieros y costes de oportunidad. Se trata de dos conceptos diferentes que, en la práctica, difieren con mucha frecuencia, por lo que se hace necesario decidir cuál de los dos se debe utilizar. Por ejemplo, si el SNS reembolsa a una compañía de seguros mediante un sistema de tarifas, ¿se ha de utilizar la tarifa para valorar los costes, aun siendo posible que dichas tarifas no reflejen los costes de oportunidad?

En segundo lugar, tenemos las listas de costes estándar que a veces se elaboran con el objetivo último de conseguir consistencia entre los estudios de evaluación económica. Las guías han de pronunciarse al respecto de si se admite (o incluso se recomienda) el uso de este tipo de listas de costes estándar.

En tercer lugar está la cuestión del horizonte temporal y su impacto sobre los costes. Los costes de oportunidad de la aplicación de una tecnología médica pueden diferir en función de si se miden en el corto o en el largo plazo. En el corto plazo, gran parte de los costes son fijos, por lo que puede existir una gran discrepancia entre los costes medios y los costes marginales. En el largo plazo, sin embargo, ambos tipos de costes tienden a coincidir. Cabe esperar que las guías sugieran si se opta por el corto o el largo plazo y, por tanto, si los estudios se basarán en los costes medios o en los costes marginales.



Por último, un concepto de difícil tratamiento es el valor del tiempo del paciente y/o de sus cuidadores informales (Tranmer et al 2005). Podemos distinguir tres tipos de tiempo perdido que ya se han mencionado más arriba: 1) el tiempo perdido del trabajo remunerado, 2) el tiempo perdido del trabajo no remunerado y 3) el tiempo de ocio. Respecto del tiempo perdido de trabajo remunerado (1), las guías han de indicar si su valoración ha de ajustarse al “enfoque del capital humano” (Johanesson y Karlson 1997) o si, por el contrario, debe adoptarse el “enfoque de los costes de fricción” (Koopmanschap et al 1995). Esta elección es importante y ha de ser justificada teóricamente, ya que los dos enfoques pueden producir resultados radicalmente diferentes. Respecto del tiempo de trabajo no remunerado (2), también existen varias alternativas: utilizar el salario de los profesionales que realizan un trabajo similar (coste de reemplazo), usar el salario de reserva del cuidador informal, o recurrir a la valoración contingente. Por último, en relación con la valoración del tiempo dedicado al ocio (3), también hay varias alternativas posibles que las guías han de tomar en consideración.

1.7. Horizonte temporal de los costes y efectos

Dentro de este apartado, las cuestiones más importantes que deben contemplar unas guías de evaluación son: la determinación del horizonte temporal recomendado y los factores que determinan dicha elección, el uso de modelos para extender datos de corto plazo a un plazo más largo y la variación del horizonte temporal en el análisis de sensibilidad.



En primer lugar, las guías tienen que indicar si existe un horizonte temporal recomendado para la evaluación económica y si este horizonte temporal es diferente según el tipo de análisis (tipo de intervención, tipo de enfermedad, etc.). El horizonte temporal tendría que ser el mismo para costes y beneficios.

Son muchos los factores que pueden determinar la elección del horizonte temporal. En general, la teoría aconseja utilizar un horizonte temporal de largo plazo, preferiblemente que abarque toda la vida de los pacientes. Esto permitiría capturar todos los costes y beneficios futuros de los individuos, sean estos directamente debidos a la intervención analizada, o sean asociados a efectos secundarios u otros eventos (Drummond, 2005; Gold, 1996).

Sin embargo, hay que tener en cuenta que no todos los programas sanitarios ni todas las enfermedades tienen el mismo perfil temporal. Por ejemplo, en el caso de enfermedades crónicas los tratamientos pueden tener un impacto durante toda la vida de un paciente y puede ser relevante medir las diferencias en calidad de vida, mortalidad, o costes entre intervenciones en un largo plazo. Hay otras situaciones, sin embargo, en las que un tratamiento tiene su principal impacto en el corto plazo (por ejemplo, en el caso de infecciones agudas), de suerte que se antoja menos relevante evaluar los efectos de dicho tratamiento durante toda la vida de un paciente.



Otro factor que se ha de tener en cuenta al fijar el horizonte temporal es la disponibilidad de datos. A menudo se dispone de datos con validez interna contrastada sólo en un corto plazo (ensayos clínicos) y es necesario utilizar datos de estudios observacionales (con las consecuentes limitaciones) e incluir supuestos u opiniones de expertos para extender el estudio a un horizonte temporal más amplio. Esto da lugar a un *trade-off* entre la calidad de los datos utilizados en un estudio y la necesidad de evaluar tecnologías sanitarias en un largo plazo. Uno de nuestros objetivos es analizar si las guías consideran estos factores en la elección del horizonte temporal, y si recomiendan diferentes enfoques según el tipo de enfermedad, intervención, etc.

Como se ha señalado un poco más arriba, es muy difícil disponer de datos clínicos de pacientes en un horizonte temporal muy amplio (por ejemplo, 30 años) y a menudo es necesario asumir supuestos sobre el impacto futuro de un tratamiento en términos de costes y beneficios. A tal fin se utilizan modelos que permitan extender la evidencia de corto plazo a un horizonte temporal más largo, incorporando datos que derivan de fuentes diferentes (ensayos clínicos, estudios observacionales, opiniones de expertos, etc.). Los modelos permiten también estimar la incertidumbre alrededor de los datos y los supuestos realizados. En particular, los modelos de Markov han sido utilizados frecuentemente en literatura en el caso de intervenciones para enfermedades crónicas (Sonnerberg, 1993). Puede que las guías recomienden métodos específicos para extender datos de corto plazo a un horizonte temporal de largo plazo.



Finalmente, la utilización de diferentes horizontes temporales puede tener un impacto sobre el coste-efectividad relativo de las tecnologías sanitarias (Hay, 1998). Esto sugiere la pertinencia de incluir en el estudio escenarios alternativos en que se consideren horizontes temporales diferentes, sobre todo en el caso de evaluaciones económicas basadas en modelos que en general incorporan gran número de supuestos. Por ejemplo, se podría utilizar un horizonte temporal que abarcara toda la vida de los pacientes en el caso base (con datos modelizados), e incluir un horizonte temporal más breve únicamente con datos primarios en un escenario alternativo.

1.8. Tasa de descuento

En relación con el descuento de las magnitudes futuras, las cuestiones más relevantes que esperamos encontrar reflejadas en las guías tienen que ver con la recomendación del descuento de costes y beneficios, la justificación de la tasa o tasas recomendadas y su variación en el análisis de sensibilidad.

Las guías tienen que indicar si consideran necesario aplicar una tasa de descuento para los costes y los beneficios futuros, y cuál es la tasa recomendada en el caso base. Son varias las razones que pueden aducirse para defender la necesidad de ajustar costes y beneficios futuros. En general, los individuos prefieren recibir los beneficios cuanto antes y retrasar los costes lo más posible. Esta característica habitual de las personas, a la que nos referimos con el término “preferencia



temporal” no es, contrariamente a lo que en principio muchas personas puedan suponer, únicamente consecuencia de la inflación (Cairns, 1994). Son otras las causas que explican esta actitud. Así, por ejemplo, algunas personas tienen un horizonte temporal limitado y otorgan por ello un mayor valor al momento presente. Además, los seres humanos contemplamos el futuro con incertidumbre, lo que nos conduce a preferir disfrutar de los beneficios en el presente (cierto), y sufrir los costes en el futuro (incierto).

Por otro lado, existe en la literatura un vivo debate en relación a si es necesario (o conveniente) aplicar la misma tasa de descuento a los costes y a los beneficios, o si, por el contrario, sería más apropiado aplicar una tasa más elevada para los primeros (Lázaro, 2002; Lázaro et al., 2002). Algunos expertos sostienen que al descontar, por ejemplo, los años de vida que se ganan en el futuro, se prima a las generaciones presentes frente a las generaciones futuras; o que no es imaginable que los individuos redistribuyan su salud intertemporalmente como lo hacen con su riqueza. También algunos sostienen que la evidencia empírica sugiere que los individuos descuentan las ganancias en salud y las ganancias monetarias a tasas diferentes (Torgerson, 1998; Smith, 2001; West 2003). Por otra parte, algunos estudiosos sostienen que si se descuentan los costes futuros pero no los beneficios (o éstos se descuentan a tasas inferiores), las conclusiones finales pueden resultar paradójicas. La “paradoja de Keeler y Cretin” muestra cómo cuando la efectividad de un tratamiento y sus costes se descuentan con tasas diferentes (menores para los beneficios) la recomendación implícita en el resultado de la evaluación puede ser la de posponer indefinidamente la



realización de un programa sanitario (Keeler, 1983). También se defiende el descuento de costes y beneficios a una misma tasa por razones de coherencia interna del análisis.

Una última cuestión tiene que ver con la posibilidad de variar las tasas de descuento tanto para los costes como para los beneficios futuros en el análisis de sensibilidad, a fin de estimar su impacto sobre los resultados del estudio.

1.9. Tratamiento de la incertidumbre

Las cuestiones que deben contemplar las guías en relación con el tratamiento de la incertidumbre son muy variadas, comenzando por la necesidad de identificar qué fuentes de incertidumbre se consideran relevantes en una evaluación económica y si se distingue entre estrategias para tratar la incertidumbre y métodos para tratar la variabilidad de los datos utilizados en el análisis (Briggs, 2001). Por ejemplo, existe incertidumbre respecto de los parámetros y los supuestos de un modelo de decisión (estructura, métodos analíticos, etc.) pero también existe variabilidad debida a la heterogeneidad de los pacientes de un estudio y a las diferencias entre prácticas clínicas, recursos utilizados o costes unitarios entre países o entre diferentes centros de un mismo país.

En segundo lugar, puede resultar de interés discriminar entre evaluaciones económicas basadas en datos que derivan de un solo estudio (por ejemplo, un ensayo



clínico) y evaluaciones económicas basadas en una síntesis de estudios diferentes (normalmente, recurriendo a un modelo de decisión). Para el primer caso se dispone de diversas alternativas para tratar la incertidumbre, como el análisis de sensibilidad determinístico y el *bootstrapping* (Campbell, 1999); alternativas que no siempre coinciden con las estrategias que cabe emprender en el segundo supuesto, al que además del análisis determinístico, se le puede aplicar el análisis de sensibilidad probabilístico con simulaciones de *Monte Carlo* (Briggs, 2000).

Las guías han de especificar qué tipo de análisis de sensibilidad se recomienda y si se sugiere la aplicación de métodos diferentes según cuál sea la fuente de incertidumbre. En general, existen dos tipos de análisis de sensibilidad: el determinístico y el probabilístico o estocástico. El análisis de sensibilidad determinístico incluye el análisis univariado (se cambia un solo parámetro cada vez), el análisis multivariado (se modifican dos o más parámetros simultáneamente), el análisis de escenarios (en el que se asumen escenarios diferentes al caso base y se estima si las conclusiones de la evaluación económica cambian significativamente o no en estos nuevos escenarios) y el análisis de umbral (se identifican los valores críticos del parámetro –o de los parámetros– por encima o por debajo de los cuales las conclusiones del estudio cambian). Podría considerarse también una modalidad del análisis de sensibilidad determinístico el análisis de subgrupos, en el que el estudio se repite para subgrupos homogéneos de pacientes.



El análisis de sensibilidad probabilístico o estocástico admite dos formas diferentes en función de las características de los datos disponibles. Cuando los datos proceden de un único ensayo clínico y los investigadores disponen de los datos individuales de los pacientes es posible aplicar el método del *bootstrapping*, mediante el cual se generan muestras repetidas que permiten la obtención de intervalos de confianza para la ratio incremental coste-efectividad (Chaudary, 1996). Cuando los datos se han obtenido de la síntesis de diferentes estudios (o a partir de la opinión de expertos) y se ha empleado un modelo de decisión, procede llevar a cabo simulaciones de *Monte Carlo*, método en virtud del cual se asignan distribuciones a todos los parámetros del modelo y se obtienen también intervalos de confianza para la ratio coste-efectividad (Hay, 1998; Kuntz, 2001).

En el supuesto de que las guías recomienden llevar a cabo un análisis de sensibilidad, pueden extender su recomendación a todos los datos y parámetros del estudio, o bien limitar la prescripción a algunos de estos parámetros (datos de efectividad, recursos consumidos, utilidades, etc.). En el caso de que las guías recomienden enfocar el análisis de sensibilidad sobre datos y parámetros específicos, es importante comprobar si se proponen métodos que permitan identificar estos parámetros y si se requiere que la elección de dichos datos y parámetros se justifique de algún modo. Asimismo, en el caso de que los valores algunos parámetros estén sometidos a una gran incertidumbre, puede que las guías demanden la obtención de más información.



Las guías deberían aclarar cómo se elige el rango sobre el que se harán variar los datos y parámetros del estudio en el análisis de sensibilidad. Por ejemplo, los rangos elegidos podrían derivar de intervalos de confianza obtenidos en un ensayo clínico, o bien ser el resultado de un meta-análisis, de la opinión de expertos o de supuestos formulados por los investigadores. En el caso del análisis de sensibilidad probabilístico, las guías podrían recomendar el uso de distribuciones particulares asociadas a parámetros específicos (Briggs, 2003; Briggs 2001).

Para terminar con la cuestión del tratamiento de la incertidumbre, el modo de presentar los resultados del análisis de sensibilidad también debería ser objeto de atención por parte de las guías de evaluación. Por ejemplo, en un análisis de sensibilidad determinístico univariado se podrían utilizar los denominados *tornado diagrams*, que representan gráficamente el impacto de cada parámetro sobre los resultados coste-efectividad. Por otra parte, en el análisis de sensibilidad probabilístico se pueden usar las “curvas de aceptabilidad” (Fenwick, 2001), donde se representa gráficamente la probabilidad de que un tratamiento sea coste-efectivo dados diferentes niveles de disposición a pagar por el resultado (por un AVAC, por ejemplo).

1.10. Uso de modelos decisionales

En relación con el uso de modelos de decisión, las guías deben prestar atención a cuestiones como las siguientes: los supuestos en los que utilizar un modelo



decisional resulta recomendable; los tipos de modelos que se aconsejan, en función del tipo de patología o de intervención; los detalles relativos a la estructura del modelo, la fuente de los datos utilizados y los parámetros del mismo; y, por último, su coherencia interna y externa.

En determinadas circunstancias puede resultar necesario desarrollar un modelo de decisión en el curso de una evaluación económica (Phillips, 2004; Weinstein, 2003). Por ejemplo, es posible que el horizonte temporal de un ensayo clínico sea inadecuado (por corto) para los objetivos del análisis económico, en cuyo caso un modelo de decisión se convierte en el instrumento apropiado para extender el análisis al largo plazo (Sanders, 1999). También encontramos situaciones en las que los datos de efectividad o de uso de recursos provienen de varias fuentes diversas, o bien escenarios en los que son muchas las tecnologías que se han de comparar. En todos estos supuestos, la flexibilidad de los modelos permite sintetizar de manera rápida y detallada toda la información identificada en las fuentes. Por otra parte, el uso de un modelo de decisión da pie a un análisis detallado de la incertidumbre que rodea a los datos clínicos y económicos, al tiempo que facilita la transferibilidad de datos entre diferentes contextos (véase también, al respecto, el apartado 1.9 sobre incertidumbre y el que más adelante se dedica a la transferibilidad (1.11).

Una primera cuestión relevante atañe a la elección del modelo, que habrá de ser diferente según cuál sea la enfermedad y/o la intervención evaluada. Por ejemplo, en determinadas situaciones puede que los árboles de decisión estándares resulten



complejos o demasiado largos (Barton, 2004). Esto ocurre, en particular, cuando estamos ante enfermedades caracterizadas por la recurrencia de los estados de salud, o cuando la evaluación se centra en tratamientos o enfermedades en el largo plazo (típicamente, todas las enfermedades crónicas). En tales situaciones, los modelos de Markov resultan ser más flexibles y menos complicados, ya que su estructura se caracteriza por estados de salud y probabilidades de transición de un estado a otro (Sonnenberg, 1994; Briggs, 1998). Una importante disyuntiva que se presenta al evaluador es la de optar entre un modelo determinístico, en el que todos los parámetros se consideran fijos y sólo cambian en el análisis de sensibilidad determinístico, y un modelo probabilístico, en el que se asigna una distribución a los parámetros del modelo (Briggs, 2006).

La estructura de un modelo ha de reflejar de manera coherente el tipo de patología analizada y las posibles consecuencias de las intervenciones que se someten a comparación. En la tarea de elegir la estructura del modelo el investigador puede encontrar orientación en diversas fuentes como la literatura clínica, los modelos publicados, la opinión de expertos, etc. Existen situaciones, no obstante, en las que la estructura del modelo se asienta principalmente en supuestos de los propios investigadores. Puesto que la estructura del modelo previsiblemente tendrá un impacto importante sobre los resultados del análisis coste-efectividad, es muy recomendable que una evaluación económica basada en un modelo de decisión sea muy transparente al definir y justificar todas las fuentes y supuestos empleados, así como las posibles limitaciones derivadas de tales supuestos (Weinstein, 2003). Una



descripción detallada de las características estructurales de un modelo permite, además, juzgar su coherencia desde el punto de vista clínico, a la vez que aumenta la probabilidad de que el modelo pueda ser adaptado a otros contextos o intervenciones. Por ejemplo, en el caso de un modelo de Markov, es fundamental indicar, como mínimo, los estados de salud considerados y la duración de un ciclo.

El otro elemento básico, junto a la estructura, que permite juzgar un modelo de decisión es la calidad de los datos que incorpora. Las fuentes de las que proceden los parámetros del modelo pueden ser de dispar calidad, como distinto puede ser el grado de validez interna de los datos (por ejemplo, los datos de eficacia pueden derivar de ensayos clínicos, estudios observacionales, opiniones de expertos, etc.) y diferente su relevancia para el contexto de interés. También puede haber variabilidad en la calidad de los métodos empleados para la síntesis de los datos obtenidos (meta-análisis, media ponderada, elección de un valor medio basada en supuestos, etc.). Las guías deberían recomendar una descripción detallada de todas las fuentes de los parámetros, así como una justificación y explicación de los métodos de síntesis (véase también el apartado 1.2 sobre la síntesis de la evidencia).

Más allá de esto, también interesa que se aporten detalles sobre el grado de incertidumbre que rodea a los parámetros utilizados, como, por ejemplo, el intervalo de confianza para el análisis de sensibilidad en el caso de un modelo determinístico o el tipo de distribución asociada a cada parámetro en el caso de modelos probabilísticos (Briggs, 2002). Por último, las guías podrían demandar una



descripción de los métodos reutilizados para transformar datos primarios en parámetros del modelo. Por ejemplo, podría ser un requisito para las guías el explicar cómo se han obtenido las probabilidades de transición de un modelo de Markov a partir de los datos de eficacia de ensayos clínicos, o qué funciones se han utilizado para estimar la mortalidad de los pacientes a largo plazo (Miller, 1994).

En lo que atañe a la coherencia de los modelos, comenzaremos señalando que por coherencia interna se entiende la coherencia lógica del modelo desde el punto de vista matemático. El modelo no puede arrojar resultados que sean irreales o contrarios a la intuición. Esto hace que sea recomendable comprobar la coherencia interna del modelo, por ejemplo utilizando valores extremos en los parámetros para verificar que no existen errores de códigos matemáticos (McCabe, 2000). Por otro lado, parece más importante aún contrastar la coherencia externa del modelo, es decir, asegurarse de que los resultados del modelo de decisión se compadecen con el resto de la evidencia clínica y económica disponible (Weinstein, 2003). El proceso en virtud del cual los datos y resultados de un modelo se validan respecto a otros datos reales (por ejemplo, estadísticas nacionales), recibe el nombre de “proceso de calibración” (Hammerschmidt, 2003). Los resultados del modelo desarrollado podrían, asimismo, compararse con los de otros modelos utilizados para la misma enfermedad e intervención, con el fin de subrayar las diferencias o similitudes.



1.11. Transferibilidad de datos y resultados

Por lo que a la transferibilidad de los resultados se refiere, son muchas y variadas las cuestiones que deberían suscitar la atención de las guías. En primer lugar, conviene señalar que existen distintos conceptos de transferibilidad. La pregunta clave que habitualmente se desea responder es si resulta posible transferir los resultados coste-efectividad de un país a otro (Barbieri et al, 2005). Sin embargo, hay otros conceptos de transferibilidad que pueden ser importantes, como el que atañe a la validez externa de un ensayo clínico (es decir, si las condiciones y los resultados de un ensayo clínico pueden representar la práctica clínica “real”) o el referido a la transferibilidad de resultados entre distintos centros de un mismo país.

En una evaluación económica se necesitan diferentes tipos de datos para calcular el coste-efectividad de un medicamento o de una tecnología: los datos de eficacia de un tratamiento, los recursos utilizados, los costes unitarios, y las utilidades asociadas a los estados de salud (en el caso de un análisis coste-utilidad). Algunos de estos elementos pueden ser más transferibles (de un país a otro o de un centro a otro) que otros. Habitualmente, se considera que los datos de eficacia (en particular, los referidos a la eficacia “relativa” de un tratamiento respecto a otro) son transferibles de un país a otro, mientras que con los datos de índole económica ocurre justo lo contrario, considerándose incorrecto transferirlos de un contexto geográfico a otro distinto. Sin embargo, en ocasiones puede ser difícil transferir de un lugar a otro incluso los datos de eficacia, debido a la existencia de diferencias de naturaleza



epidemiológica, del *case-mix* de los pacientes, de prácticas clínicas, etc. (Drummond, 2001). Las guías deberían identificar con precisión qué elementos se pueden obtener a partir de estudios realizados en otros países, y qué datos han de ser estrictamente específicos del propio país.

En el debate sobre la transferibilidad, puede resultar conveniente discriminar entre evaluaciones económicas basadas en datos que derivan de un sólo estudio (por ejemplo, un ensayo clínico) y evaluaciones económicas basadas en una síntesis de estudios diferentes (normalmente, utilizando un modelo de decisión). Los problemas de transferibilidad asociados a los dos tipos de análisis son diferentes, como también lo son los métodos propuestos para su tratamiento (Drummond, 2005).

En el caso de las evaluaciones económicas que se basan en micro-datos de pacientes (normalmente obtenidos a partir de ensayos clínicos), se han desarrollado recientemente nuevos métodos estadísticos para solucionar el problema de la transferibilidad de resultados en. En particular, ha recibido gran atención el caso de las evaluaciones económicas basadas en estudios clínicos multinacionales y/o multicéntricos, cuya principal ventaja nace de la posibilidad de reclutar un amplio número de pacientes en diferentes países. Desde el punto de vista del análisis económico dos son los posibles enfoques para un estudio multinacional.

De un lado, podrían considerarse transferibles los datos de efectividad y de recursos consumidos entre los diferentes centros y países del ensayo clínico. En este caso, se



podría multiplicar la *media* de los recursos consumidos por todos los pacientes del ensayo clínico por los costes unitarios de un país, para obtener los costes totales por paciente de aquel país. Análogamente, se podría estimar la efectividad *media* de un tratamiento para todos los pacientes del ensayo y considerar que este valor es apropiado para todos los países.

El enfoque alternativo consistiría en no considerar transferibles los datos de efectividad y de recursos utilizados. En este caso, cada país tendría que estimar la efectividad y los recursos consumidos por un tratamiento sólo a partir de datos de los pacientes enrolados en el propio país (Schulman 1998; Cook, 2003). Aunque este enfoque parece el más apropiado dada la heterogeneidad que presumiblemente se encontrará entre pacientes de diferentes países, presenta el problema de que el tamaño de las muestras de pacientes en algunos países puede ser demasiado pequeño, con los consiguientes problemas de significación estadística que ello plantea (Sculpher, 2004). Por esta razón se han propuesto algunos métodos que pretenden hacer frente a este problema. Por ejemplo, se podría aplicar un test de heterogeneidad entre los pacientes de diferentes países para estimar qué datos son transferibles y cuáles no (Hoch, 2002). También se podrían aplicar métodos de regresión para separar los datos del análisis y adaptar algunos elementos al lugar específico. Por último, los “modelos multi-nivel” (*multi-level models*) permiten estimar la variabilidad entre diferentes centros y/o países de un estudio y calcular una ratio coste-efectividad específica por cada lugar (Manca, 2005).



En el caso de los análisis basados en modelos, si bien el uso de éstos permite sintetizar datos de diferentes fuentes, los estudios a los que se acude para obtener los parámetros de un modelo difieren normalmente en su diseño (ensayos clínicos, estudios observacionales, etc.), en las características de los pacientes, en su calidad y en el lugar de realización (Wilke, 2003). Puede que las guías recomienden métodos para seleccionar los estudios y sintetizar la información, así como que mencionen explícitamente el problema de la transferibilidad en relación con el análisis de sensibilidad.

Otra de las cuestiones que deben tratarse es la relativa a los problemas de validez externa de los ensayos clínicos. Cuando una evaluación económica se basa en un ensayo clínico algunos elementos pueden diferir de la práctica clínica “real”. Por ejemplo, puede que algunos recursos consumidos dependan de las condiciones especiales del estudio y no se encuentren en la práctica habitual (*protocol-driven costs*). También pueden existir diferencias en el nivel de experiencia de los profesionales que participan al estudio, en relación con la media de los profesionales sanitarios, así como diferencias en las características de los centros que participan en el ensayo respecto de la media de los centros del país. Existen procedimientos para solucionar estos problemas, como el uso simultáneo de datos obtenidos de fuentes diferentes al ensayo clínico para comparar los resultados.



1.12. Presentación de los resultados

La presentación de los resultados de una evaluación económica constituye un factor fundamental del análisis, en la medida en que puede orientar la toma de decisiones sobre las tecnologías analizadas. Además, cuanto más clara y transparente sea la información, mayores serán las posibilidades de revisión o ampliación del análisis en investigaciones futuras. De ahí que sea habitual que las guías recomienden un caso de referencia (*referente case*) o una *checklist* que los investigadores deben seguir en la presentación de sus resultados (Drummond, 1996). Con carácter general, los resultados deberían ofrecer información sobre costes y efectos agregados y desagregados, incluir un análisis incremental de costes y efectos, tener en cuenta la incertidumbre en los resultados y presentar en detalle (por medio de tablas y gráficos) los análisis de sensibilidad, de subgrupos, de escenarios, etc.

La presentación de costes y efectos de forma desagregada permite obtener más información sobre algunos aspectos potencialmente importantes del análisis (Siegel, 1996). A menudo, no basta con referir los costes y los efectos totales asociados a unas tecnologías, así como su ratio coste-efectividad. En particular en el caso de los costes, parece necesario presentar la información para cada tecnología desglosada por categorías (por ejemplo, medicamentos, servicios hospitalarios, visitas, pruebas, etc.), descontados y no descontados, separados por subgrupos de la población de estudio (cuando se realiza análisis de subgrupos), según la perspectivas utilizadas (cuando se utiliza más de una perspectiva), etc. Este grado de detalle en la



presentación de los costes permite, por ejemplo, estimar el impacto de una nueva tecnología sobre los diferentes agentes sanitarios, pacientes e instituciones.

La información sería aún más rica si los efectos de las tecnologías se presentaran también de forma desagregada. Por ejemplo, se podrían referir los AVAC ganados en cada uno de los años que integran el horizonte temporal del análisis con el fin de estimar si los beneficios de las tecnologías se producen en el corto o en el largo plazo. También podrían presentarse los resultados por subgrupos de pacientes, o dar cuenta del tiempo estimado que los pacientes pasarán en cada estado de salud (en el caso de utilizar modelos de Markov).

De otro lado, una evaluación económica se basa en el análisis comparativo entre dos o más tecnologías, de lo que se deduce que la ratio coste-efectividad ha de ser “incremental”, esto es, ha de estar calculada como cociente entre el aumento en costes que supone una alternativa y la ganancia en efectividad que proporciona, en ambos casos respecto a la tecnología utilizada como término de comparación (“comparador”). Una ratio coste-efectividad ‘media’ (es decir, costes totales divididos por beneficios totales de una tecnología) no proporciona una información adecuada, pues únicamente da cuenta de la comparación entre la tecnología y un hipotético programa con costes y beneficios nulos. En consecuencia, una evaluación económica tiene que calcular ratios coste-efectividad incrementales entre las alternativas comparadas (Drummond, 2005).



Cuando la comparación tiene lugar entre más de dos estrategias alternativas es importante no calcular los ratios incrementales coste-efectividad para las alternativas que resultan “dominadas” por otras (más caras y menos efectivas que otras tecnologías) o por combinaciones lineales de otras (supuesto de “dominancia extendida”). En estos casos basta con identificar dicha tecnología como una estrategia dominada, toda vez que los ratios coste-efectividad no tienen ninguna relevancia.

La presentación de los resultados del análisis puede hacerse mediante tablas (costes y beneficios incrementales y ratios incrementales) o por medio de gráficos. Una forma habitual de presentar gráficamente los resultados de una evaluación económica es el “plano coste-efectividad”, cuyos ejes de abscisas y ordenadas miden, respectivamente, los beneficios y los costes incrementales de los programas respecto del comparador, que se sitúa en el origen de coordenadas.

Ya se ha puesto de manifiesto la importancia de un análisis detallado de la incertidumbre y la variabilidad que afecta a los datos y supuestos del estudio (véase el apartado 1.9). También es importante tener en cuenta la incertidumbre en los resultados finales del análisis (costes, efectos y ratios coste-efectividad). Por tanto, parece apropiado que los resultados no se presenten únicamente en forma de estimaciones puntuales, sino que se muestren de forma detallada todos los resultados del análisis de sensibilidad, de subgrupos, de escenarios, etc (Manning, 1996). A tal fin se pueden utilizar las “curvas de aceptabilidad”, en el caso de un análisis



probabilística, donde se representa gráficamente la probabilidad de que un tratamiento sea coste-efectivo según diferentes niveles de disposición a pagar (Fenwick, 2001). Por otro lado, en el caso de análisis de subgrupos, los resultados se pueden presentar en tablas.

Por último, junto a la estructura que debe seguir la presentación de resultados, las guías podrían sugerir cómo interpretar los resultados coste-efectividad obtenidos. Por ejemplo, puede resultar recomendable hacer explícitas las limitaciones del análisis, subrayando los elementos más inciertos que requieren de investigaciones futuras. También sería interesante comparar los resultados obtenidos con los de otras evaluaciones publicadas sobre la misma tecnología o sobre otras tecnologías utilizadas para el mismo problema de salud. Por último, se podrían explicitar unos criterios de decisión para estimar cuándo una ratio coste-efectividad es aceptable. En relación con esto último, resulta prácticamente imposible invocar un criterio claro de toma de decisiones cuando la medida de efectividad utilizada es una medida intermedia. Cuando se utilizan los AVAC o los años de vida ganados como medida de beneficios, es más habitual encontrar reglas de decisión en la literatura. Por ejemplo, muchos estudios sugieren la cifra de €50,000 por año de vida ganado (o AVAC) como el umbral máximo aceptable para las nuevas tecnologías evaluadas (Mason, 1993; Tengs, 1995).



2. Revisión de las guías (NICE, Canadá, Australia y Suecia)

2.1. Perspectiva de la evaluación económica

Las guías del NICE establecen que la perspectiva, en lo que se refiere a los costes, debe ser la del SNS (el *National Health Service*) y los servicios sociales (SS). Por lo que atañe a los resultados, éstos se limitan a los efectos sobre la salud del paciente principalmente y, en caso de que fuera relevante, la de otros individuos. Las guías establecen que esta perspectiva es congruente con el objetivo de maximizar la salud con los recursos de que disponen el SNS y los SS. Sin embargo, también se señala que, si las tecnologías médicas tienen características que son valoradas por los individuos, pero que no se traducen en mejoras en la salud, éstas deben ser mencionadas en el informe. Lo mismo ocurre por el lado de los costes, esto es, las guías proponen que se incluyan, como estimaciones secundarias, los costes que recaen fuera del presupuesto del SNS y de los SS.

En las guías **canadienses** también se establece que la perspectiva ha de ser la del SNS en el caso base, admitiéndose que esta perspectiva se pueda ampliar a los servicios sociales. Al igual que en el caso anterior, se abre la posibilidad a realizar un análisis secundario tomando en consideración los costes que recaen fuera del SNS,



en el supuesto de que tuvieran un impacto significativo sobre los resultados. Por el lado de los resultados, las guías son menos explícitas en lo que respecta a la perspectiva. No obstante, se señala que el caso base utilizará como medida de resultados la mejora en la salud del paciente, excluyéndose explícitamente, las consecuencias sobre la salud de sus cuidadores informales. La guía prevé que estas consecuencias sobre cuidadores informales u otras características de las tecnologías valoradas por los pacientes, puedan incluirse entre los resultados en el curso de un análisis secundario.

Las guías **australianas** de 2006 adoptan una perspectiva claramente social. Por el lado de los recursos consumidos, se considera que las evaluaciones económicas deben recoger todos los costes, con independencia de quien los soporte (entidades públicas, pacientes, compañías de seguro, etc.). Análogamente, en relación con los beneficios, se contempla la inclusión de los beneficios no sanitarios o que recaen sobre terceros.

Las guías **suecas** también estiman que el análisis debe realizarse desde una perspectiva social. En consecuencia, recomiendan la inclusión de todos los costes, con independencia de si recaen sobre una jurisdicción estatal o local, sobre el



paciente o sobre terceros. Por el lado de los beneficios, las guías abren la puerta al uso del método de la disposición a pagar, por lo que, en principio y en congruencia con la perspectiva social, se pueden medir beneficios que no se limiten a la mejora en la salud del paciente. Cabe interpretar, por tanto, que las guías suecas, aunque no lo digan explícitamente, asumen que el objetivo de la evaluación económica es maximizar el bienestar de la población (no únicamente su salud) a partir de los recursos de que dispone la sociedad en su conjunto.

2.2. Revisión sistemática de la literatura

Las guías del **NICE** exigen la realización de una revisión sistemática de la evidencia clínica. En una primera fase se habrán de identificar todos los estudios potencialmente relevantes sobre las tecnologías sanitarias comparadas, siendo posible, en una segunda fase, excluir algunos estudios, cuando no se consideren relevantes para el contexto de Reino Unido, o no se estimen adecuados en términos de calidad, por las medidas clínicas utilizadas, los comparadores elegidos, etc.

No se requiere de manera explícita una información específica sobre los métodos de búsqueda y los criterios de inclusión/exclusión de estudios. Sin embargo, las guías subrayan de modo reiterado la necesidad de referir de forma transparente los métodos utilizados para identificar y seleccionar los estudios relevantes. Por ejemplo, si se excluye un estudio identificado en la revisión de la literatura, es preciso ofrecer una justificación clara para su exclusión. Se recomienda que, tanto la revisión de la



literatura como la selección de los estudios, se realice por más de un investigador, pararon el fin garantizar la validez del proceso.

Las guías del NICE exigen presentar una lista detallada de todos los artículos elegidos para obtener la evidencia clínica. Algunos de los detalles que se han de incluir sobre cada estudio son las características de los pacientes (edad, sexo, nivel de gravedad de la enfermedad), las alternativas comparadas, las medidas y los resultados de eficacia y tolerabilidad de las alternativas comparadas, el contexto del análisis (centro y país) y el año de publicación. Además es obligado hacer un análisis crítico sobre la calidad de cualquier estudio clínico elegido. Las guías indican que los ensayos clínicos basados en comparaciones directas entre las tecnologías estudiadas representan las mejores fuentes de datos. Sin embargo, se asume que no siempre es posible disponer de tal información, siendo necesario incluir ensayos clínicos con comparaciones indirectas o estudios observacionales. En estos casos, el investigador ha de subrayar las posibles limitaciones asociadas al diseño u otras características de los estudios.

Se sugiere el uso de meta-análisis cuando se disponga de un número suficiente de estudios relevantes y con medidas de efectividad comparables. En caso contrario, será suficiente presentar los resultados de cada estudio, explicando las potenciales limitaciones. Las guías señalan la conveniencia de estimar el impacto de los diferentes resultados de los estudios en el análisis de sensibilidad y de subgrupos. Así mismo, subrayan la importancia de distinguir entre el efecto absoluto de una



tecnología y el riesgo relativo, no sólo en el caso de comparaciones indirectas, sino también cuando se pretenda adaptar datos de estudios de otros países al contexto del Reino Unido.

El NICE mucho énfasis en el análisis de la heterogeneidad entre los estudios seleccionados. En general, las cuestiones de heterogeneidad, como las de incertidumbre y las de variabilidad de los resultados clínicos, aparecen en diferentes apartados de las guías. Cuando hay heterogeneidad en los resultados estadísticos de los estudios analizados, la recomendación es utilizar un meta-análisis de efecto aleatorio (*random effect*), antes que un meta-análisis de efectos fijos (*fixed effect*). La variabilidad en las características de los pacientes de los diferentes estudios es otro aspecto fundamental del análisis y tiene que ser evaluado por medio de análisis de subgrupos o de sensibilidad. Por último, cuando hay heterogeneidad en la calidad de los estudios incluidos se sugiere realizar análisis de sensibilidad variando el número de estudios incluidos en el meta-análisis. No hay en la guía del NICE un apartado específico referido a la heterogeneidad en las comparaciones indirectas, pero resulta siempre necesario presentar el riesgo relativo y el valor absoluto de la eficacia de las tecnologías comparadas.

En las guías de **Canadá** se establece claramente que la evidencia clínica para el análisis económico tiene que derivar de una revisión sistemática de la literatura, para



no excluir estudios potencialmente relevantes. El investigador tendrá que justificar adecuadamente la ausencia de revisión sistemática de la literatura (por ejemplo, cuando sólo exista un ensayo clínico relevante que compara las tecnologías de interés). Se requiere la presentación detallada de los métodos de búsqueda y de los criterios de inclusión y exclusión de estudios, aunque las guías no ofrecen información adicional sobre este punto.

Las guías canadienses indican que es necesario presentar información relativa a las características de los estudios utilizados para obtener datos de eficacia, efectividad y eventos secundarios pero no explicitan qué tipo de información. Sí se establece que los investigadores habrán de estimar la calidad de los estudios utilizados, subrayando su validez interna y externa y explicando las limitaciones de cada uno de ellos.

Las guías afirman que el uso del meta-análisis para sintetizar los datos de eficacia obtenidos por diferentes estudios aumenta la credibilidad de la evaluación económica. Se considera importante presentar de forma transparente cualquier método de síntesis utilizado, justificar su elección, así como referir en detalle todos los resultados. Se destaca, asimismo, la importancia de separar el efecto absoluto de una tecnología y el riesgo relativo de dicha tecnología respecto a un comparador en el caso de comparaciones indirectas y también para adaptar datos de estudios de otros países al contexto canadiense.



El asunto de la heterogeneidad entre los estudios clínicos parece un elemento clave, como lo sugiere el hecho de que es tratado en muchos apartados de las guías canadienses (incertidumbre, transferibilidad, modelos, etc.). Se indica la necesidad de realizar análisis de subgrupos y de sensibilidad para tener en cuenta la variabilidad en las características de los pacientes o de los contextos en que se han realizados los estudios. En el caso de comparaciones indirectas, se sugiere el uso de meta-regresiones para tener en cuenta la potencial heterogeneidad de los estudios utilizados. En general, la variabilidad y las diferencias en los datos clínicos tienen que ser tratadas adecuadamente en el análisis de la incertidumbre.

Las guías **australianas** apuestan claramente por la necesidad de abordar una revisión sistemática de la literatura con el objetivo de identificar al menos todos los ensayos clínicos que comparan directamente las tecnologías analizadas. Si no se llevase a cabo una revisión sistemática de la evidencia clínica, existiría el riesgo de excluir algunos estudios relevantes para la evaluación económica, reduciendo la credibilidad del análisis y aumentando la posibilidad de sesgos en los resultados.

Se exige una presentación detallada de los métodos de búsqueda de la literatura. En particular se considera necesario referir: las bases de datos utilizadas, las fechas iniciales y finales de la búsqueda para cada base de datos, todos los términos utilizados y la metodología, la fecha en que se ha realizado la búsqueda, así como



cualquier búsqueda adicional de artículos no publicados (por ejemplo, presentaciones de conferencias, páginas Web, etc.). Igualmente, se recomienda la inclusión de una descripción completa de los criterios de inclusión y exclusión de estudios. Las guías ofrecen algunos criterios generales para la inclusión de estudios. En cualquier circunstancia es necesario incluir todos los ensayos clínicos aleatorizados que comparen directamente las tecnologías estudiadas. Cuando no se dispone de comparaciones directas, se admite, también, la inclusión de ensayos clínicos que comparen indirectamente las tecnologías mediante un comparador común. En el supuesto de que no existan tampoco ensayos con comparaciones indirectas, se acepta que se incluyan datos procedentes de otros tipos de estudios (por ejemplo, observacionales), aunque ello implique que la validez interna del análisis se reduzca de forma notable. Como regla general, no obstante, si existe evidencia clínica suficiente obtenida por ensayos clínicos con comparaciones directas, se sugiere excluir del análisis económico el resto de estudios disponibles. Los resultados de la búsqueda se habrán de presentar en tablas que contengan el número de estudios identificados en cada base de datos (o en otras fuentes), el número y el diseño de los estudios seleccionados y excluidos, y las razones para la exclusión.

En Australia se exige un alto nivel de detalle en la descripción de las características y de la calidad de los estudios elegidos para la evidencia clínica. Por ejemplo, en el caso de ensayos clínicos con comparaciones directas es necesario incluir los siguientes datos: a) tecnologías comparadas, con detalles sobre las dosis y los métodos de administración en el caso de fármacos, frecuencia y duración del



tratamiento etc.; b) periodo de seguimiento; c) método de aleatorización; d) número de pacientes totales y para cada grupo; e) características demográficas y clínicas de los pacientes, incluyendo los resultados estadísticos de las potenciales diferencias entre grupos; f) resultados del estudio y posibles diferencias estadísticas (se sugieren formas diferentes de presentación de los resultados y de su significación estadística según el tipo de medida clínica). Además, se establece la necesidad de aportar un listado de los estudios incluidos y un breve resumen de cada estudio destacando los elementos más relevantes. Las guías indican que no basta con describir cada estudio elegido, si no que se requiere un análisis crítico de las características de cada estudio, explicando las principales limitaciones. Por ejemplo, en un ensayo clínico el tamaño de la muestra podría ser demasiado pequeño para el análisis estadístico, o podría tener un periodo de seguimiento inadecuado. Por otro lado, los estudios observacionales presentan limitaciones en el diseño mismo del estudio (en particular, *el confounding* cuando existe heterogeneidad en los pacientes de los diferentes grupos). Para cada tipo de estudio las guías australianas proponen una lista de preguntas que el investigador tiene que contestar con el fin de valorar la calidad de la evidencia clínica incluida.

Se recomienda el uso de meta-análisis cuando la evidencia clínica deriva de dos o más ensayos clínicos aleatorizados de calidad comparable. El meta-análisis permite obtener resultados más claros y podría generar diferencias significativas en términos estadísticos entre las alternativas comparadas no reveladas individualmente en los ensayos clínicos, debido tal vez a un tamaño de la muestra inadecuado. En caso de no



realizar el meta-análisis para la síntesis de la evidencia clínica, el investigador tendrá que explicar las razones de tal proceder, y explicar cualquier método alternativo utilizado. También es obligado referir de forma detallada los resultados del meta-análisis, presentando los resultados de cada estudio y los resultados de la combinación de los mismos. Las guías australianas proporcionan mucha información acerca de los métodos de combinación de los ensayos clínicos con comparación directa, sugiriendo diferentes técnicas estadísticas según el tipo de medida clínica encontrada. En cualquier caso, se ha de ofrecer información sobre el valor absoluto de la eficacia de las alternativas comparadas, el riesgo relativo y los intervalos de confianza. Por último, las guías contienen un apartado específico dedicado a las comparaciones indirectas, en el que se fija la necesidad de calcular el riesgo relativo (y el intervalo de confianza) de las alternativas comparadas respecto al comparador común. De nuevo, en este caso se proponen diferentes técnicas estadísticas según el tipo de medida clínica utilizada en los estudios originales. Así mismo, se ofrecen algunos modelos de tablas para facilitar la presentación de los resultados de la síntesis de la evidencia clínica.

La heterogeneidad en la calidad de los estudios seleccionados y las potenciales diferencias en las características de los pacientes de dichos estudios se aborda en varios apartados de las guías australianas. Con carácter general, se recomienda realizar un meta-análisis de efectos aleatorios (*random effect*) para tener en cuenta la heterogeneidad en los resultados estadísticos de los estudios analizados. También se proponen análisis de subgrupos y de sensibilidad (probabilístico) en el supuesto de



variabilidad entre pacientes, contextos y otras características de los estudios. Cuando existen diferencias entre pacientes en ensayos clínicos con comparaciones indirectas, se advierte sobre la necesidad de aplicar técnicas estadísticas adecuadas. Por ejemplo, se sugiere la posibilidad de recurrir a meta-regresiones o modelos lineales generalizados (*generalised linear models*) en el caso de que exista un número suficiente de estudios.

Las guías **suecas** se limitan a señalar la importancia de ofrecer información adecuada sobre los datos clínicos utilizados, las fuentes y los métodos de síntesis, pero no suministran ningún otro detalle.

2.3. Población objetivo

Las guías del **NICE** recomiendan ofrecer detalles sobre las características demográficas de los pacientes que pueden beneficiarse de un tratamiento (edad, sexo, clase socioeconómica) y distinguir entre diferentes niveles de gravedad de la enfermedad, si se considera necesario.

El análisis de subgrupos representa un factor fundamental en la evaluación económica para el NICE. Se requiere un análisis detallado de los resultados de efectividad y de costes para subgrupos homogéneos de la población objetivo. Es



también obligado informar sobre las características demográficas y biológicas de los subgrupos, explicando las razones que llevan a predecir resultados diferentes entre dichos subgrupos. Los métodos de análisis tienen que estar justificados. Se recomienda realizar análisis de sensibilidad para los parámetros más inciertos, ya que puede existir más incertidumbre alrededor de los datos de subgrupos de pacientes que respecto de los de la población total de un estudio. Las guías ponen énfasis también en las posibles diferencias entre grupos en lo que se refiere a efectos secundarios de los tratamientos.

Se subrayan las posibles diferencias entre la población de un ensayo clínico y los pacientes que se pueden beneficiar de un tratamiento en la práctica clínica real. Se sugiere realizar un análisis de subgrupos incluyendo sólo los datos de los pacientes que completan el estudio para reflejar mejor lo que podría ocurrir en la práctica real (contrariamente al caso base, donde se recomienda utilizar el método del *intent-to-treat*). Es necesario estimar el impacto en términos de costes totales que una nueva tecnología puede tener para el SNS del Reino Unido, incluyendo datos epidemiológicos sobre la enfermedad y el número de pacientes a los que se podría destinar la tecnología.

Las guías **canadienses** afirman que la elección de la población objetivo puede ser fundamental para determinar el coste-efectividad de un tratamiento, razón por la cual



se exige proporcionar todos los detalles relevantes de dicha población, que incluyen las características demográficas (edad, sexo, clase socioeconómica) y los datos sobre la enfermedad de dichos pacientes (gravedad de la enfermedad, posibles enfermedades concomitantes, niveles de riesgo de efectos secundarios, etc.). También se prevé la posibilidad de aportar información adicional sobre la población objetivo, como el centro donde se recibe el tratamiento (por ejemplo, hospital), la tasa de cumplimiento, el tipo de tratamiento que reciben en la práctica habitual, etc.

Se requiere un análisis de subgrupos cuando hay diferencias entre pacientes en efectividad, preferencias (utilidades) o costes del tratamiento. En este sentido, las guías canadienses señalan la conveniencia de presentar resultados para grupos de pacientes homogéneos. El investigador tendrá, también, que considerar la posibilidad de que un tratamiento aprobado sea utilizado para grupos de pacientes o indicaciones diferentes al caso base, y deberá realizar un análisis incluyendo dichas desviaciones de la práctica estándar. Las guías ponen énfasis en el posible riesgo diferente asociado a pacientes con características particulares (por ejemplo, mujeres embarazadas).

A este respecto, se subrayan las dificultades que conlleva definir de manera apropiada y detallada los candidatos a recibir una nueva tecnología sanitaria y las posibles diferencias entre los pacientes de un ensayo clínico y la población objetivo en la práctica real. Por eso se recomienda tener en cuenta estos aspectos en el análisis de subgrupos y de sensibilidad. Las guías exigen estimar el impacto presupuestario



de una nueva tecnología, en el que se consideren los datos epidemiológicos de la enfermedad y la tasa de utilización esperada de dicha tecnología.

Las guías **australianas** subrayan la importancia de una definición muy detallada de la población objetivo. En particular, consideran necesario proporcionar información sobre las siguientes características: edad, sexo, gravedad de la enfermedad, tratamientos recibidos anteriormente (si son relevantes), así como cualquier criterio específico para la elección de los pacientes. También se recomienda informar sobre otros aspectos como el centro donde se recibe el tratamiento, los posibles efectos secundarios de los tratamientos, la tasa de cumplimiento, etc.

Estas guías advierten de la necesidad de tener mucha cautela al realizar un análisis de subgrupos, siendo preciso justificar con test estadísticos la elección de dichos subgrupos. Las guías australianas se centran en el análisis de las diferencias entre pacientes en términos de eficacia clínica. Por ejemplo, cuando los datos de eficacia derivan de ensayos clínicos, la heterogeneidad entre pacientes tiene que demostrarse por medio de “tests de interacción”, que prueben y cuantifiquen la asociación entre el efecto del tratamiento y las características de los pacientes. Para el caso de variables continuas se puede realizar un análisis de sensibilidad estándar.



Se recomienda presentar las posibles diferencias entre la población objetivo de la evaluación económica y la población analizada en los ensayos clínicos. Se advierte de que la población objetivo en la práctica real es, en general, más amplia y más heterogénea que a la población de los estudios clínicos. Se estima necesario también considerar el número total de personas que sufren la enfermedad en el país, el número de pacientes que se pueden beneficiar del tratamiento, y la tasa de utilización que se espera para la nueva tecnología. Se ofrecen muchos detalles sobre los métodos de análisis de impacto presupuestario (que veremos en su apartado específico).

En las guías de **Suecia** no se da información sobre el nivel de detalle necesario para definir la población objetivo. Mientras el caso base de la evaluación económica debe tener en cuenta a toda la población que se puede beneficiar de una tecnología sanitaria, las guías recomiendan desarrollar análisis de subgrupos cuando se presume que el coste-efectividad de un tratamiento puede variar según las características de los pacientes. Por ejemplo, se pueden realizar análisis de subgrupos por sexo, edad, nivel de riesgo y gravedad de la enfermedad. Se indica que la población total a la que se destina la tecnología evaluada puede ser diferente de la población analizada en un ensayo clínico y los modelos son los instrumentos más adecuados para orientar el análisis a un contexto real. También se recomienda realizar una estimación del número de pacientes que se pueden beneficiar de la tecnología evaluada en Suecia.



2.4. Término de comparación de la tecnología evaluada

Las guías del NICE recomiendan considerar, en primera instancia, todas las posibles alternativas a la tecnología de estudio para una indicación específica. Después, el análisis se podría limitar a la alternativa más efectiva y a la práctica vigente, si no coincidieran ambas.

Las guías subrayan las dificultades que emergen en algunas ocasiones en la tarea de elegir los comparadores ideales, ya que la práctica clínica puede no ser consistente en diferentes centros sanitarios de Inglaterra y Gales, o entre el Reino Unido y otros países. Los elementos clave para seleccionar los comparadores son la efectividad del tratamiento y su difusión en el Reino Unido (práctica habitual). En algunos casos, la tecnología estudiada y los comparadores tienen que ser evaluados a lo largo de una secuencia de tratamientos. No se proporcionan detalles explícitos sobre los métodos de selección de los comparadores.

La práctica habitual se define como la práctica más comúnmente utilizada en el Reino Unido. No se ofrece más detalle al respecto. Se admite que, en algunos casos, el comparador pueda ser “no hacer nada” o “intervención mínima” (*supportive care*), pero no se precisa más.



Las guías **canadienses** recomiendan tener en cuenta en la fase inicial todas las alternativas aprobadas para la indicación de la tecnología objeto de estudio. Conviene después, a juicio de las guías, seleccionar las alternativas más apropiadas a través de un proceso de eliminación en el que los expertos clínicos pueden ser de gran ayuda para identificar los comparadores relevantes. En general, los comparadores más apropiados incluyen la práctica vigente y la estrategia más efectiva (si no coincide con la primera). En el caso de fármacos, se considera como comparador también la alternativa más barata y más comúnmente utilizada.

Las guías asumen la dificultad de incluir en el análisis todas las alternativas existentes para la tecnología de estudio, de ahí que resulte necesario seleccionar los comparadores sobre la base de criterios como, por ejemplo, la disponibilidad de datos o el tiempo disponible para realizar el análisis. También se considera importante, a la hora de elegir un comparador, tener en cuenta la variabilidad en las prácticas clínicas en diferentes áreas de Canadá, así como las diferencias entre subgrupos de pacientes. La práctica vigente (que se recomienda como comparador) podría, en tal caso, variar entre pacientes y centros. Se señalan situaciones en las que el nuevo tratamiento debe ser comparado a lo largo de una secuencia de tratamientos, en particular en el caso de fármacos. Por último, se advierte de que el investigador debería tener en cuenta las nuevas tecnologías, más baratas, que pueden entrar en el mercado durante el horizonte temporal del análisis, como es el caso de fármacos genéricos cuando se sabe que la patente de un medicamento va a caducar en breve.



La opinión de expertos clínicos es fundamental en la elección de los comparadores más apropiados.

Las guías canadienses definen como práctica vigente la más comúnmente utilizada para la indicación de estudio. Esto significa que la práctica vigente es la estrategia utilizada para el mayor número de pacientes, y se puede identificar por medio de bases de datos o a través de las opiniones de los expertos. Cuando no exista una única práctica habitual, se habrá de comparar la nueva tecnología con todas las vigentes. En algunos casos, las guías estiman que puede ser apropiado comparar la nueva tecnología también con la opción “no hacer nada” o con la “intervención mínima” (*watchful waiting*), aunque no se dan más detalles.

Las guías **australianas** se centran en el caso de los medicamentos. Cuando el nuevo fármaco pertenece a una clase terapéutica donde ya existen medicamentos aprobados por el PBS, el comparador más apropiado será el fármaco perteneciente a dicha clase más comúnmente utilizado en Australia. En el caso que el medicamento se incluya en una nueva clase terapéutica, el comparador más apropiado será el fármaco más utilizado para la indicación de estudio. Finalmente, en el caso de que no exista ningún fármaco aprobado para dicha indicación, el comparador será la práctica clínica más utilizada. Las guías identifican circunstancias en las que conviene incluir más de un comparador (véase párrafo siguiente).



Las guías afirman que, normalmente, no existe desacuerdo acerca del comparador principal de un nuevo fármaco, si bien su elección es fundamental para un correcto análisis coste-efectividad. El criterio general es elegir como comparador el fármaco más utilizado en la práctica, y su selección dependerá de la clase terapéutica del nuevo medicamento. En el caso que el nuevo fármaco sea, por ejemplo, una vacuna, el comparador será la vacuna ya existente candidata a ser sustituida por la nueva y, de no existir vacuna previa, se utilizará como referencia la práctica clínica habitual. La guía recomienda tener en cuenta también las diferencias entre pacientes. Por ejemplo, si un nuevo tratamiento puede sustituir diferentes fármacos para subgrupos de pacientes, todos estos fármacos deberán ser incluidos como comparadores. También hay que tener en cuenta los datos disponibles. Las guías aconsejan utilizar datos que procedan de ensayos clínicos con comparaciones directas entre fármacos. Por último, se indica que la opinión de expertos puede ser importante a la hora de elegir el comparador.

La práctica vigente se define como la estrategia utilizada para el mayor número de pacientes, y se puede identificar por medio de opiniones de expertos. En otros términos, la práctica vigente será la estrategia que el nuevo fármaco va a sustituir con más probabilidad. Cuando no exista una única práctica habitual, es conveniente comparar la nueva tecnología con todas las vigentes. Admiten las guías que, en algunos casos, el comparador puede ser “no hacer nada” o “intervención mínima” (*active placebo*), sobre todo cuando éste hubiera sido el comparador utilizado en el



ensayo clínico principal para el nuevo fármaco. No se ofrecen, sin embargo, otros detalles al respecto.

En **Suecia** se recomienda comparar la nueva tecnología con la intervención más apropiada para el contexto sueco, o sea, la práctica habitual. No se proporcionan detalles acerca de los elementos clave para seleccionar los comparadores. La práctica vigente se define como la práctica más comúnmente utilizada, sin que conste en las guías más información. Finalmente, las opciones “no hacer nada” o “intervención mínima” pueden ser incorporadas en el análisis cuando representen la práctica más comúnmente utilizada.

2.5. Medida de resultados preferida

Las guías del **NICE** no excluyen de forma directa la utilización de variables intermedias, sin embargo, insisten en que el resultado principal ha de ser clínicamente relevante, esto es, ha de medir los efectos sobre la salud y los efectos adversos que sean importantes para los pacientes. Aunque estas guías no aclaran lo que se entiende por resultados “clínicamente relevantes”, el contexto general permite deducir, como veremos a continuación, que la utilización de variables intermedias no se contempla.



Las guías establecen claramente que, para el caso de referencia, se requiere un instrumento genérico (no específico) estandarizado y validado que cuantifique los efectos de las tecnologías en términos de CVRS de los pacientes. Esto no excluye el uso de instrumentos específicos, pues se admite la posibilidad de que se presenten análisis adicionales (aparte del caso de referencia) donde la CVRS se haya medido utilizando instrumentos específicos para una enfermedad, siempre que pueda justificarse la necesidad de recurrir a instrumentos específicos.

Las guías no tratan específicamente la cuestión de si las medidas de CVRS han de estar basadas en la psicometría o en preferencias (utilidades), aunque cabe deducir que parecen inclinarse por estas últimas. Así, se puede leer “...el valor de los cambios en la CVRS de los pacientes (esto es, utilidades)...”.

Las guías del NICE se inclinan claramente por la población general como fuente de obtención de las preferencias. Insisten en que se deben valorar los efectos de las tecnologías médicas sobre la CVRS de manera que reflejen las preferencias de la población general. Esto, por otra parte, parece excluir los perfiles de salud basados en la psicometría a la hora de evaluar la CVRS. No obstante, también se acepta que se puedan presentar análisis adicionales basados en valores de utilidad procedentes de pacientes, si se puede justificar adecuadamente y si los resultados se modifican sustancialmente respecto de los del caso base.



El NICE establece que los métodos utilizados para obtener utilidades de la población general han de estar basados en elecciones entre alternativas (*choice-based*). En consecuencia, las guías aceptan tanto la “Compensación de Tiempos” (*Time Trade-Off*, TTO) como la “Lotería Estándar” (*Standard Gamble*, SG), excluyendo métodos no basados en elecciones como la “Escala Visual Analógica” (*Visual Analogue Scale*, VAS).

Las guías consideran necesario utilizar un sistema de clasificación que describa la CVRS de los pacientes a través del tiempo. Reconocen la existencia de diversos sistemas de clasificación cuyos resultados no siempre son comparables. Sin embargo, y pese a identificar el Euroqol como el instrumento más utilizado en el Reino Unido, no se inclinan por ningún instrumento de medida de CVRS en particular, limitándose a señalar que el evaluador debe ofrecer razones que justifiquen el uso de uno concreto.

Las guías del NICE establecen que el valor de los efectos sobre la salud debe expresarse en términos de AVAC para el horizonte temporal apropiado. No obstante, también admiten que, en ocasiones, los supuestos que subyacen al modelo AVAC (intercambio proporcional constante, independencia aditiva) pueden no ser adecuados, por lo que se acepta el uso de alguna medida alternativa, como los años de salud equivalente (HYE), siempre y cuando se justifique que el uso de los AVAC no resulta apropiado en el contexto.



Por último, el NICE no niega que la valoración monetaria de la salud pueda jugar algún papel en la evaluación económica; especialmente, cita el caso de los efectos de los tratamientos que no se plasman en cambios en la salud. En dichos casos, se abre la puerta a la utilización de métodos basados en la disposición a pagar con el fin de poder valorar todas las consecuencias en términos monetarios.

Las guías de **Canadá** muestran una clara preferencia por el uso de variables de resultados finales. Asumiendo que, en muchas ocasiones, surgen problemas derivados de la escasez de evidencia en términos de variables de resultados finales, las guías animan a la utilización de modelos que permitan predecir los resultados finales a partir de datos de variables intermedias. En el caso de que se utilicen variables intermedias, las guías reclaman que la variable intermedia sea altamente predictiva de un resultado importante para el paciente. No obstante, en tal supuesto, las guías recuerdan la conveniencia de modelizar con el fin de estimar el impacto de las variables intermedias sobre los resultados.

Las guías canadienses hacen una distinción entre perfiles que, implícitamente, difiere de la realizada por el NICE. Distinguen entre perfiles específicos, perfiles genéricos y perfiles basados en preferencias. Los específicos y los genéricos se engloban dentro de lo que en este informe hemos denominado medidas basadas en la psicometría. Por el contrario, en el documento del NICE algunos perfiles como el Euroqol serían



considerados como medidas genéricas. Las guías canadienses son muy tajantes en este punto y señalan que las medidas basadas en preferencias proporcionan el único enfoque adecuado para su uso en el análisis coste-utilidad. Puede resultar conveniente aclarar que el hecho de que las guías del NICE acepten los perfiles genéricos y las canadienses no, simplemente se debe a un problema de denominación de perfiles basados en preferencias. Según las guías del NICE, estos perfiles se pueden denominar como “genéricos” pero según las guías canadienses, este término se restringe para perfiles procedentes de la psicometría como el SF-36. Más allá de esta discrepancia, ambas guías coinciden en este punto: en las evaluaciones económicas únicamente han de usarse perfiles basados en preferencias.

Las guías canadienses aconsejan el uso de preferencias basadas en una muestra representativa de la población general, adecuadamente informada acerca de los estados de salud que se valoran. El argumento para ello es que los miembros de la población general son los financiadores últimos del sistema público de salud y, además, son pacientes potenciales. No obstante, las guías también consideran que los pacientes que tienen experiencia directa con los estados de salud relevantes pueden ser una fuente aceptable de las preferencias. La recomendación consiste en usar las preferencias de la población general en el caso de referencia y las de los pacientes en un análisis de sensibilidad, si bien se advierte de que esto puede ser “innecesario o no posible”. En suma, la postura de las guías canadienses es menos clara que la del NICE.



En coincidencia con las guías del NICE, las canadienses recomiendan el uso del la “Compensación de Tiempos” (TTO) y la “Lotería Estándar” (SG). Consideran que la “Escala Visual Analógica” (VAS) no debe utilizarse como único método de obtención de preferencias, aunque tampoco aclaran del todo si debe jugar algún papel o no cuando se utiliza junto con otros métodos. Las guías muestran su preferencia por el uso de perfiles de salud para el cálculo de utilidades en las evaluaciones económicas, ya que son fáciles de obtener, comparar e interpretar. Sin embargo, tampoco excluyen la medida directa de preferencias mediante TTO o SG, señalando que dicho modo de proceder puede ser apropiado en algunas circunstancias.

Las guías canadienses apuestan por los AVAC como medida de resultados en las evaluaciones económicas. Sin embargo, parecen ser algo más críticas con las limitaciones de los AVAC que el NICE, pues demandan que los cambios en la esperanza de vida y en la calidad de vida se presenten de manera separada. De igual modo, exigen que se explicita el método utilizado para combinar estos dos resultados en un único indicador. Es decir, aunque la recomendación consiste en emplear los AVAC para medir los resultados, el nivel de exigencia es superior al de las guías del NICE en lo que se refiere al detalle con el que los evaluadores deben presentar cómo se han calculado los AVAC a partir de información sobre la duración y calidad de vida.

No se excluye el ACB ni los métodos de valoración monetaria de la salud aunque se advierte de que, por no haber sido estos métodos ampliamente usados en el contexto



de decisiones de asignación de recursos en el sector sanitario, resulta necesario investigar más con el fin de validar el uso de dichos métodos en el ámbito de la evaluación económica.

Las guías de **Australia** manifiestan su preferencia por las variables de resultados finales. Sin embargo, se muestran muy sensibles al hecho de que pocos ensayos de medicamentos son lo suficientemente grandes para medir cambios en los resultados finales, por lo que aceptan el uso de variables intermedias. Establecen, no obstante, algunos criterios para pasar de variables intermedias a variables finales, en particular a AVAC.

Las guías contemplan el uso de perfiles genéricos y específicos. Dado que se favorece el uso de AVAC como medida de resultados, los perfiles genéricos son los preferidos. Sin embargo, no se excluyen los específicos debido a las limitaciones que pueden afectar a los perfiles genéricos (posible insensibilidad, falta de representatividad de los participantes en el estudio, etc.).

La preferencia por las medidas de calidad de vida que proporcionan utilidades es clara en las guías australianas. Sugieren obtener las preferencias mediante perfiles de salud multiatributo, donde los estados de salud provengan de pacientes y las utilidades de los perfiles se hayan obtenido de la población general. Las guías hacen



especial hincapié en la conveniencia de que los estados de salud (no así las utilidades) provengan de los pacientes. Desaconsejan, por tanto, cálculos de AVAC basados en entrevistas a la población general en las que se les pide que valoren la enfermedad cuyas utilidades se desean obtener. En suma, se desea evitar que las utilidades se estimen “ex post”, esto es, una vez la evaluación económica está terminada. Se explicita en las guías que los métodos obtenidos par obtener preferencias han de estar basados en elecciones, esto es, se admiten métodos como el TTO o el SG. Por otra parte, las guías señalan el Análisis Coste-Utilidad como el preferido; por tanto, los AVAC son la medida de resultados recomendada.

En las guías australianas se permite la valoración monetaria de los estados de salud, pero se señala que ésta se utilizará de forma suplementaria a la valoración mediante AVAC. Se advierte de que, si se utiliza el método de la valoración monetaria, se deberá justificar la razón para ello. Sin embargo, se ha de hacer constar que en esta cuestión se ha producido un cambio muy nítido entre las guías de 2002 y las del 2006. Aunque sigue siendo una forma de valoración secundaria, las guías dedican bastante espacio a dar indicaciones sobre cuestiones que habrán de tener en cuenta quienes utilicen la valoración monetaria. Todo esto abre la puerta a la utilización del ACB en el caso de Australia.



Las guías **suecas** contemplan la posibilidad de usar variables intermedias pero, en tal caso, exigen la inclusión de un modelo que relacione dichas variables con variables de resultados finales. No comentan nada al respecto del uso de perfiles de salud genéricos o específicos y, aunque no se hace explícito, las guías suecas no parecen contemplar la posibilidad de usar medidas de CVRS basadas en la psicometría.

Por lo que atañe al origen de las preferencias, las guías suecas se separan de la norma general, ya que establecen que las utilidades han de estar basadas en valoraciones de las personas que sufren el estado de salud en cuestión, antes que en ‘pesos’ calculados a partir de valoraciones medias de la población sobre estados de salud hipotéticos (p.ej. la “tarifa social” del Euroqol-5D). Una solución puede ser utilizar en las evaluaciones ‘pesos’ obtenidos de estudios previos con pacientes.

Las guías suecas establecen que las utilidades deben basarse, en primer lugar, en métodos como la “Lotería Estándar” o la “Compensación de Tiempos” y, en segunda instancia, en el método de la “Escala Visual”.

Se recomienda el uso de los AVAC como medida de resultado. En el supuesto de que el efecto del tratamiento afecte principalmente a la esperanza de vida, las guías establecen que el resultado se deberá expresar tanto en AVAC como en años de vida ganados. Sea cual fuere la forma en que se midieran los resultados en los estudios de los que se toman los datos, con variables intermedias o con variables finales (número



de ataques de corazón o de migraña), el resultado se ha de traducir en todo caso a AVAC.

Las guías suecas son las únicas que otorgan plena validez a la valoración monetaria de la salud mediante el procedimiento de la disposición a pagar. Dicho método se considera, de hecho, el más apropiado en los casos en que resulta difícil calcular AVAC (p.ej. dolor intenso sufrido durante un corto periodo de tiempo).

2.6. Costes incluidos en el análisis

Las guías del **NICE** establecen que los costes que se han de considerar en el caso base son los relacionados con aquellos recursos que caen bajo el control del SNS y de los SS y que, además, sean distintos para las diferentes tecnologías que se evalúan. Se acepta la inclusión, aparte del caso de referencia, de los costes que recaen sobre otras partes de la sociedad (pacientes, cuidadores informales). Las guías indican, de forma genérica, que la evaluación económica ha de demostrar que los costes se han identificado de forma sistemática. Indican expresamente que el IVA no debe considerarse como un coste.

No se encuentran apenas indicaciones sobre la forma de medir los recursos en las guías del NICE, más allá de la recomendación genérica de que, si los costes proceden de la literatura, los métodos utilizados para identificar las fuentes de información



estén bien definidos y que, en el caso de que haya varias posibles fuentes alternativas, ha de justificarse la elección de unas frente a otras.

Las guías consideran que los recursos han de ser valorados de acuerdo con los precios que son relevantes para el SNS y los SS. En concreto, señalan que si existe una lista pública de precios (productos farmacéuticos y algunos instrumentos médicos) deben usarse tales precios, con independencia de que el precio real haya resultado ser mayor o menor. Esto es así pese a que se reconoce explícitamente que estos precios oficiales pueden no reflejar los costes de oportunidad. Se hace referencia, aunque de forma un tanto indirecta, a la estandarización de costes, pues se afirma que los costes unitarios y los precios han de ser consistentes entre estudios. De ahí la recomendación de que se utilicen precios oficiales cuando exista alguna lista publicada por el *Department of Health*.

Las guías **canadienses** establecen que los costes han de clasificarse en categorías que sean relevantes para el financiador público. Por un lado, deben clasificarse los costes en función de los diferentes entes públicos (atención primaria, hospitales,...) que vayan a verse afectados por los resultados de la evaluación económica. Por lo que respecta a qué costes incluir y cuáles no, las guías señalan que, con carácter general, todos los costes, presentes y futuros, consecuencia de la intervención han de ser tenidos en cuenta. Entre los costes que no deben incluirse están los propios del



protocolo del ensayo clínico y las transferencias. Por lo que respecta a los costes no relacionados con el tratamiento sanitario que se evalúa y que se producen durante los años de vida ganados a consecuencia de la intervención (*survivor costs*), las guías indican que tales costes pueden incluirse en la evaluación económica, dentro de un análisis de sensibilidad, en el caso de que se estime que su impacto es importante. Finalmente, en relación con los costes en tiempo, se admite la inclusión del tiempo perdido tanto de trabajo remunerado, como no remunerado, no así del tiempo de ocio perdido por los pacientes, ya que se entiende incluido en los AVAC.

Las guías no se inclinan ni por el método *top down* ni por el “microcoste” a la hora de medir los recursos consumidos, aunque señalan que los conceptos que más influencia puedan tener en las estimaciones finales, han de ser valorados con mayor precisión. Señalan que las técnicas utilizadas para repartir costes generales o para calcular amortizaciones han de ser explicitadas, si bien, no se decantan por una metodología concreta. Por lo que se refiere a las fuentes, se admite el uso de datos procedentes de fuentes diversas, que variarán en función de su calidad y de su posible aplicación al ámbito canadiense.

Los recursos, según las guías canadienses, han de valorarse según su coste de oportunidad. Se han de incluir todos los costes (de capital, generales, etc.). Por tanto, se recomienda el uso de los costes medios totales como coste de referencia. Se recomienda el uso de precios de mercado, donde existan, así como la utilización de costes estándar para aquellos casos en que exista disponibilidad. Si los costes se



calculan explícitamente, han de incluir todo el coste económico de los recursos consumidos. Si los costes son compartidos por diversos agentes (p.ej. co-pagos en medicamentos), el analista debe indicar, en un análisis de sensibilidad, las consecuencias para el análisis de posibles variaciones en la parte financiada por el sector público. Por lo que a la valoración del tiempo de trabajo remunerado se refiere, se recomienda el “método de los costes de fricción” (Brouwer y Koopmanschap, 2005). No se inclinan por ningún método en particular a la hora de valorar el tiempo perdido de trabajo no remunerado.

Las guías de **Australia** son muy parcas en sus indicaciones sobre la forma de estimar los costes en las evaluaciones económicas y se refieren, de modo reiterado, al “Manual de Costes asociado” (Commonwealth Department of Health and Ageing, 2002.). Dicho Manual es muy claro y detallado a la hora de explicar la metodología que se debe seguir para calcular los costes de los recursos sanitarios, incluso presenta estimaciones de muchos costes estándar.

Las guías señalan que los costes principales son los que se producen al proporcionar los servicios sanitarios. Afirman que pueden considerarse los costes que se producen más allá del sector sanitario (recursos no sanitarios) aunque advierten de que no serán tan relevantes. El “Manual de Costes” describe, de forma detallada, los recursos sanitarios cuyos costes han de medirse. Desaconsejan la inclusión de costes



indirectos (tiempo de trabajo remunerado perdido) por dos razones: desde el punto de vista de la mejora en el bienestar del paciente, consideran que ese efecto ya debe de estar recogido en la medida de resultados (las utilidades); en cuanto a los beneficios sociales derivados de una más pronta incorporación al trabajo, las guías consideran que esta es una cuestión poco clara. No obstante, ofrecen algunas indicaciones sobre cómo incluir estos costes si es que se considera necesaria su inclusión. Así, consideran que el “método de los costes de fricción” es preferible al “método del capital humano” aunque afirman que incluso aquél sobreestima los beneficios de una posible reducción del número de días de trabajo perdido. Por último, las guías excluyen los costes sanitarios futuros generados por otras enfermedades que los pacientes tendrán en el futuro debido a la mayor esperanza de vida causada por el tratamiento.

El “Manual de costes” describe, caso por caso, la unidad de medida de los recursos sanitarios utilizados y, por lo que respecta a la valoración de dichos recursos, proporciona una lista de costes unitarios estándar que debe utilizarse en una evaluación económica. Todos estos costes se refieren a costes directos sanitarios (o de servicios sociales). Las guías contemplan la posibilidad de que haya conceptos cuyos costes unitarios no estén en el Manual y, por tanto, hayan de ser incorporados por el analista de otras fuentes o bien estimados empíricamente. Las guías indican que tales procedimientos de cálculo serán objeto de evaluación individualizada, caso por caso.



Las guías **suecas** se limitan a indicar que el estudio de costes ha de hacerse bajo una perspectiva social, por lo que deben incluirse todos los costes relevantes. Consideran que deben incluirse los costes indirectos (tiempo perdido de trabajo remunerado) y establecen que la valoración de tales costes ha de basarse en el “método del capital humano”. Las guías también indican que, si el tratamiento prolonga la vida, deben incluirse los costes asociados a la prolongación de la vida en términos netos (consumo total menos mejoras en la producción en los años de vida ganados).

2.7. Horizonte temporal de los costes y efectos

Las guías del **NICE** recomiendan elegir un horizonte temporal suficientemente largo que permita reflejar cualquier diferencia en costes o beneficios entre las tecnologías comparadas. El horizonte temporal puede, en consecuencia, variar según el tipo de enfermedad o de intervención, aunque la recomendación general es que los análisis consideren toda la vida de los pacientes.

Asumen las guías que muchas intervenciones tiene un impacto sobre costes y beneficios durante toda la vida de los pacientes. El caso más evidente es el de los tratamientos de enfermedades crónicas (diabetes, enfermedades cardiovasculares, algunos tipos de cáncer), donde resulta apropiado estimar costes y efectos hasta la muerte del paciente. También, en general, se sugiere utilizar el horizonte temporal de



la vida de los pacientes cuando diferentes tecnologías pueden tener un impacto diverso sobre la mortalidad de los individuos. El investigador tendrá que justificar, en todo caso, la elección de un horizonte temporal más corto. Por ejemplo, tendrá que demostrar que los costes y los beneficios de un tratamiento se producen sólo en un breve plazo, o que no hay ningún impacto sobre la mortalidad de los pacientes.

Se recomienda utilizar modelos para extender los datos primarios de corto plazo a un horizonte temporal más largo (preferiblemente, toda la vida de los pacientes). Como el desarrollo de un modelo de decisión requiere de muchos supuestos, las guías recomiendan justificar y presentar en detalle cualquier supuesto relativo a la estructura del modelo, así como explicitar claramente todos los datos incluidos. Dada la incertidumbre que rodea a muchos de los valores y supuestos del modelo, se sugiere elaborar diferentes escenarios, en particular para los beneficios futuros de un tratamiento.

En el supuesto de que se extrapolen datos de corto plazo a un horizonte temporal más largo, se aconseja utilizar diferentes escenarios para evaluar el impacto de dichos horizontes alternativos sobre los resultados en términos de coste-efectividad.

Las guías **canadienses** recomiendan elegir un horizonte temporal tal que refleje cualquier diferencia en costes o beneficios entre las tecnologías comparadas, lo que



conduce a la sugerencia de utilizar un horizonte temporal que incluya toda la vida de los pacientes. Costes y beneficios tienen que evaluarse a lo largo del mismo horizonte temporal para dar coherencia al análisis.

La recomendación de las guías es utilizar un horizonte temporal que abarque toda la vida de los pacientes en el caso de tratamientos para enfermedades crónicas (artritis reumatoide o diabetes, se citan como ejemplo), así como cuando las alternativas analizadas tienen diferentes impactos sobre la mortalidad de los pacientes. Por otro lado, no se considera necesario extender el horizonte temporal del análisis en los casos en que no se registran diferencias importantes entre costes y beneficios de alternativas en el largo plazo. El investigador tendrá que justificar la elección de un horizonte temporal más corto que el que abarca la vida de los pacientes.

Como es muy difícil obtener datos primarios de largo plazo, los datos que derivan, por ejemplo, de ensayos clínicos se habrán de extrapolar a un horizonte temporal más largo por medio de modelos. Las guías consideran fundamental describir todos los supuestos del modelo, en particular, la relación entre los datos de eficacia de corto plazo y los de largo plazo. Se advierte de que la extrapolación lineal de los costes puede no ser apropiada.

En el caso de enfermedades crónicas se recomienda adoptar diferentes horizontes temporales respecto al caso base (1 año y 5 años, por ejemplo). Estos horizontes alternativos deben presentarse por separado respecto al caso base.



En el caso de **Australia**, el análisis tiene que referirse a un horizonte temporal en el que todos los costes y los beneficios del tratamiento estudiado y de su comparador sean objeto de estimación. El horizonte temporal apropiado está, pues, relacionado con la historia natural de la enfermedad y con el tipo de intervención.

No existe para las guías australianas un único horizonte temporal adecuado, dependiendo la elección de dos factores clave: el tipo de enfermedad y el tratamiento analizado. Por ejemplo, en el caso de enfermedades crónicas, el horizonte temporal tiene que ser muy largo (por ejemplo, úlcera péptica) o abarcar la totalidad de la vida de los pacientes (por ejemplo, cáncer). En otras circunstancias el horizonte temporal puede ser muy corto (por ejemplo, 15-20 días para el tratamiento con antibióticos de infecciones urinarias). El investigador tendrá siempre que justificar el horizonte temporal elegido.

Se recomienda recurrir a modelos cuando sea necesario extender datos primarios de corto plazo a un horizonte temporal más largo. Se mencionan los modelos de Markov como método más apropiado para realizar evaluaciones económicas de largo plazo, porque permite verificar de manera transparente la evolución temporal de los costes y beneficios de diferentes tratamientos. En consonancia con esta sugerencia, se recomienda proporcionar todos los detalles posibles sobre la estructura de los



modelos de Markov (duración de cada ciclo, número de ciclos, estados de salud, etc.). En particular, se recomienda dar información sobre el tipo de probabilidades de transición, esto es, acerca de si estas probabilidades se consideran constantes o no durante el horizonte temporal del modelo. Las guías sostienen que para muchas enfermedades puede no resultar apropiado suponer probabilidades de transición constantes a lo largo de la vida de los pacientes.

Las guías explicitan la importancia de modificar el horizonte temporal del análisis, sobre todo en el caso de evaluaciones económicas basadas en modelos, sin que se mencionen más detalles al respecto.

Las guías de **Suecia** recomiendan utilizar un horizonte temporal que cubra los costes y beneficios más relevantes asociados a una intervención. En consecuencia, el horizonte temporal puede variar según el análisis, aunque en algunas circunstancias puede ser necesario calcular costes y beneficios a lo largo de toda la vida de los pacientes.

Las guías suecas exigen utilizar el horizonte temporal de la vida de los pacientes en el caso de intervenciones que afectan la supervivencia de los individuos, así como en el supuesto de enfermedades crónicas. Esto no significa necesariamente que la duración del tratamiento haya de coincidir con la vida de los pacientes.



La extrapolación de datos primarios de corto plazo (de ensayos clínicos en particular) a horizontes temporales más largos tiene que llevarse a cabo por medio de modelos. Por último, en el caso de enfermedades crónicas las guías aconsejan variar la duración del tratamiento de 1 a 5 años.

2.8. Tasa de descuento

El **NICE** recomienda en sus guías descontar costes y beneficios futuros a una tasa del 3.5% en el caso base. La tasa de descuento elegida se basa en las recomendaciones del *UK Treasury*. Se aconseja variar las tasas de descuento entre el 0% y el 6% para costes y beneficios cuando se estime que tendrán un fuerte impacto sobre los resultados finales.

Las guías de **Canadá** recomiendan descontar costes y beneficios futuros a una tasa del 5% en el caso base. La tasa de descuento elegida representa, según las guías, las preferencias temporales de la sociedad canadiense. En el análisis de sensibilidad se estima necesario considerar el caso de una tasa de descuento del 0% para costes y beneficios para estimar el impacto del ajuste temporal. También se recomienda adoptar una tasa del 3% por razones de comparabilidad con las evaluaciones económicas de otros países.



En **Australia** se recomienda descontar costes y beneficios futuros (horizonte temporal mayor de 1 año) a una tasa del 5% en el caso base. Proponen las guías variar la tasa de descuento entre el 0% y 5%, tanto para los costes como para los beneficios en el análisis de sensibilidad. También se recomienda recalcular los resultados utilizando una tasa del 5% para los costes y del 0% para los beneficios. No ofrecen las guías ninguna justificación explícita para la elección de tales tasas.

Las guías **suecas** recomiendan descontar costes y beneficios futuros a una tasa del 3% en el caso base, sin dar justificación para tal propuesta. Se recomienda variar la tasa de descuento entre el 0% y 5% para costes y beneficios. También se sugiere recalcular los resultados coste-efectividad utilizando una tasa del 3% para los costes y del 0% para los beneficios. Las guías suecas no ofrecen justificación alguna para las tasas propuestas.

2.9. Tratamiento de la incertidumbre

Las guías del **NICE** especifican diferentes fuentes de incertidumbre, distinguiendo entre la incertidumbre en los parámetros del análisis y otras formas de incertidumbre.



Estas otras formas incluirían la incertidumbre acerca de los supuestos y la estructura del modelo de decisión, la elección de los estudios en los que conviene incluir un meta-análisis, la extrapolación de datos de corto a largo plazo, etc. También hacen las guías una distinción implícita entre incertidumbre y variabilidad de los datos, subrayando la importancia del análisis de subgrupos para tener en cuenta la potencial heterogeneidad entre pacientes.

Las guías sugieren utilizar diferentes métodos para tratar la incertidumbre según el tipo de datos que se incluyan en las evaluaciones económicas. Así, en el supuesto de utilizar micro-datos de pacientes se recomienda el análisis estocástico por medio de *bootstrapping*, mientras que en los casos en que se hayan utilizado modelos se propone el análisis probabilístico con simulaciones de *Monte Carlo*.

Las guías recomiendan en todo caso el análisis de sensibilidad probabilístico o estocástico, dependiendo la modalidad concreta del tipo de datos disponible. La razón es que un análisis de sensibilidad determinístico o de escenario no puede incorporar la incertidumbre en más de dos o tres parámetros simultáneamente y, por tanto, no constituye una forma óptima de cuantificar la incertidumbre en dichos parámetros. Igualmente, se recomienda la realización de análisis de subgrupos, dada la variabilidad entre pacientes y los diferentes beneficios y costes que éstos pueden recibir de un nuevo tratamiento.



Según el NICE es conveniente realizar un análisis de sensibilidad probabilístico para todos los parámetros de un modelo de decisión o para todos los *inputs* de una evaluación económica, si ésta utiliza micro-datos de pacientes. También se recomienda llevar a cabo un análisis de sensibilidad para los supuestos y la estructura del modelo de decisión, los estudios incluidos en los meta-análisis, o la extrapolación de datos de corto a largo plazo. En el caso de análisis de subgrupos las guías exigen justificar la elección de dichos subgrupos sobre la base de diferencias identificadas entre pacientes, en lo que respecta a los costes o beneficios asociados a un tratamiento. Por último, se sugiere evaluar el nivel de incertidumbre alrededor de cada parámetro e input del análisis por medio del método del “valor esperado de la información”.

No existe en las guías un apartado explícito relativo a la elección de los rangos para el análisis de sensibilidad. Sin embargo, como la recomendación de realizar un análisis de sensibilidad probabilístico es clara, todos los parámetros se deberán modificar sobre la base de las distribuciones asignadas.

Sostienen las guías que “los métodos más apropiados para presentar los resultados del análisis de sensibilidad son los *scatter plots* y las *confidence ellipses* en el plano coste-efectividad, y las “curvas de aceptabilidad”. Todas estas representaciones gráficas derivan del análisis de sensibilidad estocástico o probabilístico, y permiten estimar la probabilidad de que un tratamiento o fármaco sea coste-efectivo dados



diferentes niveles de disposición a pagar por los resultados por parte del SNS del Reino Unido.

Las guías **canadienses** describen con detalle las diferentes formas de incertidumbre y variabilidad que pueden aparecer en una evaluación económica. En particular, se distingue entre incertidumbre en los parámetros de un modelo e incertidumbre del modelo en sí mismo (estructura, supuestos, métodos analíticos y fuentes de datos); además se establece una clara distinción entre la variabilidad debida a la heterogeneidad de los pacientes de un estudio y la variabilidad debida a las diferencias en las prácticas clínicas entre áreas geográficas.

Las guías se centran en las evaluaciones económicas basadas en modelos (las más frecuentes), aunque mencionan brevemente la posibilidad de que las evaluaciones se realicen con micro-datos de pacientes (recomendando, en este caso, un análisis estocástico).

Para tratar la incertidumbre en los parámetros del modelo de decisión se recomienda, como mínimo, un análisis de sensibilidad determinístico que incluya un análisis univariado, multivariado, y de escenarios extremos o de umbral. El análisis de sensibilidad probabilístico con simulaciones de *Monte Carlo* se considera el método más apropiado y es el tipo de análisis que recomiendan las guías para el caso de



incertidumbre alrededor de los parámetros. Sin embargo, se admite que pueda tratarse la incertidumbre del modelo en sí mismo recurriendo al análisis determinístico. El uso de análisis de subgrupos se sugiere para tratar la variabilidad entre los pacientes de un estudio. En este caso las guías proponen presentar los resultados de cada subgrupo con detalle, recomendando que la elección de los subgrupos tenga lugar en el curso del estudio de efectividad (por ejemplo, un ensayo clínico) y no posteriormente, durante el estudio coste-efectividad.

En el caso de que se realice un análisis de sensibilidad determinístico univariado, se considera imprescindible incluir todos los parámetros del modelo, los supuestos más importantes y los métodos analíticos clave (por ejemplo, los supuestos asumidos en relación con la duración del efecto de un tratamiento, el horizonte temporal, la tasa de descuento, o el método para calcular los costes indirectos), así como la elección de las fuentes para los inputs (por ejemplo, los estudios incluidos en un meta-análisis). Además, se exige la realización de un análisis multivariado sobre los parámetros que tengan un mayor impacto en los resultados del estudio. Las guías aconsejan tratar la incertidumbre que afecta a los métodos analíticos clave también por medio de un análisis de escenarios. Finalmente, en el caso de realizar un análisis probabilístico, se recomienda tener en cuenta todos los parámetros del modelo. En este caso, se estima que el recurso a “métodos Bayesianos” permite evaluar adecuadamente el nivel de incertidumbre alrededor de cada parámetro e input del análisis (“valor esperado de la información”).



Las guías canadienses señalan que, cuando se utiliza un análisis determinístico univariado, conviene modificar los parámetros según sus valores extremos. Esto resulta particularmente importante en el caso de que los datos deriven de la opinión de expertos. También se acepta que puedan variar según sus intervalos de confianza, cuando los datos provengan de ensayos clínicos o meta-análisis. El investigador, siempre debe justificar la elección del rango sobre el que se modifican los parámetros, lo que también es de aplicación en el caso de análisis probabilísticos, en los que el investigador ha de identificar y justificar el tipo de distribución asociada a cada parámetro, el rango utilizado para cada distribución y el número de simulaciones de *Monte Carlo* realizadas.

Los resultados de un análisis de sensibilidad determinístico se pueden presentar por medio de tablas, pero las guías canadienses prefieren las representaciones gráficas. En particular se recomiendan los *tornado diagrams* en el caso de análisis univariado, y los *scatter plots* en el plano coste-efectividad y las “curvas de aceptabilidad” en el caso de un análisis probabilístico.

Las guías de **Australia** identifican diferentes formas de incertidumbre que pueden aparecer en una evaluación económica. En particular, se distingue entre incertidumbre en los parámetros del modelo e incertidumbre en los supuestos del



modelo mismo. Además se menciona la incertidumbre debida a la variabilidad de los pacientes, en el caso de evaluaciones basadas en ensayos clínicos.

Las guías discriminan de modo explícito entre los métodos aptos para tratar la incertidumbre en el caso de evaluaciones económicas con micro-datos de pacientes (por ejemplo, análisis estocástico por medio de *bootstrapping*) y los apropiados para el caso de evaluaciones basadas en modelos (análisis determinístico o análisis probabilístico con simulaciones de *Monte Carlo*). En el caso de evaluaciones basadas en datos procedentes de ensayos clínicos, el investigador está obligado a justificar la elección de métodos paramétricos (por ejemplo, el “método de Fiellers”) o no paramétricos (*bootstrapping*).

Se admiten diversas modalidades de análisis de sensibilidad, entre ellas, el análisis determinístico univariado, el multivariado, el análisis de escenarios extremos y el análisis probabilístico. No se requiere explícitamente un tipo específico de análisis de sensibilidad, aunque las guías muestran su preferencia por los métodos probabilísticos, sobre todo cuando la evaluación económica se basa en modelos. En el caso de utilizar simulaciones de *Monte Carlo* se considera imprescindible justificar el número de simulaciones realizadas.

Se requiere llevar a cabo un análisis de sensibilidad univariado sobre todos los parámetros del modelo. También se recomienda el análisis univariado respecto de algunos supuestos claves del modelo. Se recomienda la realización posterior de un



análisis multivariado sobre aquellos parámetros que tienen un mayor impacto sobre los resultados finales del análisis coste-efectividad. Por último, en el caso de realizar un análisis probabilístico, se recomienda tener en cuenta todos los parámetros del modelo.

En el análisis univariado se recomienda modificar los parámetros utilizando unos valores extremos plausibles, correspondiendo al investigador la tarea de elegir dichos extremos y justificar la elección. A modo de ejemplo, si los datos proceden de ensayos clínicos se recomienda utilizar los valores extremos del intervalo de confianza encontrado en el estudio original. En el supuesto de que se haya utilizado una tasa de descuento para costes o beneficios, se recomienda someter a variación dicha tasa, incluyendo un valor del 0%. En los análisis probabilísticos, se exige justificar el tipo de distribución asociada a todos los parámetros así como los valores utilizados.

El análisis de sensibilidad univariado se puede presentar por medio de tablas y mediante *tornado diagrams*, para especificar el impacto de cada parámetro sobre los ratios coste-efectividad. Los resultados del análisis multivariado se pueden presentar también por medio de tablas o gráficos. En el caso del análisis de sensibilidad probabilístico se pueden utilizar las “curvas de aceptabilidad” o, alternativamente, puntos sobre el plano coste-efectividad.



Las guías **suecas** no proporcionan ningún detalle sobre los métodos para el tratamiento de la incertidumbre. Se limitan a subrayar la importancia de realizar un análisis de sensibilidad sobre los parámetros de la evaluación económica.

2.10. Uso de modelos decisionales

Las guías del **NICE** requieren el desarrollo de un modelo de decisión para cualquier evaluación económica y, en particular, para conseguir los siguientes objetivos: a) extender al largo plazo datos de corto plazo; b) sintetizar la información de diferentes fuentes; c) conseguir que el análisis sea relevante para la población del Reino Unido; d) transformar medidas de resultados intermedias en medidas de resultados finales. Adicionalmente, las guías subrayan la importancia de los modelos de decisión en el análisis de la incertidumbre y el de subgrupos. No se recomienda de manera explícita algún tipo de modelo para enfermedades o intervenciones específicas, si bien las guías del NICE exigen el uso de modelos probabilísticos.

Por lo que se refiere a la estructura de los modelos, las guías se limitan a señalar la necesidad de que se suministre toda la información y documentación utilizadas para construir el modelo de decisión. También se exige justificar todos los supuestos sobre la estructura del modelo. El NICE pone especial énfasis en la importancia de presentar de manera detallada todas las características de un modelo cuando este



sirve para extender al largo plazo datos de corto plazo y los investigadores se han visto obligados a realizar supuestos sobre el futuro.

Las guías establecen la obligación de proporcionar toda la información y documentación sobre las fuentes de datos del modelo. En particular, se estima fundamental ofrecer detalles acerca de los métodos de síntesis de los datos, recomendándose algunos métodos específicos (véase el apartado sobre la síntesis de la evidencia). De igual modo, es requisito para las guías que se describan de manera detallada todos los parámetros y las distribuciones asociadas a cada uno de ellos.

No se ha encontrado en las guías del NICE una referencia específica a la coherencia interna y externa de los modelos.

Las guías **canadienses** identifican dos situaciones en las que es imprescindible recurrir a un modelo de decisión: para sintetizar datos que proceden de fuentes diversas o se basan en supuestos; y para extender a un horizonte temporal más apropiado los datos de corto plazo.

Las guías no muestran preferencia alguna por algún tipo de modelo en particular en función del tipo de enfermedad o intervención. Cualquier modelo, afirman, debe ser relevante para la finalidad del estudio y para la población objetivo. Asimismo, se



establece que los resultados del modelo han de ser relevantes en términos de medida de beneficios, comparadores, etc. Las guías permiten la utilización de modelos determinísticos y probabilísticos, aunque se explicita que los segundos proporcionan más información sobre la incertidumbre de los resultados.

Las guías canadienses subrayan la importancia de la estructura del modelo por su impacto sobre los resultados finales. En consonancia con ello, exigen la descripción detallada de dicha estructura, así como la justificación de todos los supuestos sobre la misma y de las fuentes utilizadas en su elección. La estructura del modelo ha de reflejar de modo coherente el tipo de enfermedad analizada y las posibles consecuencias de las intervenciones comparadas, al tiempo que debe ser relevante para el objetivo del estudio. Otra recomendación de las guías es la de desarrollar modelos flexibles, que permitan la incorporación de datos futuros (por ejemplo, efectos secundarios), o nuevas tecnologías. Esta flexibilidad que se demanda ayuda, además, a transferir datos de un contexto a otro. En el caso de los modelos de Markov las guías canadienses subrayan la necesidad de informar sobre la duración de los ciclos y justificarla según criterios clínicos. Finalmente, se hace referencia a dos *checklists* para juzgar la calidad de la estructura del modelo (Weinstein et al 2003, Phillips et al 2004).

Las guías de Canadá aceptan cualquier fuente de datos para los parámetros del modelo, aunque se advierte de que la calidad del modelo dependerá de la de los datos que incorpore. Por esta razón, es preciso dar cuenta de todas las limitaciones



asociadas a las fuentes principales de datos. Se recomienda incluir la opinión de expertos sólo en el supuesto de que no existan otras fuentes disponibles. Las guías exigen una descripción detallada (población de estudio, tipo de estudio, horizonte temporal, etc.) de las fuentes de datos incluidas en el modelo. Se recomienda, asimismo, revisar la literatura de manera sistemática y sintetizar los datos por medio de meta-análisis. Aunque se admiten otros métodos de síntesis de datos, éstos han de ser justificados, toda vez que existe un riesgo de sesgo en el análisis cuando se excluyen algunos estudios sin justificación. También deben describirse con detalle los métodos empleados para extender datos primarios de corto a largo plazo. Advierten las guías de que la extrapolación lineal de los costes puede no ser apropiada, debido por ejemplo a la existencia de economías de escala. Por último, se requieren detalles sobre todos los valores de los parámetros y sus distribuciones en el caso de modelos probabilísticos. Las guías hacen referencia a las dos *checklists* antes citadas (Weinstein et al 2003, Phillips et al 2004) para juzgar la calidad de los datos del Modelo.

Las guías canadienses recomiendan, por último, comprobar la coherencia interna y externa del modelo, pues se considera éste un elemento importante para juzgar la calidad del estudio. Para comprobar la coherencia interna del modelo se sugiere utilizar valores extremos, poner todos los parámetros a cero y realizar análisis de escenarios. Las guías consideran más importante aún validar los resultados de los modelos con datos reales, de manera que cualquier diferencia relevante debe ser analizada y explicada. Además, se aconseja comparar los resultados del modelo



construido con los obtenidos de otros modelos publicados para la misma enfermedad e intervención, y justificar las posibles diferencias. Sugieren las guías que, cuando sea posible, la coherencia interna y externa del modelo sea comprobada por investigadores independientes.

Las guías **australianas** recomiendan el uso de un modelo de decisión con diversas finalidades, a saber: a) extender al largo plazo datos de corto plazo; b) incrementar la validez externa de un ensayo clínico; c) conseguir que el análisis sea relevante para la población de Australia; d) transformar medidas de beneficios intermedias en medidas de beneficio finales; e) sintetizar la información de diferentes fuentes; f) transferir datos de otros países al contexto australiano; g) estimar la incertidumbre sobre los datos utilizados y los resultados finales.

No se recomiendan modelos específicos según el tipo de enfermedad o intervención, pero se aconseja el uso de modelos de Markov en los estudios de largo plazo en los que haya recurrencia de eventos. En el caso particular de la evaluación de vacunas, se recomienda el uso de modelos dinámicos respecto a modelos estáticos. Las guías no muestran explícitamente sus preferencias por modelos determinísticos o probabilísticos, pero se aportan muchos detalles acerca del análisis probabilístico y sus ventajas en términos de tratamiento de la incertidumbre.



Se considera imprescindible facilitar toda la información disponible sobre la estructura del modelo, incluyendo una representación gráfica y, si es posible, una versión electrónica del modelo. Cualquier supuesto sobre la estructura del modelo ha de ser debidamente justificado y, en todo caso, dicha estructura tiene que reflejar de manera coherente el tipo de enfermedad analizada y las intervenciones comparadas. En el caso de los modelos de Markov se exige que consten, al menos, los siguientes detalles: estados de salud, duración de los ciclos (y su justificación), probabilidades de transición (y justificación del uso de probabilidades fijas o variables), horizonte temporal, e inclusión o exclusión de la “corrección de medio-ciclo” (*half-cycle correction*). También piden las guías que se justifiquen los métodos utilizados para desarrollar el análisis y los cálculos del modelo de Markov (por ejemplo, análisis determinístico de cohorte, simulaciones individuales de *Monte Carlo* de primer o segundo orden, etc.). En general, se prefieren modelos flexibles que permitan la introducción de nuevos datos o nuevas alternativas.

Las guías australianas demandan una presentación detallada de todas las fuentes de datos en forma de tablas, admitiéndose la posibilidad de enviar copia de las fuentes de datos en un documento separado. Las guías proporcionan muchos detalles sobre los métodos que pueden emplearse para seleccionar y sintetizar los estudios primarios (véase el apartado sobre la síntesis de la evidencia). Se exige, asimismo, que se describan los métodos para extender datos primarios de corto a largo plazo. Para cada uno de los parámetros se requiere proporcionar, al menos, el nombre, el valor medio y las fuentes principales. Además, se recomienda explicar y justificar



cualquier correlación (o ausencia de la misma) entre parámetros. Por último, las guías solicitan que se aporten detalles sobre el análisis de la incertidumbre, incluyendo los valores alternativos de cada parámetro y las distribuciones asumidas en el caso de un modelo probabilístico.

La conveniencia de comprobar la coherencia externa del modelo por medio de un proceso de calibración es una prescripción explícita en las guías australianas. Por otro lado, se recomienda implícitamente evaluar la coherencia interna del modelo, ya que se establece como requisito el envío de la versión electrónica del modelo de decisión que permita comprobar si existen errores de códigos matemáticos.

Las guías **suecas** indican algunas circunstancias en las que se aconseja utilizar un modelo de decisión a) para aumentar la validez externa de los ensayos clínicos; b) para extender a un largo plazo datos de corto plazo; c) para sintetizar la información de muchas fuentes; d) para transferir datos de otros países al contexto sueco. No se proporciona, sin embargo, ninguna información sobre los tipos de modelos recomendados, como tampoco se requieren explícitamente detalles sobre la estructura del modelo, ni sobre las fuentes de datos y los parámetros del modelo.

Las guías recomiendan la validación interna y externa del modelo, pero no ofrecen información acerca de los métodos para llevar a cabo dicha validación.



2.11. Transferibilidad de datos y resultados

Las guías del **NICE** consideran de manera explícita diferentes conceptos de transferibilidad y distintos problemas relacionados con la transferibilidad de los datos y resultados de un estudio de evaluación económica. En particular, se subraya la importancia de que cualquier estudio remitido al NICE para su consideración tenga relevancia para el contexto de Inglaterra y Gales. También se tiene en cuenta la variabilidad entre diferentes centros de un mismo país y la posible heterogeneidad de los pacientes enrolados en estos centros. Por último, se menciona la cuestión de la validez externa de los estudios basados en ensayos clínicos.

Las guías analizan cada elemento de las evaluaciones económicas por separado y ofrecen recomendaciones concretas al respecto de su transferibilidad, en los siguientes términos:

- a. *Efectividad*: las guías distinguen entre el efecto ‘absoluto’ de un tratamiento y el efecto ‘relativo’ (por ejemplo, comparado con placebo). Se advierte de que la eficacia absoluta de un tratamiento puede variar entre países o entre centros, debido a las diferencias potenciales entre pacientes, prácticas clínicas, etc. Por esta razón las guías recomiendan adaptar el efecto absoluto de un tratamiento al contexto de Inglaterra y Gales. La eficacia



relativa de los tratamientos se considera en cambio transferible de un contexto a otro.

- b. *Recursos utilizados y costes unitarios*: unos y otros tienen que ser relevantes para el SNS de Inglaterra y Gales. No se consideran transferibles los datos económicos de un país a otro, y se previene sobre la necesidad de utilizar precios unitarios y recursos consumidos en el Reino Unido. En el caso de que se utilicen datos de recursos consumidos de otros países, el investigador tendrá que justificar su elección y tener en cuenta la variabilidad en estos datos por medio de un análisis de sensibilidad.
- c. *Utilidades*: se recomienda utilizar instrumentos para los cuales se dispone de preferencias de la población del Reino Unido. La elección recomendada es el Euroqol-5D con datos de la población inglesa.

Las guías parecen concentrarse más en el caso de evaluaciones económicas basadas en modelos, por lo que no ofrecen detalles sobre las evaluaciones basadas en microdatos de pacientes. No existe en las guías del NICE un apartado específico sobre la transferibilidad de las evaluaciones económicas basadas en datos de pacientes individuales. No obstante, en el análisis de la incertidumbre se recomienda utilizar un análisis de sensibilidad estocástico por medio del *bootstrapping*, con el fin de tener en cuenta la variabilidad entre subgrupos de pacientes en diferentes centros y contextos (véase el apartado sobre el tratamiento de la incertidumbre).



Por lo que respecta a los problemas de transferibilidad en análisis basados en modelos, las guías ponen gran énfasis en los métodos de selección de las fuentes de datos de un árbol de decisión. En particular, se recomienda tener en cuenta, de manera explícita, la heterogeneidad entre estudios en lo que se refiere a pacientes, calidad, práctica clínica en diferentes países, etc. Se recomienda sintetizar los estudios por medio de meta-análisis, cuando ello sea posible. En tal caso, se considera importante utilizar un modelo de efectos aleatorios para tener en cuenta la heterogeneidad entre estudios. También se sugiere variar los estudios incluidos en el meta-análisis, en el análisis de sensibilidad, para estimar el impacto sobre los resultados. En general, las guías sostienen que el análisis de sensibilidad probabilístico garantiza una correcta estimación de la incertidumbre y la variabilidad de los datos utilizados en el modelo.

Las guías del NICE mencionan de manera sucinta los problemas de validez externa de algunos ensayos clínicos y la dificultad de trasladar directamente los resultados de estos estudios al escenario de la práctica clínica habitual. Se sugiere sostener los resultados de ensayos clínicos con datos de otros estudios, preferiblemente con un horizonte temporal de largo plazo.



Las guías **canadienses** identifican claramente los diferentes aspectos relacionados con la transferibilidad. Hay apartados específicos en las guías para los problemas de validez externa de los ensayos clínicos y para la transferibilidad de resultados de otros países al contexto canadiense, otorgándose también gran importancia a la posible variabilidad entre regiones del país.

Se analizan todos los elementos del análisis farmacoeconómico para juzgar su transferibilidad de manera separada:

- a. *Efectividad*: afirman las guías que la transferibilidad de datos de efectividad depende de las circunstancias concretas, y se necesita un test de heterogeneidad para evaluar si es apropiado utilizar datos de otros países para el contexto canadiense.
- b. *Recursos utilizados y costes unitarios*: ambos tienen que ser relevantes para el contexto canadiense. No se consideran transferibles los datos económicos de un país a otro, y se recomienda utilizar precios unitarios y recursos consumidos en Canadá. En el caso de que se utilicen datos de recursos consumidos de otros países, el investigador tiene que ajustar estos datos a la realidad del país. Se subraya, asimismo, la importancia de presentar los datos económicos de manera muy transparente (por ejemplo, separando los costes unitarios de los recursos utilizados).



- c. *Utilidades*: las guías advierten de que los valores de utilidad obtenidos por poblaciones de países diferentes pueden no ser en general transferibles a Canadá, debido a las diferencias en aspectos culturales entre países.

Se realiza una distinción explícita entre evaluaciones económicas basadas en microdatos de pacientes y basadas en modelos, proporcionándose detalles acerca de cómo tratar el asunto de la transferibilidad en cada una de las dos situaciones

Hay un apartado específico en las guías para el caso de evaluaciones económicas basadas en ensayos clínicos realizados en diferentes países y centros. En general, se acepta el uso de los datos de eficacia globales de los ensayos clínicos (una media de los datos de todos los países y centros), aunque se estima necesario realizar un test de homogeneidad. Por otro lado, no se considera admisible utilizar la media de recursos utilizados en los centros del ensayo multinacional, porque éstos serán sistemáticamente diferentes entre países. Aunque se reconozca el potencial problema de la escasa significatividad estadística en el caso de que la muestra de pacientes canadienses en el estudio sea demasiado pequeña, se aconseja utilizar datos económicos específicos para Canadá. Se sugieren algunos métodos para tratar la transferibilidad de datos en estudios multinacionales, como los “modelos multi-nivel”, las “meta-regresiones” o el “método empírico Bayesiano”.

Las guías identifican los modelos como un instrumento clave que permite tratar los problemas de transferibilidad de datos por medio del análisis de sensibilidad (en



particular, el análisis probabilístico). También se recomienda seleccionar las fuentes de los parámetros por medio de una revisión sistemática de la literatura y justificar la inclusión de cada estudio concreto.

Las guías advierten acerca de los problemas de validez externa de los ensayos clínicos subrayando la diferencia entre eficacia y efectividad. Los encargados de tomar las decisiones están interesados a la efectividad de un tratamiento en la práctica clínica “real”, y los datos utilizados en la evaluación económica deberían centrar su atención en la efectividad de los tratamientos. Por eso se exige presentar los datos en manera transparente para permitir que los decisores evalúen su relevancia en el contexto clínico habitual canadiense.

Aunque no exista un apartado específico dedicado a este tema, las guías de **Australia** consideran implícitamente las diferentes formas de transferibilidad, enfatizando la importancia del uso de datos aplicables al contexto australiano, y los problemas relacionados con la validez externa de los ensayos clínicos. Se hacen las siguientes consideraciones sobre la transferibilidad de los diferentes elementos de la evaluación económica:

- a. *Efectividad*: las guías señalan que conviene tener en cuenta las posibles diferencias entre los pacientes de estudios realizados en otros países y la



población de interés del contexto australiano. Por ejemplo, pueden existir diferencias epidemiológicas entre pacientes de diferentes países y la efectividad de un tratamiento obtenida en un contexto diferente puede no ser aplicable a Australia.

- b. *Recursos utilizados y costes unitarios*: tienen que ser específicos para Australia, pues los datos económicos no son transferibles entre países debido a las diferencias existentes en la práctica clínica, los sistemas de financiación, los incentivos, etc.
- c. *Utilidades*: las guías australianas no hacen mención a la posible transferibilidad de los valores de utilidades.

No se contiene una distinción explícita en las guías entre evaluaciones económicas basadas en micro-datos de pacientes y basadas en modelos, pero se señala la importancia de los modelos para aumentar la transferibilidad de datos obtenidos en contextos diferentes a la realidad australiana.

Las guías proporcionan bastantes detalles acerca del modo de tratar la posible heterogeneidad de los datos de pacientes. Por ejemplo, se sugiere el uso de meta-regresiones para separar las características de los pacientes de un estudio, y se destaca la importancia del análisis de sub-grupos. No se menciona expresamente el caso de los estudios multinacionales, pero en diferentes partes de las guías se alude a



la posibilidad de utilizar datos clínicos de otros países y a la imposibilidad de transferir los datos económicos.

Al igual que las guías canadienses, las de Australia señalan los modelos como un instrumento útil para tratar los problemas de transferibilidad de datos. El análisis de sensibilidad probabilístico es el método preferido por las guías para tener en cuenta de la variabilidad en los parámetros del modelo. Se ofrecen igualmente muchos detalles al respecto de la forma de elegir las fuentes de datos para el modelo y sobre cómo sintetizar los estudios incluidos (meta-análisis). Asimismo, se proporciona una escala jerárquica de la calidad de los estudios clínicos candidatos a formar parte de las fuentes para el análisis económico.

Por último, se recomienda el uso de estudios no aleatorizados y de largo plazo para aumentar la validez externa de los ensayos clínicos. En particular, se advierte sobre la necesidad de excluir los costes debidos al protocolo para reflejar la realidad clínica. También se sugieren algunos métodos para reducir el sesgo potencial de los estudios no aleatorizados.

Las guías **suecas** se limitan a señalar la importancia de utilizar datos clínicos y económicos relevantes para el contexto sueco. Adicionalmente, subrayan la importancia de los modelos como herramienta para aumentar el grado de validez



externa de un ensayo clínico, así como su utilidad para adaptar datos obtenidos en otros países a la realidad sueca en términos de costes, población de estudio y práctica clínica habitual.

2.12. Presentación de los resultados

Las guías del **NICE** recomiendan proporcionar información acerca de algunos elementos fundamentales en los resultados del análisis, como mínimo: costes y efectos totales por separado, costes desagregados por categorías, ratio coste-efectividad incremental, análisis de sensibilidad y de subgrupos. Se aconseja utilizar tablas y gráficos para facilitar la lectura e interpretación de los resultados.

Los costes y beneficios totales para cada tecnología analizada han de presentarse separadamente. Además, es necesario identificar por separado los costes asociados a diferentes agentes sanitarios (Sistema Nacional de Salud, hospital, atención primaria) y las diferentes categorías de recursos (fármacos, estancias hospitalarias, visitas médicas, etc.). Por último las guías del NICE consideran un elemento clave la realización de análisis de subgrupos y su presentación detallada, separando también en este caso los costes y los efectos totales.

Es obligatorio para el NICE realizar un análisis incremental de costes y beneficios y presentar ratios incrementales coste-efectividad. Conviene identificar los casos en que una o más estrategias están dominadas por otras o por combinaciones de éstas.



Los resultados del análisis probabilístico se han de presentar en un plano coste-efectividad.

Las guías del NICE conceden gran importancia al tratamiento de la incertidumbre, lo que ha de reflejarse necesariamente en la presentación de los resultados del análisis. Se exige que se informe sobre la media y la desviación estándar de los costes y efectos totales así como de las ratios coste-efectividad. Estas ratios también se presentarán gráficamente como puntos en el plano coste-efectividad y como curvas de aceptabilidad. Por último, se recomienda proporcionar detalles acerca de las características de cada grupo de pacientes y suministrar datos sobre los costes y efectos por separado en el análisis de subgrupos.

Se recomienda identificar las limitaciones del análisis, sobre todo en relación con la incertidumbre, y apuntar sugerencias sobre la necesidad de estudios futuros. El NICE requiere que la medida de efectividad utilizada en el informe sean los AVAC y se hacen explícitos algunos criterios de decisión. En particular, las guías afirman que (6.2.6.10): a) “cuando la ratio es inferior a £20,000 por AVAC, la aceptabilidad de una tecnología ha de basarse sobre todo en criterios coste-efectividad”; b) “cuando la ratio se encuentra entre £20,000 y £30,000 por AVAC hay que considerar otros elementos además del coste-efectividad, como el nivel de incertidumbre sobre los resultados, el tipo de enfermedad, el impacto presupuestario y si una tecnología es innovadora”; c) “cuando la ratio es superior a £30,000 por AVAC conviene justificar



la aceptabilidad de la tecnología por medio de otros elementos distintos de su coste-efectividad”.

Las guías en **Canadá** disponen de un formato estándar para la presentación de los resultados de la evaluación económica. Éste incluye todos los aspectos relevantes para una presentación clara y transparente del análisis, como la realización de un análisis incremental de costes y efectos, la presentación de costes y efectos agregados y desagregados, la presentación de los resultados de cada perspectiva utilizada por separado, la descripción detallada de los análisis de sensibilidad y de subgrupos, o la utilización de tablas y gráficos para representar de forma más fácilmente interpretable los resultados coste-efectividad.

Las guías recomiendan presentar, en una primera fase, los costes y beneficios desagregados y descontados. El investigador tendrá que explicar cualquier procedimiento de agregación de costes y beneficios que, además, se sugiere incluir sólo al final del análisis. Cuando la evaluación económica utiliza diferentes perspectivas, los resultados de cada perspectiva han de presentarse por separado, así como cuando se utilizan varios horizontes temporales. Respecto a los efectos de la tecnología, se recomienda ofrecer los resultados en unidad naturales, y explicar el proceso de agregación utilizado para obtener los AVAC u otras medidas.



Se exige la realización de un análisis incremental de costes y beneficios y el cálculo de una ratio incremental coste-efectividad, pues, se afirma, la ratio coste-efectividad media da una información incorrecta. Las guías recomiendan identificar los casos de dominancia y dominancia extendida. Se sugiere presentar los resultados coste-efectividad en una tabla en la que las tecnologías ocupen las filas en orden creciente de efectividad. Los resultados del análisis probabilístico se deben presentar como “puntos” en un plano coste-efectividad.

Las guías canadienses se muestran claramente partidarias de las evaluaciones económicas basadas en análisis probabilísticos. De ahí que, a la hora de presentar los resultados de los análisis de sensibilidad, se recomiende la utilización de “curvas de aceptabilidad” y “puntos” en el plano coste-efectividad (resultados de las simulaciones de *Monte Carlo*). Los cambios en las ratios incrementales coste-efectividad resultado de la modificación de los parámetros se puede representar por medio de *tornado diagrams*. Por último, las guías consideran apropiado aportar los resultados de los análisis de subgrupos en una tabla, subrayando aquellas características de los pacientes que tienen un mayor impacto sobre las ratios coste-efectividad.

El informe de evaluación debe incluir una discusión en la que se destaquen las limitaciones del análisis (en métodos, supuestos, datos, etc.), una comparación con otros estudios que evalúen la tecnología analizada (métodos y resultados) e indicaciones acerca de cómo desarrollar trabajos futuros. No explicitan las guías



canadienses criterio de decisión alguno, si bien se espera que el investigador justifique adecuadamente sus conclusiones relativas a la aceptabilidad de la tecnología, basándose en los resultados coste-efectividad.

Las guías **australianas** definen con detalle los aspectos más importantes para una presentación correcta de los resultados del análisis, pero no disponen de un formato estándar o una *checklist*. Entre los elementos relevantes se destacan el análisis incremental de costes y efectos, la presentación de costes y efectos desagregados, la importancia de la variabilidad e incertidumbre y la necesidad de proporcionar detalles sobre cada una de las fases de la evaluación económica que conducen a la obtención de los resultados finales.

Se recomienda presentar los resultados de costes y efectos desagregados en una primera fase del análisis y proporcionar los valores agregados sólo al final. Si es posible, se sugiere ofrecer una descripción detallada de los resultados para cada categoría de costes y separar recursos utilizados y costes unitarios. También se recomienda presentar los resultados de efectividad en unidad naturales, y explicar el proceso de agregación utilizado para obtener los AVAC u otras medidas. Además de los costes totales se considera de interés que se informe sobre el coste por paciente/año. En el caso de utilizar un modelo de Markov, se recomienda presentar costes y efectos para cada estado de salud y separados por ciclos. Por último, todos



los resultados del análisis de subgrupos se habrán de presentar por separado, enfatizando las características más relevantes para el coste-efectividad de la tecnología. Las guías proporcionan bastantes detalles sobre la forma de presentar costes y efectos desagregados y agregados (por medio de tablas y gráficos).

Se exige realizar un análisis incremental de costes y beneficios y calcular las ratios incrementales entre dos o más alternativas. Es interesante subrayar que las guías australianas aconsejan no utilizar el término “dominada” para identificar una tecnología, a menos que la condición de dominancia se confirme en los análisis de sensibilidad. Se sugiere presentar los resultados del análisis probabilístico como “puntos” en un plano coste-efectividad.

Los resultados de la evaluación económica han de presentarse, según las guías, no sólo como estimaciones puntuales sino informando sobre la desviación estándar de costes, efectos y ratios incrementales coste-efectividad. El análisis de sensibilidad probabilístico se puede presentar por medio de “puntos” en el plano coste-efectividad o de “curvas de aceptabilidad”. También consideran importante las guías australianas que los resultados de análisis univariados, se presenten incluyendo valores descontados de costes y efectos.

El investigador tendrá que justificar cualquier conclusión sobre el coste-efectividad de la tecnología analizada. En particular, las guías australianas sugieren algunos criterios básicos. Por ejemplo, si las medidas de efectividad utilizadas en el análisis



no incluyen los AVAC o los años de vida ganados, se recomienda comparar los resultados obtenidos con resultados de otros estudios publicados sobre la misma tecnología u otras similares. Es probable que, a la luz de esta comparación, una nueva tecnología para la que se dispone de alternativas ya analizadas y caracterizadas por ratios coste-efectividad más favorables, no sea aceptada. El análisis de subgrupos también puede facilitar las decisiones sobre la aceptabilidad o no de una tecnología, que podría ser más rentable en un particular grupo de pacientes. Otros elementos como el impacto presupuestario y la disponibilidad de otras tecnologías deben ser, igualmente, tenidos en cuenta. Aunque las guías no identifican un umbral coste-efectividad en el caso de evaluaciones basadas en AVAC o años de vida ganados, sugieren comparar los resultados con otros estudios, por ejemplo, recurriendo a *league tables* (tablas en que se presentan las ratios incrementales coste-efectividad de varias tecnologías y enfermedades en orden creciente).

En este aspecto, como en casi todos los analizados, las guías **suecas** son las más parcas en detalles. Se limitan a señalar la necesidad de realizar un análisis coste-efectividad incremental, presentando por separado costes y efectos incrementales. No dan información sobre el resto de cuestiones específicas.



3. Comentarios sobre las guías y conclusiones

3.1. Perspectiva de la evaluación económica

Aun no existiendo unanimidad entre las guías en relación con la perspectiva, todas coinciden en la posibilidad (en algunos casos en forma de análisis secundario) de incluir costes y beneficios que asuman una perspectiva social. Existen, sin embargo, contradicciones entre las guías que proponen como primera opción adoptar la perspectiva del SNS. Por ejemplo, las guías del NICE admiten que se incorpore la salud del cuidador informal, mientras que las canadienses la excluyen explícitamente. Esto es consecuencia, en nuestra opinión, de la ausencia de un fundamento teórico que sustente la adopción de la perspectiva del financiador.

Desde el punto de vista normativo, parece difícil justificar otra perspectiva que no sea la social cuando el financiador es el sector público, pues éste, al menos en teoría, debe maximizar el bienestar de la población (Johannesson, 1995). Cuesta trabajo admitir que el sector público pueda ignorar las consecuencias que las enfermedades y los tratamientos pudieran tener, por ejemplo, sobre la productividad del sistema económico de un país. Por otro lado, la denominada “perspectiva del financiador” no está exenta de contradicciones. Así, por ejemplo, ninguna de las guías que recomienda esta perspectiva sugiere que el precio de los medicamentos que se ha de utilizar en la evaluación económica, se limite al porcentaje pagado por el sector



público. Como mucho, esta posibilidad se contempla para el análisis del impacto presupuestario. Sin embargo, para ser consecuente con los argumentos que apoyan esta perspectiva restringida, la evaluación económica debería considerar como coste del fármaco únicamente la parte del precio que paga el SNS, lo que haría que el coste-efectividad dependiera del porcentaje de co-pago. No conocemos ninguna guía que contemple esta posibilidad. De igual modo, si nos ciñéramos al enfoque del financiador, sería discutible computar entre los costes los correspondientes a años futuros, habida cuenta de que la vigencia temporal de los presupuestos públicos está limitada a un año.

También por el lado de los beneficios resulta difícil admitir que un sistema público no considere beneficio una mejora en la comodidad del paciente a la hora de recibir un tratamiento (la denominada “utilidad del proceso”), sin entrar a discutir si la magnitud de ese beneficio es grande o pequeña comparada con la de otros beneficios que implican mejoras en la salud.

En suma, desde el punto de vista normativo consideramos que la perspectiva social es la correcta, siendo la perspectiva del financiador, a nuestro juicio, una opción básicamente estratégica. Una solución de compromiso puede ser presentar los resultados utilizando ambas perspectivas (Brouwer et al., 2006).



3.2. Revisión sistemática de la literatura

Existen notables diferencias en el detalle que alcanzan las guías en lo que se refiere a la revisión sistemática de la literatura. Mientras las guías australianas dedican varias secciones a la búsqueda de la literatura, la calidad de los estudios incluidos y los métodos de síntesis, las guías del NICE y las canadienses contemplan estos asuntos con un detalle significativamente menor. Sin embargo, las tres guías concuerdan en los aspectos más importantes. Todo lo contrario cabe decir de las guías suecas, que por no dar ninguna información detallada sobre el tema, no serán incluidas en los comentarios que siguen. En general, parece que las recomendaciones de las tres guías citadas son conformes con las recomendaciones teóricas acerca de la revisión de la literatura.

Así, las tres guías recomiendan realizar una revisión sistemática de la literatura para evitar el riesgo de un sesgo a favor o en contra de las tecnologías comparadas. La obtención de toda la evidencia clínica disponible se revela como un factor imprescindible para aumentar la credibilidad del análisis económico. En particular, mientras las guías canadienses y del NICE proponen identificar en una primera fase toda la evidencia clínica (sin restricciones según la calidad del estudio) para, en una segunda fase, excluir los estudios no relevantes, las guías australianas recomiendan incluir sólo los ensayos clínicos con comparaciones directas entre las tecnologías estudiadas. Se admite la inclusión de comparaciones indirectas y estudios observacionales únicamente cuando no existen datos de ensayos clínicos con



comparaciones directas disponibles. Todas las guías concuerdan con la necesidad de justificar las circunstancias en las que se omite el requisito de realizar una revisión sistemática de la literatura clínica.

Las guías subrayan de manera unánime la importancia de describir de manera transparente los métodos de búsqueda y los criterios de selección de los estudios. Las guías del NICE y las canadienses no proporcionan recomendaciones explícitas sobre la forma más adecuada de presentar estos datos, en tanto que las de Australia ofrecen numerosos detalles y restricciones al respecto. En todos los casos se requieren algunos elementos específicos de la búsqueda de la literatura y se proponen algunas orientaciones para la selección de los estudios clínicos. En particular, se presenta una escala jerárquica de la evidencia clínica, con prioridad para los ensayos clínicos con comparaciones directas, seguidos por las comparaciones indirectas y, finalmente, los estudios observacionales.

Las guías del NICE y australianas incluyen una lista detallada de las características de los estudios que es necesario hacer explícitas, como, por ejemplo, el diseño del estudio, las alternativas comparadas, el número de pacientes, sus características demográficas y el nivel de gravedad de su enfermedad para cada grupo, el periodo de seguimiento, el país o los países en que se ha desarrollado el estudio y el año en que se ha realizado (o publicado), los resultados de eficacia, etc. Las guías australianas añaden, además, una serie de elementos importantes según el tipo de estudio seleccionado (por ejemplo, el método de randomización en el caso de ensayos



clínicos aleatorizados). Por su parte, las guías canadienses se limitan a señalar la necesidad de proporcionar información sobre las características de los estudios utilizados para obtener datos de eficacia, efectividad y eventos secundarios sin indicar explícitamente qué tipo de información. Todas las guías exigen incluir también algunos comentarios sobre la calidad de los estudios seleccionados. De nuevo, mientras las guías canadienses y las del NICE dejan bastante libertad al investigador para que éste exprese su opinión acerca de las limitaciones de los estudios elegidos, las guías australianas hacen muchas precisiones útiles sobre los aspectos que se consideran fundamentales para evaluar la calidad de los estudios.

Las guías coinciden también en este punto y recomiendan el uso de un meta-análisis, en el caso de disponer de un número suficiente de estudios. El meta-análisis aumenta la credibilidad de la evaluación económica y podría poner de manifiesto diferencias significativas en términos estadísticos entre las alternativas comparadas que los ensayos individuales no son capaces de revelar. En el caso de que no se realice un meta-análisis, el evaluador tiene que especificar las razones de su elección y explicar en detalle cualquier método alternativo que se haya empleado. En el caso de comparaciones indirectas las guías sugieren distinguir de manera adecuada entre riesgo relativo de las tecnologías respecto a un comparador común y riesgo absoluto de dichas tecnologías, destacando la necesidad de considerar el riesgo relativo en estas circunstancias. Las guías australianas dedican un apartado específico a las comparaciones indirectas, destacando sus potenciales limitaciones y sugiriendo algunos métodos estadísticos para solucionar los problemas que plantean.



El análisis de la heterogeneidad entre los estudios seleccionados es objeto de particular atención por parte de las tres guías, tratándose el asunto transversalmente en varias secciones. Se subraya la necesidad de utilizar un meta-análisis de efectos aleatorios cuando hay heterogeneidad en los resultados estadísticos de los estudios analizados. Existe consenso también en señalar los análisis de subgrupos y de sensibilidad como los métodos más apropiados para tener en cuenta la variabilidad de los pacientes de los estudios y las posibles diferencias entre contextos y países. Las guías del NICE también sugieren realizar análisis de sensibilidad excluyendo algunos estudios utilizados en el caso base. Las guías australianas y canadienses proponen algunos métodos estadísticos de reciente desarrollo (Ades, 2006; Nixon, 2006) con el fin de adaptar las diferencias en las características de los pacientes, en el caso de comparaciones indirectas. En general, los métodos recomendados por las guías resultan conformes con los postulados teóricos de la evaluación económica (Drummond, 2005).

3.3. Población objetivo

Mientras las guías suecas no dan apenas información respecto de la definición de la población objetivo, las otras tres recomiendan ofrecer detalles sobre las características demográficas de la población objetivo, así como acerca de las características de la enfermedad (gravedad, posibles enfermedades concomitantes, etc.). Las guías canadienses y australianas también sugieren aportar información



adicional que puede resultar útil para establecer los costes y la efectividad de un tratamiento en la práctica real, como, por ejemplo, el centro donde se recibe el tratamiento, la tasa de cumplimiento esperada, los tratamientos recibidos con anterioridad, etc. En general, todas las guías subrayan explícita o implícitamente la necesidad de definir de manera adecuada la población objetivo y consideran que la elección de la población objetivo es fundamental para determinar el coste-efectividad de un tratamiento.

En términos generales, las guías recomiendan análisis de subgrupos cuando se identifican diferencias entre pacientes en términos de efectividad y/o costes de un tratamiento. En particular, las guías del NICE y las canadienses consideran los análisis de subgrupos como un factor fundamental de la evaluación económica y recomiendan realizarlas en todas las circunstancias o, al menos, al tener en cuenta la variabilidad de los resultados de los pacientes en el análisis de sensibilidad. En relación con el análisis de subgrupos, se considera imprescindible justificar la elección de los subgrupos de pacientes y explicar los métodos utilizados en el análisis. Las guías australianas ponen más énfasis en la posible heterogeneidad de los datos de efectividad obtenidos en los ensayos clínico y, en consecuencia, recomiendan llevar a cabo tests estadísticos para probar la asociación entre el efecto del tratamiento y las características de los pacientes. También el uso de modelos puede ser un instrumento útil para evaluar los resultados del coste-efectividad de un tratamiento, al variar las características de los pacientes. Las recomendaciones de las



guías coinciden con la teoría reciente que subraya la importancia de la variabilidad entre pacientes en la evaluación económica de una nueva tecnología (Coyle, 2003).

En todas las guías se indican las posibles diferencias entre la población de un ensayo clínico y los pacientes que se pueden beneficiar de un tratamiento en la práctica clínica real y se proponen métodos para tener en cuenta dicha diferencia. Por ejemplo, las guías del NICE sugieren realizar un análisis de subgrupos incluyendo sólo los datos de los pacientes que completan el estudio para reflejar más fielmente lo que podría ocurrir en la práctica real, mientras las guías suecas sugieren utilizar modelos. El tamaño de la población objetivo también es un elemento importante del análisis, y todas las guías recomiendan estimar, al menos, el número de personas que sufren de la enfermedad en el país y el número de pacientes que pueden beneficiarse del tratamiento. Estos datos sirven para estimar el impacto presupuestario que puede tener la nueva tecnología ya que resulta importante a la hora de tomar la decisión de reembolsar o no el tratamiento.

3.4. Término de comparación de la tecnología evaluada

Las guías presentan bastantes aspectos coincidentes, aunque con un nivel de detalle dispar. En particular, las guías canadienses ofrecen mucha más información en relación con los elementos claves que se han de tomar en consideración al elegir un comparador, en tanto que las guías suecas y las del NICE apenas proporcionan



informaciones básicas. Las guías australianas, por su parte, se concentran en el caso de los fármacos.

Todas las guías recomiendan comparar el nuevo tratamiento, al menos, con la práctica habitual, esto es, con el tratamiento más comúnmente utilizado en su respectivo contexto geográfico. Las guías del NICE y las canadienses sugieren la oportunidad de considerar, en la fase inicial del estudio, todas las alternativas aprobadas para la indicación de la nueva tecnología, para, en una fase posterior, seleccionar las alternativas más relevantes. Junto con la práctica habitual, estas alternativas pueden incluir también la estrategia más efectiva y la estrategia más barata. Esta es la razón por la cual las guías de Canadá y las del NICE parecen contemplar la posibilidad de que la práctica habitual no sea siempre la más apropiada, subrayando que, en algunas circunstancias, puede resultar necesario recurrir a otros comparadores.

Las guías canadienses y las del NICE subrayan las dificultades en la elección de los comparadores cuando hay muchas alternativas posibles, mientras que las guías australianas sostienen, por el contrario, que es habitual el consenso en dicha elección (las guías suecas no dan ningún detalle sobre esta cuestión). Tal diferencia se debe probablemente a los diferentes criterios de selección inicial, ya que las guías canadienses y británicas requieren la inclusión de todos los comparadores potenciales en una primera fase. Los elementos clave para seleccionar los comparadores son, en todas las guías, la variabilidad entre pacientes y entre prácticas clínicas en diferentes



centros del país, y la disponibilidad de datos. La opinión de los expertos se sugiere como el método más apropiado para seleccionar los comparadores.

La práctica vigente se define como la práctica más comúnmente utilizada, esto es, la estrategia utilizada para el mayor número de pacientes en cada país. Todas las guías (excluyendo las suecas, que no precisan nada al respecto) recomiendan comparar la nueva tecnología con todas las prácticas habituales cuando no sea posible identificar una práctica dominante o cuando puedan existir diferencias entre centros y/o zonas del país. Las guías australianas definen la práctica vigente como el fármaco que el nuevo medicamento va a sustituir con más probabilidad.

Las guías coinciden en la necesidad de incluir el “no hacer nada” o “intervención mínima” como comparador en algunas circunstancias. Sin embargo, no se ofrecen en general muchos detalles sobre cuáles son estas circunstancias, y no quedan claras las razones para la inclusión de estos términos de comparación. Existen, asimismo, diferencias en la definición de “intervención mínima”. Sólo las guías australianas, indican claramente una situación donde es necesario incluir “placebo” como comparador. Por último las guías suecas aconsejan la inclusión de la opción “no hacer nada” únicamente en el supuesto de que ésta sea la práctica vigente.



3.5. Medida de resultados preferida

Las guías se decantan unánimemente por las variables de resultados finales. El problema es, tal y como señalan las guías australianas, que muchos ensayos clínicos no ofrecen resultados así medidos, por lo que las guías han de adoptar una decisión clave acerca de cómo afrontar este problema. Dos son las posibilidades: dar como válidas las variables intermedias que se usan en muchos de los ensayos clínicos (que parece ser la opción de las guías australianas), o pedir la traducción de variables intermedias a variables finales mediante el uso de modelos.

La razón para elegir una alternativa u otra radica en el papel que se considere que ha de desempeñar la evaluación económica en la toma de decisiones sobre asignación de recursos sanitarios. La evaluación económica puede considerarse como una herramienta que se limita a generar información sistematizada y ponerla a disposición de los que toman las decisiones, siendo éstos quienes han de tener la capacidad para procesar toda esa información y decidir en consecuencia. Esto es, si aceptamos que el resultado de una evaluación económica puede ser que “el medicamento A reduce más la hipertensión que el medicamento B pero cuesta dos veces más” estamos admitiendo que quienes toman las decisiones sabrán que hacer con dicha información. En ese caso, quizás la forma más aceptable de evaluación económica (si es que puede llamarse así) sería el denominado “análisis coste-consecuencia”. Se proporcionaría a los decisores información acerca de las consecuencias sobre la salud y el consumo de recursos asociado a diversas



alternativas para que ellos tomaran una decisión informada. Sin embargo, si se quiere que la evaluación económica llegue un poco más lejos y oriente algo más a los responsables de tomar decisiones sobre asignación de recursos, las medidas de resultado basadas en variables intermedias no son adecuadas.

Las guías analizadas se inclinan por el uso de perfiles genéricos. El argumento principal que sustenta esta opción es que los perfiles de salud genéricos permiten establecer comparaciones entre la efectividad de tecnologías médicas para patologías diversas, lo cual no cabe si se utilizan perfiles específicos. El posible recurso a perfiles específicos en la evaluación económica depende, de nuevo, del papel que se espera que juegue ésta en la toma de decisiones. Si se estima que afirmaciones del estilo “el medicamento A respecto del B mejora en 3 puntos la calidad de vida de un 20% de pacientes en la escala Y”, pueden ayudar a la toma de decisiones de asignación de recursos sanitarios, entonces podrían ser incluidas estas escalas específicas en las evaluaciones económicas. En tal supuesto, habría de quedar claro que los resultados ni son comparables entre patologías, ni tienen una interpretación objetiva en relación con el coste, ni (incluso) permiten la comparación de forma clara entre diversas ganancias de salud para una misma patología. Con esto se pretende advertir de que no es correcto utilizar reglas de decisión como el “coste por punto de la escala ganado”, ya que implica que cada punto de la escala tiene el mismo valor. En nuestra opinión, el uso de estas escalas antes que contribuir a mejorar la toma de decisiones de asignación de recursos sanitarios, genera confusión en el proceso. Las guías, de hecho, no aconsejan esta perspectiva.



Las escalas basadas en la psicometría, aunque tienen la ventaja de que producen resultados numéricos susceptibles de ser comparados entre patologías, presentan problemas similares a los de las escalas específicas en su aplicación a la evaluación económica. Por esta razón no es de extrañar que la mayor parte de las guías no contemplen el uso de perfiles genéricos como el SF-36 u otros similares en las evaluaciones económicas. En conclusión, si el objetivo de la evaluación económica es proporcionar reglas de decisión que sean claras y tengan un cierto fundamento en valoraciones sociales (preferencias), las herramientas de medida de la CVRS que se deben emplear son las utilidades. De nuevo, aunque la mayor parte de guías adoptan esta perspectiva, algunas como las australianas admiten la utilización de perfiles psicométricos en los informes farmacoeconómicos.

Por lo que atañe al origen de las preferencias, la mayoría de las guías propone utilizar valores procedentes de la población general. En este sentido, las guías suecas constituyen una singularidad. En realidad, no creemos que puedan esgrimirse argumentos incontestables a favor de una u otra opción (pacientes *vs* población general). Parece claro que la población general, tanto por el hecho de financiar los servicios sanitarios como por el de constituir el colectivo de pacientes potenciales, ha de ver reflejadas sus preferencias en el proceso de asignación de los recursos sanitarios. Por otro lado, la literatura ofrece evidencia de que los valores de los pacientes no coinciden con los de la población general (Ubel et al 2003) o, al menos, de que existen razones para suponer que hay discrepancias entre las preferencias de



los pacientes y las de la población general. Si estas discrepancias se debieran a que la población general valora los problemas de salud sobre la base de estereotipos que poco tienen que ver con la realidad de la enfermedad, dichas preferencias serían cuestionables. Incluso podrían estos valores conducir a decisiones sobre asignación de recursos que acabarían perjudicando a los propios individuos que han contribuido a su formación cuando, en un futuro, pasasen a ser enfermos. No creemos, por tanto, que la literatura permita adoptar una postura tajante a este respecto. En esta línea, nos parece adecuada la sugerencia de las guías canadienses en el sentido de que las medidas directas de utilidades de los pacientes puedan jugar un papel importante en ciertas ocasiones, junto con la recomendación de utilizar de manera sistemática utilidades procedentes de instrumentos tales como el *Euroqol* o el *HUI*. Así mismo, también parece oportuna la propuesta de las guías australianas de que los estados de salud provengan de medidas directas de los pacientes.

Por lo que respecta a los métodos de obtención de utilidades, en la actualidad no parece que exista una razón teórica definitiva para elegir entre la “Compensación Temporal” (TTO) o la “Lotería Estándar” (SG). Sí parece, por el contrario, que hay argumentos suficientes para excluir a la “Escala Visual Analógica” (VAS) de los métodos apropiados para obtener utilidades, o, al menos, para no utilizar dicho método en exclusiva (Bleichrodt y Johannesson, 1997). En este sentido, nos parece correcta la decisión de la mayor parte de guías de excluir este método y considerar tanto el SG como el TTO como métodos que gozan de similar validez teórica.



En la actualidad, los AVAC son la medida de resultados sanitarios que presenta menos limitaciones en su uso para evaluar económicamente tratamientos sanitarios. Por tanto, no es de extrañar que la mayor parte de guías se inclinen por los AVAC como medida prioritaria. No por ello deja de resultar comprensible la actitud precavida de muchas de las guías en relación con los AVAC, pues existe evidencia suficiente de que algunos (cuando no todos) los supuestos sobre los que se asientan los AVAC no reflejan de forma adecuada las preferencias de los individuos. Respecto a este punto, la postura que creemos más correcta en el momento actual se puede sintetizar del modo siguiente:

- a. Los AVAC no son una representación exacta de las preferencias sociales. Sería absurdo pretender negar este hecho, habida cuenta de que los AVAC son un modelo y, como todos los modelos, conllevan una simplificación de la realidad. Por consiguiente, habrá situaciones en las que la realidad sea mucho más complicada y el modelo no la describa de forma adecuada.
- b. No existe ningún modelo mejor que el AVAC (incluido el de los Años de Salud Equivalentes - HYE). Por tanto, en nuestra opinión, la mejor opción consiste en aceptar el AVAC, por defecto.
- c. Resulta necesario identificar aquellas situaciones en las que el modelo se comporta peor y sugerir alternativas para esos supuestos que, de momento, serán casos aislados. En este sentido, las guías suecas aciertan cuando señalan



una situación concreta en la que los AVAC claramente han puesto de manifiesto sus carencias (cambios importantes pero de muy corta duración en el estado de salud) y, además, proponen un método alternativo para solucionar estas limitaciones de los AVAC. Esta actitud de las guías suecas nos parece particularmente elogiada.

Finalmente, por lo que respecta a la cuestión de la medida de resultados, las guías no consideran, en general, el ACB y la disposición a pagar como una opción metodológica al mismo nivel que el ACU desde el punto de vista de su aplicación en la práctica. Sin embargo, no excluyen a priori la posibilidad de recurrir al método de la disposición a pagar. En nuestra opinión, no existe experiencia suficiente como para poner el ACB a la misma altura que el ACU como método que oriente la toma de decisiones de asignación de recursos. No obstante lo cual, y en línea con lo que se acaba de señalar en el punto anterior, el ACB puede ser una alternativa interesante en los casos en los que existan razones evidentes para pensar que el ACU falla.

3.6. Costes incluidos en el análisis

La revisión de las cuatro guías pone de manifiesto notables discrepancias entre ellas en el tema de los costes. Así, por ejemplo, las guías del NICE y las australianas parecen asumir un enfoque más pragmático y exigen que su caso base o de referencia se lleve a cabo desde la perspectiva del financiador público. En oposición a ello, las guías canadienses y suecas adoptan la perspectiva social. Ciertamente, desde un



punto de vista normativo, como antes se ha señalado, parece difícil justificar otra cosa que no sea la perspectiva social. Esta es la razón por la cual, incluso aquellas guías que adoptan la perspectiva del financiador, sobre todo si es público, aconsejan realizar un análisis de sensibilidad asumiendo la perspectiva social.

Las guías no son demasiado precisas a la hora de recomendar unos métodos u otros de medida de los recursos. De nuevo parece que estamos ante un problema de “teoría vs práctica”. Esto es, lo ideal es disponer de una medida lo más pormenorizada y exacta de los recursos utilizados (microcoste o “bottom-up”), pero eso suele estar reñido con la posibilidad de disponer en la práctica de medidas muy precisas de todos los recursos consumidos. En principio, sería deseable algún tipo de indicación acerca de cuándo un recurso es lo suficientemente importante como para demandar una medida detallada de su consumo aunque, posiblemente, cualquier umbral que se utilizase acabaría siendo, más o menos arbitrario.

En el apartado de valoración monetaria del uso de recursos, consideramos oportuno destacar que tanto las guías de NICE, como las canadienses y las australianas, hacen referencia a la utilización de costes estándar. Las principales ventajas de tener costes estándar es la comparabilidad entre evaluaciones y la transparencia de las mismas. El inconveniente radica en que dichos costes, en ocasiones, pueden no reflejar de forma adecuada los costes de oportunidad. Como hemos visto, este es un peaje que las guías están dispuestas a pagar, dadas las ventajas que representa poder disponer de unos costes homogéneos de referencia.



Por otra parte, resulta un tanto decepcionante el total desacuerdo que se pone de manifiesto entre las guías en relación con el tratamiento de los costes de productividad. En las cuatro guías revisadas encontramos prácticamente todas las alternativas metodológicas posibles, a saber: no tomar en cuenta estos costes en absoluto (Australia), tomarlos en consideración en forma de análisis complementario (NICE), incluirlos en el análisis utilizando el método de los costes de fricción (Canadá), y valorarlos, pero recurriendo al método del capital humano (Suecia). En lo que sí parece existir consenso, siquiera parcial, es en que estos costes se deben tener en cuenta de una forma u otra, y es que, desde el punto de vista normativo, es difícil justificar su exclusión. En lo que respecta al procedimiento de valoración, sin embargo, la literatura no permite dar argumentos concluyentes a favor del método de fricción o del método de capital humano.

Por último, respecto de los costes relacionados con las mejoras en la esperanza de vida causadas por las tecnologías médicas (“survivor costs”), únicamente las guías suecas se pronuncian claramente a favor de su inclusión.

3.7. Horizonte temporal de los costes y efectos

En las cuestiones referidas al horizonte temporal de las evaluaciones las guías analizadas muestran un acuerdo casi total. Así, todas las guías recomiendan que el análisis se desarrolle en un horizonte temporal suficientemente largo como para que



sean tenidos en cuenta todos los costes y beneficios asociados a todas las intervenciones comparadas. En muchos casos, los costes y efectos relacionados directa o indirectamente con una tecnología se manifiestan durante toda la vida de los pacientes y las guías aconsejan, en consecuencia, utilizar este horizonte temporal. No se obliga, sin embargo, a utilizar dicho horizonte temporal en todas las circunstancias. Por otra parte, aunque sólo las guías canadienses dicen explícitamente que el horizonte temporal ha de ser el mismo para costes y beneficios, esta prescripción parece estar implícita en el resto de guías.

En todas las guías el factor clave para elegir el horizonte temporal es el tipo de enfermedad. Las guías coinciden en la necesidad de estimar costes y beneficios de los tratamientos durante toda la vida de los pacientes en el caso de enfermedades crónicas. En general, se subraya la importancia de elegir dicho horizonte temporal en el caso de que las intervenciones comparadas tengan un impacto diferente sobre la mortalidad de los pacientes. También las guías del NICE, canadienses y australianas indican que es necesario justificar la elección de un horizonte temporal más corto.

De la revisión de las guías aflora un procedimiento común a todas para extender datos primarios de corto plazo a un horizonte temporal de largo plazo, a saber, el desarrollo de modelos de decisión. Las guías suecas no proporcionan detalles acerca de estos modelos, mientras que las otras tres advierten de la necesidad de incorporar muchos supuestos en la extrapolación de datos. Esta es la principal razón por la cual se recomienda expresamente la presentación de los modelos de manera transparente,



explicando todos los supuestos asumidos e identificando las fuentes de datos. Además, el análisis de sensibilidad tendrá que indicar el nivel de incertidumbre que rodea a cada uno de los parámetros y a la propia estructura del modelo. Es interesante señalar que las guías del NICE sugieren utilizar diferentes escenarios para los efectos de un tratamiento respecto del cual no se dispone de datos primarios, que parece el elemento más relevante de la extrapolación. Las guías australianas sostienen, correctamente en nuestra opinión, que los modelos de Markov son el instrumento más adecuado para extender datos primarios, sobre todo en el caso de enfermedades crónicas.

También coinciden las guías en la necesidad de variar el horizonte temporal en un análisis de sensibilidad o en escenarios alternativos, sobre todo cuando la evaluación económica se basa en modelos. En particular, en el caso de enfermedades crónicas, las guías canadienses sugieren aplicar un horizonte temporal de 1 y 5 años, mientras las guías suecas aconsejan variar la duración del tratamiento de 1 a 5 años.

3.8. Tasa de descuento

Todas las guías recomiendan descontar tanto los costes como los beneficios futuros. La tasa de descuento propuesta varía entre el 3% (Suecia) y el 5% (Australia y Canadá). Estas tasas resultan coherentes con las tasas comúnmente utilizadas en las evaluaciones económicas que se encuentran en literatura.



En general, se dice poco en las guías acerca de las razones que conducen a elegir la tasa de descuento en el caso base. Sin embargo, todas las guías optan por aplicar la misma tasa a costes y beneficios futuros, probablemente para asegurar la coherencia del análisis y evitar el riesgo de situaciones paradójicas, como antes se señaló.

En todas las guías se refleja la necesidad de variar la tasa de descuento en el análisis de sensibilidad para calibrar su impacto sobre los resultados finales. Es interesante hacer ver cómo todas las guías sugieren construir un escenario en el que no se aplica ningún ajuste temporal a costes y beneficios (tasa del 0%). Por último, las guías suecas y australianas, recomiendan explícitamente incluir un análisis de sensibilidad en que los beneficios se descuentan a una tasa menor que los costes.

3.9. Tratamiento de la incertidumbre

Las guías difieren por lo que respecta al grado de detalle y la importancia que se otorga al tratamiento de la incertidumbre. En particular, dos guías (NICE y Canadá) analizan con gran detalle todos los elementos relevantes en el tratamiento de la incertidumbre y la variabilidad en la evaluación, en tanto que las guías australianas analizan sólo algunos aspectos de la cuestión. Finalmente, las guías suecas sólo aluden de modo sucinto al análisis de sensibilidad y por esta razón no se consideran en los comentarios que siguen.



Las guías del NICE y las de Canadá coinciden en distinguir entre fuentes de incertidumbre y de variabilidad. También refieren con gran detalle todas las formas relevantes de incertidumbre en parámetros y características de un modelo, y por otro lado las fuentes de variabilidad de los datos. Las guías australianas, por su parte, hacen referencia únicamente a la incertidumbre que afecta a los parámetros y, parcialmente, a la estructura de los modelos.

De la lectura de las guías se puede concluir que todas ellas prestan más atención a las evaluaciones económicas basadas en modelos (y, por tanto, basadas en síntesis de datos de fuentes diferentes), por ser las más frecuentes en la literatura. No obstante, las guías del NICE y las canadienses consideran la posibilidad de que las evaluaciones se realicen a partir de datos de pacientes, en cuyo caso recomiendan explícitamente el uso del análisis estocástico.

Las guías del NICE exigen la realización de un análisis de sensibilidad probabilístico para tratar la incertidumbre alrededor de los parámetros de los modelos. También las guías canadienses aconsejan un análisis probabilístico en este caso, aunque no con carácter obligatorio. En ambos casos, se considera este tipo de análisis el más apropiado para cuantificar la incertidumbre sobre los inputs de un modelo y sobre las interacciones entre ellos. Por otro lado, las guías australianas no contemplan la posibilidad de utilizar estos métodos más avanzados, y recomiendan métodos más tradicionales, como el análisis univariado o bivariado. En el caso de la incertidumbre sobre las características de los modelos, todas las guías sugieren el uso del análisis



determinístico, aunque las guías canadienses y las del NICE proporcionan muchos más detalles acerca de los posibles elementos de incertidumbre. Por último, estas dos mismas guías aconsejan presentar resultados para subgrupos de pacientes, así como tener en cuenta la posible heterogeneidad de los costes y beneficios para individuos de diferentes edades, sexo, comorbilidades, etc.

Todas las guías recomiendan realizar al menos un análisis de sensibilidad univariado sobre todos los parámetros de un modelo y los supuestos claves del modelo en sí mismo. Las guías canadienses y australianas, proponen un análisis multivariado sólo sobre los parámetros de mayor impacto y sobre los resultados coste-efectividad. Por su parte, las guías de NICE y las canadienses recomiendan además el análisis probabilístico, y enfatizan la importancia de llevar a cabo este análisis sobre todos los inputs y parámetros incluidos en el estudio. Finalmente, las guías del NICE y las canadienses mencionan la posibilidad de adoptar el método del “valor esperado de la información” para valorar la necesidad de obtener más información sobre algunos de los parámetros (Claxton, 1999; Sculpher y Claxton, 2005).

Las guías que recomiendan (australianas) o aceptan (canadienses) el análisis de sensibilidad determinístico, proponen variar los parámetros según sus valores extremos, así como tener en cuenta todos los escenarios posibles. No obstante, las guías canadienses ofrecen bastantes más detalles en relación con el rango en el que pueden moverse estos parámetros, y sugieren rangos diferentes según las fuentes de los datos. En particular, las guías canadienses recomiendan que la variación de los



parámetros esté siempre justificada por el analista. Las guías del NICE, que aceptan sólo los análisis probabilísticos, recomiendan presentar todas las distribuciones asociadas a los parámetros con detalle. Por último, tanto estas guías como las canadienses exigen que el análisis de subgrupos incluya grupos homogéneos de pacientes en términos de costes y beneficios.

Aunque la utilización de tablas se considera aceptable por todas las guías, se estima más apropiado el uso de representaciones gráficas. Para el caso del análisis univariado, las guías canadienses y australianas sugieren el uso de *tornado diagrams*, que permiten visualizar el impacto que tiene sobre los resultados cada parámetro incierto. En el caso del análisis probabilístico, las guías canadienses y las del NICE aconsejan el uso de *scatter plots* en un plano coste-efectividad o de “curvas de aceptabilidad”. Estas representaciones gráficas proporcionan mucha más información a los responsables de tomar las decisiones que la que se deriva de una única ratio coste-efectividad, ya que permiten cuantificar la probabilidad de que un nuevo tratamiento sea coste-efectivo según diferentes disposiciones a pagar (Fenwick, 2001). De este modo, cada institución puede decidir si considera conveniente o no el reembolso de la tecnología o tratamiento evaluados en función de cuál sea su restricción presupuestaria (Stinnett, 1998).



3.10. Uso de modelos decisionales

Las guías revisadas coinciden en la mayoría de los elementos importantes que tienen que ver con el uso de modelos, pero alcanzan un nivel de detalle muy dispar. Por ejemplo, las guías australianas y canadienses se caracterizan por una descripción detallada de todos los aspectos fundamentales de los modelos, en tanto que las guías del NICE y las suecas son menos precisas al respecto. Habría que subrayar, no obstante, que en el caso de la guías del NICE, algunos detalles relativos a los modelos están recogidos (explícita o implícitamente) en otros apartados como el dedicado al tratamiento de la incertidumbre o el que describe los métodos de síntesis de la evidencia. En general, parece que las recomendaciones de las guías se ajustan a lo que establece la teoría en relación con el uso de modelos en evaluaciones económicas.

Todas las guías coinciden en que la principal virtualidad de los modelos está en su capacidad para extender datos de corto plazo a un horizonte temporal más largo y en su utilidad como método de sintetizar datos procedentes de muchas fuentes distintas. Algunas guías identifican otros objetivos importantes, como conseguir que el análisis sea relevante para la población de estudio, transformar medidas de beneficios intermedias en medidas de beneficio finales, transferir datos de otros países al contexto geográfico propio y estimar la incertidumbre que gravita sobre los parámetros. Para las guías del NICE el desarrollo de un modelo decisional es obligado en todas estas situaciones.



En general, las guías no proporcionan demasiados detalles sobre el tipo de modelo que resulta más apropiado para cada tipo de enfermedad e intervención, y la decisión se deja en manos del investigador. Sin embargo, sí se advierte en las guías de que el modelo elegido tiene que reflejar la condición de estudio y ser relevante para la población objetivo y la finalidad del análisis. Por otro lado, todas las guías (excluyendo las suecas que no dan detalles al respecto) muestran su preferencia por el uso de modelos probabilísticos; preferencia que, en el caso de las guías del NICE se convierte en requisito obligatorio. Este enfoque resulta plenamente coherente con las nuevas teorías sobre el tratamiento de la incertidumbre que rodea a los parámetros del modelo (Sculpher et al 2000). Las guías australianas, por su parte, expresan su preferencia por los modelos dinámicos para el caso particular de las vacunas.

La transparencia del modelo parece un elemento clave. Todas las guías (excluyendo, de nuevo, a las suecas, que no dan detalles) demandan la descripción detallada y la justificación de todos los supuestos sobre la estructura del modelo, así como de las fuentes utilizadas en su elección. Las guías australianas requieren además que se aporte una versión electrónica del modelo. Por otro lado, tanto las guías australianas como las canadienses recomiendan desarrollar modelos flexibles, con la finalidad de que sea posible la incorporación de datos futuros de nuevas tecnologías. La flexibilidad se revela, además, como un elemento importante para transferir datos de un país a otro. Las guías australianas proporcionan, además, muchos detalles sobre el



tipo de información necesaria en el caso de modelos de Markov y tratan de forma muy exhaustiva los elementos clave de la estructura de los modelos.

En lo relativo a los detalles sobre la procedencia de los parámetros del modelo, también se pone mucho énfasis por parte de las guías en la necesaria transparencia del modelo por lo que se refiere a la información relativa a los datos utilizados, sus fuentes y los métodos para sintetizar toda la información. Las guías coinciden en la exigencia de aportar cierta información mínima sobre los parámetros, como el valor medio, su intervalo de variación, las distribuciones asumidas, etc. Se exige, asimismo, describir los métodos para extender datos primarios de corto a largo plazo. Como se ha señalado antes, las guías australianas y canadienses tratan de forma más exhaustiva los detalles que se recomienda aportar en relación con los datos y los parámetros del modelo.

Las guías del NICE son las únicas entre las revisadas que no exigen de forma explícita la comprobación de la coherencia interna y externa de los modelos. Las tres restantes recomiendan utilizar todos los métodos estándares para comprobar la coherencia del modelo.

3.11. Transferibilidad de datos y resultados

Las guías coinciden en la mayoría de las cuestiones clave sobre el tema de la transferibilidad, aunque, como se ha señalado en otras partes de esta revisión, el



grado de detalle que alcanzan es ciertamente dispar. En particular, las guías canadienses contienen un apartado específico dedicado a la generalización de los resultados, separando explícitamente las diferentes dimensiones de la transferibilidad. Las guías del NICE y las australianas también ofrecen numerosos detalles sobre la cuestión, sin que ello merezca un apartado específico en las guías (gran parte de las disposiciones relativas a la transferibilidad se encuentran en las secciones dedicadas al tratamiento de la incertidumbre). Las guías suecas sólo mencionan la necesidad de adaptar algunos datos al contexto sueco y la utilidad de los modelos para tener en cuenta de la variabilidad de los datos, sin profundizar en el asunto, por lo que no se hará referencia a ellas en los comentarios que siguen.

Las tres guías (excluimos las suecas) consideran los diferentes aspectos relevantes en relación con la cuestión de la transferibilidad. Se subrayan, de un lado, los problemas de transferibilidad de datos entre países y, por otro, los problemas de la validez externa de los datos de eficacia y recursos consumidos procedentes de ensayos clínicos. Las guías de NICE y canadienses mencionan además la posible heterogeneidad de los datos procedentes de diferentes centros del país, así como las potenciales diferencias entre regiones de un mismo país.

Todas las guías coinciden en la imposibilidad de utilizar datos económicos generados en otros países. Se exige en todo caso que los costes unitarios y los datos sobre recursos consumidos se refieran al propio país. En el supuesto de utilizar datos de recursos procedentes de otros países, los investigadores tendrán que adaptarlos a las



realidades locales, justificar la elección de los valores, y someterlos a variación en un análisis de sensibilidad. Por otro lado, todas las guías consideran transferibles los datos de eficacia y efectividad, aunque sólo en determinadas circunstancias. Por ejemplo, el NICE sostiene que mientras los resultados de eficacia relativa de un tratamiento son transferibles de un país a otro, los datos de eficacia absoluta no lo son y, en consecuencia, procede su adaptación al contexto del Reino Unido. Las guías canadienses y australianas sugieren llevar a cabo tests de heterogeneidad y analizar la comparabilidad de los pacientes y de las realidades epidemiológicas antes de utilizar datos de eficacia procedentes de otros países. Por último, las guías del NICE y las canadienses aconsejan el uso de utilidades obtenidas de la población del Reino Unido y de Canadá, respectivamente, mientras las guías australianas no hacen referencia explícita a posibles problemas de transferibilidad de los valores de utilidades.

En general, las guías reflejan el consenso en la literatura teórica sobre evaluación económica, a saber, que los datos económicos no son transferibles entre países, mientras que los datos de efectividad y beneficios son transferibles sólo en algunas situaciones, siempre que exista homogeneidad en los pacientes, las características epidemiológicas y las prácticas clínicas (Welte et al 2004)

Las tres guías distinguen entre entre evaluaciones económicas basadas en microdatos de pacientes y evaluaciones basadas en modelos, aunque las del NICE y las



australianas proporcionan más detalles al respecto de los análisis basados en modelos.

Las guías canadienses abordan de manera explícita la cuestión de la transferibilidad de los datos de pacientes individuales y, en particular, el caso de los ensayos clínicos multinacionales. Los canadienses aceptan el uso de los datos de eficacia agregados de los ensayos, aunque recomiendan realizar un test de homogeneidad. Por otro lado, consideran inadmisibles utilizar la media agregada por lo que se refiere a los recursos consumidos en el ensayo, pues se considera que pueden diferir sistemáticamente entre países. Para solucionar este problema se sugieren diversas estrategias, y se da cuenta de todos los métodos estadísticos recientemente desarrollados a tal fin. Las guías australianas aconsejan también el uso de técnicas estadísticas (meta-regresiones, test de homogeneidad) para tener en cuenta las posibles diferencias entre pacientes de un estudio y el contexto australiano. Por último, el NICE recomienda el uso del *bootstrapping* para tener en cuenta de la variabilidad y las diferencias entre pacientes de un estudio y la población del Reino Unido.

Todas las guías subrayan la importancia de los modelos para estimar la incertidumbre, la variabilidad y la transferibilidad de los datos obtenidos por varias fuentes. Las guías analizadas concuerdan en algunos puntos clave. Todas coinciden en la necesidad de justificar la inclusión de los estudios utilizados para obtener los parámetros del modelo y consideran que el meta-análisis es el método de síntesis más apropiado. Así mismo es unánime la recomendación de tener en cuenta la



homogeneidad entre los estudios incluidos y los contextos del análisis, así como se comparte la convicción de que el análisis de sensibilidad probabilístico resulta útil no sólo para estimar la variabilidad de los parámetros sino también para juzgar su transferibilidad.

El problema que se plantea con las condiciones especiales en las que tiene lugar un ensayo clínico y la dificultad que ello representa de cara a su transferibilidad a la práctica clínica real se reconoce y se trata en todas las guías. En general, la recomendación consiste en comparar con estudios alternativos (por ejemplo, observacionales), preferiblemente a largo plazo, para confirmar o modificar algunos datos de los ensayos clínicos, además de excluir algunos elementos no transferibles a la práctica habitual como, por ejemplo, los costes del protocolo.

3.12. Presentación de los resultados

Las guías del NICE, australianas y canadienses muestran una plena coincidencia en todos los aspectos teóricos clave de la presentación de los resultados. De nuevo, las guías suecas apenas ofrecen información y serán, en consecuencia, excluidas de estos comentarios.

Todas las guías se extienden en detalles acerca de la información necesaria para presentar los resultados del análisis, y coinciden en los aspectos más relevantes, como el análisis incremental, la presentación de costes y efectos desagregados y la



importancia del tratamiento de la incertidumbre. Las guías comparten también la recomendación consistente en presentar los resultados en forma de gráficos y tablas para facilitar la interpretación. Además, las guías del NICE y las canadienses disponen de un formato estándar que permite a los investigadores ajustarse a las demandas de información de las instituciones.

Se considera, en general, fundamental presentar costes y efectos de forma desagregada en la primera parte del informe. En el caso de los costes, se sugiere presentar los resultados totales para cada tecnología separados por categorías (todas las guías), descontados y no descontados (todas las guías), según la perspectiva (Canadá y Australia) y según el agente sanitario (NICE). Las guías australianas y canadienses sugieren también presentar los resultados de efectividad en unidades naturales y explicar el proceso de agregación utilizado para obtener los AVAC u otras medidas de resultados finales. Las guías australianas indican además las formas más apropiadas de ofrecer los resultados en el caso de evaluaciones económicas basadas en modelos. Por último, todas las guías sugieren presentar por separado y en detalle todos los resultados de los análisis de subgrupo y destacar los aspectos o características que tienen más impacto sobre el coste-efectividad de las tecnologías.

Todas las guías exigen la realización de un análisis incremental, subrayando la información equívoca que ofrecen las ratios medias coste-efectividad. Mientras las guías del NICE y canadienses sugieren especificar (también por medio de tablas) las estrategias dominadas en el caso base, las guías australianas aconsejan extremar la



cautela al excluir estrategias por dominancia (estricta o extendida), que en todo caso debe ser confirmada por los resultados del análisis de sensibilidad. Todas las guías muestran su preferencia (en el caso del NICE es un exigencia) por el análisis probabilístico, estimándose que el mejor modo de presentar las ratios incrementales es mediante puntos (simulaciones) en el plano coste-efectividad.

Las guías coinciden en que el análisis de sensibilidad y de subgrupos representa un aspecto fundamental de la evaluación económica y ha de presentarse en detalle. Se recomienda en general utilizar las curvas de aceptabilidad para el análisis de sensibilidad probabilístico y los “diagramas tornado” para el univariado. Las guías admiten la presentación de los análisis de subgrupo por medio de tablas muy detalladas.

Las guías consideran imprescindible que el informe incluya una discusión en la que se pongan de manifiesto las limitaciones del análisis, se compare el estudio con otros que evalúen la misma tecnología o similares, y se oriente acerca de los valores más expuestos a incertidumbre y sobre los cuales se precisa investigación adicional en el futuro. Sólo las guías del NICE explicitan reglas de decisión al respecto de la aceptabilidad de las tecnologías analizadas (conformes con los umbrales identificados en la literatura). Las guías canadienses y australianas se conforman con sugerir que los resultados se comparen con otros estudios para poder formar un juicio acerca de su coste-efectividad. En general, todas las guías coinciden en la importancia de considerar otros aspectos además del análisis coste-efectividad a la



hora de aceptar un programa, como el nivel de incertidumbre en los resultados, el impacto presupuestario o si existen alternativas similares para la enfermedad estudiada.



Referencias bibliográficas

Ades AE, Lu G, Higgins JPT. The interpretation of random-effects meta-analysis in decision models. *Med Decis Making* 2005 ; 25: 646-654

Ades AE, Sculpher M, Sutton A, Abrams A, Cooper N, Welton N, et al. Bayesian methods for evidence synthesis in cost-effectiveness analysis. *Pharmacoeconomics* 2006;24:2-29

Barbieri M, Drummond M, Willke R et al. Variability of cost-effectiveness estimates for pharmaceuticals in Western Europe: lessons for inferring generalizability. *Value Health* 2005; 8(1): 10-23

Bleichrodt H, Johannesson M. An experimental test of a theoretical foundation for rating-scale valuations. *Med Decis Making* 1997 Apr-Jun;17(2):208-216.

Brazier J, Deverill M, Green C, Harper R, Booth A. A review of the use of health status measures in economic evaluation. *Health Technol Assess* 1999; 3(9).

Briggs A, Claxton K, Sculpher M. *Decision Modelling for Health Economic Evaluation*. Oxford: Oxford University Press 2006

Briggs A, Goeree R, Blackhouse G, O'Brien B. Probabilistic analysis of cost-effectiveness models: choosing between treatment strategies for gastroesophageal reflux disease. *Med Decis Making* 2002; 22: 290–308

Briggs A, Sculpher M, Buxton M. Uncertainty in the economic evaluation of health care technologies. The role of sensitivity analysis. *Health Econ* 1994; 3: 95-104



Briggs A, Sculpher M. An introduction to Markov modelling for economic evaluation. *Pharmacoeconomics* 1998; 13: 397–409

Briggs A. Handling uncertainty in economic evaluation and presenting the results. In Drummond M, McGuire A, editors. *Economic evaluation in health care: merging theory with practice*. Oxford: Oxford University Press; 2001. Chapter 8

Briggs AH, Ades AE, Price MJ. Probabilistic sensitivity analysis for decision trees with multiple branches: use of the Dirichlet distribution in a Bayesian framework. *Med Decis Making* 2003; 23:341–50

Briggs AH. Handling uncertainty in cost-effectiveness models. *Pharmacoeconomics* 2000; 17:479–500

Brouwer WB, Koopmanschap MA. The friction-cost method : replacement for nothing and leisure for free? *Pharmacoeconomics*. 2005;23(2):105-11.

Brouwer WB, van Exel NJ, Baltussen RM, Rutten FF. A dollar is a dollar is a dollar-or is it? *Value Health*. 2006 Sep-Oct;9(5):341-7.

Cairns J. Valuing future benefits. *Health Econ* 1994; 3 (4): 221-229

Caldwell DM, Ades AE, Higgins JPG. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* 2005; 331: 897-900

Campbell MK, Torgerson DJ. Bootstrapping: estimating confidence intervals for cost-effectiveness ratios. *Q J Med* 1999; 92: 177-82

Chaudhary MA, Stearns SC. Estimating confidence intervals for cost-effectiveness ratios: an example from a randomized trial. *Stat Med* 1996; 15:1447–58



Clark HD, Wells GA, Huet C et al. Assessing the quality of randomised trials: reliability of the Jadad scale. *Control Clin Trials* 1999; 20 (5) 448-52

Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. *J Health Econ* 1999; 18:341–64

Commonwealth Department of Health and Ageing, *Manual of resource items and their associated costs for use in major submissions to the Pharmaceutical Benefits Advisory Committee involving economic analyses*, Commonwealth of Australia 2002.

Cook JR, Drummond M, Glick H, Heyse JF. Assessing the appropriateness of combining economic data from multinational clinical trials. *Stat Med* 2003; 22: 1955-76

Coyle D, Buxton MJ, O'Brien BJ. Stratified cost-effectiveness analysis: a framework for establishing efficient limited use criteria. *Health Econ* 2003; 12(5): 421-7

Drummond MF, Manca A, Sculpher MJ. Increasing the generalisability of economic evaluations: recommendations for the design, analysis and reporting of studies. *Int J Technol Assess Health Care* 2005; 21: 165-171

Drummond MF, Pang F. Transferability of economic evaluation results. In: Drummond MF, McGuire A, eds., *Economic Evaluation in Health Care: Merging Theory with Practice*. Oxford: Oxford University Press, 2001

Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. *Methods for the economic evaluation of health care programmes. Third edition*: Oxford: Oxford University Press; 2005



Drummond, M, Jefferson, T. Guidelines for authors and peer reviewers of economic submissions to the BMJ. *BMJ* 1996; 313 (7052): 275-283

Fenwick E, Claxton K, Sculpher M. Respresenting uncertainty: the role of cost-effectiveness acceptability curves. *Health Econ* 2001; 8: 779-787

Gold MR., Siegel JE, Russell LB, Siegel JE, Weinstein MC. *Cost-effectiveness analysis in health and medicine*, Oxford: Oxford University Press;1996.

Gyrd-Hansen D. Willingness to pay for a QALY. *Health Econ* 2003;12(12):1049-60.

Hammerschmidt T, Goertz A, Wagenpfeil S et al. Validation of health economics models: the example of EVITA. *Value Health* 2003; 6 (5): 553

Hay JW. Economic modeling and sensitivity analysis. *Value Health* 1998; 1: 187-93

Hoch JS, Briggs AH, Willan AR. Something old, something new, something borrowed, something blue: a framework for the marriage of health econometrics and cost-effectiveness analysis. *Health Econ* 2002; 11(5): 415-30

Jacobs P, Golmohammadi K, Longobardi T. Lifetime costs for medical services: a methodological review. *Int J Technol Assess Health Care* 2003; 19 (2) 278-86.

Jacobs P, Ohinmaa A, Brady B. Providing systematic guidance in pharmaco-economic guidelines for analysing costs. *Pharmacoeconomics* 2005; 23(2):143-53.

Johannesson M, Karlsson G. The friction cost method: a comment. *J Health Econ* 1997; 16(2):249-55; discussion 257-9.



Johannesson M, O'Connor RM. Cost-utility analysis from a societal perspective. *Health Policy* 1997 Mar; 39(3): 241-53.

Johannesson M. A note on the depreciation of the societal perspective in economic evaluation of health care. *Health Policy* 1995; 33: 59-66

Keeler EB, Cretin S. Discounting of life-saving and other nonmonetary effects. *Management Science* 1983; 29: 300-306

Kind P. Measuring quality of life in evaluating clinical interventions: an overview. *Ann Med* 2001; 33(5): 323-7.

Koopmanschap M, Burdorf A, Jacob K, Meerding WJ, Brouwer W, Severens H. Measuring productivity changes in economic evaluation: setting the research agenda. *Pharmacoeconomics* 2005; 23(1): 47-54.

Koopmanschap MA, Rutten FF, van Ineveld BM, van Roijen L. The friction cost method for measuring indirect costs of disease. *J Health Econ* 1995; 14(2): 171-89

Koopmanschap MA, Rutten FFH. Indirect costs: the consequences of production loss or increased costs of production. *Med Care* 1996; 34: DS59-68.

Kuntz K, Weinstein M. Modelling in economic evaluation. In Drummond M, McGuire A, editors. *Economic evaluation in health care: merging theory with practice*. Oxford: Oxford University Press; 2001

Lázaro A. Theoretical arguments for the discounting of health consequences: where do we go from here? *Pharmacoeconomics* 2002; 20(14): 943-61



Lázaro A., Barberán R., Rubio E. Why discount health more than monetary consequences in the economic evaluation of health care programmes? *Appl Econ*, 2002; 34 (3): 339-350

Manca A, Rice N, Sculpher MJ, Briggs AH. Assessing generalisability by location in trial based cost effectiveness analysis: the use of multilevel models. *Health Econ* 2005;14: 471-485

Mandelblatt JS, Fryback DG, Weinstein MC et al. Assessing the effectiveness of health interventions. In Gold, Siegel, *Cost-effectiveness in Health and Medicine*, Oxford University Press, 1996

Manning W, Fryback D, Weinstein M. Reflecting uncertainty in cost-effectiveness analysis. In Gold, Siegel, *Cost-effectiveness in Health and Medicine*, Oxford University Press, 1996

Mason J., Drummond M., Torrance G. Some guidelines on the use of cost effectiveness league tables. *BMJ* 1993; 306 (6877); 570-572.

McCabe C, Dixon S. Testing the validity of cost-effectiveness models. *Pharmacoeconomics* 2000; 17: 501-13

Nixon RM, Bansback N and Brennan A. Using mixed treatment comparisons and meta-regression to perform indirect comparisons to estimate the efficacy of biologic treatments in rheumatoid arthritis. *Stat Med* Aug 2006

Nord E. Severity of illness versus expected benefit in societal evaluation of healthcare interventions. *Expert Rev Pharmacoecon Outcomes Res* 2001; 1(1): 85-92

Olsen JA, Richardson J. Production gains from health care: what should be included in cost-effectiveness analyses? *Soc Sci Med* 1999; 49: 17-26.



Philips Z, Ginnelly L, Sculpher M et al. Review of guidelines for good practice in decision analytic modelling in health technology assessment. *Health Technol Assess* 2004; 8(36):1-172.

Rasanen P, Roine E, Sintonen H, Semberg-Konttinen V, Ryyanen OP, Roine R. Use of quality-adjusted life years for the estimation of effectiveness of health care: A systematic literature review. *Int J Technol Assess Health Care* 2006; 22(2): 235-41.

Revicki DA, Kaplan RM. Relationship between psychometric and utility-based approaches to the measurement of health-related quality of life. *Qual.Life Res* 1993; 2(6): 477-487.

Rothwell. External validity of randomised controlled trials: “to whom do the results of trials apply”. *Lancet* 2005; 165: 13-4

Sanders GD, Hlatky MA, Heidenreich PA et al. Use of decision models to extrapolate results from clinical trials: cost effectiveness of the implantable cardioverter defibrillator (ICD) *Med Decis Making* 1999; 19: 519

Schulman K, Burke J, Drummond M et al. Resource costing for multinational neurologic clinical trials: methods and results. *Health Econ* 1998; 7(7): 629-38

Sculpher M, Claxton K. Establishing the cost-effectiveness of new pharmaceuticals under conditions of uncertainty: when is there sufficient evidence? *Value Health* 2005; 8; 433-446

Sculpher M, Fenwick E, Claxton K. Assessing quality in decision analytic cost-effectiveness models: a suggested framework and example of application. *Pharmacoeconomics* 2000; 17: 461–77.



Siegel JE, Weinstein MC, Torrance G. Reporting cost-effectiveness studies and results. In Gold, Siegel, *Cost-effectiveness in Health and Medicine*, Oxford University Press, 1996

Smith DH, Gravelle H. The practice of discounting in economic evaluations of healthcare interventions. *Int J Technol Assess Health Care* 2001; 17(2):236-43.

Sonnenberg FA, Beck JR. Markov models in medical decision making: a practical guide. *Med Decis Making* 1993; 13 (4): 322-28

Sonnenberg FA, Roberts MS, Tsevat J et al. Toward a peer review process for medical decision analysis models. *Med Care* 1994; 32: JS52-64.

Sutton AJ, Abrams KR, Jones, DR et al. *Methods for meta analysis in medical research*. London: John Wiley; 2000.

Sutton AJ, Abrams KR, Jones, DR et al. Systematic reviews of trials and other studies. *Health Technol Assess* 1999; 2(19): 1-276

Tengs T, Pliskin J, Safran D et al. Five-hundred life-saving interventions and their cost-effectiveness. *Risk Analysis* 1995; 15 (3): 369-390.

Torgerson, D, Raftery, J. Discounting. *BMJ*, 1999; 319: 914- 915

Torrance G, Siegel JE, Luce BR. Framing and designing the cost-effectiveness analysis. In Gold, Siegel, *Cost-effectiveness in Health and Medicine*, Oxford University Press, 1996

Torrance GW. Measurement of health state utilities for economic appraisal. *J Health Econ* 1986; 5(1): 1-30.



Tranmer JE, Guerriere DN, Ungar WJ, Coyte PC. Valuing patient and caregiver time: a review of the literature. *Pharmacoeconomics* 2005; 23(5): 449-59.

Trueman P, Drummond M, Hutton J. Developing guidance for budget impact analysis. *Pharmacoeconomics* 2001; 19(6): 609-621

Tsuchiya A, Brazier J, Roberts J. Comparison of valuation methods used to generate the EQ-5D and the SF-6D value sets. *J Health Econ* 2006; 25(2): 334-46.

Ubel PA, Loewenstein G, Jepson C. 2003. Whose quality of life? A commentary exploring discrepancies between health state evaluations of patients and the general public. *Qual Life Res* 12: 599-607

Ubel PA, Richardson J, Menzel P. Societal value, the person trade-off, and the dilemma of whose values to measure for cost-effectiveness analysis. *Health Econ* 2000 Mar;9(2):127-136.

Weinstein MC, O'Brien B, Hornberger J et al. Principles of good practice for decision analytic modeling in health-care evaluation: report of the ISPOR Task Force on Good Research Practices--Modeling Studies. *Value Health* 2003; 6(1): 9-17

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(13): 857-76

West RR, McNabb R, Thompson AG et al. Estimating implied rates of discount in healthcare decision-making. *Health Technol Assess* 2003; 7(38): 1-60

Williams A. Intergenerational equity. an exploration of the "fair innings" argument. *Health Econ* 1997; 6: 117-32.



Willke RJ. Tailor-made or off-the-rack? The problem of transferability of health economic data . *Expert Rev Pharmacoecon Outcomes Res* 2003; 3(1): 1-4.