



Modelos analíticos de decisión en evaluación económica: tipos, metodología, análisis y comunicación de los resultados

Coordinadores del Proyecto:

Javier Soto
Miguel Angel Casado
Itziar Oyagüez

Modelos analíticos de decisión en evaluación económica: tipos, metodología, análisis y comunicación de los resultados

Coordinadores del Proyecto:

Javier Soto

Miguel Angel Casado

Itziar Oyagüez

Modelos analíticos de decisión en evaluación económica: tipos, metodología, análisis y comunicación de los resultados

© Fundación Porib

© Coordinadores del Proyecto: Javier Soto, Miguel Angel Casado, Itziar Oyagüez

Diseño y maquetación: Esstudio Ediciones

Primera edición: diciembre, 2023

ISBN: 978-84-19781-33-8

Todos los derechos reservados. Queda prohibida, sin autorización de los titulares del copyright, cualquier forma de reproducción total o parcial de este trabajo, distribución, comunicación pública o transformación de esta obra, salvo excepción prevista por la ley, así como la distribución de ejemplares por medio de alquiler o préstamo público. La infracción de estos apartados puede constituir un delito contra la Propiedad Intelectual.

Índice

Prólogo	5
Javier Soto, Miguel Ángel Casado, Itziar Oyagüez	
Capítulo 1. Modelos analíticos de decisión: Fundamentos y necesidad, tipos, aspectos claves en su diseño y evaluación	8
Raquel Domínguez-Hernández, Alberto de la Cuadra-Grande, Miguel Ángel Casado	
Capítulo 2. Resultados en salud a incorporar en los modelos analíticos de decisión	39
Raúl Parra Garcés, Antia Gómez Fernández, Francisco Abad Santos	
Capítulo 3. Tipos de costes a incluir en los modelos analíticos de decisión	69
Ana Ortega Eslava	
Capítulo 4. Obtención del valor de utilidad de las alternativas en evaluación en el modelo analítico de decisión	111
José María Abellán Perpiñán	
Capítulo 5. El manejo de la incertidumbre en los modelos de análisis de decisiones	148
David Epstein	
Capítulo 6. Análisis, presentación e interpretación de los resultados en modelos analíticos de decisión	171
José Manuel Rodríguez Barrios	
Capítulo 7. Evaluación de la calidad metodológica de un modelo analítico de decisión a través de una lista-guía	201
Javier Soto Álvarez	

Capítulo 8. Elaboración de un modelo analítico de decisión de novo: pasos a seguir	244
María Mareque, Itziar Oyagüez	
Capítulo 9. Modelos analíticos de decisión determinísticos basados en árboles de decisión	258
Javier Soto Álvarez	
Capítulo 10. Modelos de Markov y modelos de supervivencia dividida	280
Antonio Olry de Labry Lima, Zuzana Špacírová	
Capítulo 11. Modelos de transición entre estados	307
Carlos Crespo Palomo	
Capítulo 12. Modelos de simulación de eventos discretos	355
Javier Mar	
Capítulo 13. Modelos dinámicos de transmisión	371
M ^a Reyes Lorente Antoñanzas, Fernando Antoñanzas Villar	
Capítulo 14. Modelos bayesianos	394
Miguel Ángel Negrín, Francisco José Vázquez Polo	
Capítulo 15. Software para análisis de decisiones en medicina	417
Francisco Javier Díez, Carmen María Yago, Jorge Pérez-Martín, Manuel Luque, Manuel Arias	
Capítulo 16. Retos futuros de los modelos analíticos de decisión	458
Marta Traperó-Bertran, Mireia Díaz Sanchis, Ana Magdalena Vargas-Martínez	
Capítulo 17. El papel de los modelos en la toma de decisiones	538
Jorge Mestre Ferrándiz	

Prólogo

En el siglo XXI resulta evidente que los recursos disponibles en nuestro Sistema Nacional de Salud (SNS) son, cada vez, más limitados para abordar todas las demandas en atención sanitaria de la población. Esta situación se debe a diferentes motivos, entre ellos, el progresivo envejecimiento de la población en nuestro país, con una de las mayores esperanzas de vida a nivel mundial, y el abordaje de la cronicidad. Todo ello, supone una carga de la enfermedad relevante para la sociedad, asociada a su diagnóstico, tratamiento y seguimiento. Además, la incorporación al arsenal terapéutico de nuevos medicamentos y tecnologías sanitarias, que incrementan de manera importante los resultados en salud de los pacientes, pero que generan un impacto presupuestario significativo, es una causa adicional que justifica la necesidad real de disponer de recursos adicionales en el SNS y de emplear los existentes de la forma más eficiente posible.

En este contexto, la evaluación y reevaluación de la eficiencia de los medicamentos y otras intervenciones sanitarias cobra, una mayor transcendencia y relevancia a la hora de destinar los recursos a financiar con cargo al erario público aquellas opciones terapéuticas que permitan obtener mejores resultados en salud en los pacientes (mayor supervivencia global, significativos incrementos en calidad de vida relacionada con la salud, menor porcentaje de efectos adversos, mayor conveniencia, etc.), con una disminución notable del coste de oportunidad para el SNS y la sociedad. Además, en nuestro país, se están poniendo en marcha distintas iniciativas llevadas a cabo por grupos de trabajo independientes y por la Administración Sanitaria que, junto con futuras normativas relacionadas con la evaluación de tecnologías sanitarias, su precio y financiación, supondrán un gran avance en la disciplina.

De todos los métodos existentes en la actualidad para efectuar una evaluación de la eficiencia de las tecnologías sanitarias, el más utilizado son los modelos analíticos de decisión (MAD). Estas herramientas permiten simular, proyectar y extrapolar a corto, medio y largo plazo, en un entorno de incertidumbre clínica y riesgo financiero, los resultados en salud alcanzados (y el consumo de recursos asociado) de diferentes alternativas para diagnosticar o tratar una enfermedad, permitiendo conocer su nivel de eficiencia. Los resultados de los MAD son muy útiles para una toma informada de decisiones en política sanitaria, pero en su diseño, desarrollo y análisis es necesario seguir unos altos estándares de calidad y objetividad, para que los resultados sean válidos y creíbles para los profesionales sanitarios y otros decisores en el SNS. Este es un tema prioritario en la actualidad en todos los países y, más concretamente en el nuestro, dado que existe una corriente que aboga por potenciar el uso sistemático de las evaluaciones económicas para llevar a cabo decisiones mejor fundamentadas y con mayor transparencia a la hora de decidir que debe financiarse con fondos públicos. Todo ello, en aras de gestionar las necesidades no cubiertas y priorizar el acceso de los pacientes a tratamientos innovadores, más seguros, eficaces, efectivos y eficientes, garantizando la propia sostenibilidad y solvencia del SNS.

En este sentido, lo que nos decidió y, realmente, nos motivó para poner en marcha este proyecto, fue la idea de que existiese un libro de referencia en castellano sobre los MAD. Esta iniciativa está dirigida a todas aquellas personas que trabajan en el mundo de la evaluación de la eficiencia y utilizan MAD como metodología para su elaboración o que tienen la necesidad de evaluar la calidad metodológica de un MAD ya disponible, como paso previo a emplear sus resultados en la toma de decisiones. A nuestro entender, la entidad y grandeza de este libro se basa, primordialmente, en que está elaborado por prestigiosos investigadores que trabajan, en nuestro país, en el mundo de las evaluaciones económicas y los MAD, con excelentes y contrastados conocimientos en estas materias y métodos, y que conocen las peculiaridades de nuestro sistema público de salud y de nuestra sociedad. Todo ello, redundando en que las enseñanzas de sus capítulos se puedan

emplear con garantías para elaborar, analizar, interpretar y comunicar los resultados de un MAD.

Para finalizar, nos gustaría dar nuestro agradecimiento de manera especial a todos los expertos que han asumido el objetivo común de impulsar este proyecto y que han colaborado en la elaboración de los distintos capítulos de este libro, lo que ha sido esencial para que el documento, que a continuación pueden disfrutar, sea una realidad. Gracias por vuestra dedicación, tiempo, esfuerzo y entusiasmo.

Javier Soto, Miguel Ángel Casado, Itziar Oyagüez



1

MODELOS ANALÍTICOS DE DECISIÓN: FUNDAMENTOS Y NECESIDAD, TIPOS, ASPECTOS CLAVES EN SU DISEÑO Y EVALUACIÓN

Raquel Domínguez-Hernández^a, Alberto de la Cuadra-Grande^a,
Miguel Ángel Casado^{a,b,c}

^a *Pharmacoeconomics & Outcomes Research Iberia (PORIB)*

^b *Fundación PORIB*

^c *Consejo de Dirección del Máster de Formación Permanente en Evaluación Sanitaria
y Acceso al Mercado (Fármaco-Economía) de la Universidad Carlos III de Madrid*



Introducción a los modelos analíticos de decisión

Los modelos analíticos de decisión son esquemas teóricos que permiten extrapolar los resultados obtenidos en estudios clínicos o a partir de datos de vida o mundo real (en inglés, *Real-World Evidence*; *RWE*), a corto, medio o largo plazo, realizando simulaciones o proyecciones para evaluar la eficacia, la efectividad y la eficiencia de una intervención sanitaria (1).

Para poder generar evidencia económica, los modelos simplifican el curso real o historia natural de una enfermedad, incluyendo aquellos eventos fundamentales asociados a la patología (1). Por ello, nunca deben ser empleados como sustitutos de los ensayos clínicos, que pretenden estudiar la eficacia y seguridad de una intervención sanitaria, ni de los estudios observacionales o de *RWE*, cuyo objetivo es registrar datos o valorar la efectividad, sin intervención del investigador, para medir y describir la evolución de la población estudiada (2).

Cuando se diseña un modelo farmacoeconómico, el fin último de este es servir como herramienta en la toma de decisiones en salud para una asignación eficiente de los recursos sanitarios en un sistema cambiante y marcado por la incertidumbre clínica y el riesgo financiero (1). En estas circunstancias, si bien la evidencia procedente de los estudios biomédicos constituye información de gran valor, por lo general, no constituye un sustento suficientemente sólido y robusto para el agente decisor, lo que justifica el diseño de modelos analíticos de decisión (Figura 1).

Entre las ventajas que ofrecen los modelos analíticos de decisión, cabe destacar la inmediatez en la obtención de resultados y el bajo coste que suponen, en comparación con los estudios clínicos. Sin embargo, hay que tener en cuenta que los modelos son simulaciones y, por ende, pueden no representar fielmente la realidad, además de la cantidad de premisas o asunciones que deben tomarse a la hora de su diseño (1).



Figura 1. Justificación del desarrollo de modelos analíticos de decisión



RWE: *Real-World Evidence*.

Fuente: Adaptada de (1) y (3).

Por otro lado, los modelos deberían cumplir una serie de características para que los resultados obtenidos sean aceptables y útiles para los gestores y profesionales sanitarios. Estos deben ser transparentes para garantizar su reproducibilidad. Resulta imprescindible que un modelo explicita las fuentes de información empleadas, así como las premisas consideradas. Tienen que mostrar validez, empleando fuentes basadas en la evidencia científica, y credibilidad, reflejando el estado actual del conocimiento. Además, los modelos no solo deben proporcionar siempre los mismos resultados, sino que también deben permitir a cualquier persona reproducirlos. Finalmente, conviene que un modelo se diseñe con flexibilidad, permitiendo la posibilidad de adaptarlos a distintos contextos, entornos o países (1).

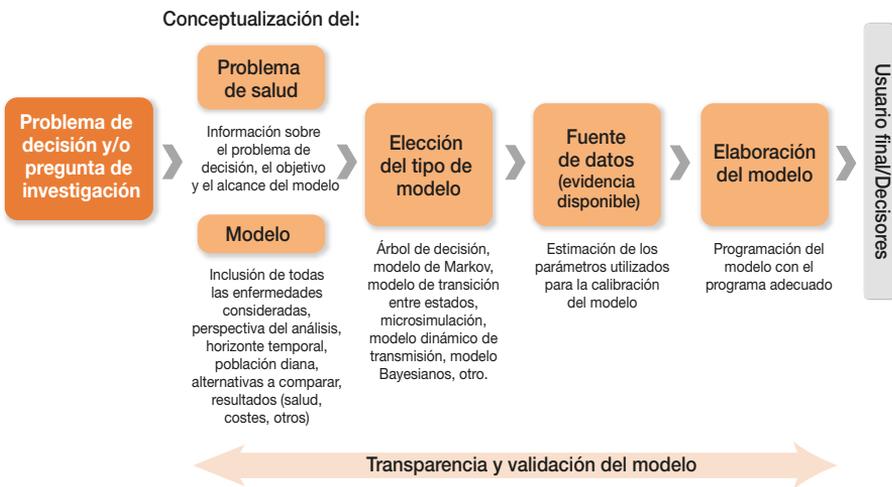


Desarrollo de un modelo de decisión

Para la construcción de un modelo analítico de decisión es necesario seguir una serie de pasos que garanticen la fiabilidad del modelo. Estos pasos se componen principalmente de la conceptualización del problema de salud y del modelo, es decir, la elección del tipo de modelo a utilizar, las fuentes de datos que “alimentarán” el modelo, la programación del modelo y su transparencia, verificación y validación (4) (Figura 2).



Figura 2. Desarrollo y construcción de un modelo analítico de decisión



Fuente: Adaptado (4).

a. Conceptualización del problema de salud y del modelo

El proceso para el diseño del modelo empieza con el planteamiento del problema de decisión o la pregunta de investigación y la conceptualización del modelo. Esto es clave para definir posteriormente todo lo relacionado con los parámetros o *inputs* en los que se basa y la forma de mostrar los resultados (4,5).

El problema debe expresar claramente la hipótesis de trabajo, el objetivo del análisis y el alcance del modelo, y justificar el motivo de su realización. Además, el modelo tiene que incluir todos los elementos necesarios para responder al problema de decisión planteado. Estos elementos se basan en la perspectiva del análisis, la población diana y sus características, las alternativas a la intervención de estudio, los resultados que se deben mostrar para facilitar la toma de decisiones, y el horizonte temporal (4,5).

Es importante que la estructura del modelo siga la línea de la pregunta de investigación inicial y no dependa de la disponibilidad de los datos (5). Lo ideal durante su desarrollo es contar con profesionales sanitarios como asesores para garantizar que refleje el proceso o avance de la enfermedad de forma teórica y que validen los elementos que lo componen.

b. Tipos de modelos, características, y elección del modelo a seguir

Los modelos deben de representar la historia natural de la enfermedad de forma simple. Estos pueden ser de diferentes tipos y con diferentes metodologías para ajustarse lo máximo posible a la realidad que se quiere simular y ofrecer resultados acordes a ello. Se pueden clasificar de forma distinta en función de su estructura, si su naturaleza es probabilística o no (determinísticos o estocásticos), y por las técnicas cuantitativas que utilice (modelos de análisis de decisiones, de extrapolación, epidemiológicos y modelos de Markov) (1).

En función de sus características, existen los modelos determinísticos basados en árboles de decisión, los modelos de transición entre estados basados en modelos de Markov y modelos de simulación individual de pacientes, los modelos de simulación de eventos discretos, los modelos dinámicos de transmisión, y los modelos bayesianos. En la Tabla 1 se incluyen los modelos más habituales para la realización de evaluaciones económicas y sus características. En los siguientes capítulos, se describirán cada uno de ellos de forma más detallada (5-9).

Tabla 1. Características principales de los modelos analíticos de decisión

Modelo	Características
Árbol de decisión	Es un modelo determinístico (utiliza fórmulas algebraicas)
	Se caracteriza por un conjunto de ramas que representan la progresión de una enfermedad, en función de probabilidades
	Se compone de nodos que describen elecciones o acciones (representadas por medio de un cuadrado), acontecimientos, consecuencias o de azar (círculo) y terminales (triángulo)
	Se usa cuando el horizonte temporal a evaluar es corto (finito)
	Los valores asignados a los parámetros que alimentan el modelo no dependen ni varían a lo largo del tiempo
Modelo de Markov	Es un modelo estocástico
	Se representa mediante un diagrama de transición de estados
	Está formado por un conjunto finito de estados mutuamente excluyentes denominados estados de salud (transitorios, recurrentes y absorbentes)
	Los individuos pueden transitar entre los estados según un conjunto de “probabilidades de transición”
	Las probabilidades de transición no dependen de la historia clínica pasada de los individuos, ni del tiempo transcurrido en un estado (propiedad markoviana)
	Las probabilidades de transición pueden ser constantes (cadena de Markov) o dinámicas a lo largo del tiempo
	Las transiciones entre estados ocurren en un periodo de tiempo discreto llamado “ciclo”
	Los pacientes solo pueden estar en único estado de salud en cada ciclo
Los pacientes no interactúan entre ellos	
Modelo de supervivencia dividida (<i>partitioned survival models</i>)	Normalmente se compone de tres estados de salud (libre de progresión, progresión y muerte)
	Cuantifica a los estados como variables independientes
	No utiliza transiciones entre estados
	Utiliza la curva de supervivencia global y la curva de supervivencia libre de progresión para determinar la proporción de pacientes en cada estado de salud
	Los pacientes solo se mueven progresivamente

Tabla 1. cont. Características principales de los modelos analíticos de decisión

Modelo	Características
Modelo de simulación individual de pacientes	Es un modelo estocástico
	Está formado por un conjunto de estados mutuamente excluyentes
	Los individuos pueden transitar entre los estados
	Las transiciones se realizan en intervalos discretos
	Los eventos ocurren a lo largo de un ciclo
	Se tiene en cuenta la historia clínica pasada de los individuos
Modelo de simulación de eventos discretos	Es un modelo estocástico
	Está formado por una serie de elementos: <ul style="list-style-type: none"> • Entidades: normalmente representan a los pacientes • Atributos: son las cualidades o propiedades de los pacientes (edad, sexo, gravedad de la enfermedad, calidad de vida, ...) • Eventos: cualquier situación o estado que puede ocurrir durante la evolución de una enfermedad (un acontecimiento adverso, ...) • Variables: definen al modelo y se mantienen a lo largo del tiempo (costes, tasa de descuento, duración de la simulación, ...) • Recursos: se definen para un periodo de tiempo determinado y conllevan un coste (visitas médicas, pruebas diagnósticas, ingresos hospitalarios, tratamientos, ...) • Acumuladores estadísticos: no participan en la simulación; se utilizan para la obtención del resultado final (costes, supervivencia, ...) • Colas: los recursos pueden dejar de estar disponible al ser utilizados por otro paciente • Retrasos: pueden ser explícitos (duración de pruebas diagnósticas) o implícitos (lista de espera para un proceso) • Ecuaciones de riesgo y generadores de números aleatorios
	Los valores de los atributos están ligados a una entidad concreta, es decir, cada entidad tiene los mismos atributos, pero con diferentes valores
	La aparición de un evento no implica el cambio de estado de un paciente
	El periodo de tiempo es continuo, los eventos pueden ocurrir en cualquier momento de la simulación sin la necesidad de ciclos de duración fija

Tabla 1. cont. Características principales de los modelos analíticos de decisión

Modelo	Características
Modelo dinámico de transmisión	Puede ser determinístico o estocástico, individual o de cohortes y estático o dinámico
	Se representa mediante un diagrama causal
	Se utiliza en poblaciones extensas/grandes
	Se crea a partir de ecuaciones finitas y ecuaciones diferenciales, en función de cómo se considere el tiempo (variable discreta o continua)
	Recoge relaciones no lineales
	Establece la influencia del estado de un individuo sobre otro, es decir, existe interacción entre pacientes
Modelo bayesiano	Utiliza métodos bayesianos para estimar la distribución de probabilidad sobre todos los valores posibles de un parámetro
	Utiliza información externa adicional (información <i>a priori</i>) que combina con los datos del modelo (información externa) para obtener el resultado final (distribución <i>a posteriori</i>)
	El resultado final se obtiene a partir de la aplicación del Teorema de Bayes

Aunque los árboles de decisión y los modelos de Markov son las herramientas más utilizadas para las evaluaciones sanitarias y para la toma de decisiones por los gestores sanitarios, existen diferentes situaciones o problemas en los que es necesario usar otro tipo de métodos (5). En ocasiones, el problema planteado puede ser representado por varios modelos o una combinación de modelos; en ese caso, la elección va a depender del que mejor se ajuste a la realidad. Así, a la hora de elegir el tipo de modelo a utilizar en la evaluación económica, es importante tener en cuenta ciertas características presentadas en el problema de decisión que afectan a su elección. Estas características se basan en la unidad de representación, es decir, si el modelo considera a los pacientes de forma individual o en cohortes homogéneas, si existe interacción entre individuos, si el periodo de tiempo a evaluar es corto o largo, si existen recursos limitados (4,5,10), si la enfermedad del análisis es infecciosa, o si se utilizan métodos bayesianos para la estimación de los parámetros (Tabla 2).

Tabla 2. Valoraciones para la elección del tipo de modelo a utilizar

Características que afectan a la elección del modelo	Modelo
Unidad de representación	<i>Grupos/Cohortes</i> <ul style="list-style-type: none"> • Árbol de decisión • Modelo de Markov
	<i>Individuos</i> <ul style="list-style-type: none"> • Modelo de simulación individual (Microsimulación) • Simulación de eventos discretos
Interacción entre individuos y otros componentes del modelo	<ul style="list-style-type: none"> • Modelos dinámicos de transmisión • Simulación de eventos discretos
Horizonte temporal y medición del tiempo	<i>Cortos</i> <ul style="list-style-type: none"> • Árbol de decisión
	<i>Largos</i> <ul style="list-style-type: none"> • Modelo de Markov • Modelo dinámico de transmisión • Simulación de eventos discretos
Recursos limitados	Simulación de eventos discretos
Enfermedad infecciosa	Modelo dinámico de transmisión
Métodos bayesianos para la estimación de parámetros	Modelo bayesiano

Además de las características descritas en la tabla anterior, hay que valorar otros aspectos comunes a las evaluaciones económicas y que afectan también a la elección del tipo de modelo. Estos se basan en el uso y el contexto en el que se empleen los resultados, el tipo de enfermedad a modelizar (aguda, grave, infecciosa), si las características clínicas previas o concomitantes de los pacientes son relevantes, el número de estados a incluir y el horizonte temporal de la simulación (Figura 3) (1).



Figura 3. Fundamentos que afectan en la elección del modelo



Los árboles de decisión son apropiados para representar problemas de decisión con horizontes temporales cortos o estructuras de valores complejas. En cambio, si es necesario representar la enfermedad a través de estados sería adecuado utilizar modelos de transición entre estados por su sencillez y por la posibilidad de evaluar la incertidumbre de los parámetros (5). En este punto, dadas las desventajas de los modelos markovianos basadas en la falta de memoria, y la complejidad en su desarrollo con el incremento de estados de salud, habría que elegir entre un modelo individual (microsimulaciones) o de cohortes (6). Cuando existen interacciones entre individuos, los modelos dinámicos de transmisión (8) y los modelos de simulación de eventos discretos son los que mejor representarían la intervención a modelizar (7). Este último, también es útil

cuando el problema de investigación implica restricciones en los recursos (5,6). A diferencia de otros, en los modelos de simulación de eventos discretos, los recursos tienen limitaciones (son finitos), es decir, siguen el concepto de coste de oportunidad (5).

Los modelos dinámicos de transmisión se utilizan cuando la intervención a evaluar se emplea en una enfermedad infecciosa. La principal diferencia con el resto de los modelos estocásticos es que existe un riesgo constante de infección entre los individuos infectados y la población, y este riesgo va a depender de la cantidad de personas infectadas en un momento dado (8).

Actualmente, se están llevando a la práctica modelos de supervivencia dividida para evaluar o informar de las decisiones de reembolso de los tratamientos oncológicos. En comparación con los modelos de transición entre estados, no incluyen un vínculo estructural entre los criterios de valoración clínicos intermedios y la supervivencia. Por otra parte, su estructura limita la realización de los análisis de sensibilidad para la evaluación de la incertidumbre en el periodo de extrapolación de la supervivencia (9).

Además, de los modelos descritos anteriormente, desde una perspectiva de salud pública, hay que hacer una mención diferencial a la realización de evaluaciones para estimar el impacto en términos en salud y económicos de vacunas o programas de vacunación orientadas a prevenir enfermedades infecciosas. Estos modelos se pueden desarrollar siguiendo las directrices de un análisis coste-efectividad (compara una nueva intervención con otra ya existente), optimización restringida (incluye el presupuesto y otras restricciones) (en inglés, *constrained optimization*) y modelos fiscales o de retorno social de la inversión (estima las consecuencias económicas y fiscales para un gobierno que invierte en salud) (en inglés, *fiscal health modeling*). La característica más destacable a la hora de desarrollar evaluaciones económicas con vacunas es que estas previenen enfermedades infecciosas, es decir, tienen efectos indirectos sobre terceros y, por

tanto, deberían ser desarrollados a partir de modelos dinámicos de transmisión (11).

Es importante señalar que la metodología utilizada en cada uno de los modelos descritos anteriormente debe ajustarse lo máximo posible a la realidad que se quiere simular y ofrecer resultados acorde a ello. Asimismo, aunque la simplicidad de un modelo es una cualidad importante para la transparencia y validación, se aconseja cierto grado de complejidad para que el modelo no pierda valor y garantice que la enfermedad o proceso reflejen de forma fiel lo que se desea analizar (4).

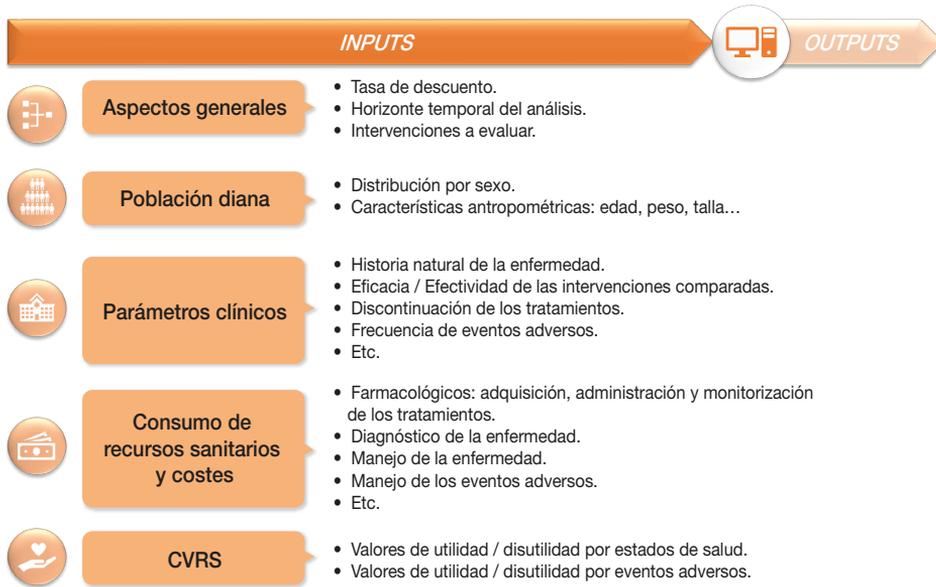
Finalmente, otras consideraciones a tener en cuenta en la elección, cuando varios modelos pueden ser potencialmente apropiados, son: evaluar el *software* disponible, las habilidades de los investigadores sobre su manejo y el tiempo necesario para el desarrollo de la simulación (10).

c. Manejo de los inputs en los modelos analíticos de decisión

Existen diferentes tipos de evaluaciones económicas en salud en las que se comparan los beneficios y/o los costes de diferentes alternativas desde distintas perspectivas (pacientes, autoridades sanitarias, financiadores, sociedad). La realización de una u otra podría condicionar el uso de un modelo u otro, así como los parámetros con los que se alimentan los modelos, que reciben el nombre de *inputs* (Figura 4).



Figura 4. *Inputs* a considerar en los modelos analíticos de decisión



CVRS: Calidad de Vida Relacionada con la Salud.

d. Fuente de datos

Variables clínicas

Un aspecto clave en los modelos es la fuente de obtención de datos. Por lo general, las variables clínicas proceden de ensayos clínicos o estudios de RWE. Con los ensayos clínicos, obtendremos resultados de eficacia, medidos en condiciones experimentales y, con los estudios observacionales o de RWE, datos de efectividad en la práctica clínica. Estos últimos resultan más útiles en evaluación económica, dado que aportan información más práctica para el decisor (12). Más allá de estos estudios, existen otros métodos de síntesis de evidencia basados en revisiones sistemáticas, como los metaanálisis, los metaanálisis en red

(NMA; en inglés *Network Meta-Analyses*) o las comparaciones indirectas, que son cada vez son más aceptados para su utilización como fuente de resultados en salud (13).

Calidad de vida relacionada con la salud (CVRS)

Existen determinadas evaluaciones económicas que, más allá de incorporar la eficacia o efectividad que presenta la intervención en estudio, también consideran el impacto de la enfermedad y de los tratamientos sobre la calidad de vida de los pacientes (14). Para ello, los modelos incorporan unos índices denominados utilidades en salud que reflejan la preferencia de la población por un estado de salud. Sus valores oscilan entre 0 (muerte o peor estado de salud imaginable para el paciente) y 1 (salud plena o mejor estado de salud imaginable) (15). Existen diversos cuestionarios validados para medir la CVRS y generar valores de utilidad, entre los que cabe destacar el cuestionario EQ-5D (*EuroQol 5 Dimensions*) ampliamente utilizado por su sencillez (16).

Consumo de recursos sanitarios y costes

Los costes suponen un aspecto esencial a estudiar en una evaluación económica. Para caracterizar los costes asociados al procedimiento sanitario estudiado se sigue un proceso dividido en tres fases: 1) Identificación; 2) Medición; y 3) Valoración.

La fase de “identificación” consiste en listar todos los recursos relevantes para la evaluación económica a realizar. Existen diversos tipos de recursos, clasificados según su ámbito (sanitarios y no sanitarios) y, a su vez, como directos e indirectos. Su inclusión dependerá principalmente de la perspectiva y el objetivo del análisis, la magnitud del coste y el horizonte temporal (17).

Una vez identificados los recursos que se incluirán en el modelo, la siguiente fase de “medición” cuantifica las unidades de cada recurso sanitario que se consumen en el proceso evaluado. Esta fase implica, por

ejemplo, determinar el número de consultas a las que acude un paciente, o la posología de cada medicamento que este recibe (17).

Por último, en la fase de “valoración” se multiplicará el número de recursos sanitarios consumidos por el coste unitario de cada uno de ellos. Los costes unitarios proceden de una amplia variedad de fuentes que comprenden bases de datos de costes, tarifas autonómicas publicadas o estudios económicos previamente realizados (17). En el caso de la pérdida de productividad existen metodologías específicas de estimación, como el coste de capital humano o los costes de fricción (18).

Tanto en la medición de los costes como de los resultados en salud, existen numerosas fuentes de incertidumbre. Además, pueden existir múltiples fuentes para un mismo coste o resultado. Asimismo, en muchas ocasiones las estimaciones empleadas consisten en un valor promedio del efecto observado para un tratamiento, o del coste que ha supuesto un procedimiento para los pacientes. Debido a ello, es esencial justificar la elección de las distintas fuentes empleando aquellas que reflejen la mejor evidencia disponible, así como estudiar el impacto que su modificación por otras fuentes alternativas produce sobre los resultados obtenidos con el modelo (17).

e. Resultados

Independientemente de su origen, los efectos que una intervención sanitaria produce sobre la salud de los pacientes se pueden estudiar por medio de variables clínicas, que son parámetros empleados por los profesionales sanitarios en su práctica clínica habitual. Normalmente, se trata de variables intermedias que únicamente miden la reducción de la morbilidad. Sin embargo, en la evaluación económica conviene utilizar variables finales que miden la reducción de la mortalidad y, por consiguiente, aportan información más útil para la toma de decisiones en salud (1,19).

Existen numerosas medidas de resultados en salud orientadas en la mortalidad de los pacientes. Entre estas, los años de vida ganados (AVG) y los años de vida ajustados por calidad (AVAC) se emplean frecuentemente en los análisis coste-efectividad (ACE) y coste-utilidad (ACU), respectivamente (14). Los AVG se calculan en función de la edad a la que fallecen los pacientes con cada alternativa. Para ello, se emplean tablas de supervivencia o se extrapolan los resultados de los estudios a lo largo de la vida de los pacientes (20). Por su parte, los AVAC miden la expectativa de vida (AVG) ajustada en función de la calidad de vida con la que viven los pacientes, medida como valor de utilidad (Tabla 3).

Tabla 3. Medidas de resultados en salud empleadas en los modelos analíticos de decisión

Medidas de Resultados en Salud			
Término en Castellano 		Término en Inglés 	
Año de Vida Ganado	AVG	<i>Life Year Gained</i>	LYG
Año de Vida Ajustado por Calidad	AVAC	<i>Quality-Adjusted Life Year</i>	QALY
Año de Vida Ajustado por Discapacidad	AVAD	<i>Disability-Adjusted Life Year</i>	DALY
Año de Vida Ajustado por Nivel de Salud	—	<i>Health-Adjusted Life Year</i>	HALY
Años Potenciales de Vida Perdidos	APVP	<i>Potential Years of Life Lost</i>	PYLL
Esperanza de Vida Ajustada por Nivel de Salud	—	<i>Health-Adjusted Life Expectancy</i>	HALE
Esperanza de Vida Libre de Discapacidad	—	<i>Disability Free Life Years / Healthy Life Years</i>	DFLE / HLY

En las evaluaciones económicas completas, que consideran costes y beneficios en salud y en la que se comparan al menos dos alternativas, el resultado final estimado es la relación coste-efectividad (RCEI) y

coste-utilidad (RCUI) incrementales. Estas permiten valorar el coste que produce una intervención frente a su comparador (costes incrementales), en relación con los beneficios o resultados en salud incrementales. La diferencia entre el RCEI y RCUI es el tipo de resultado que miden AVG o AVAC, respectivamente (14):

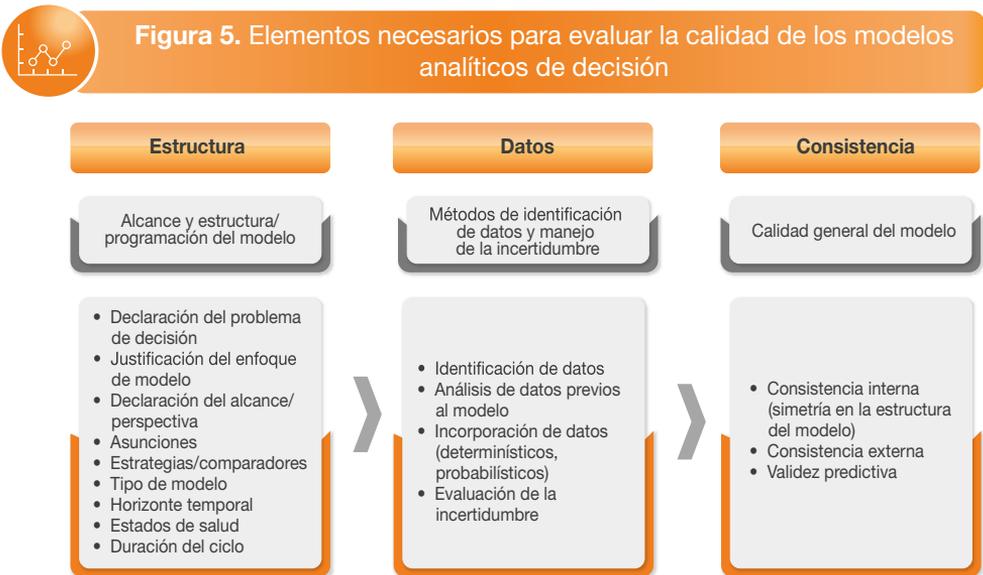
$$\text{RCEI / RCUI} = \frac{(\text{Coste Intervención Evaluada}) - (\text{Coste Comparador})}{(\text{AVG / AVAC Intervención Evaluada}) - (\text{AVG / AVAC Comparador})}$$

La RCEI y/o RCUI aporta información sobre el coste que supone la ganancia de una unidad de resultados incremental (con la alternativa que produce mejores resultados en salud), es por ello que los resultados se utilizan de forma rutinaria para la toma de decisiones en política sanitaria. La interpretación de estos resultados con objeto de determinar si la adopción de una alternativa tiene un coste adicional razonable en función del incremento de eficacia o efectividad que produce en los pacientes, requiere la identificación de un criterio o umbral de eficiencia que establezca cuál es el coste máximo que la sociedad o sus individuos están dispuestos a pagar por cada unidad de efectividad adicional lograda con una opción terapéutica frente a otras. De esta forma, se puede establecer si una alternativa terapéutica es considerada eficiente o no. Por ejemplo, el NICE (*National Institute for Health and Care Excellence*) estableció un umbral de disponibilidad a pagar de 20.000 - 30.000 £/AVAC, de forma que las intervenciones que generen un RCUI por debajo de este valor se consideran coste-efectivas (21). En España, de forma tradicional y no oficial se ha utilizado un valor de 30.000 euros por AVAC. Sin embargo, desde 2015, un estudio del Servicio Canario de Salud ha establecido un umbral de eficiencia de 22.000 - 25.000 euros por AVAC (22), aunque hay controversia sobre este valor de referencia, que según otros autores podría situarse entre los 25.000 a 60.000 euros por AVAC.

f. Transparencia, credibilidad, verificación y validación de los modelos

Los modelos de análisis de decisión deben simular la enfermedad o el proceso para reflejar de la forma más precisa la realidad en la práctica clínica. Para ello, se debe comprobar la capacidad que tiene el modelo para predecir eventos futuros. Además, se debe establecer si el modelo es lo suficientemente válido o preciso para el fin que se está construyendo (23,24).

En la Figura 5 se muestra una relación de los elementos incluidos en las guías metodológicas sobre el uso de modelos que son necesarios tener en cuenta a la hora de evaluar la calidad de estos. Aunque existen discrepancias sobre en qué medida o cómo evaluarlos, la propuesta incluye elementos comunes que forman parte de la estructura, de los parámetros y de la consistencia del modelo (25).



Fuente: Adaptada de (25).

Transparencia

Es importante que la información sea clara y comprensible para ayudar a los usuarios y decisores a entender qué hace el modelo y cómo trabaja. Todos los modelos deberían de tener un manual sencillo para los usuarios sin conocimientos de modelización, en el que se describa la metodología y la estructura del modelo, se listen los parámetros utilizados, los resultados y la forma de validación. Del mismo modo, también debería de contar con un informe (en inglés, *Technical Report o User Guide*) más técnico y detallado, orientado a los especialistas en modelización y a los revisores para que estos modelos puedan ser evaluados y reproducidos por personas con suficientes conocimientos en la disciplina (5,23). Es importante señalar que, dado que los modelos son complejos y se destina mucho tiempo y recursos a su construcción, habrá documentación confidencial que proteja la propiedad intelectual (23).

Credibilidad

Evalúa en qué medida el modelo se ajusta a la realidad. Los modelos, como cualquier otra fuente de información, tienen sesgos y limitaciones que se generan a lo largo de todo el proceso. Además, ante la falta de información o con el fin de simplificar la programación, se realizan una serie de asunciones o premisas que podrían afectar a los resultados y a la fiabilidad.

Existen diferentes cuestionarios con una serie de preguntas orientadas a su desarrollo que intentan dar credibilidad a los modelos, favorecer las prácticas en su diseño y ayudar a los evaluadores de las tecnologías sanitarias (2). Además, su uso puede ayudar a los decisores a establecer la confianza que transmiten los resultados del modelo.

Verificación

En la verificación es importante garantizar que no haya problemas en la programación y que se siguen las reglas de construcción de cada

modelo, por ejemplo, en la creación de estados de salud, la aplicación de cada parámetro, si se utilizan las fórmulas o ecuaciones correctas, etc. Además, se debe comprobar que cualquier modificación o variación de algún parámetro provoca un cambio coherente en los resultados del análisis (6,24).

Uno de los cuestionarios más utilizados para la verificación, sobre todo a la hora de enviar los trabajos a una publicación científica es el CHEERS (*Consolidated Health Economic Evaluation Reporting Standards*). Está compuesto por una serie de ítems (28 en total) relacionados con el diseño, la estructura y la metodología del modelo y de la evaluación económica, y unas recomendaciones sobre como describir o detallar cada uno de ellos y sobre como proporcionar la información necesaria para el análisis (26).

Validación

La validación de un modelo es un conjunto de métodos que evalúan su precisión cuando se efectúan cambios importantes en los parámetros que lo componen. Existen diferentes tipos de validación descritos en la siguiente tabla (Tabla 4) (2,5,23).

Tabla 4. Tipos de validación

Tipos	Descripción/Aplicación
Validez descriptiva	El modelo se construye en base a la mejor evidencia disponible
Validez interna (verificación)	El comportamiento o programación del modelo es correcto
Validez cruzada	Compara los resultados de un modelo con otro para comprobar que son similares utilizando diferentes métodos
Validez externa	Comprueba la capacidad de un modelo para estimar resultados en práctica real
Validez predictiva	Compara los resultados extrapolados del modelo con los resultados en práctica real, a futuro

Aunque se realicen validaciones, siempre existe incertidumbre en el modelo. Así, los análisis de sensibilidad ayudan a evaluar el impacto que tiene el cambio de un valor sobre los resultados, pero no sustituyen a la validación (23).



Evaluación de la incertidumbre inherente a los modelos

La incertidumbre en los modelos analíticos de decisión es un factor inherente que puede presentarse en la metodología, en los parámetros y en los resultados del análisis. Tal y como se describe en la Tabla 5, las fuentes de incertidumbre pueden ser clasificadas en distintos tipos (27).

Tabla 5. Clasificación de los tipos de incertidumbre en los modelos analíticos de decisión

Tipo de incertidumbre	Descripción	Otros términos análogos
Estocástica	Variabilidad aleatoria en los resultados para los mismos parámetros	Variabilidad / Error de Monte Carlo / Incertidumbre de primer orden
Paramétrica	Incertidumbre en la estimación del parámetro de interés	Incertidumbre de segundo orden
Heterogeneidad	Variabilidad entre pacientes que puede ser atribuida a sus características	Variabilidad
Estructural	Variabilidad causada por las asunciones tomadas para el desarrollo del modelo	Incertidumbre del modelo

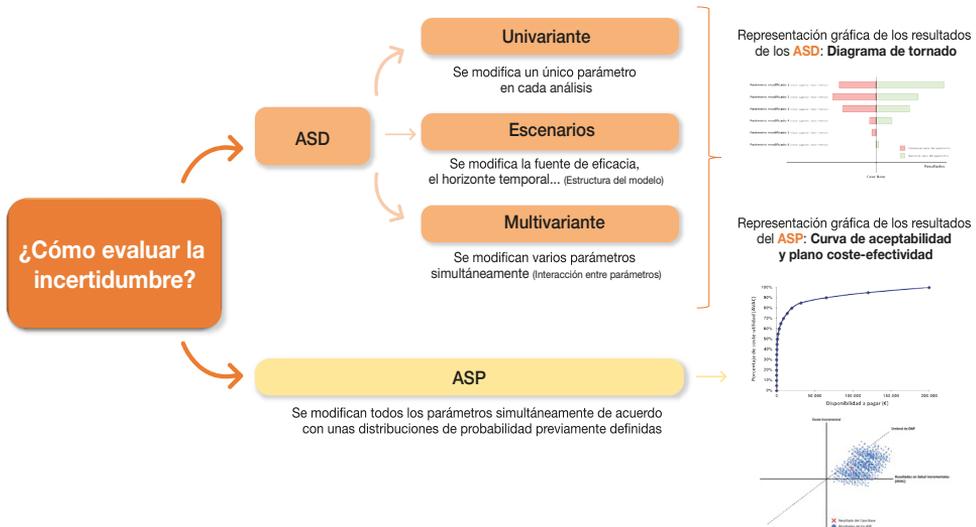
Fuente: Adaptado de (27).

Ante esta realidad, resulta esencial evaluar y declarar todas las fuentes de incertidumbre identificadas en una evaluación económica (27, 28). Existen herramientas que orientan sobre el manejo adecuado de esta incertidumbre, como la herramienta TRUST (*TRansparent Uncertainty ASsessment*), que permite identificar, evaluar y declarar las fuentes de incertidumbre inherentes en un modelo de forma sistemática (28).

Por otro lado, existen distintas técnicas para estudiar la incertidumbre en los estudios de modelización. Estas se denominan análisis de sensibilidad (Figura 5), y deben acompañar siempre a los resultados obtenidos (27).



Figura 6. Análisis de sensibilidad para evaluar la incertidumbre en torno a los resultados de un estudio de modelización



ASD: Análisis de Sensibilidad Determinístico; ASP: Análisis de Sensibilidad Probabilístico.

Fuente: Adaptada de (27).

Debido a las limitaciones que presentan los análisis de sensibilidad determinísticos (ASD) convencionales, diversos autores han propuesto nuevos enfoques para su realización. Una de las principales críticas que reciben estos análisis de sensibilidad es que las RCEI resultantes de distintos ASD no siguen una distribución normal. Por ello, se propone representar las RCEI resultantes en un plano coste-efectividad para visualizar como se distribuyen los resultados. Asimismo, también se puede presentar el beneficio neto incremental, tanto en términos monetarios como en beneficios en salud. Ambas alternativas solucionan los problemas matemáticos asociados a la RCEI (29).

Los ASD convencionales presentan otra limitación importante a la hora de determinar la tendencia con la que varían los resultados cuando se cambia un mismo parámetro por múltiples valores, que, por lo general, se asume lineal. Ante esta realidad, se han desarrollado los ASD por pasos, que ponen de manifiesto las variaciones no lineales de los resultados en diversos ASD con un mismo parámetro (29).

Por último, tradicionalmente se asume que un mismo parámetro puede tomar los distintos valores que conforman su rango de variabilidad con la misma probabilidad. Esto representa una importante carencia en los análisis convencionales, que trata de ser solventada mediante los ASD por percentiles (29).

Más allá de las aproximaciones descritas, una alternativa para abordar las limitaciones de los ASD consiste en realizar análisis de sensibilidad probabilísticos (ASP). En estos análisis se modifican todas las variables del modelo simultáneamente, tomando cada uno valores aleatorios de acuerdo con una distribución estadística específica. Este proceso se repite sucesivas veces, almacenando el resultado de cada iteración, de forma que el ASP otorga un resultado promedio y, si se desea, su medida de dispersión (27).

Los grandes esfuerzos orientados a desarrollar nuevos métodos con los que evaluar la incertidumbre ponen de manifiesto la importancia de realizar estos análisis, justificando adecuadamente su elección e interpretando sus resultados para describir las fuentes de incertidumbre existente en torno a un modelo y su impacto sobre los resultados obtenidos.



Presentación, interpretación y comunicación de los resultados

A la hora de llevar a cabo una evaluación económica basada en un modelo analítico de decisión, no solo es preciso seguir las pautas de buenas prácticas para el diseño y realización del estudio, sino también para informar sobre los resultados obtenidos. En este contexto, existen diversas guías de ámbito nacional (30) e internacional (31), que orientan sobre la estructura y aspectos a incluir en cada apartado de un informe de evaluación económica.

En la Tabla 5 se incluyen las pautas básicas que deben estar presentes en cualquier informe de resultados de una evaluación económica. Aunque la tabla se organiza en base las secciones de un artículo científico, se trata de una estructura aplicable a cualquier tipo de informe de evaluación económica.

Tabla 6. Pautas para informar sobre los resultados obtenidos en una evaluación económica a partir de un modelo analítico de decisión

¿Cómo presentar, comunicar e interpretar los resultados obtenidos en una evaluación económica realizada mediante un modelo analítico de decisión?

1. Introducción

- Relevancia clínica de la enfermedad:
 - Contexto epidemiológico.
 - Pronóstico y evolución de la enfermedad.
- Relevancia económica de la enfermedad.
- Descripción del arsenal terapéutico disponible para el tratamiento de la enfermedad evaluada:
 - Descripción de las alternativas terapéuticas a comparar y sus resultados.
 - Patrón de prescripción característico de la región donde se realiza la evaluación económica.
 - Evaluaciones económicas previamente realizadas para la alternativa terapéutica que se pretende evaluar.
- Objetivo del estudio y su aportación en la toma de decisiones.

Tabla 6 cont. Pautas para informar sobre los resultados obtenidos en una evaluación económica a partir de un modelo analítico de decisión

2. Materiales y métodos

- Descripción del modelo:
 - Tipo de modelo.
 - Estructura.
 - Especificación del *software* empleado para su desarrollo.
 - Estudios de calibración y validación (si procede).
- Descripción de la población diana.
- Especificación del horizonte temporal y tasa de descuento aplicada.
- Descripción de los resultados en salud:
 - Datos y fuentes (incluyendo valores de utilidad si procede).
 - Declaración de las aproximaciones y asunciones.
- Descripción de los costes:
 - Indicación de la perspectiva del análisis.
 - Listado del consumo de recursos sanitarios y sus fuentes.
 - Listado de costes unitarios y sus fuentes.
 - Declaración de las aproximaciones y asunciones.
- Descripción de los análisis de sensibilidad:
 - Justificación de la elección de los análisis realizados.
 - Información de los parámetros modificados incluyendo los nuevos valores para el ASD, o las distribuciones empleadas en el ASP.

3. Resultados

- Presentación de costes y resultados en salud (para todas las medidas de resultados en salud consideradas en los análisis) obtenidos en el caso base:
 - Desagregados para cada alternativa terapéutica incluida en la evaluación económica.
 - Totales para cada alternativa terapéutica incluida en la evaluación económica.
 - Incrementales (desagregados y totales) de la alternativa terapéutica evaluada frente a cada comparador.
- Comunicación de la ratio coste-efectividad incremental obtenido para cada medida de resultados considerada en el caso base.
- Comunicación del beneficio neto sanitario obtenido en el caso base.
- Descripción de los resultados obtenidos en los análisis de sensibilidad
- Inclusión de gráficos en caso de haber incluido ASP: plano coste-efectividad y curva de aceptabilidad.

Tabla 6 cont. Pautas para informar sobre los resultados obtenidos en una evaluación económica a partir de un modelo analítico de decisión

4. Discusión

- Comparación de los resultados con otros estudios previamente publicados para validarlos.
- Comparación de los resultados con opiniones de expertos (si procede).
- Discusión de las discrepancias y limitaciones identificadas en el modelo y evaluación económica realizada.
- Discusión de las fortalezas observadas en el modelo y evaluación económica realizada.
- Mención a la necesidad de diseñar futuros modelos o estudios alternativos.
- Manifestación de la utilidad de los resultados para la toma de decisiones en el contexto donde se realizó la evaluación económica.

5. Otras secciones

- Autoría: Nombre, filiación y contribución de los autores del modelo, así como de otros colaboradores que han tomado parte en la ejecución de la evaluación económica.
- Conflicto de intereses de los autores, fuente de financiación del estudio y detalle de la relación entre ellos, así como los contratos entre ambas partes.
- Referencias:
 - Inclusión de todas las referencias empleadas para el diseño y alimentación del modelo.
- Apéndices:
 - Especificación de las asunciones.
 - Indicación de las fuentes de datos no publicadas: *Data on File* u opiniones de expertos.
 - Listado de fórmulas empleadas para el diseño del modelo.
 - Evaluación de la calidad metodológica del modelo y del informe de resultados según *checklist* publicadas (Ej. CHEERS 2022).

ASD: Análisis de Sensibilidad Determinísticos; ASP: Análisis de Sensibilidad Probabilísticos;
CHEERS: *Consolidated Health Economic Evaluation Reporting Standards*.

Fuente: (32-35). Adaptado de (36).

Existen algunas iniciativas que han desarrollado listados de ítems a incluir en los informes de resultados de las evaluaciones económicas. Entre ellas, el Grupo de Buenas Prácticas de la ISPOR (*International Society of Pharmacoeconomics and Outcomes Research*) ha publicado diversas actualizaciones de la lista de verificación CHEERS (34,35), cuyo uso se encuentra cada vez más expandido a la hora de evaluar la calidad metodológica de los estudios económicos. Otras alternativas para evaluar esta calidad incluyen la herramienta CHEQUE (*Criteria for Health Economic Quality Evaluation*), que permite cuantificar la calidad metodológica mediante una puntuación basada en los 24 ítems a evaluar, entre los que destacan numerosos aspectos relacionados con el diseño de los modelos analíticos de decisión (37).

Los modelos pueden aportar información valiosa para la toma de decisiones. Sin embargo, es crucial cumplir con las pautas incluidas en las listas de verificación descritas previamente para garantizar la transparencia y reproducibilidad de estos. Solo así, una institución sanitaria podrá evaluar críticamente la evidencia económica de la que dispone para tomar sus decisiones, en base a los resultados obtenidos por medio de los modelos analíticos de decisión.



Bibliografía

1. Rodríguez-Barríos JM. Papel de los modelos en las evaluaciones económicas en el campo sanitario. *Farm Hosp.* 2004;28:231-42.
2. Caro JJ, Eddy DM, Kan H, et al. Questionnaire to Assess Relevance and Credibility of Modeling Studies for Informing Health Care Decision Making: An ISPOR-AMCP-NPC Good Practice Task Force Report. *Value Health.* 2014;17:174-82.
3. Tappenden P, Chilcott J, Brennan A, et al. Whole Disease Modeling to Inform Resource Allocation Decisions in Cancer: A Methodological Framework. *Value Health.* 2012;15:1127-36.
4. Roberts M, Russell LB, Paltiel AD, et al. Conceptualizing a Model: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-2. *Value Health.* 2012;15:804-11.
5. Caro JJ, Briggs AH, Siebert U, et al. Modeling Good Research Practices Overview: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-1. *Value Health.* 2012;15:796-803.
6. Siebert U, Alagoz O, Bayoumi AM, et al. State-Transition Modeling: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-3. *Value Health.* 2012;15:812-20.
7. Karnon J, Stahl J, Brennan A, et al. Modeling using Discrete Event Simulation: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-4. *Value Health.* 2012;15:821-7.
8. Pitman R, Fisman D, Zaric GS, et al. Dynamic Transmission Modeling: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-5. *Value Health.* 2012;15:828-34.
9. Woods BS, Sideris E, Palmer S, et al. Partitioned Survival and State Transition Models for Healthcare Decision Making in Oncology: Where Are We Now? *Value Health.* 2020;23:1613-21.
10. Brennan A, Chick SE, Davies R. A taxonomy of model structures for economic evaluation of health technologies. *Health Econ.* 2006;15:1295-310.

11. Mauskopf J, Standaert B, Connolly MP, et al. Economic Analysis of Vaccination Programs: An ISPOR Good Practices for Outcomes Research Task Force Report. *Value Health*. 2018;21:1133-49.
12. Revicki DA, Frank L. Pharmacoeconomic evaluation in the real world. Effectiveness versus efficacy studies. *PharmacoEconomics*. 1999;15:423-34.
13. Egger M, Smith GD, Altman D. *Systematic Reviews in Health Care: Meta-Analysis in Context*. 2nd Ed. London, BMJ Books; 2008.
14. Drummond MF, Sculpher MJ, Claxton K, et al. *Methods for the economic evaluation of health care programmes*. 4th Ed. New York Oxford University Press; 2015.
15. Gold MR, Stevenson D, Fryback DG. HALYS and QALYS and DALYS, Oh My: similarities and differences in summary measures of population Health. *Annu Rev Public Health*. 2002;23:115-34.
16. EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. *Health Policy*. 1990;16:199-208.
17. Soto Álvarez J. Tipos de costes a incluir en una evaluación económica. En: *Evaluación económica de medicamentos y tecnologías sanitarias: Principios, métodos y aplicaciones en política sanitaria*. 1^a Ed. Madrid: Springer Healthcare Iberica; 2012. p. 37-52.
18. Koopmanschap MA, Rutten FF. A practical guide for calculating indirect costs of disease. *PharmacoEconomics*. 1996;10:460-6.
19. Soto Álvarez J. Medición y evaluación de los resultados en salud en las evaluaciones económicas. En: *Evaluación económica de medicamentos y tecnologías sanitarias: Principios, métodos y aplicaciones en política sanitaria*. 1^a Ed. Madrid: Springer Healthcare Iberica; 2012. p. 53-62.
20. Bonneux L. How to measure the burden of mortality? *J Epidemiol Community Health*. 2002;56:128-31.
21. National Institute for Health and Care Excellence (NICE). *NICE health technology evaluations: the manual 2022*; <https://www.nice.org.uk/process/pmg36/resources/nice-health-technology-evaluations-the-manual-pdf-722867792> (acceso verificado el 17 Oct 2023)
22. Vallejo-Torres L, García-Lorenzo B, Rivero-Arias O, et al. Disposición a pagar de la sociedad española por un Año de Vida Ajustado por Calidad. Ca-

- narias: Servicio de Evaluación del Servicio Canario de Salud (SESCS); 2016; https://www3.gobiernodecanarias.org/sanidad/scs/content/c6b59111420f-11e7-952b-a987475f03d3/SESCS_2016_DAP_AVAC.pdf (acceso verificado el 18 Oct 2023)
23. Eddy DM, Hollingworth W, Caro JJ, et al. Model Transparency and Validation: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-7. *Value Health*. 2012;15:843-50.
 24. Dasbach EJ, Elbasha EH. Verification of Decision-Analytic Models for Health Economic Evaluations: An Overview. *PharmacoEconomics*. 2017;35:673-83.
 25. Philips Z, Bojke L, Sculpher M, et al. Good Practice Guidelines for Decision-Analytic Modelling in Health Technology Assessment: A Review and Consolidation of Quality Assessment. *PharmacoEconomics*. 2006;24:355-71.
 26. Augustovski F, García Martí S, Espinoza MA, et al. Estándares Consolidados de Reporte de Evaluaciones Económicas Sanitarias: adaptación al español de la lista de comprobación CHEERS 2022. *Value Health Reg Issues*. 2022;27:110-4.
 27. Briggs AH, Weinstein MC, Fenwick EAL, et al. Model Parameter Estimation and Uncertainty: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-6. *Value Health*. 2012;15:835-42.
 28. Grimm SE, Pouwels X, Ramaekers BLT, et al. Development and Validation of the TRansparent Uncertainty ASsessmentT (TRUST) Tool for Assessing Uncertainties in Health Economic Decision Models. *PharmacoEconomics*. 2020;38:205-16.
 29. Vreman RA, Geenen JW, Knies S, et al. The Application and Implications of Novel Deterministic Sensitivity Analysis Methods. *PharmacoEconomics*. 2021;39:1-17.
 30. Ortega Eslava A, Marín Gil R, Fraga Fuentes MD, et al. Guía de evaluación económica e impacto presupuestario en los informes de evaluación de medicamentos. Guía práctica asociada al programa MADRE v. 4.0; Madrid: Sociedad Española de Farmacia Hospitalaria (SEFH) (Ed), 2017. <http://grupodetrabajo.sefh.es/genesis> (acceso verificado el 18 Oct 2023)

31. Canadian Agency for Drugs and Technology in Health (CADTH) Ottawa 202. [https:// www.cadth.ca/](https://www.cadth.ca/) (acceso verificado el 18 de Oct 2023)
32. López Bastida J, Oliva J, Antoñanzas F, et al. A proposed guideline for economic evaluation of health technologies. *Gac Sanit.* 2010;24:154-70.
33. Sullivan SM, Wells G, Coyle D. What Guidance are Economists Given on How to Present Economic Evaluations for Policymakers? A Systematic Review. *Value Health.* 2015;18:915-24.
34. Husereau D, Drummond M, Petrou S, et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS)—Explanation and Elaboration: A Report of the ISPOR Health Economic Evaluation Publication Guidelines Good Reporting Practices Task Force. *Value Health.* 2013;16:231-50.
35. Husereau D, Drummond M, Augustovski F, et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS) 2022 Explanation and Elaboration: A Report of the ISPOR CHEERS II Good Practices Task Force. *Value Health.* 2022;25:10-31.
36. Soto Álvarez J. Diseño y realización de evaluaciones económicas a través de modelos analíticos de decisión. En: *Evaluación económica de medicamentos y tecnologías sanitarias: Principios, métodos y aplicaciones en política sanitaria.* 1ª ed. Madrid: Springer Healthcare Iberica; 2012. p. 171-216.
37. Kim DD, Do LA, Synnott PG, et al. Developing Criteria for Health Economic Quality Evaluation Tool. *Value Health.* 2023;26:1225-34.



2

RESULTADOS EN SALUD A INCORPORAR EN LOS MODELOS ANALÍTICOS DE DECISIÓN

Raúl Parra Garcés, Antia Gómez Fernández, Francisco Abad Santos

Servicio de Farmacología Clínica

Hospital Universitario de la Princesa, IIS-Princesa, Madrid

Facultad de Medicina, Universidad Autónoma de Madrid

Las evaluaciones económicas de medicamentos han cobrado importancia para lograr una utilización más racional de los recursos existentes. La necesidad de establecer prioridades en el gasto sanitario ha hecho que hayan alcanzado mayor difusión durante los últimos años y que actualmente se recomienden o exijan evaluaciones económicas de medicamentos para la toma de decisiones sobre su posible financiación pública, precio y utilización (1).

En cualquier evaluación económica, la correcta evaluación de los resultados en salud obtenidos tras la administración de las opciones terapéuticas a estudio es uno de los pilares fundamentales (junto con los costes generados) para poder obtener conclusiones válidas, creíbles y relevantes para los responsables de la toma de decisiones (2).



Estimación de la eficacia/efectividad y seguridad comparada de las alternativas evaluadas

Las decisiones de asignación relativas a la priorización de los recursos sanitarios entre las intervenciones que compiten entre sí implican la evaluación del impacto tanto en los costes como en los resultados sanitarios. Sin embargo, uno de los factores limitantes es que los estudios utilizan medidas diferentes de resultados en salud para demostrar el efecto de un tratamiento. Por ejemplo, un estudio puede informar sobre las tasas de supervivencia, mientras que otro puede centrarse en la incidencia de las úlceras por presión y los días sin dolor. Cuando nos enfrentamos a este tipo de medidas de resultados tan diversos que surgen de las posibles intervenciones, es difícil determinar hacia dónde deben dirigirse los recursos sanitarios de forma eficiente (3).

Durante el siglo xx, las estadísticas de mortalidad y supervivencia habían sido los mejores indicadores para valorar el estado de salud de la población y las intervenciones sanitarias, dado que las enfermedades, predominantemente infecciosas, producían una alta mortalidad. Los indicadores de morbilidad (incidencia y prevalencia de enfermedades) se desarrollaron posteriormente, como consecuencia de la aparición de los tratamientos y la mejora de las condiciones sanitarias, dando paso a las enfermedades crónicas y a un aumento de la esperanza de vida. Todo este proceso ha venido del interés de los individuos por la calidad de vida. Sin embargo, esta calidad de vida no se refleja con los indicadores clásicos de morbilidad y de mortalidad. Así, en los últimos años se han desarrollado nuevos indicadores que buscan medir la calidad de vida y el estado de salud (1).

Los resultados en salud obtenidos se van a poder expresar tanto en unidades clínicas derivadas de evitar morbimortalidad, como en unidades humanísticas derivadas del incremento de la calidad de vida y el nivel de satisfacción de los pacientes (3), siendo siempre necesario razonar la mejor forma en la que se miden estos resultados. En la Tabla 1 se pueden ver algunos ejemplos de variables de resultados utilizados en diferentes patologías.

Tabla 1. Ejemplos de variables de resultados para medir la eficacia del tratamiento en algunas patologías.

Enfermedad	Variables intermedias	Variables finales
Hipertensión arterial	Presión arterial	Mortalidad cardiovascular Infarto agudo de miocardio Ictus
Enfermedad renal crónica	Filtrado glomerular Hemoglobina	Calidad de vida Mortalidad
Diabetes mellitus	Hemoglobina glicosilada	Ingresos hospitalarios Complicaciones cardiovasculares
Dislipidemia	Perfil lipídico: colesterol total, LDL, HDL	Mortalidad cardiovascular
Osteoporosis	Densidad mineral ósea	Fracturas de cadera Fracturas vertebrales Calidad de vida
Enfermedad pulmonar obstructiva crónica	Función pulmonar: VEMS, capacidad vital	Ingreso por exacerbaciones Mortalidad Calidad de vida

VEMS: volumen espirado máximo en el primer segundo de la espiración forzada

Si los resultados se van a expresar en unidades clínicas (parámetros que los profesionales sanitarios utilizan diariamente para valorar la eficacia de los tratamientos administrados), éstas podrán derivarse de la disminución de la morbilidad, tales como: porcentaje de pacientes curados, complicaciones evitadas, eventos evitados (infartos de miocardio evitados, ictus evitados, etc.), tiempo libre de enfermedad, recurrencias evitadas, entre muchas otras.

Los resultados también se pueden medir a partir de la disminución de la mortalidad, en cuyo caso se emplean variables del tipo de año de vida ganado, muerte evitada, vida salvada, años de vida ajustados por calidad, supervivencia global, etc.

En algunos casos podemos utilizar una variable subrogada o intermedia (Tabla 1), que se define como «una medida de laboratorio o signo físico que se usa en ensayos terapéuticos como sustituto de una variable clínicamente significativa que es una medida directa sobre lo que siente un paciente, sus funciones o su supervivencia y que se espera que prediga el efecto de la terapia» (4). Aunque en ocasiones se emplean estas variables para valorar los resultados en salud en las evaluaciones económicas (por ejemplo: mmHg reducidos en la cifra de tensión arterial, disminución de mg de colesterol plasmático, número de lesiones ulcerosas en endoscopia, etc.), en la actualidad hay un amplio consenso en la necesidad de que siempre se empleen variables finales derivadas de la reducción de la morbilidad a la hora de valorar los resultados en salud obtenidos tras la aplicación de cualquier alternativa terapéutica. Esto es debido a que van a ser mucho más útiles a la hora de comparar adecuadamente las evaluaciones económicas para fijar prioridades en política de salud (5).

Independientemente de la unidad clínica elegida para valorar los resultados en salud, éstos se podrán expresar de diferente forma por lo que es preciso definir claramente los siguientes términos:

- a) **Eficacia:** se refiere a los efectos de los fármacos cuando se utilizan en las condiciones experimentales de los ensayos clínicos, donde, de forma protocolizada y con unos criterios de inclusión estrictos, un grupo de investigadores estudia una muestra pequeña y muy seleccionada de pacientes, utilizando normalmente variables de valoración intermedias, pero también variables finales en la fase III de desarrollo clínico.
- b) **Efectividad:** hace referencia a los efectos de los fármacos cuando se utilizan en las condiciones de la práctica clínica habitual, como sucede en los ensayos clínicos pragmáticos y los diseños observacionales. Normalmente, la efectividad es más baja que la eficacia porque hay más variabilidad en el tipo de pacientes, y el seguimiento y el cumplimiento con el tratamiento no son tan rigurosos.

c) **Eficiencia:** mide los efectos de los fármacos en relación con los recursos que debemos invertir para conseguirlos; mientras que los dos conceptos anteriores únicamente hacen referencia a los efectos de un fármaco, el término eficiencia incluye además su coste y siempre debe evaluarse en comparación a otra alternativa. En este caso, la pregunta a la que se intenta responder es si la utilización de un determinado fármaco compensa desde un punto de vista económico (6).

Siempre que sea posible, habrá que intentar emplear resultados en salud que reflejen la efectividad de los tratamientos administrados, ya que van a ser datos mucho más útiles para la toma de decisión en política de medicamentos y tecnologías sanitarias (7).

Resultados en salud medidos en unidades clínicas derivadas de la disminución de la morbilidad y de la mortalidad

La valoración de los resultados en salud derivados de una reducción de la morbilidad o de la mortalidad de la enfermedad diana va a ser diferente si se trata de variables categóricas o variables cuantitativas. En el caso de **variables categóricas** (por ejemplo: aparición de un evento como puede ser la muerte o la curación), la medida del efecto de un tratamiento administrado se efectúa empleando diferentes técnicas analíticas, tales como el riesgo absoluto (RA) y el riesgo relativo (RR), la reducción del riesgo absoluto (RRA), la reducción del riesgo relativo (RRR), el *odds ratio* (OR) y el número necesario a tratar (NNT) (5,8).

El **riesgo absoluto (RA)** de cada opción terapéutica evaluada representa la proporción de pacientes tratados en los que aparece el resultado valorado (habitualmente, un evento), por lo que valores más bajos reflejan los mejores resultados. La **reducción del riesgo absoluto (RRA)** es la diferencia entre el riesgo absoluto del grupo en tratamiento con la

nueva opción a valorar y el riesgo absoluto del grupo control y, cuanto más elevado sea, mejores resultados presenta la nueva alternativa terapéutica.

El **riesgo relativo (RR)** es el cociente entre el riesgo del grupo en tratamiento con la nueva opción terapéutica y el grupo en tratamiento con el control y, habitualmente, se expresa como el cociente entre el riesgo del tratamiento que presenta un mayor riesgo y el que muestra uno menor (9). La **reducción del riesgo relativo (RRR)**, también llamado fracción atribuible (FA), es el cociente entre la reducción del riesgo absoluto de ambos grupos de tratamiento (nueva opción en estudio y control) y el riesgo absoluto en el grupo control, y nos va a proporcionar información del beneficio de la nueva opción terapéutica en términos relativos. Su valor siempre va a ser mucho más elevado que la RRA, por lo que suele emplearse mucho más a la hora de comunicar los resultados de los estudios de investigación clínica, tanto en los informes finales como en las publicaciones. Una de las limitaciones de utilizar el valor de la RRR en lugar del valor de la RRA es que al eliminar la magnitud del riesgo sin tratamiento (riesgo basal), se puede sobreestimar o infraestimar el impacto de la nueva alternativa, si el riesgo sin tratamiento fuera muy bajo o muy alto, respectivamente (10).

El **odds ratio (OR)** (a veces traducido como razón de ventajas) es una forma de calcular el riesgo relativo en los estudios de casos y controles donde no conocemos la incidencia; representa el cociente entre el *odds* del grupo donde se emplea la nueva opción a evaluar y el *odds* del grupo control, es decir, el cociente entre la probabilidad de que el resultado evaluado ocurra y la probabilidad de que no ocurra y, habitualmente, se estima por el cociente entre el número de veces que ha ocurrido el resultado evaluado y el número de veces que no ha ocurrido. Si no hay efecto de los tratamientos administrados, el OR es 1, mientras que cuanto mayor sea el efecto, el valor de OR más se aleja de 1 (en positivo o en negativo) (5).

El **número necesario a tratar (NNT)** representa el número de pacientes que deberían recibir la nueva opción terapéutica para lograr evitar el resultado evaluado en un paciente, cuando se compara con el grupo control, y se calcula como el inverso de la RRA. Tanto el OR como el NNT representan el tamaño del efecto cuando la variable de respuesta es categórica. Por ejemplo, en un estudio donde se evaluase el efecto protector de un medicamento anti-hipertensivo en la aparición de ictus, un NNT de 15 significaría que sería necesario tratar a 15 pacientes con el medicamento para evitar la aparición de un ictus. Cuanto mayor sea el tamaño del efecto del tratamiento administrado, menor será el NNT. Si no se encontrasen diferencias en la eficacia entre los grupos en evaluación, la RRA sería cero y el NNT infinito (11).

En la actualidad, se recomienda que, a la hora de notificar los resultados de los estudios clínicos, se comuniquen siempre empleando una medida relativa y una absoluta, y se aporten los intervalos de confianza de los estándares utilizados (12).

En el caso de **variables cuantitativas**, la medida del efecto de un tratamiento administrado se realiza a través del cálculo del tamaño del efecto para dichas variables cuantitativas (por ejemplo, reducción del peso). Este tamaño del efecto mide la magnitud derivada de utilizar el tratamiento A o el tratamiento B en dos grupos diferentes de pacientes (11).

El más común de los estadísticos empleados para calcular el tamaño del efecto con variables cuantitativas es la **d de Cohen** (a veces, también denominada diferencia de medias estandarizada), parámetro designado para la situación en la que las respuestas de los tratamientos A y B presentan distribuciones normales con desviaciones estándar equivalentes. El cálculo mediante la d de Cohen se realiza a partir de la siguiente fórmula: $d = (M_1 - M_2) / DE_{ponderada}$; donde $(M_1 - M_2)$ es la diferencia de las medias de los grupos comparados y $DE_{ponderada}$ es la desviación estándar ponderada. Nos informa de cuántas desviaciones típicas de diferencia hay entre los resultados de los dos grupos que se comparan (grupo experimental y grupo de control) (13).

Los resultados que se obtienen a partir de variables de mortalidad son los que más pueden ayudar a los decisores sanitarios a la hora de establecer prioridades y de asignar los recursos existentes de manera apropiada. Por esta razón, en las evaluaciones económicas se recomienda, siempre que se pueda, emplear resultados que reflejen la disminución en mortalidad. A continuación, se explican las diferentes formas de analizarla.

El **número de muertes evitadas (*deaths avoided*) o de vidas salvadas (*lives saved*)**, pese a ser un parámetro de interés clínico, en la evaluación económica raramente se utiliza como resultado, ya que es un parámetro demasiado agregado y su utilización no facilita la toma de decisiones (13).

Los **años de vida ganados (*life-years gained*)** es uno de los parámetros más utilizados en los análisis coste-efectividad como medida de los resultados alcanzados por las opciones terapéuticas a estudio. Para su cálculo, va a ser necesario comparar la edad a la que los pacientes fallecen en los grupos evaluados y calcular los años de vida ganados en cada uno de ellos a partir de: a) las tablas de mortalidad en el caso de estudios prospectivos en los que se proyecta la supervivencia de los pacientes durante toda su expectativa de vida; o bien, b) simular los resultados que aparecerán en los grupos en estudio a lo largo de toda su expectativa de vida, a través de modelos de simulación efectuados con la información obtenida en los estudios prospectivos (14).

Los **años potenciales de vida perdidos (APVP) (*potential years of life lost*)** es un indicador de la mortalidad prematura, basado en una esperanza de vida predeterminada, donde se otorga una mayor importancia a las causas de muerte que ocurren a edades tempranas de la vida que a las que suceden en edades más avanzadas. Refleja el número de años que teóricamente una persona deja de vivir si la muerte se presenta de manera prematura. Este indicador se utiliza para identificar las principales causas de mortalidad prematura, comparar los resultados

obtenidos por edad, sexo y raza, y establecer las estrategias de prevención (15,16).

Teniendo en cuenta que es posible padecer enfermedades con una elevada morbilidad y una baja mortalidad y viceversa, es útil y recomendable considerar ambos factores en combinación a la hora de evaluar los resultados de distintas alternativas terapéuticas. Como resultado de esta combinación, van a existir parámetros o medidas que evalúen expectativa de vida ajustada por el nivel de salud, o discapacidad, que se explican a continuación.

La **esperanza de vida ajustada por el nivel de salud (*health-adjusted life expectancy*)** es un indicador que ajusta la expectativa de vida de los sujetos acorde a su nivel y estado de salud y, por lo tanto, incorpora mortalidad y estado de salud en un único parámetro. Representa el número equivalente de años en estado de salud ideal que un sujeto puede esperar vivir durante su vida si estuviese expuesto cada año a las condiciones más predominantes de morbilidad (que empeoren el estado de salud) existentes en ese momento. Para poder cuantificar este parámetro en una población definida, se requieren dos datos: a) disponer de la tasa de mortalidad ajustada por la edad, y b) tener una medición del estado de salud ajustado por la edad (17).

La **esperanza de vida libre de discapacidad (*disability-free life expectancy*)** es una medida intermedia de la esperanza de vida ajustada al nivel de salud explicada previamente, y difieren en la manera en que los años de vida vividos son manejados. Este indicador define un umbral acorde al estatus de discapacidad, de tal manera que los años de vida vividos con un estado de salud por encima de este valor umbral son contados totalmente, y los que están por debajo de este valor umbral no son tenidos en cuenta (18).

El **año de vida ajustado por el nivel de salud (*health-adjusted life years o HALY*)** es una medida de la salud de la población que permite combinar el impacto de la mortalidad y la morbilidad de manera simultánea, y que permite hacer comparaciones entre diferentes enfermedades.

A la hora de efectuar su cálculo, es necesario seguir tres pasos: a) describir el nivel de salud (como un estado de salud o enfermedad); b) desarrollar valores para corregir ese estado, evaluando la calidad de vida (valorada a través de una variable denominada utilidad), y c) combinar los valores de diferentes estados de salud con la estimación de la esperanza de vida (3).

Sin embargo, los indicadores más empleados en la actualidad son los años de vida ajustados por discapacidad y los años de vida ajustados por calidad, que se revisan en detalle a continuación.

Los años de vida ajustados por discapacidad (AVAD) (*disability-adjusted life years o DALY*) combinan los años perdidos por una enfermedad y los años vividos con una discapacidad, y valoran la pérdida derivada de vivir una vida corta con una discapacidad. Los AVAD se calculan combinando cuatro elementos: a) la expectativa de vida, b) el valor de la vida a diferentes edades, c) el valor de tiempo futuro, y d) el valor de evitar la discapacidad. Los AVAD son la suma de los años de vida perdidos y los años de vida vividos con discapacidad y cómo impactan los diferentes tratamientos a la hora de reducir la aparición de los mismos. De esta manera, los AVAD representan la suma de los valores presentes en los años futuros perdidos debido a una mortalidad prematura y el valor presente de los años futuros de vida ajustados por la gravedad media (frecuencia e intensidad) de cualquier discapacidad física o mental causada por una enfermedad (18).

Por lo tanto, los AVAD son una medida de pérdida, y lo que se busca con los tratamientos administrados es evitar que aparezcan AVAD, de tal manera que cuando se emplea este indicador en los estudios coste-efectividad el resultado se expresa como el coste/AVAD evitado (19). Cuando se emplea este indicador en una evaluación económica lo que se hace es calcular los AVAD producidos en dos escenarios diferentes: cuando se trata a los pacientes con la opción terapéutica a evaluar, y cuando ésta no se utiliza. En este cálculo van a

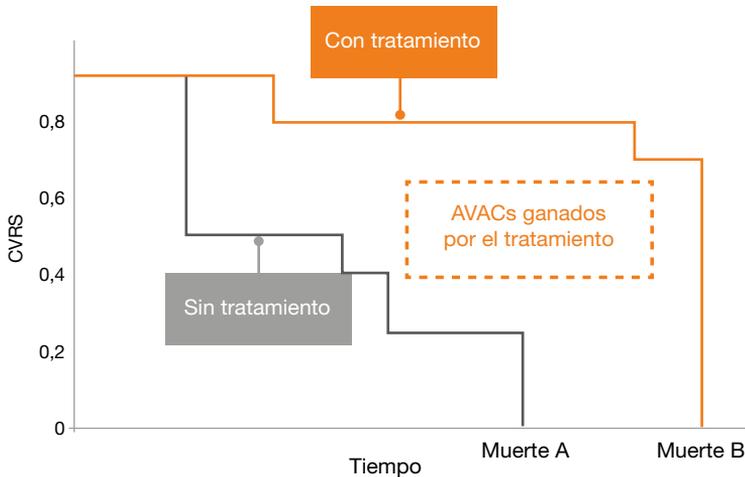
intervenir distintos factores, tales como: la edad de muerte de los pacientes, la esperanza de vida de los pacientes en el momento de su muerte, una tasa de descuento, un peso de ajuste a la edad, un peso de ajuste de la discapacidad, la edad del comienzo de la enfermedad y la duración de la discapacidad. La estimación de estos factores no es tarea fácil, por lo que su uso como indicador en evaluaciones económicas ha sido tradicionalmente escaso (20). En la actualidad, tampoco se recomienda su empleo a la hora de asignar prioridades en política sanitaria, ya que se ha visto que no produce una medida útil de la salud de la población (21).

Los **años de vida ajustados por calidad (AVAC) (*quality-adjusted life year o QALY*)** es una medida elaborada para evaluar los efectos de una alternativa terapéutica en términos de expectativa de vida (años de vida ganados) y de calidad de vida medida como utilidad. La utilidad refleja la preferencia del paciente por el estado de salud que presenta y adopta un valor entre cero y uno, siendo cero el peor estado de salud posible, y uno el mejor. Este indicador es el empleado en los análisis coste-utilidad (3).

Valoración de los años de vida: el concepto de AVAC

Los AVAC permiten combinar los efectos de las intervenciones sanitarias sobre la mortalidad y la morbilidad en un único índice (3), proporcionando así una medida común que permite realizar comparaciones entre diferentes áreas de enfermedad. A lo largo del tiempo, los individuos experimentan diferentes estados de salud, que se ponderan en función de las puntuaciones de utilidad asociadas a ellos. El concepto de AVAC, que combina la supervivencia de un individuo con su calidad de vida relacionada con la salud (CVRS), puede verse en la Figura 1.

Figura 1. AVAC ganados por tratamiento



Se observa que el sujeto B (con tratamiento) tiene mayor supervivencia (eje tiempo) y con mejor calidad de vida relacionada con la salud (eje CVRS) que el sujeto A (sin tratamiento).

La Figura 1 muestra los AVAC que puede ganar un individuo si recibe un tratamiento en lugar de no recibirlo. El área bajo la curva equivale al valor total de AVAC. La línea inferior muestra el perfil de salud si no se recibe ningún tratamiento: la CVRS del individuo se reduce con el tiempo, hasta que muere (Muerte A). Sin embargo, si se recibe un tratamiento, el individuo sigue la línea superior: su CVRS se mantiene en un nivel más alto durante más tiempo, además de vivir más tiempo (Muerte B). Por tanto, la diferencia en el área total entre las dos curvas indica el número de AVAC ganados por el uso del tratamiento (22).

Los AVAC se calculan multiplicando la duración del tiempo transcurrido en un estado de salud por el peso de la CVRS (es decir, la puntuación de utilidad) asociada a ese estado de salud. Por lo tanto, se incorporan los dos elementos clave: la CVRS y la supervivencia.

Así, el enfoque de los AVAC convencionales, incluiría los siguientes supuestos:

- a) La salud se define como el tiempo ponderado por el valor (AVAC) en el horizonte temporal correspondiente.
- b) El valor se mide en términos de preferencia (deseabilidad).
- c) Las preferencias medidas entre individuos pueden agregarse y utilizarse para el grupo.
- d) Los AVAC pueden agregarse entre individuos, es decir, un AVAC es un AVAC independientemente de quién lo gane o lo pierda (23).

La mayoría de las guías existentes para realizar evaluaciones económicas aconsejan efectuar un ajuste temporal, tanto a los costes como a los resultados en salud futuros cuando el horizonte temporal es mayor de un año (24,25). Los AVAC que se producen en el futuro se descuentan a valores actuales, para incorporar la idea de que las personas prefieren recibir los beneficios de la salud en el presente (es decir, preferencia temporal positiva). Sin embargo, el descuento sigue siendo objeto de controversia; por ejemplo, se sigue debatiendo si los AVAC deben descontarse a una tasa inferior a la de los costes. Además, diferentes personas pueden tener diferentes preferencias temporales para los resultados de salud (22).

Alternativas a los AVAC

El **equivalente de año sano (HYE) (*healthy year equivalent*)** es una medida de calidad de vida que se basa en un procedimiento de dos etapas que utiliza un cuestionario para obtener preferencias. El HYE ha sido criticado por la dificultad de su aplicación, porque requiere que los encuestados evalúen un gran número de posibles combinaciones de estados de salud y duraciones de estados de salud (26).

La **disposición a pagar (DAP) (*willingness to pay*)** utilizado en un marco de análisis coste-beneficio consiste en obtener valoraciones de los beneficios para la salud en términos monetarios preguntando a los individuos cuánto estarían dispuestos a pagar para obtener o evitar los efectos sobre la salud. La principal crítica al marco coste-beneficio es que tiene un fundamento individualista porque se basa en la obtención de la DAP del individuo por el beneficio de la salud. Dado que la DAP está estrechamente asociada a la capacidad de pago, una valoración del estado de salud basada en la DAP perjudicará sistemáticamente a las personas con menores ingresos al vincular directamente los efectos sobre la salud a los recursos económicos de una persona (23,27).

Métodos existentes para obtener los resultados en salud

La mayoría de los instrumentos modernos de medición se basan en la evaluación de la salud percibida desde la perspectiva de los individuos. La utilización de estos indicadores permite conocer el efecto de los tratamientos o de las intervenciones sanitarias en el bienestar de los pacientes, conocido como la calidad de vida relacionada con la salud (CVRS) (9).

El estudio de la CVRS permite: a) medir la eficacia y efectividad de las intervenciones médicas; b) mejorar las decisiones clínicas; c) evaluar la calidad de la asistencia sanitaria; d) estimar las necesidades de la población; y e) conocer las causas y consecuencias de las diferencias en el estado de salud entre individuos o grupos de individuos (9).

Conceptualización de la calidad de vida relacionada con la salud

La CVRS ha sido definida como el valor asignado a la duración de la vida que se modifica por la deficiencia, el estado funcional, la percepción y la oportunidad social debidas a una enfermedad, accidente, tratamiento o decisión política.

Instrumentos de medida

No existe un único instrumento de medida de la CVRS que pueda ser aplicado a todas las áreas de estudio. Los instrumentos de medida suelen ser cuestionarios que contienen las instrucciones para su cumplimiento. Las dimensiones de la CVRS a evaluar (dolor, movilidad, etc) incluyen un número determinado de ítems que describen las categorías jerárquicas o no de cada una de esas dimensiones (no tengo dolor, tengo moderado dolor, tengo mucho dolor). Además, los cuestionarios más completos tienen un sistema de puntuación basado en una escala de intervalo o de razón que dota de sensibilidad al instrumento.

Estos instrumentos deben cumplir con una serie de propiedades psicométricas: a) validez (el instrumento mide aquello que pretende medir); b) fiabilidad (el instrumento produce los mismos resultados en mediciones repetidas bajo las mismas condiciones); y c) sensibilidad a los cambios (el instrumento produce diferentes resultados en mediciones repetidas si las condiciones cambian). Esta última propiedad es esencial para incorporar la medida de la CVRS en ensayos clínicos con medicamentos (10).

A continuación, se destacan algunos instrumentos de medición: los instrumentos de medición de la valoración funcional, los instrumentos de medición de la salud mental, los perfiles de salud, las medidas de utilidad y los instrumentos específicos.

Índices de valoración funcional

Están basados en la repercusión funcional de la enfermedad en el individuo y se centran en la medida de la incapacidad, particularmente en relación con la valoración de enfermedades crónicas, rellenas por un profesional sanitario. Algunos ejemplos:

- a) Índice de Katz, mide las actividades básicas e instrumentales de la vida diaria (28).

- b) Criterios de la *New York Heart Association*, valora el grado funcional de los pacientes cardiopatas (29).
- c) Índice de Karnofsky, valora el grado funcional y el pronóstico de enfermos de cáncer y de los candidatos a trasplante (30).

Índices de salud mental

Estos instrumentos se utilizan en la práctica clínica psiquiátrica, aunque son aplicables a pacientes en otros ámbitos (10). Algunos ejemplos:

- a) Escala de Depresión de Hamilton para detectar depresión psiquiátrica (31).
- b) Cuestionario de Salud General de Goldberg que detecta potenciales casos psiquiátricos en la población general (32).

Perfiles de salud

Son aplicables a una amplia variedad de poblaciones y cubren un amplio espectro de dimensiones de la calidad de vida (física, mental y social). Proporcionan el perfil de salud de un individuo o de una población, según las diferentes dimensiones incluidas en el perfil. Sus aplicaciones principales son la evaluación de tratamientos médicos (comparando los resultados del perfil antes y después del tratamiento), como medida de comparación entre grupos (enfermos crónicos y estado de salud de la población general), como fuente de información complementaria del paciente en la práctica clínica y como instrumento de ayuda en la toma de decisiones en política sanitaria (10). Algunos ejemplos:

- a) Perfil de las Consecuencias de la Enfermedad (*Sickness Impact Profile*): mide la disfunción provocada por la enfermedad en doce actividades de la vida diaria o categorías, y está basado en los

cambios en la conducta de los individuos. Su aplicación se centra exclusivamente en la población enferma (33).

- b) Perfil de Salud de Nottingham (*Nottingham Health Profile*): puede ser aplicado tanto en estudios comunitarios como clínicos (34).
- c) Cuestionario de Salud SF-36: contiene 36 ítems y está especialmente pensado como perfil genérico para incorporar a los ensayos clínicos con medicamentos (35). Ofrece una perspectiva general del estado de salud de la persona con la ventaja de que es fácil y rápido de rellenar, a la vez que también es sencillo de evaluar.

Los perfiles de salud ofrecen algunas ventajas en la evaluación de medicamentos. Por ejemplo, permiten determinar los efectos del tratamiento en diferentes aspectos de la CVRS sin necesidad de utilizar múltiples instrumentos; también pueden utilizarse para comparar diferentes tratamientos en diferentes enfermedades (10).

La principal limitación es que el contenido de los perfiles puede no centrarse en aquellos aspectos de la CVRS más importantes de la enfermedad que se estudia y, por tanto, puede ser un instrumento poco sensible a los cambios antes y después del tratamiento (10).

Medidas de utilidad

Son índices que se basan en las preferencias o utilidades que los individuos asignan a diferentes estados de salud. Las utilidades se obtienen por medio de diferentes técnicas. Normalmente, la escala de medida se valora desde 0 (peor estado de salud imaginable, a veces muerte) a 1 (mejor estado de salud imaginable). Algunos ejemplos:

- a) Matriz de Rosser y Kind: formada por dos dimensiones, con 8 niveles de discapacidad y cuatro niveles de sufrimiento/dolor, que forman un conjunto de 29 estados de salud. Cada combinación

de estados obtiene una puntuación que se agrega en un índice global (36).

- b) Escala de Bienestar (*Quality of Well-Being Scale*): es un indicador de discapacidad y necesidad de atención sanitaria que puede ser aplicado a individuos o poblaciones con cualquier enfermedad. Combina en un solo índice cuatro atributos: a) movilidad; b) actividad física; c) actividad mental; y d) síntomas/problemas complejos. Se compone de 43 niveles de función y 36 síntomas complejos. Cada estado de salud tiene un valor que va de 0 (muerte) a 1 (bienestar completo) que representa la utilidad asignada a un atributo de salud. Este índice, junto con los años de vida, forma el índice de años de bienestar debidos a un tratamiento o tecnología sanitaria (37).
- c) EuroQol: índice genérico que se compone de dos partes. La primera es una descripción del estado de salud en cinco dimensiones: a) movilidad; b) cuidado personal; c) actividades cotidianas; d) dolor/malestar; y e) ansiedad/depresión. Cada una de las dimensiones tiene tres ítems que definen tres niveles de gravedad. En total se pueden formar 243 estados de salud distintos. En la segunda parte, el individuo puntúa su estado de salud en una escala visual analógica que va de 0 (peor estado de salud imaginable) a 100 (mejor estado de salud imaginable) (38).

Instrumentos específicos

Incluyen aspectos importantes de una determinada enfermedad (por ejemplo, artritis reumatoide), de una población (por ejemplo, ancianos), o un determinado síntoma clínico (por ejemplo, dolor) (10).

Tienen la ventaja de presentar una alta sensibilidad ante el problema específico de salud que se está evaluando. Su mayor desventaja es que no permiten comparaciones entre diferentes condiciones, y por lo tanto no son útiles para conocer la eficiencia relativa de diferentes intervenciones sanitarias.



Identificación de las fuentes disponibles y evaluación de su calidad metodológica

En cualquier evaluación económica es esencial disponer de resultados válidos, precisos y de calidad. Para llegar a conocer los resultados derivados del uso de las distintas opciones en evaluación, existen diferentes métodos (todos ellos complementarios entre sí), cada uno con sus ventajas e inconvenientes.

Los **ensayos clínicos** se consideran como la metodología más válida y con menos sesgos a la hora de disponer de los datos de resultados (eficacia), aunque también presentan sus limitaciones e inconvenientes (39).

Los **diseños observacionales** presentan la gran ventaja de su mayor validez externa (efectividad), aunque los resultados obtenidos con este método están sujetos a una menor validez interna y precisión, ya que el control de posibles sesgos y factores de confusión es complicado. Aún con estas limitaciones, en muchas ocasiones los resultados de este tipo de estudios resultan de gran valor y son totalmente complementarios a los resultados de los ensayos clínicos (40).

Con el fin de poder evaluar diferentes ensayos clínicos y estudios observacionales publicados en la literatura, y poder obtener los resultados de todos ellos de manera agregada, cada vez se realizan más revisiones sistemáticas de la literatura y meta-análisis (5,41). Otro de los métodos empleados para conocer los resultados de distintas intervenciones sanitarias son los modelos analíticos de decisión, diseños que permiten poder extrapolar los resultados obtenidos en estudios clínicos prospectivos a horizontes temporales más extensos que el período de seguimiento original, datos esenciales a la hora de asignar los recursos disponibles y fijar prioridades en política de medicamentos y tecnologías sanitarias (41).

Evaluaciones económicas dentro de ensayos clínicos

Las principales ventajas de las evaluaciones económicas incorporadas a ensayos clínicos, habitualmente durante la fase III, derivan del rigor científico del propio método, que permite obtener resultados con una alta validez interna, y de la posibilidad de generar información muy importante para la toma de decisiones en el momento de la comercialización; por ejemplo, sobre la financiación pública del fármaco, o sobre su inclusión en guías terapéuticas (42-43).

El principal problema de las evaluaciones económicas a partir de los ensayos clínicos se refiere a la escasa validez externa que frecuentemente tienen estos estudios. Tanto los efectos farmacológicos como los costes se recogen en condiciones experimentales, condiciones que pueden ser muy diferentes a las que luego tendrán lugar en la práctica clínica diaria. Ni los médicos ni los pacientes que se incluyen son siempre representativos de los que luego prescribirán y recibirán los tratamientos, respectivamente (44-45).

El número de pacientes y el periodo de seguimiento de los ensayos clínicos suelen ser muy limitados, por lo que resulta necesario establecer variables de valoración subrogadas que pueden no correlacionarse completamente con las variables finales (46-47).

Evaluaciones económicas a partir de estudios observacionales

Desde el punto de vista metodológico, los resultados de los estudios observacionales son muy representativos de lo que sucede a la mayoría de los pacientes, pero a expensas de una gran pérdida de validez interna, debido a la ausencia de asignación aleatoria, que pueden dar lugar a sesgos de selección por asignación a tratamiento (48).

Existen técnicas estadísticas que sirven para minimizar este problema, aunque pueden existir factores desconocidos para el investigador que mantengan el sesgo. Precisamente el sesgo de asignación al

tratamiento es una de las limitaciones más importantes de los estudios realizados a partir de bases de datos de práctica clínica. Otros problemas que habrá que tener en cuenta al utilizar estas bases de datos son la posible baja calidad de la información, la ausencia de datos sobre los efectos de una determinada terapia, y la presencia de datos incompletos, mal codificados o erróneos (5,49).

A pesar de sus problemas, muchas veces los estudios realizados mediante bases de datos han sido el único abordaje metodológico posible para aproximarse al estudio de los resultados en salud y obtener información imposible de conseguir mediante ensayos clínicos (49).

Modelos económicos de análisis de decisión

Puesto que los ensayos clínicos no suelen generar toda la información necesaria que se precisa en una evaluación económica, es necesario recurrir a otras fuentes como registros hospitalarios, opiniones de expertos, revisión de historias clínicas, estudios epidemiológicos, etc. Generalmente, en los modelos de análisis de decisión, se parte de los resultados de algún ensayo clínico o meta-análisis que luego se completan con información procedente de las otras fuentes indicadas (50).

La principal ventaja de los modelos económicos es que permiten combinar distintos tipos de datos para dar lugar, de forma relativamente rápida y barata, a una información de suma utilidad para la toma de decisiones. Con los modelos se intenta simular todas las posibilidades que pueden ocurrir durante un periodo de tiempo determinado, tras el inicio de los tratamientos comparados. Otras ventajas de los modelos de análisis de decisión respecto a las evaluaciones realizadas en ensayos clínicos son el menor coste y la rapidez con la que puede disponerse de información, aunque quizás la mayor ventaja es que, debido a su mayor validez externa, sus conclusiones pueden tener mayor importancia para la toma de decisiones de política sanitaria (51).

A pesar de estas ventajas, los estudios basados en modelos farmacoeconómicos suelen generar cierto rechazo, y es frecuente que sus resultados sean examinados con cierta desconfianza. La necesidad de recurrir a suposiciones y a opiniones de expertos para algunos parámetros es otro de los inconvenientes que argumentan los detractores de los modelos. Con frecuencia, los resultados han sido muy sensibles a algunas de las suposiciones empleadas y la opinión de los grupos de expertos puede reflejar la opinión de determinados especialistas cuya experiencia puede no ser representativa de la práctica médica habitual (5).

Comparaciones indirectas, meta-análisis en red y comparaciones indirectas ajustadas por emparejamiento

Los métodos estándar para las **comparaciones indirectas y el meta-análisis en red** se basan en datos agregados, partiendo del supuesto de que no hay diferencias entre los ensayos en la distribución de las variables modificadoras del efecto (52).

Los métodos que forman «comparaciones indirectas ajustadas a la población» son cada vez más comunes para algunas agencias como el Instituto Nacional de la Salud y la Excelencia Asistencial (NICE).

La **comparación indirecta ajustada con emparejamiento** es una metodología desarrollada para la evaluación de nuevos tratamientos frente a sus alternativas cuando no se dispone de comparación directa mediante un ensayo clínico aleatorizado y controlado. Estas comparaciones son de especial interés en enfermedades graves como el área de la hemato-oncología, en la que la incertidumbre en la toma de decisiones sobre la inclusión de nuevos fármacos se ve frecuentemente acentuada tanto por la gravedad de la enfermedad como por el elevado coste del tratamiento. La incertidumbre está particularmente acentuada cuando se trata de indicaciones terapéuticas con tratamientos nuevos (52).

Lo ideal sería disponer de datos de pacientes individuales (DPI) de todos los estudios para ajustar completamente las diferencias de los pacientes usando una red de meta-regresión. Sin embargo, un escenario muy común es cuando una empresa tiene DPI de su propio ensayo, pero sólo datos públicos agregados de su competidor. Los métodos de ajuste de la población utilizan los DPI disponibles para ajustar los desequilibrios de la distribución de las covariables observadas (53-54). En estos casos, utilizando métodos de regresión o métodos de emparejamiento mediante índices de propensión (*propensity score*) permiten realizar la comparación, mitigándose los sesgos derivados de las diferencias entre las poblaciones de ensayos.

Denominamos **comparaciones indirectas naïve** (no ajustadas) a aquellas en las que la evaluación se realiza comparando directamente los brazos de cada alternativa en estudios separados como si se tratara de brazos de un mismo estudio sin tener en consideración para su realización los brazos de control con los que se comparan ambas alternativas (55). Frente a estas, las **comparaciones indirectas ajustadas** son aquellas en las que, a partir de dos estudios que valoran por separado las eficacias de dos tratamientos de interés frente a un comparador común, se puede comparar de manera indirecta la eficacia de éstos. Las comparaciones indirectas asumen que existe similitud clínica y metodológica entre los estudios evaluados. Es decir, que las características basales de los pacientes, los seguimientos y la medición de resultados han sido similares en los estudios sobre los que se realiza la comparación. En ocasiones, la elección de diferentes medidas del efecto relativo (riesgo relativo, *odds ratio*, diferencia de riesgos) puede sesgar los resultados y llevar a conclusiones equivocadas sobre la efectividad comparada (56).

Este método nos permite reducir la sensibilidad a las distintas medidas de efecto y, adicionalmente, hace posible resolver las diferencias en la definición de variables de resultado de los estudios y la comparación de las diferentes dosis con relevancia clínica empleadas en los ensayos clínicos (57-59).

Información proveniente de expertos

Cuando no hay datos disponibles de estudios previos para algunos parámetros de resultados en salud, las opiniones de expertos son una buena fuente de información. Para ello, es de vital importancia que queden claramente documentados los métodos utilizados. Algunos expertos sugieren que se debe definir un consenso antes de ver las opiniones de los expertos y que las estimaciones finales de los parámetros que se utilicen tendrían que basarse en la opinión de la mayoría, y no en la individual. Sin embargo, otros autores (60) abalan que el consenso no debe forzarse, ya que puede haber mucha variabilidad de opiniones, basadas en las propias experiencias de los expertos. Además, Soto (61) recomienda que los valores derivados de la opinión de los expertos deben ser validados. Sin embargo, la validez de las opiniones de los expertos no puede comprobarse sin datos con los que compararla y debe haber una justificación sobre cualquier elección que se haya hecho, prestando especial atención en la identificación de los datos de los parámetros de los que los resultados del modelo pueden ser más sensibles. Esto se debe a que la identificación de todos los datos para cada parámetro puede no representar un buen uso de los recursos de investigación (62).

Teniendo en cuenta las ventajas y limitaciones de cada una de las fuentes de resultados en salud, parece claro que lo más productivo es utilizarlos de forma complementaria, evaluando en cada situación concreta cual sería el tipo de estudio más conveniente en función de los recursos disponibles, de la urgencia en encontrar una respuesta al problema, o simplemente de la oportunidad de llevarlo a cabo.



Bibliografía

1. Lewis D. Economic evaluation of health care programs. *Aust Econ Rev.* 2004;37:350–8.
2. Hanratty B, Craig D, Nixon J, et al. Are the best available clinical effectiveness data used in economic evaluations of drug therapies? *J Health Serv Res Policy.* 2007;12:138–41.
3. Gold MR, Stevenson D, Fryback DG. HALYs and QALYs and DALYs, oh my: similarities and differences in summary measures of population health. *Annu Rev Public Health.* 2003;23:115–34.
4. Johnson JR, Ning YM, Farrell A, et al. Accelerated approval of oncology products: The food and drug administration experience. *J Natl Cancer Inst.* 2011;103:636–44.
5. Soto Alvarez J. Medición y evaluación de los resultados en salud en las evaluaciones económicas. En: *Evaluación económica de medicamentos y tecnologías sanitarias: Principios, métodos y aplicaciones en política sanitaria.* 1ª Ed. Springer Healthcare, Madrid. 2012 p.52-74.
6. Haynes B. Can it work? Does it work? Is it worth it? *BMJ.* 1999;319:652–3.
7. Revicki DA, Frank L. Pharmacoeconomic Evaluation in the Real World. *PharmacoEconomics.* 1999;15:423–34.
8. Schechtman E. Odds ratio, relative risk, absolute risk reduction, and the number needed to treat—which of these should we use? *Value Health.* 2002;5:431–6.
9. Etches V, Frank J, di Ruggiero E, et al. Measuring population health: A review of indicators. *Annu Rev Public Health.* 2006;27:29–55.
10. Neumann PJ, Goldie SJ, Weinstein MC. Preference-based measures in economic evaluation in health care. *Annu Rev Public Health.* 2000;21:587– 611.
11. Casado A, Prieto L, Alonso J. El tamaño del efecto de la diferencia entre dos medias: ¿estadísticamente significativo o clínicamente relevante?. *Med Clin (Barc).* 1999;112:584–8.

12. Bleichrodt H. Health utility indices and equity considerations. *J Health Econ.* 1997;16:65–91.
13. Kraemer HC, Kupfer DJ. Size of treatment effects and their importance to clinical research and practice. *Biol Psychiatry.* 2006;59:990–6.
14. Bonneux L. How to measure the burden of mortality? *J Epidemiol Community Health.* 2002;56:128–31.
15. Guy GP, Ekwueme DU. Years of potential life lost and indirect costs of melanoma and non-melanoma skin cancer. *PharmacoEconomics.* 2011;29:863–74.
16. Burnet NG, Jefferies SJ, Benson RJ, et al. Years of life lost (YLL) from cancer is an important measure of population burden —and should be considered when allocating research funds. *Br J Cancer.* 2005;92:241–5.
17. Manuel DG, Schultz SE. Health-related quality of life and health-adjusted life expectancy of people with diabetes in Ontario, Canada, 1996–1997. *Diabetes Care.* 2004;27:407–14.
18. Mont D. Measuring health and disability. *The Lancet.* 2007;369:1658–63.
19. Anand S, Hanson K. Disability-adjusted life years: a critical review. *J Health Econ.* 1997;16:685–702.
20. Murray CJL, Acharya AK. Understanding DALYs. *J Health Econ.* 1997;16:703–30.
21. Lyttkens CH. Time to disable DALYs? On the use of disability-adjusted life-years in health policy. *Eur J Health Econ.* 2003;4:195–202.
22. Whitehead SJ, Ali S. Health outcomes in economic evaluation: the QALY and utilities. *Br Med Bull.* 2010;96:5–21.
23. Weinstein MC, Stason WB. Foundations of cost-effectiveness analysis for health and medical practices. *N Engl J Med.* 1977;296:716–21.
24. Lazaro A. Theoretical arguments for the discounting of health consequences. *PharmacoEconomics* 2002;20:943–61.
25. Severens JL, Milne RJ, Severens H. Discounting health outcomes in economic evaluation: the ongoing debate. *Value Health.* 2004;7:397–401.
26. Care LM-M. Measuring effects without randomized trials? Options, problems, challenges. *Med Care.* 1995; 33 (4 Suppl): AS8-AS14

27. Klose T. A utility-theoretic model for QALYs and willingness to pay. *Health Econ.* 2003;12:17–31.
28. Wallace M. Katz Index of Independence in Activities of Daily Living (ADL). *Urol Nurs.* 2007;27:93–94.
29. Fisher JD. New York Heart Association classification. *Arch Intern Med.* 1972;129:836.
30. Mor V, Laliberte L, Morris JN, et al. The Karnofsky performance status scale an examination of its reliability and validity in a research setting. *Cancer.* 1984;53:2002–7.
31. Zimmerman M, Martinez JH, Young D, et al. Severity classification on the Hamilton depression rating scale. *J Affect Disord.* 2013;150:384–8.
32. Goldberg DP, Hillier VF. A scaled version of the General Health Questionnaire. *Psychol Med.* 1979;9:139–45.
33. The Sickness Impact Profile: development and final revision of a health status measure. *Med Care.* 1981;19: 787–805.
34. Wiklund I. The Nottingham Health Profile--a measure of health-related quality of life. *Scand J Prim Health Care Suppl.* 1990;1:15–8.
35. Ware JE Jr, Gandek B. The SF-36 Health Survey: development and use in mental health research and the IQOLA project. *Int J Ment Health.* 2015;23:49–73.
36. Rosser R, Kind P. A scale of valuations of states of illness: is there a social consensus? *Int J Epidemiol.* 1978;7:347–58.
37. Kaplan RM, Anderson JP, Ganiats TG. The Quality of Well-being Scale: rationale for a single quality of life index. In: *Quality of Life Assessment: Key Issues in the 1990s.* Ed. Stuart RW, Rosser RM. Kluwer Academic Publishers, 1993;p. 65–94.
38. EuroQol - a new facility for the measurement of health-related quality of life. *Health Policy.* 1990;16:199–208.
39. Klimt CR. The conduct and principles of randomized clinical trials. *Control Clin Trials.* 1981;1:283–93.
40. Moses LE. Measuring effects without randomized trials? Options, problems, challenges. *Med care.* 1995;33(4 Suppl):AS8–14.

41. Egger M, Davey Smith G, Altman DG. Systematic reviews in health care: Meta-Analysis in context. London. BMJ Books 2008; p. 47-69.
42. Casas M. Los sistemas de clasificación de pacientes. Conceptos básicos. En Jiménez Jiménez J (ed.) Manual de gestión para jefes de servicio. Madrid: Díaz de Santos, 1997; p. 289-301.
43. Mauskopf J, Schulman K, Bell L, et al. A strategy for collecting pharmacoeconomic data during phase II/III clinical trials. *PharmacoEconomics*. 1996;9:264-77.
44. Baltussen R, Leidl R, Ament A. Real world designs in economic evaluation. Bridging the gap between clinical research and policy-making. *PharmacoEconomics*. 1999;16:449-58.
45. Thompson SG, Barber JA. How should cost data in pragmatic randomised trials be analysed? *BMJ*. 2000;320:1197-2000.
46. Ridyard CH, Hughes DA. Methods for the collection of resource use data within clinical trials: A systematic review of studies funded by the UK Health Technology Assessment program. *Value Health*. 2010;13:867-72.
47. Glasziou PP; Simes Rj, Hall J. Design of a cost-effectiveness study within a randomized trial: the LIPID trial for secondary prevention of IHD. *Control Clin Trials* 1997,18:464-76.
48. McSham WF, Al M, Doshi A, et al. The ISPOR good practices for quality improvement of cost-effectiveness research task force report. *Value Health*. 2009;12:1086-99.
49. Doshi JA, Glick HA, Polsky D. Analyses of cost data in economic evaluations conducted alongside randomized controlled trials. *Value Health*. 2006;9:334-40.
50. Mihaylova B, Briggs A, O'Hagan A, et al. Review of statistical methods for analyzing healthcare resources and costs. *Health Econ*. 2011;20:897-916.
51. Backhouse ME. Use of randomized controlled trials for producing cost-effectiveness evidence. Potential Impact of design choices on sample size and study duration. *PharmacoEconomics*. 2002;20:1061-77.
52. Ortega Eslava A, Fraga Fuentes MD, Alegre Del Rey EJ, et al. Comparaciones indirectas. *Farm Hosp*. 2012;36:173-5.

53. Ortega Eslava A, Fraga Fuentes MD, Puigventós Latorre F, et al. Comparaciones indirectas en los informes de evaluación de medicamentos en la web del grupo GENESIS de la SEFH. *Farm Hosp.* 2012;36:176-9
54. Rubio Cebrián S, Rubio González B. Conceptos e indicadores básicos en economía. Madrid: Escuela Nacional de Sanidad; 2012. <https://www.impact-hta.eu/work-package-2> (acceso verificado el 28 de Feb 2023)
55. Ortega A, Fraga MD, Alegre-del-Rey EJ, et al. A checklist for critical appraisal of indirect comparisons. *Int J Clin Pract.* 2014;68:1181-9.
56. Bucher HC, Guyatt GH, Griffith LE, et al. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol.* 1997;50:683-91.
57. Stanley TD, Jarrell SB. Meta-Regression Analysis: A Quantitative Method of Literature Surveys. *J Econ Surv.* 2005;19:299 -308.
58. Signorovitch JE, Sikirica V, Erder MH, et al. Matching adjusted indirect comparisons: a new tool for timely comparative effectiveness research. *Value Health.* 2012;15:940-7.
59. Phillippo DM, Ades AE, Dias S, et al. Methods for population-adjusted indirect comparisons in submission in Health Technology Appraisal. *Med Decis Making* 2018; 38:200-11.
60. 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.
61. Soto J. Health economic evaluations using decision analytic modeling: principles and practices. Utilization of a checklist to their development and appraisal. *Int J Technol Assess Health Care.* 2002; 18: 94–111.
62. Philips Z, Bojke L, Sculpher M, et al. Good Practice Guidelines for Decision-Analytic Modelling in Health Technology Assessment. *PharmacoEconomics.* 2006;24: 355-371.



3

TIPOS DE COSTES A INCLUIR EN LOS MODELOS ANALÍTICOS DE DECISIÓN

Ana Ortega Eslava

Directora. Servicio de Farmacia. Clínica Universidad de Navarra
Profesora titular. Departamento de Ciencias Farmacéuticas. Facultad de Farmacia
y Nutrición. Universidad de Navarra



Introducción

La evaluación económica constituye una herramienta de ayuda en **el uso eficiente de los recursos**. Dado que existe una limitación de recursos y, por tanto, que los recursos empleados en una alternativa no van a estar disponibles para ser empleados en otra, el uso eficiente de los recursos implica obtener los mejores resultados con el uso de los recursos disponibles. En el campo sanitario, la **evaluación económica** puede contribuir a este uso eficiente de recursos ya que consiste en la comparación entre varias alternativas en cuanto a los efectos sobre la salud y en cuanto al uso de los recursos.

En una evaluación económica es relevante el **coste de oportunidad** que hace referencia al beneficio que se deja de obtener por no utilizar los recursos en la mejor alternativa. Por lo tanto, el coste de una estrategia no es un valor fijo, depende de aquello que se deja de realizar por dedicar los recursos a esa estrategia. Subraya la idea de que la puesta en marcha de un programa por un evaluador o decisor, con el coste asociado que supone, beneficiará a un grupo determinado de la población e implicará el sacrificio de otros individuos que no podrán recibir unos servicios sanitarios beneficiosos (1). Cuando se elige una alternativa, se renuncia a destinar esos recursos a otra alternativa y por tanto, se renuncia a los beneficios que esas otras alternativas podrían aportar.

En las evaluaciones económicas será por tanto relevante analizar tanto los costes como los efectos sobre la salud de varias alternativas, siendo **los costes uno de los componentes claves a analizar**.

En consecuencia, las evaluaciones económicas deberían tratar de valorar todos los recursos consumidos en términos de su coste de oportunidad, aunque por razones prácticas lo usual es valorar los recursos en términos de los precios observables en el mercado (2). De modo que **los costes se calculan multiplicando los recursos utilizados por el**

coste unitario de cada recurso. Es importante diferenciar entre estos dos conceptos y presentarlos de forma diferenciada.

Desde el **punto de vista social**, el coste es el valor de los recursos empleados (**no exclusivamente monetarios**) en la prestación de un servicio (1). La **valoración monetaria de los recursos** consumidos va a representar la estimación de los costes atribuibles al empleo de la alternativa terapéutica y al comparador estudiados en la evaluación económica (2).

En una evaluación económica, interesa el **coste incremental** de una alternativa frente a otra, la diferencia entre el coste de las alternativas, y además en un análisis de coste-efectividad o coste-utilidad lo que interesa es el **coste marginal** que hace referencia al coste extra de producir una unidad adicional de resultado. En ocasiones, se calcula lo que algunos autores denominan el coste medio por unidad de resultado de una única alternativa en lugar de los costes anteriormente descritos, pero este no es de interés en las evaluaciones económicas. En los análisis de coste-efectividad o coste-utilidad se calcula la **ratio coste-efectividad incremental**, el coste adicional de una alternativa frente a otra dividido por el beneficio adicional de una alternativa frente a otra. Por tanto, aquellos costes que sean comunes entre alternativas se podrían obviar, ya que se va a realizar la diferencia de costes.



Tipos de costes

Existen **diferentes formas de clasificar** los costes a incluir en una evaluación económica. Los costes se podrían clasificar según su naturaleza en sanitarios o no sanitarios; según su comportamiento respecto a la actividad productiva en fijos, variables o mixtos; según la imputabilidad a la actividad en directos o indirectos; según la

capacidad del mercado de determinar su valor en tangibles e intangibles; según la naturaleza en costes de personal, de equipamiento, de material, generales; según el momento en que se dan en futuros o no futuros, etc. (3). O también se pueden clasificar según quien los soporta en costes para el sistema sanitario, para los pacientes y los familiares y para otros.

Una clasificación frecuente de los costes diferencia los costes en: costes sanitarios, costes no sanitarios, costes para el paciente/las familias, y los costes asociados a la pérdida de productividad (4). La clasificación más adecuada no está clara y el uso de estos términos no es homogéneo en las diferentes fuentes. La tipificación más clásica y conocida que dividía los costes en directos, indirectos e intangibles está cambiando y actualmente se clasifican, en muchas ocasiones, los costes en costes sanitarios, no sanitarios y otros costes (1). En la Tabla 1 se muestra una clasificación de los costes adaptada de Casado MA y col. (1).

Tabla 1. Tipos de costes en evaluaciones económicas de intervenciones sanitarias (adaptado de Casado MA y col (1))

	Sanitarios	No sanitarios
Directos	Cuidados hospitalarios, tratamiento farmacológico, etc	Gastos de desplazamiento del paciente, cuidados domiciliarios, servicios sociales, cuidados informales, etc
No directos	Costes futuros: consumo de servicios sanitarios en los años de vida ganados como consecuencia de la intervención	Pérdida de productividad, etc

1. Clasificación de los costes según su naturaleza en sanitarios y no sanitarios

1.1. Costes sanitarios

Los costes sanitarios o asociados con la intervención o programa sanitario incluyen: los costes directos, los costes sanitarios futuros y otros costes.

1.1.1. Costes sanitarios directos

Hacen referencia a los costes sanitarios directamente relacionados con la alternativa valorada. Incluyen: costes de adquisición, preparación, distribución, administración y monitorización de los medicamentos, acontecimientos adversos de los fármacos, costes del tratamiento de rescate o del fracaso, equipamiento y material sanitario, hospitalización, visitas médicas (en el hospital, consultas externas, por especialidades, etc) o de enfermería, tiempo y salarios del personal sanitario, pruebas diagnósticas y terapéuticas, procedimientos radiológicos, análisis de laboratorio, etc. (1).

En este apartado se incluyen los costes directamente relacionados con la alternativa independientemente de quien los soporte o de cuando ocurran.

1.1.2. Costes sanitarios futuros

Incluyen los costes sanitarios atribuibles a la mejora en la esperanza de vida de los individuos como resultado del proceso sanitario. Existe controversia sobre los costes a incluir en este apartado.

Los costes futuros se pueden clasificar según si corresponden a problemas de salud que están o no relacionados con el problema de salud inicial objeto de la evaluación (5). O también se pueden clasificar

en función de si corresponden a años de vida ganados o bien a los que el individuo habría vivido sin la intervención. En general, se argumenta que los costes relacionados y producidos durante los años en los cuales también hubiera vivido el individuo deben ser incluidos, mientras que los no relacionados deben ser excluidos (5).

Por el simple hecho de que un tratamiento prolongue la vida del paciente respecto a otra alternativa, este paciente va a tener costes porque tiene la posibilidad al estar vivo de desarrollar otras enfermedades. Los costes que son directamente una consecuencia, por ejemplo, de un programa de prevención se deberían considerar, pero esto no es a lo que nos referimos en este apartado, sino costes no directamente relacionados con el programa evaluado en periodos posteriores, ej en los años de vida añadidos. Aunque la realidad es que en algún estudio en los que se han incorporado no han supuesto una diferencia importante en el cálculo de coste-efectividad incremental respecto a no haberlos incluido (4), en parte puede ser debido a aplicar las tasas de descuento a costes a largo plazo, por ello en muchos casos pueden ser ignorados sin afectar de forma importante a los resultados. Una opción adecuada puede ser analizar el efecto de estos en el análisis de sensibilidad (4).

1.2. Costes no sanitarios

1.2.1. Costes no sanitarios asumidos por los pacientes y familiares consecuencia directa de la intervención sanitaria

Entre estos se incluyen los siguientes (1):

- Recursos consumidos por el paciente y la familia en la atención sanitaria evaluada que por su naturaleza no son costes sanitarios como por ejemplo el desplazamiento (por ejemplo, autobús, tren, taxi o automóvil).

- Tiempo invertido por el paciente, familiares o amigos para asumir el papel de cuidadores del paciente, durante el desplazamiento, el tratamiento, la participación en el programa sanitario, o el suministro de cuidados informales.
- Otros costes asumidos por los pacientes (por ej.: adaptación de una habitación para asistencia domiciliaria).

1.2.2. Recursos procedentes de sectores no sanitarios

Suministrados por organismos públicos (servicios sociales, ayuda domiciliaria, servicios residenciales para ancianos, etc) o sectores voluntarios (1).

1.2.2. Pérdidas de productividad

Incluyen la pérdida de productividad como consecuencia de una condición sanitaria. Algunos autores, los denominan costes indirectos.

Se dividen en: costes de morbilidad (incapacidad, pérdida o restricción de la capacidad laboral atribuible a la enfermedad, no evitados por la alternativa terapéutica) y costes de mortalidad (pérdida de productividad a causa de la muerte prematura) (1). Estos costes pueden abarcar tanto la disminución de la productividad como los días perdidos en el trabajo.

1.2.4. Costes no sanitarios futuros: su inclusión está muy discutida.

1.3. Otros costes

Entre ellos se pueden incluir los costes de transferencias. Los costes de transferencias son flujos monetarios de un grupo de individuos a otros, subsidios, etc. (1), sin que ello realmente implique consumo de recursos

para la sociedad en su conjunto y, por ello, en general, se recomienda su exclusión de las evaluaciones económicas (5). Aunque su inclusión depende de la perspectiva del análisis.

2. Otras formas de clasificar o agrupar los costes

Otra forma de agrupar los costes, sin cambiar la clasificación, sería comenzar la clasificación por **directos/no directos**. Entre los costes directos se encontrarían sanitarios (medicamentos, hospitalización, etc) y no sanitarios (transporte, cuidados informales). Por otro lado, estarían las pérdidas laborales y otros costes.

Los **costes intangibles** harían referencia a la pérdida de bienestar por parte de los pacientes y de sus familiares, así como los derivados de aspectos subjetivos que sufre el paciente y sus seres allegados, tales como el dolor, sufrimiento, el temor, la ansiedad, etc. (2). Su cuantificación es difícil, aunque se puede aplicar la metodología de la disponibilidad a pagar (2). Existe además controversia acerca de la incorporación de estos costes en el apartado de costes. Una de las principales razones, además de la dificultad en su determinación, es el riesgo de su doble cuantificación, ya que se incorporan en muchas ocasiones en la medida de la efectividad, por ejemplo, en la medición de la calidad de vida del paciente.

Otra posible forma de clasificación sería agrupar los costes **según quien los soporta** sector sanitario, paciente y familia, y otros sectores. Algunos autores hacen referencia a los costes no sanitarios como los costes financiados directamente por el paciente o su familia/amigos y no sufragados por el sistema sanitario (2), queriendo hacer referencia con “no sanitarios” a “no soportados por el sistema sanitario”, no a si el coste es de naturaleza sanitario o no, por ejemplo, el coste de un medicamento es de naturaleza sanitario, pero puede ser soportado en su totalidad o en parte por el paciente.

Por otro lado, los costes se pueden clasificar en **costes fijos y costes variables**. Los costes fijos serían aquellos que no varían al variar el nivel de actividad y los variables son los que aumentan al aumentar el nivel de producción (2).

Entre otros costes se pueden incluir también los costes de capital, o los de estructura (2). Los **costes de capital** hacen referencia a los empleados para adquirir los grandes activos fijos necesarios para producir resultados en salud. Por ejemplo, equipamiento técnico, edificios, terrenos, etc., que representan la inversión en un activo que se va a utilizar a lo largo del tiempo. Rara vez se incluyen en la evaluación económica, salvo que estén incluidos en un coste general, o si están directamente relacionados con la alternativa valorada. Los **costes de estructura**, generalmente se incluyen en costes de procesos, por ejemplo, en el coste de una prueba de laboratorio está incluido el coste del aparato, los reactivos, etc (2).



Identificación, medición y valoración de los costes

La identificación, medición y valoración de los costes son procesos diferenciados, que en ocasiones se realizan o presentan de forma conjunta, pero que deberían presentarse y realizarse de forma diferenciada.

1. Identificación de los costes

Identificar los costes consiste en listar los costes que es necesario considerar en una evaluación económica concreta.

Es necesario **decidir el grado de detalle** y precisión del análisis. Es deseable que sea lo más detallado posible. Es importante identificar todos los costes, identificar los relevantes e incluir estos últimos. Sin embargo, es necesario **explicar en caso de excluir** algunos costes cuáles se excluyen y el motivo por el que se excluyen, y mostrar

que los resultados no se modificarían de forma significativa por su inclusión o no.

Los costes a incluir dependen de múltiples factores, a continuación, se comentan algunos de los más relevantes.

1.1. Factores que influyen en los costes a incluir en una evaluación económica

1.1.1. La perspectiva del análisis

Los tipos de costes a incluir en una evaluación económica dependen de la perspectiva desde la que se realice el análisis. Los costes a incluir serán diferentes si la perspectiva es la de una parte del sistema sanitario, por ejemplo, el hospital, la del sistema sanitario (que incluiría el hospital, atención primaria, la consulta de especializada, etc), la del sistema público o el gobierno (que incluiría además de los costes para el sistema sanitario los costes relacionados con otros ministerios), la de una compañía aseguradora, la de un paciente individual o la de la sociedad. En la Tabla 2 se muestra una posible agrupación de los costes más frecuentes a incluir según la perspectiva del análisis (1).

La perspectiva más amplia será la de la sociedad, ya que incluye tanto el sistema sanitario como otros servicios públicos, los servicios privados y los propios pacientes. En general, se recomienda la utilización de la perspectiva de la sociedad con la inclusión de todos los costes con independencia de quién es el pagador. Pero se recomienda **presentar de forma diferenciada los costes desde las diferentes perspectivas**, por ejemplo, desde la perspectiva del sistema sanitario y desde la de la sociedad. Una intervención puede ser costo-efectiva desde una perspectiva y no serlo desde otra. Por tanto, la perspectiva es clave en la estimación de los costes.

Tabla 2. Tipos de costes más frecuentes a incluir según la perspectiva del análisis (6)

Costes de la intervención			
Perspectiva de la sociedad			
Perspectiva del financiador público		Perspectiva del paciente	
Perspectiva del proveedor o financiador sanitario			
Costes directos			
Costes directos sanitarios	Costes directos no sanitarios financiados públicamente	Costes directos para el paciente	Pérdidas de productividad laboral
Atención sanitaria: <ul style="list-style-type: none"> • Atención primaria • Atención especializada • Enfermería • Servicios de urgencia • Hospital de día • Atención sanitaria domiciliaria • Otros servicios sanitarios Medicación Hospitalizaciones Pruebas diagnósticas Transporte medicalizado	Cuidados profesionales: <ul style="list-style-type: none"> • Cuidados personales formales • Servicios sociosanitarios 	<ul style="list-style-type: none"> • Copagos • Seguros privados • Atención sanitaria privada • Material sanitario • Transporte privado al centro sanitario • Cuidados profesionales • Cuidados personales informales • Adaptación de infraestructuras 	Costes en que incurre el paciente/ sociedad: <ul style="list-style-type: none"> • Por mortalidad prematura • Por morbilidad (incapacidad, absentismo, presentismo, jubilación anticipada) • Subsidios por enfermedad*

*Los subsidios por enfermedad se considerarán desde la perspectiva del financiador público.

Todo tipo de evaluación económica, e independientemente de la perspectiva desde la que se realiza, debe incluir los costes directos sanitarios, aunque es cierto que si la perspectiva del análisis es la del hospital sólo incluirá los costes para el hospital y no incluirá, por ejemplo, los costes para

atención primaria, si es la del sistema sanitario incluirá los costes para el sistema sanitario, independientemente de si son asumidos por el hospital o por atención primaria, y si la perspectiva es la del paciente, incluirá los costes para el paciente, tanto los directos sanitarios como los no sanitarios que sean asumidos por el paciente.

1.1.2. El objetivo del análisis

Si el objetivo de la evaluación económica es comparar exclusivamente un número limitado de alternativas conocidas, se pueden eliminar los costes comunes a las alternativas, ya que como se ha comentado anteriormente los costes relevantes son los diferenciales entre las alternativas comparadas y los comunes se anularían al hacer la diferencia de costes entre las alternativas. Sin embargo, si el análisis de costes de las alternativas analizadas se quiere utilizar después para comparar con otras alternativas, en otro momento, o en otro análisis, sería necesario recoger todos los costes porque no se sabe con antelación cuáles serán los costes comunes con estas otras alternativas.

Por costes comunes se entiende que tienen que ser los mismos tanto en concepto como en cuantía entre las alternativas. Así, por ejemplo, si se están comparado dos alternativas y ambas requieren terapia de soporte, esta no se debe obviar si, por ejemplo, la cantidad de terapia de soporte varía de una a otra, o si ambas alternativas se administran en el hospital de día, pero la duración de la administración, o los días que requieren administración son diferentes, no se debe obviar este coste.

1.1.3. La magnitud del coste

Algunos costes pueden ser insignificantes frente a otros, recogerlos puede no suponer un cambio **relevante** en los costes y mucho esfuerzo, y no ser práctica su recogida. En caso de no recoger algunos costes es importante **justificar** por qué se excluyen estos costes.

1.1.2. El horizonte temporal

En el contexto de toma de decisiones, se debe considerar el horizonte temporal (el tiempo) que **recoja todos los efectos relevantes** diferentes entre las alternativas, entre estos efectos se encuentran los costes, sin introducir sesgos, para el decisor o usuario (4), no se debe sesgar en favor de una alternativa u otra. En general, se acepta que se recojan todos los costes relacionados con la salud. Ello significa que en algunos casos es necesario considerar los costes a lo largo de toda la vida del paciente, si bien es verdad que los costes a más largo plazo, salvo que sean altos su repercusión será menor debido a aplicar el ajuste temporal (la tasa de descuento) que se comentará más adelante.

Otro aspecto temporal que puede cambiar en la consideración de los costes es que algunos costes pueden cambiar a lo largo de tiempo porque haya una **curva de aprendizaje**, por ejemplo, si se va mejorando la eficiencia en el uso de tecnologías o siendo más eficiente en las dosis, administración o aprovechamiento de las cantidades sobrantes, por ejemplo, de los viales, el seguimiento y gestión de efectos adversos, etc.

Un aspecto importante también es que el horizonte temporal para costes y efectos sobre la salud en una evaluación económica debe ser el mismo, y también debe ser el mismo en las diferentes alternativas que se están comparando.

2. Medición de los costes

Medir significa **cuantificar**. En una evaluación económica los recursos consumidos se pueden cuantificar mediante una **medición o, si esto no es posible, se puede realizar una estimación** de la cantidad de recursos que se consumen de cada uno de los tipos de costes que se han identificado. Esta cuantificación o estimación se realiza en **unidades del recurso**, no en unidades monetarias, por ejemplo, número de comprimidos de un

medicamento utilizado, número de una determinada prueba de laboratorio, número de días de hospitalización, número de comprimidos de un medicamento o de miligramos de un producto, tiempo invertido por los pacientes, número de visitas al médico, etc. Esta medición o estimación se deberá acompañar de una medición o estimación de la **incertidumbre** en torno a esta cuantificación.

En muchas ocasiones, la forma de cálculo de los costes viene determinada también por la **posibilidad de obtención de los datos**. Por ejemplo, cuando se toman los costes directamente de datos agregados o bases de datos o estudios previos que dan la información de forma agregada, y no es posible saber exactamente el número de unidades de cada recurso de forma diferenciada. Incluso, en ocasiones, no es posible obtener de forma diferenciada los datos de recursos consumidos de los de costes.

Algunos autores (5) clasifican los métodos de medición en métodos sintéticos, que utilizan fuentes de información secundaria como bases de datos administrativos, revisiones retrospectivas, etc., y los basados en datos primarios, que se basan en fuentes de información específicas para el estudio, obtenidas en muchos casos de forma prospectiva en el marco de un ensayo clínico o del estudio específico de evaluación. Por otro lado, recomiendan no basarse en promedios generales sin tener en cuenta el case-mix de los pacientes (5). Una de las ventajas de la medida prospectiva es poder recoger información para asociar características de los pacientes con costes, esta información es más difícil de recoger en estudios de datos retrospectivos.

Cuando los recursos consumidos, y por tanto los costes, no se miden o cuantifican en un estudio clínico experimental u observacional, se estiman en otro tipo de enfoque no empírico que habitualmente se denomina “**modelo económico**” (7). El modelo económico pretende representar la realidad de forma simplificada a partir de la síntesis de datos que pueden provenir de diferentes fuentes. La principal ventaja de los modelos radica en la flexibilidad y posibilidad de generalizar. Los

supuestos del modelo pueden ajustarse de tal manera que reflejen la idiosincrasia de un escenario. Sin embargo, es esta naturaleza incierta de los supuestos y el uso de datos no empíricos lo que puede introducir **sesgos** que amenacen la validez del modelo (7).

En la obtención de datos de la **bibliografía** es necesario considerar que es difícil que los costes calculados en otros países/entornos se puedan trasladar directamente a otro país/entorno. No solamente porque difieren los costes unitarios, sino que también **pueden diferir los recursos consumidos**, por ejemplo, cuándo empieza la terapia, la duración de la misma, la terapia usada, los recursos considerados, por ejemplo, cuáles se incluyen en el coste de una sesión de hospital de día, si se incluyen los costes comunes o no, etc.

Por lo que hace referencia a la **dosis del medicamento**, es necesario considerar la dosis que se va a utilizar. La dosis que se ha usado en el ensayo clínico en ocasiones no coincide con la incluida en ficha técnica o la realmente utilizada después por los pacientes en la práctica clínica, que se verá afectada además por la adherencia al tratamiento, las reducciones de dosis por toxicidad o por interacciones, las características de los pacientes que al final reciben el tratamiento que puede no coincidir con las del ensayo, por ejemplo en la población real puede existir un porcentaje diferente de pacientes con fallo renal o hepático que pueden requerir ajuste de dosis, la dosis modificada en función de monitorización farmacocinética, etc. Es necesario tener en consideración este aspecto en la evaluación económica. Por una parte, va a depender de si se han obtenido los datos tanto de efectividad comparada como de costes de la práctica clínica habitual, si esto fuera así, se utilizarían para ambos conceptos los datos de la práctica real, porque con ese consumo de recursos se consiguen esos resultados. Sin embargo, lo habitual es que los datos de la efectividad comparada se tomen de uno o varios ensayos clínicos y entonces sería necesario tomar los costes, al menos los del consumo del fármaco de los pacientes del ensayo clínico, pero no de lo planificado, sino de lo real que han

acabado recibiendo los pacientes, porque con esos recursos se han conseguido esos resultados. Y en el caso de otros recursos, diferentes del medicamento, se tendrán que excluir los recursos consumidos en el ensayo clínico y que no se consuman en la práctica clínica, siempre que su retirada no influya de forma significativa en el resultado. Sin embargo, si el uso de esos recursos previsiblemente afecta a los resultados será necesario considerar en el apartado de los resultados o efectos sobre la salud que con esos recursos los efectos sobre la salud no serán los mismos. Además, **si en la práctica clínica se espera un uso de recursos diferente al ensayo clínico es necesario mencionar este aspecto y analizar su posible efecto** tanto en costes como en resultados en el análisis de sensibilidad.

Otro concepto diferente es **el aprovechamiento de viales**, si en la práctica se van a poder aprovechar los viales, esto se puede tener en consideración, pero será un ajuste local, ya que depende del volumen de uso de ese medicamento a nivel local. Algunos autores recomiendan analizar el efecto del aprovechamiento o no de los viales, aquello que sea menos probable, en el análisis de sensibilidad. Y cuando existan varias **presentaciones** habría que elegir aquellas que más se aproximen a la realidad de uso. Una forma en la que se puede tener este aspecto en consideración es modificando el modo de imputación del coste del medicamento, en un escenario en que es posible el aprovechamiento de viales se introduce el coste unitario como coste por cantidad de medicamento, por ejemplo, el coste por miligramo, y en un escenario en el que el aprovechamiento de viales no es posible, se establece como unidad los viales, utilizando el coste por vial independientemente de si se utiliza uno, medio o un tercio del vial para cada paciente o preparación.

2.1. Micro-coste vs coste por proceso, métodos de abajo-arriba vs de arriba-abajo

El grado de detalle o precisión con el que es necesario recoger un coste en una evaluación económica dependerá en gran medida de la importancia de ese coste en esa evaluación económica. Recoger de forma detallada cada uno de los recursos consumidos y su coste unitario para calcular el coste puede ser muy laborioso y requerir mucho tiempo y recursos. Por ello, en muchas ocasiones se recurre a cuantificar las unidades consumidas de un determinado proceso y calcular o recoger un coste unitario de ese proceso para así hacer la valoración en unidades monetarias de ese coste.

Micro-coste significa recoger la cantidad consumida de cada recurso unitario de forma detallada y un cálculo posterior de los costes valorando cada recurso en unidades monetarias utilizando los costes unitarios de estos conceptos. Por ejemplo, contar unidades de medicamentos, de sueros, de días de hospitalización, de cada tipo de test de laboratorio, etc.

Drummond y col. (4) recogen los niveles de precisión de las diferentes formas de recoger los costes hospitalarios de más precisos a menos precisos:

1. Micro-coste: se recoge o estima el uso de cada recurso (ej, test de laboratorio, días de hospitalización, medicamentos, etc.) y el coste unitario de cada recurso.
2. Grupo case-mix: Se da el coste para cada tipo de caso o tipo de paciente. Considera la duración de la estancia. La precisión depende del grado de detalle en especificar los tipos de casos.
3. Específico por enfermedad per diem (o coste diario): da un coste diario medio para el tratamiento en cada categoría de enfermedad. Puede ser bastante general/amplio (ej cirugía ortopédica).
4. Media por diem (o coste por día): Media per diem para todas las categorías. Disponible en la mayoría de los hospitales.

Los puntos 2 a 4 podrían ser incluidos en **coste por proceso o coste bruto**.

La elección de uno u otro va a depender de la importancia de cada categoría de costes, en general cuando estamos hablando de evaluación de diferentes alternativas de tratamiento puede ser necesario recoger el micro-coste al menos para los efectos más sensibles y recoger, por ejemplo, otras formas de coste más generales para efectos adversos de los tratamientos, etc. E inclusive dentro del micro-coste se pueden aplicar diferentes niveles de precisión.

Cuando se utilizan costes por Grupo Relacionado con el Diagnóstico (GRD), por ejemplo, es importante conocer qué costes se han considerado en cada uno de ellos. Los costes por GRD pueden no ser siempre apropiados por no reflejar adecuadamente el cambio en los recursos por las intervenciones valoradas.

Spacirova Z y col. han publicado recientemente una revisión de la literatura en lengua española e inglesa entre 2005 y 2018 de las metodologías de cálculo de costes para las evaluaciones económicas (8).

Los diferentes métodos se diferencian en validez, que hace referencia a si la metodología es apropiada para la cuestión a resolver (validez externa) y que el cálculo está bien fundamentado (validez interna), y la precisión, que hace referencia a que la medida es fiable (mediciones repetidas darían los mismos resultados) y que aumenta con el tamaño de muestra. Existen diferentes métodos según nivel de agregación, micro-coste o coste-bruto, y según el método de medición de recursos y valoración de costes de arriba-abajo o de abajo-arriba.

En los métodos de **arriba-abajo**, el objetivo es dar un coste medio para un set de productos y servicios producidos por la organización en un periodo dado. El punto de partida suele ser el coste actual de los recursos consumidos en una organización en un periodo que después son asignados o repartidos en los productos y servicios. Generalmente se usan para tarifas o reportes financieros usando sus datos de contabilidad. Muchos no reflejan el coste de oportunidad y por ello no son los

mejores para las evaluaciones económicas. Generalmente son costes medios. Los costes fijos correspondientes a gastos generales se deben incluir en las evaluaciones económicas según la pregunta planteada, si para la cuestión evaluada se necesitan más costes de este tipo entonces es razonable considerarlos. Por otro lado, en general se usan costes por GRD, pero estos fueron diseñados para tarifas en general, y pueden no tener el detalle necesario para una evaluación económica. Si se usan datos retrospectivos también pueden no ser válidos.

Coste basado en la actividad. Se puede clasificar como un método de arriba-abajo micro. Requiere recoger datos detallados de actividad, generalmente mediante entrevistas u observación directa por el investigador.

En los métodos **de abajo-arriba** los costes son medidos en una muestra de pacientes en unidades naturales y después valorados en unidades monetarias. Para las evaluaciones económicas estos son más adecuados. Cuando para algunos costes se usa el método de arriba-abajo, por ejemplo, el coste por día de hospitalización y para el resto de abajo-arriba es importante saber que **no se cuentan doble** algunos costes. También es importante saber que las tarifas pueden no coincidir con los costes reales. Y que muchas veces no se calculan los costes unitarios, sino que se cogen de bases de datos, en este caso es importante valorar la representatividad para el entorno en cuestión.

En conclusión, identifican que no está estandarizada y que es importante estandarizar los conceptos, ninguna metodología puede ser considerada el estándar de oro. En general no es posible calcular todos los costes con métodos de abajo-arriba, micro-coste, por ejemplo, es difícil para gastos generales y por ello se necesita utilizar métodos de arriba-abajo. La convergencia de ambos métodos es un tema a trabajar (8).

Cuando existen **costes que son generales o comunes** a varias estrategias, es necesario hacer una distribución de los costes entre las estrategias. Por ejemplo, en un hospital los gastos de electricidad, limpieza, los costes administrativos, etc. No hay un método estándar de cómo

realizar esta distribución, pero es importante buscar un criterio razonable de cómo hacer esta distribución de plasmarlo (4).

Los costes comunes primero es necesario ver si se ven alterados por usar una u otra alternativa, si es así habrá que asignarlos y el reparto de estos se puede hacer según algún criterio que se pueda considerar relacionado con la intensidad de uso de los recursos, ej número de camas, de consultas, m², etc.

2.2. Recogida de datos individuales por paciente

La recogida de **datos individuales** de pacientes es en general más precisa siempre y cuando se cumpla que el estudio esté bien realizado **evitando sesgos e imputaciones erróneas**.

La recogida de consumo de recursos de estudios con datos individuales se puede realizar en los **mismos estudios en los cuales se evalúa los resultados de eficacia/efectividad** y seguridad o no. La ventaja de realizarlo simultáneamente es que para el análisis de coste efectividad se sabrá que con ese consumo de recursos del estudio se obtienen esos resultados del estudio. En este caso, puede no ser necesario realizar estimaciones. Aunque es difícil que en un único estudio se puedan recoger todos los datos necesarios para una evaluación económica, y, por ello, en la mayoría de las evaluaciones económicas es necesario realizar algún tipo de estimación y analizar en el análisis de sensibilidad el efecto en los resultados de la incertidumbre en torno a esas estimaciones.

Cuando se obtienen los datos de estudios, estos pueden ser estudios observacionales o ensayos clínicos. En cualquiera de los casos es importante realizar un diseño adecuado y un análisis adecuado para **evitar la influencia de posibles sesgos o factores de confusión**. Por tanto, es necesario comprobar su validez interna y también la validez externa para el fin para el cual se quieren utilizar los datos del estudio. En el caso de estudios observacionales y de ensayos clínicos no alea-

torizados que incluyan las alternativas a comparar en la evaluación económica, será necesario corregir por los posibles factores de confusión en el diseño del estudio o en el análisis de los datos mediante diferentes estrategias, como por ejemplo el índice de propensión.

Si los recursos consumidos provienen de un **ensayo clínico**, es necesario considerar cuánto de diferentes van a ser en la práctica clínica respecto al ensayo clínico. No solamente por la validez externa del estudio, lo cual afecta tanto a estudios observacionales como a ensayos clínicos y hace referencia a la posibilidad de aplicar los resultados del estudio a un entorno, unos pacientes, que pueden ser en la práctica clínica diferentes a los incluidos en el estudio, que en general suelen ser pacientes seleccionados de acuerdo a unos criterios, en particular en los ensayos clínicos. Sino también, en el caso de los ensayos clínicos, en general el uso de recursos es diferente a la práctica clínica, por ejemplo, el seguimiento es mayor, los controles y visitas más frecuentes, etc. Por ello en estos casos es necesario considerar cuántos de estos recursos no se consumirían en la práctica clínica, a la hora de tenerlo en consideración para la realización de la evaluación económica.

Otro aspecto a considerar hace referencia a si la obtención de datos es prospectiva o retrospectiva. Si los datos se obtienen de forma **retrospectiva** existirán las limitaciones propias de esta forma de obtención de datos. Por ejemplo, si la fuente es la historia clínica la información dependerá de la información registrada, el grado de detalle y la calidad de los datos.

Algunos datos, sólo pueden ser obtenidos de **entrevistas** a pacientes, y la forma de diseñar esa recogida puede tener influencia en los resultados que se encuentren, por lo que se debe hacer con la misma rigurosidad que otro tipo de estudios.

Cuando los datos provienen de pacientes es necesario considerar que los costes normalmente **no siguen una distribución normal**. La mayoría de los pacientes tienen unos costes en torno a unos valores, pero hay pacientes que requieren muchos más recursos y tienen muchos más costes, generalmente son pacientes que se complican, requie-

ren ingreso o este se alarga, tienen consecuencias con alto coste, etc, pero también puede ser que en tratamientos mantenidos en el tiempo haya unos pocos que requieran más tratamiento. Esta distribución de los costes podría seguir una distribución gamma. Su análisis requerirá, por tanto, la utilización de estrategias adecuadas como análisis no paramétricos o técnicas de remuestreo como bootstrap para poder estimar adecuadamente diferencias entre tratamientos o incertidumbre en torno a los valores o sus intervalos de confianza.

Es precisamente esta distribución no gaussiana de los recursos consumidos la que hace que en ocasiones haya surgido la cuestión de si utilizar la **media o la mediana** para estimar los recursos consumidos por paciente. En una distribución no normal la mediana es el mejor estimador de la tendencia central o el valor central de ese parámetro. Por ejemplo, si estamos hablando del número de ciclos de tratamiento o cantidad de medicamento consumido por el paciente, la mediana nos indica que el 50% de los pacientes han recibido menos del valor de la mediana y el 50% más de ese valor, pero no indica cuánto han recibido los pacientes. Por ejemplo, supongamos que la mediana es de 12 dosis y la media de 18 dosis. Una mediana de 12 dosis no indica nada respecto a si los pacientes recibían valores próximos a 12 o lejanos a este valor, sin embargo, con el valor de la media, 18 en este caso, sabemos que multiplicada por el número de individuos es el número total de dosis de medicamento que se ha consumido por todos los pacientes y es más informativo respecto a los recursos consumidos en el total de pacientes. En las evaluaciones económicas interesan los recursos consumidos por todos los pacientes con las diferentes alternativas y los beneficios obtenidos por todos los pacientes, solo que para facilitar el cálculo se trabaja con datos por paciente, y por lo tanto se utiliza el dato de la media. Por otro lado, si se dispone de los recursos de la media multiplicado por el número de pacientes, se dispondrá de recursos suficientes, en ese concepto, para hacer frente a los costes de elegir esa alternativa, cosa que si se dispone de los recursos de la mediana multiplicado por el número de pacientes no se dispondrá de los recursos suficientes para el tratamiento de

los pacientes. La media representa mejor los recursos necesarios si se elige esa alternativa. Si no se dispone del dato de la media, se estima, y se analiza el efecto de la incertidumbre en este parámetro en el análisis de sensibilidad.

Otra consideración es que los estudios en los que se analizan costes y efectos sobre la salud, en muchas ocasiones están diseñados de modo que el cálculo del **tamaño de muestra** está realizado para detectar diferencias en resultados en salud, en general eficacia o efectividad, pero no tienen poder estadístico suficiente para detectar diferencias en recursos consumidos o en costes, que en muchas ocasiones tienen una variabilidad mayor. Por ello, es necesario analizar la potencia estadística de los estudios para detectar estas diferencias o dar una estimación de los recursos consumidos o los costes con una precisión concreta.

Otra fuente de obtención de los datos de recursos consumidos o costes son las fuentes secundarias, es decir la bibliografía, paneles de expertos, etc. Una de las limitaciones de las estimaciones sintéticas o de **datos agregados** es que no es posible establecer relaciones entre consumos y características de los pacientes, y las estimaciones obtenidas corresponden entonces a estimaciones medias.

3. Valoración de los costes

El último paso hace referencia a la valoración de esos recursos consumidos. Esta valoración se realiza multiplicando cada recurso consumido por su coste unitario. Estos costes unitarios también pueden tener asociada una incertidumbre, que es importante recoger o estimar, reflejarla y analizar su efecto en los resultados en el análisis de sensibilidad. De esta forma se valoran todos los costes en una misma unidad, generalmente la unidad monetaria, por ejemplo, el euro.

En la Tabla 3 se muestra un ejemplo de cómo se pueden presentar los datos de recursos consumidos, costes unitarios y la valoración de los costes multiplicando los dos anteriores por cada tipo de coste y en total.

Tabla 3. Ejemplo de tabla tomada del guía GENESIS sobre costes a incluir y cómo presentar de forma diferenciada los recursos consumidos, los costes unitarios, los costes valorados en euros y la diferencia de costes (9)

	Medicamento A		Medicamento B		Diferencia de costes (o costes incrementales**)
	Cantidad	Coste * (€)	Cantidad	Coste * (€)	

Costes del medicamento

Coste adquisición del medicamento

Coste de preparación

Coste distribución de farmacia

Costes de administración

Coste de administración por enfermería

Costes de monitorización

Coste de monitorización para prevenir toxicidad

Coste de monitorización para asegurar correcta efectividad

Nota: Si el coste para las dos opciones comparadas es el mismo en un ítem, este se puede dejar fuera del análisis ya que no modificará el coste diferencial global. En caso de que se vaya a comparar con otras alternativas puede ser necesario incluir otros costes diferenciales frente a las nuevas alternativas valoradas.

*Calculado como el coste unitario x Cantidad; **Coste de A – Coste de B

Tabla modificada de Soto Álvarez, 2012.

Tabla 3 cont. Ejemplo de tabla tomada del guía GENESIS sobre costes a incluir y cómo presentar de forma diferenciada los recursos consumidos, los costes unitarios, los costes valorados en euros y la diferencia de costes (9)

	Medicamento A		Medicamento B		Diferencia de costes (o costes incrementales**)
	Cantidad	Coste * (€)	Cantidad	Coste * (€)	

Costes efectos adversos

Coste medicamentos para prevenir efectos adversos

Coste del tratamiento de los efectos adversos

Otros costes asistenciales

Costes días de hospitalización

Costes de Hospital de Día

Costes pruebas complementarias

Costes visitas médicas, farmacéuticas, enfermería

Costes de transporte en ambulancia o medicalizado

Costes de atención sanitaria domiciliaria

TOTAL GLOBAL

Nota: Si el coste para las dos opciones comparadas es el mismo en un ítem, este se puede dejar fuera del análisis ya que no modificará el coste diferencial global. En caso de que se vaya a comparar con otras alternativas puede ser necesario incluir otros costes diferenciales frente a las nuevas alternativas valoradas.

*Calculado como el coste unitario x Cantidad; **Coste de A – Coste de B

Tabla modificada de Soto Álvarez, 2012.

Como se ha comentado, aunque se debería considerar el coste de oportunidad, para esta función de valoración de los costes, en la práctica se utilizan los costes de mercado. Es evidente que los precios de mercado no reflejan en muchos casos costes de oportunidad. En el sector sanitario, puede ser más adecuado que el precio público del mercado para los servicios el coste de realización de ese servicio. Además, algunos costes no tienen un valor en el mercado, como puede ocurrir con el tiempo de los voluntarios o tiempo de cuidado por parte de los familiares.

El problema surge cuando se quiere comparar alternativas o estudios en los que se han recogido los costes de diferente manera, en unos son costes reales, en otras estimaciones, en otros cargos, en algunos hay micro-coste y en otros costes por GRD, tarifas de comunidades autónomas y pueden no ser comparables. En un apartado más adelante se dan algunos ejemplos de referencias donde localizar algunos de estos costes.

En relación a los costes unitarios de los medicamentos, existen diferentes aspectos que es necesario abordar. Generalmente se utiliza el precio del medicamento, pero hay que definir qué **precio de los medicamentos** es necesario considerar. Lo más adecuado es utilizar aquél que representa el coste para el agente desde el que se realiza el análisis (perspectiva) y lo más próximo al coste real posible. Así por ejemplo desde el punto de vista de la sociedad para medicamentos dispensados en la oficina de farmacia, lo adecuado sería utilizar el PVP con el IVA. Para la perspectiva del Sistema Nacional de Salud (SNS) para los medicamentos dispensados en oficina de farmacia será necesario considerar el PVP con IVA menos la aportación del paciente y en el caso del entorno hospitalario el precio de compra, que será el PVL con IVA, salvo aquellos medicamentos que tengan un precio financiado (el cual no es público y por tanto es difícil de transparentarlo) y si tienen un descuento aplicando ese descuento. El problema es que este descuento tampoco es pú-

blico y varía de unos lugares a otros y de unos entornos a otros. Por la dificultad que entraña hacer públicos precios que no son públicos, algunas organizaciones como la Sociedad Española de Farmacia Hospitalaria, en sus informes GENESIS, utilizan el PVL con IVA en el análisis basal y utilizan otros posibles precios en el análisis de sensibilidad (9). También será necesario considerar si existen reducciones de los precios de aplicación obligatoria en base a la legislación. Por ejemplo, para los medicamentos con cargo al SNS, la aplicación de las deducciones del Real Decreto 8/2010 (10). Un aspecto adicional a considerar es que los costes también pueden variar con el tiempo, no solo por el ajuste temporal sino también porque cambian los precios unitarios por la aparición de genéricos, pérdidas de patentes, nuevas indicaciones, etc.

Los costes que no son medicamentos se debe indicar el año y ajustar por el índice de precios al consumidor (IPC).

3.1. Fuentes de obtención de costes unitarios

Algunos ejemplos de fuentes para la obtención de los costes unitarios de medicamentos se recogen más adelante en el apartado 6.

Para obtener costes de pruebas, costes por GRD, costes de procesos, etc, existen bases de datos (algunas indicadas en el apartado 6) o tarifas publicadas por las Comunidades Autónomas. También se pueden obtener de los propios hospitales o de publicaciones. Por lo que se refiere a las tarifas de las CCAA, son en general tarifas para financiadores privados, frecuentemente superiores a los costes reales.

Es importante saber cómo se han calculado cada uno de estos costes o tarifas, lo cual puede ser complejo, ya que la metodología de cálculo no es homogénea y no siempre transparente, y no todos los datos son de fácil acceso o no lo son en el grado de detalle necesario para conocer qué conceptos incluyen.

Oliva J y col. (11) analizaron las ventajas e inconvenientes de disponer de un repositorio oficial y centralizado de costes unitarios frente a realizar valoraciones específicas por centro o por Comunidad Autónoma, donde se observa la contraposición entre validez externa y validez interna. Resaltan la importancia de estandarizar los métodos de valoración de los costes unitarios y establecer controles de calidad de los mismos y presentar recursos separados de costes unitarios, como ya hemos comentado. Indican que un repositorio oficial de costes unitarios sería de gran utilidad en el proceso técnico de las evaluaciones económicas. No obstante, debería analizarse la traducción práctica de la tensión entre validez interna y externa de los trabajos.

3.2. Coste de la pérdida de productividad

En muchas ocasiones se hace referencia a los costes de pérdida de productividad como costes indirectos, aunque este término es confuso y puede incluir otros costes también. Por ello, es preferible llamarlos costes derivados de la pérdida de productividad.

Generalmente hacen referencia a la imposibilidad de participar en actividades con productividad por parte del paciente o familiares como consecuencia de la enfermedad. Sólo se deberían considerar aquellas pérdidas de productividad que son realmente una consecuencia de la alternativa valorada.

Cuando la evaluación económica utiliza medidas de resultado como los años de vida o los años de vida ajustados por calidad (AVAC), los costes de mortalidad no deben ser valorados, ya que este impacto ya es tenido en cuenta en la medida de efectividad (5).

En el caso de que se trate de productividad, pero no de absentismo, se puede calcular la equivalencia en días de absentismo de esa pérdida de productividad.

Existen dos métodos más extendidos para su valoración, el método del capital humano y el de los costes de fricción o costes coyunturales (2).

El enfoque del **capital humano**, utiliza los salarios como medida de la pérdida de producción para la sociedad y el enfoque de los **costes de fricción** tiene en cuenta el desempleo y la sustitución del trabajador, proporcionando una medida más reducida de los costes (5).

En el primero de ellos, se cuantifican la reducción de las ganancias del paciente y viene determinado por el salario del paciente y, por tanto, perjudica a las personas con salarios menores o los que no reciben salario. El segundo se calcula en función del coste que la empresa necesita para sustituir al trabajador enfermo. En este caso el coste también será más elevado cuanto mayor es la especialización del trabajador. Por ello, en la práctica se utiliza en muchas ocasiones el **salario medio** por la pérdida de productividad remunerada o no del enfermo.

Oliva J y col. (11) analizaron las ventajas e inconvenientes de las aproximaciones alternativas a la valoración monetaria de las pérdidas laborales, en concreto del método del capital humano y del método de los costes de fricción, dando ambos enfoques estimaciones diferentes. Resaltaron que se debe tener especial cuidado en valorar tanto el tiempo de trabajo remunerado, como el no remunerado y el tiempo de ocio desplazados por las tareas de cuidado. Concluyen que se debería presentar los costes de tiempo remunerado diferenciando el tiempo de trabajo perdido por ausencias laborales, por productividad reducida (presentismo) y el tiempo de trabajo que la persona utiliza para recibir servicios sanitarios. El método de su estimación aún no está claro cuál es el más adecuado.

Huben K y col. (12), indican que existen múltiples instrumentos para medir la pérdida de productividad, pero la mayoría no sirven para medir la pérdida de productividad para incorporarla a una evaluación económica desde la perspectiva social. Identificaron 42 instrumentos, la mayoría no

proveen información sobre la discapacidad a largo plazo infra-estimando la pérdida de productividad a lo largo de la vida del paciente, por lo tanto, limitando su utilidad en las evaluaciones económicas. Otra limitación frecuente es que se basan en recordar hechos aumentando el riesgo de error o sesgo, cuanto más largo plazo más riesgo de errores en las estimaciones. En la mayoría de los casos eran los propios pacientes quienes recogían los datos. Consideran que son necesarias guías nacionales e internacionales en la elección del mejor instrumento de modo que las estimaciones sean completas y puedan servir para comparaciones y transferencia de resultados.

Drummond M y col. (4) identifican 4 cuestiones de preocupación o controversia al respecto y dan unas recomendaciones para su manejo. Las cuestiones de controversia hacen referencia a:

- 1) La estimación del cambio en la productividad y el uso de salarios medios. Se puede sobreestimar por diferentes motivos, parte del trabajo se recupera cuando el trabajador vuelve al trabajo o es asumido por sus compañeros, algunos trabajos son de valor inferior al salario medio, si no pierde el trabajo la productividad también puede ser menor, etc.
- 2) Contar doble. La mayor parte de las veces cuando se mide la calidad de vida no se informa a los entrevistados de que no consideren los efectos sobre la pérdida de productividad, y pueden tenerlo en consideración, por lo que, si además se considera aparte, se estarían contando por doble.
- 3) Los objetivos y la perspectiva del análisis. En muchas ocasiones se realiza el análisis desde el punto de vista o perspectiva del sector sanitario y no es relevante su inclusión
- 4) Consideración de la equidad. La inclusión de estos costes, puede hacer que se destinen más recursos al sector sanitario.

Sus sugerencias son:

- 1) Reportar los cambios en productividad de forma separada.
- 2) Reportar cantidades separadas de precios.
- 3) Considerar si las ganancias/salarios realmente reflejan el valor de la pérdida de producción o si es más válido considerar el coste de restaurar la producción.
- 4) Prestar atención a las implicaciones en la equidad, se puede hacer un análisis de sensibilidad incluyendo estos aspectos.
- 5) Considerar si se puede estar contando algo por doble.
- 6) Considerar las recomendaciones oficiales del lugar en donde se realiza el análisis.

3.3. Coste de cuidados informales

Se denominan cuidados informales (2) a la prestación de cuidados a personas dependientes por parte de la familia, amigos, vecinos u otras personas (que no reciben ninguna compensación económica por la ayuda que proporcionan) o bien por parte de cuidadores especializados, a los que les paga la familia del paciente. Estos costes se dividen en 5 componentes: cuidado del hogar, cuidado personal (vestirse, alimentarse, etc), ayudas para movilidad, ayudas con tareas administrativas, y ayudas para desenvolverse en el ámbito social.

El coste de los cuidados informales es difícil de valorar. Algunos podrían argumentar que es cero, otros equipararlo al salario medio, otros al salario de horas extras, otros lo que costaría contratar a alguien que hiciera esa labor, otros con lo que habría que compensar a una persona por quitarle ese tiempo (4). El valor de este tiempo dependería de lo que se deja de hacer en términos de trabajo remunerado, trabajo no remunerado u ocio. En muchas ocasiones, en las evaluaciones económicas no se está incorporando ese valor, pero al menos sería deseable que se identificara y se tuviera en consideración al tomar las decisiones. Otra aproximación diferente sería considerar el efecto en el cuidador como

una disminución en la calidad de vida del cuidador, y es posible que esta merma en el cuidador no sea capturada por los AVAC.

Existen dos métodos más utilizados para su valoración, el de la preferencia declarada y el de la preferencia revelada (2). El método de la **preferencia revelada**, generalmente calculada mediante el coste de oportunidad (identificar y valorar el coste que supone para el cuidador dedicar su tiempo a cuidar al enfermo, en lugar de trabajar u ocio), y el coste de sustitución (el precio de un servicio sustitutivo). Dentro de los métodos de la **preferencia declarada**, los métodos más empleados son el de la valoración contingente (la mínima cantidad que el cuidador está dispuesto a pagar por el servicio o dispuesto a percibir por estar dispuesto a prestar el servicio) y el análisis de conjunto (estimar las preferencias de uno o varios sujetos a partir de una clasificación ordinal proporcionada por técnicas de encuesta) (2).

En cuanto a estos costes, Oliva J y col. (11) indican que existen muchas preguntas abiertas sobre los mejores métodos de medición del tiempo de cuidado, de efectos positivos (sobre cuidadores y pacientes) y negativos (sobrecarga de cuidadores) (11). Existen 3 métodos: coste oportunidad, método de coste de reemplazo, y valoración contingente (disposición a pagar/disposición a ser compensado) (11). También apuntan que es necesario generar más información sobre su relevancia. La valoración del tiempo del cuidado es solo una parte, los métodos de valoración contingente serían los más sólidos en cuanto a la teoría económica, sin embargo, no resulta sencillo separar la valoración del tiempo de cuidado *per se* de la valoración que hacen los cuidadores de otros aspectos positivos o negativos asociados a las propias tareas de cuidado. Los costes de oportunidad y costes de reemplazo son alternativas razonables para valoración del tiempo de cuidado. Los autores proponen líneas de mejora o avance.

Un aspecto importante es no contabilizar por doble, por ejemplo, como pérdida de productividad por parte del cuidador y como coste de cuidado informal.

3.4. Costes de equipamiento e instalaciones

Son costes que requieren una inversión, además tanto el equipamiento como las instalaciones se van a usar a lo largo del tiempo y su valor se va depreciando con el tiempo. Por un lado, está el coste de oportunidad de esta inversión, lo que se ha invertido en ello no está disponible para ser utilizado en otra alternativa. Y el otro componente es la depreciación con el tiempo. El método más extendido para su consideración es el de anualizar la inversión inicial en los años de vida útil del equipamiento o lo que se trate, incorporando en este tanto el coste de oportunidad como la depreciación (4). Si lo valorado se usa para varios programas, el coste se distribuirá entre ellos del mismo modo que se hacen los costes generales o comunes.

3.5. Costes compartidos entre varios programas o alternativas

En general se recomienda el uso de los costes marginales. Ello requiere analizar si alguno de estos costes cambiaría si se añade un programa o alternativa de uso o la cantidad de actividad realizada. La distribución de costes comunes se puede hacer en función del uso que se hace de ese bien (4). Una aproximación frecuente es, por ejemplo, contabilizarlo en el coste diario del hospital, ejemplo por día de ingreso, hora en hospital de día, etc. Cuando se requiere mayor detalle, existen diferentes métodos para su asignación (4). El esfuerzo en recogerlos de forma detallada va a depender del impacto que tenga en el análisis que se esté realizando. Es importante, en este caso también, **no contar por doble**, si unos costes se han considerado por ejemplo en el coste de un día de hospitalización no volver a contarlos de forma separada. Por ello, es importante cuando se incluyen costes que engloban varios aspectos conocer realmente los componentes considerados en ese coste.



Ajuste temporal

Dado que los costes de las alternativas no ocurren todos en el mismo momento en el tiempo es necesario, para poder compararlos, **calcular el valor actual** de todos los costes, para ello se realiza el ajuste temporal.

Los costes futuros se reducen o descuentan para calcular su valor actual, los costes futuros no pesan tanto como los costes o ahorros actuales, esto se debe a que hay una **preferencia temporal**, como individuos y como sociedad preferimos retener los recursos liberados ahora para poder sacar beneficio de ellos (4). Este hecho responde a que la sociedad prefiere retasar los costes.

Por ello se calcula el valor actual de los costes futuros aplicando una tasa de descuento.

$$VA = VF / (1+r)^t$$

Siendo VA el valor actual, VF el valor futuro, r la tasa de descuento y t el tiempo en que ocurre ese coste (tiempo desde el momento actual hasta que sucede ese coste).

La tasa de descuento ha de reflejar la tasa de preferencia social (2). Las tasas de descuento recomendadas oscilan entre el 0 y el 5%, se seleccionará una para el análisis basal, por ejemplo, el 3%, y en el análisis de sensibilidad se emplearán otras en este rango.



Metodologías en las guías europeas

García-Mocho L y col. (13) han revisado recientemente las metodologías utilizadas en las guías europeas. Analizaron 41 documentos de 26 países europeos que correspondían a guías de evaluación económica o a guías de costes. Observaron que existe una gran disparidad entre las recomendaciones de estimación de los costes directos para ser incluidos en las evaluaciones económicas. Esta falta de estandarización influye en la arbitrariedad en la selección de los costes y en la toma de decisiones. Y es un reto poder identificar los factores que influyen en la variabilidad de los costes unitarios entre países. Existen divergencias tanto en la estimación del uso de recursos como en la valoración monetaria. En general, el reporte de los costes es pobre en Europa. Los costes indirectos o costes sociales son en general excluidos.

Estos autores apuntan que en relación a los costes futuros hay consenso en que los costes generados por enfermedades no relacionadas con el servicio sanitario dado, los costes que emergen en la vida prolongada por la terapia pero que no se deben a la enfermedad examinada en el análisis, u otros costes indirectos, se pueden no incluir o, en casos justificados, se pueden incluir en análisis adicionales. Los costes evitados o desplazados solo se mencionan en la guía irlandesa.

Identificaron dos metodologías para la identificación de los costes: micro-coste (a nivel muy detallado) o coste bruto (a nivel muy agregado). La medida de los recursos se puede obtener de dos tipos de fuentes: primaria (han sido recogidos específicamente para ese fin) o secundarias (recogidos para otro fin). Los datos en general se obtienen de bases de datos, registros de pacientes, ensayos clínicos, estudios observacionales y datos de la práctica clínica habitual. Si los datos primarios no están disponibles se pueden obtener de fuentes secundarias, como sistemas de información financiera de las instituciones sanitarias o bases de datos

de aseguradoras. En algunos países abren también la posibilidad de usar paneles de expertos, aunque en algunos solo como información complementaria y en otros estos paneles se tienen que evitar.

En cuanto a la valoración de estos recursos, indican que se pueden usar dos aproximaciones: de abajo-arriba (los recursos consumidos por cada paciente son multiplicados por los costes unitarios para obtener el coste total por paciente) o de arriba-abajo (el coste total hospitalario es dividido en costes de objetos usando una métrica definida como guiador del coste, en general coste por GRD o coste medio per diem). Entre las posibles formas de asignar los costes unitarios: coste-oportunidad (preferible pero difícil), precios en el mercado, tarifas, tasas administrativas, medida directa, cálculo de precios sombra, etc., la mayoría de los países están de acuerdo que los precios de mercado, en mercados perfectamente competitivos, refleja los costes de oportunidad, pero los mercados no son perfectos por ello se usan diferentes formas de valorar. Las guías españolas señalan diferentes fuentes de costes unitarios, pero no indican el orden de preferencia.

En cuanto a las tasas de descuento, tanto los costes como las consecuencias a futuro (más de un año) se tienen que descontar para reflejar la preferencia temporal de la sociedad. La mayor parte de los países utilizan tasas de descuento entre el 3 y 5% (13).

En cuanto al método para actualizar costes para el año y moneda relevante en el análisis, la forma más simple es aplicar la inflación y los ajustes de cambio de moneda (13). Añadir que, en España, debido a como está regulado el precio de los medicamentos, esta actualización no se aplica a los costes de los medicamentos, pero sí a otros costes.

Los autores (13) resaltan que el cálculo de los costes consume mucho tiempo y recursos, y que disponer de listas nacionales de costes unitarios usando metodología aceptada internacionalmente sería muy útil. Las guías deben recomendar de forma muy clara la metodología a seguir. Y se deben estandarizar también las definiciones.



Fuentes de datos de costes

1. Precios de medicamentos

- BotPlus: <https://botplusweb.portalfarma.com/>

Se trata de la base de datos del Consejo General de Colegios farmacéuticos. Incluye información sobre medicamentos. En ella se puede obtener, entre otra información, los precios de venta del laboratorio (P.V.L.) y los precios de venta al público (P.V.P.) de los medicamentos.

- Nomenclator: <https://www.sanidad.gob.es/profesionales/nomenclator.do>

Información disponible en la web del Ministerio de sanidad sobre los productos incluidos en la prestación farmacéutica del SNS (dispensables a través de oficinas de farmacia). Proporciona, entre otra información el PVP.

2. Costes y/o tarifas de procesos

- eSalud: <http://esalud.oblikue.com/>

Es una base de datos que requiere una cuota de suscripción para acceder al contenido. Incluye, entre otra información, costes y tarifas publicadas, tarifas de Comunidades Autónomas, costes por GRD, costes por proceso, costes por paciente y patología, costes indirectos, incluyendo la fuente y el año y la actualización al año de búsqueda.

- Boletines oficiales de las CCAA
- Publicaciones, por ejemplo, se pueden identificar en motores de búsqueda, generales, como PubMed (<https://pubmed.ncbi.nlm.nih.gov/>) que

permite acceso a bases de datos de bibliografía médica como Medline, la base página de acceso CRD de la Universidad de York, Center Review and Dissemination (<https://www.crd.york.ac.uk/CRDWeb/>) que incluye acceso a diferentes bases de datos como DARE, NHS EED y HTA, que recogen publicaciones con datos de costes. Es importante hacer una evaluación crítica de las publicaciones tanto de su validez interna como externa.

- Página Web del Ministerio de Sanidad. Dispone de diferentes datos de costes. Por ejemplo, costes hospitalarios y ambulatorios. Un ejemplo de costes por GRD a nivel hospitalario: <https://www.sanidad.gob.es/estadEstudios/estadisticas/inforRecopilaciones/anaDesarrolloGDR.htm>; otro ejemplo <https://pestadistico.inteligenciadegestion.sanidad.gob.es/publicoSNS/C/rae-cmbd/rae-cmbd/diagnosticos-principales/diagnosticos-principales>

3. Otros costes

Costes del salario medio en España, el índice de precios al consumo, etc. Se pueden encontrar en la página web del Instituto Nacional de Estadística (<https://www.ine.es/>).



Incertidumbre

Es importante conocer la incertidumbre en torno a los costes para después incorporarla en el **análisis de sensibilidad**. Como se ha comentado anteriormente los costes, entendiendo por tal la valoración monetaria de los recursos consumidos, no siguen en general una distribución normal.

Por ello, **si se dispone de los datos de costes individuales de una muestra de individuos**, para calcular la diferencia de costes entre dos alternativas y la incertidumbre en torno a esta diferencia, generalmente se utilizan métodos no paramétricos. Entre ellos, uno de los más utilizados, como se ha señalado en un apartado anterior, es el uso de técnicas de remuestreo, como **bootstrap**. Esta técnica, genera un número de muestras elevado a partir de los costes individuales de los pacientes con un número de datos en general igual al de la muestra de partida, la selección de los datos en cada muestra es aleatoria y existe la posibilidad de repetición. Se genera así una distribución empírica del parámetro en cuestión. La distribución de la media de los costes o diferencia de costes de las muestras generadas seguirá una distribución normal y podrán utilizarse técnicas paramétricas para su análisis estadístico.

Si no se dispone de los datos individuales, como ocurre en la mayoría de los modelos que parten de datos publicados, generalmente se dispone de algunos valores de su distribución. Se puede definir qué tipo de distribución siguen los datos de costes, en el caso que nos ocupa. En muchas ocasiones se asume que los costes siguen una distribución gamma (14). Cuando existe poca información en ocasiones se asume una distribución triangular partiendo de unos valores indicados por expertos en torno a los valores entre los que se podrían encontrar los valores reales. Pero el uso de esta última técnica de estimación se debería evitar e intentar buscar datos para fundamentar las estimaciones. Posteriormente para la comparación de los datos de costes entre diferentes alternativas, como no siguen distribuciones normales, se utilizan técnicas como las **simulaciones de Montecarlo**. Las simulaciones de Montecarlo permitirán la estimación de parámetros y su distribución de probabilidad. Cuando se utilizan técnicas de simulación es necesario comprobar que los datos resultantes sean reales, por ejemplo, no haya resultados negativos en los costes.



Transferibilidad de los costes entre países

La transferibilidad de costes entre países es difícil. En general, hay más diferencia entre países en los costes unitarios que en los recursos consumidos, particularmente entre países con sistemas sanitarios similares. En el caso de que se utilizaran costes de otros países, se debería justificar la comparabilidad de los costes.



Bibliografía

1. Casado Gómez MA, Ortega Eslava A, Sacristán del Castillo JA, et al. Capítulo 33. Evaluación económica de medicamentos. En: Ruiz Gayo M. y Fernandez-Alfonso M, editores. Fundamentos de Farmacología Básica y Clínica. 2ªEd. Madrid: Médica Panamericana; 2013;p. 443-56.
2. Soto Alvarez J. Tipos de costes a incluir en una evaluación económica. En: Soto Alvarez J. Evaluación económica de medicamentos y tecnologías sanitarias: Principios, métodos y aplicaciones en política sanitaria. 1ª Ed. Madrid: Springer Health Care Iberica; 2012. p. 37-52
3. Casado Gómez MA. Análisis parciales de evaluación económica de medicamentos: descripción de costes, estudios de coste de la enfermedad, análisis coste-consecuencias, análisis de costes. Análisis de minimización de costes. En: Dominguez-Gil Hurlé A, Soto Alvarez J, coordinadores. Farmacoeconomía e investigación de resultados en salud: principios y práctica. Situación actual y perspectivas futuras en España. 1ª edición. Real Academia Nacional de Farmacia. Europa Artes Gráficas; 2002.
4. Drummond MF, Sculpher MJ, Claxton K, et al. Methods for de economic evaluation of health care programmes. 4ª Ed. New York: Oxford University Press; 2015.
5. Puig-Junoy J, Ortún-Rubio V, Pinto-Prades JL. Los costes en la evaluación económica de tecnologías sanitarias. Aten Primaria. 2001;27:186-9.
6. Ortega Eslava A, Fraga Fuentes MD, López Briz E. Lectura crítica de evaluaciones económicas en salud. En: Cabello López JB por CASPe. Lectura crítica de evidencia clínica. 2ª Ed. Barcelona: Elsevier; 2021. p. 196-212.
7. Prieto L, Sacristan JA, Pinto JL, et al. por grupo ECOMED. Análisis de costes y resultados en la evaluación económica de las intervenciones sanitarias. Med Cin (Barc). 2004;122:423-9.
8. Spacirova Z, Epstein D, García-Mochon L, et al. A general framework for classifying costing methods for economic evaluation of health care. Eur J Health Econ. 2021;21:529-42.

9. Ortega Eslava A, Marín Gil R, Fraga Fuentes MD, et al. (GENESIS-SEFH). Guía de evaluación económica e impacto presupuestario en los informes de evaluación de medicamentos. Guía práctica asociada al programa MADRE v 4.0. Madrid: SEFH (ed.); 2016. https://gruposde-trabajo.sefh.es/genesis/genesis/Documents/GUIA_EE_IP_GENESIS-SEFH_19_01_2017.pdf (acceso verificado el 25 Sept 2023)
10. Real Decreto-ley 8/2010, de 20 de mayo, por el que se adoptan medidas extraordinarias para la reducción del déficit público. Jefatura del Estado «BOE» núm. 126, de 24 de mayo de 2010 Referencia: BOE-A-2010-8228. <https://www.boe.es/buscar/pdf/2010/BOE-A-2010-8228-consolidado.pdf> (acceso verificado el 25 Sept 2023)
11. Oliva J, Brosa M, Espin J, et al. Cuestiones controvertidas en evaluación económica (I): perspectiva y costes de intervenciones sanitarias. *Rev Esp Salud Pública*. 2015;89:5-14.
12. Huben K, Krol M, Coast J, et al. Measurement instruments of productivity loss of paid and unpaid work: a systematic review and assessment of suitability for health economic evaluations from a societal perspective. *Value Health*. 2021;24:1686-99.
13. García-Mochón L, Spacirova Z, Espin J. Costing methodologies in european economic evaluation guidelines: commonalities and divergences. *Eur J Health Econ*. 2022;23:979-91.
14. Briggs A, Claxton K, Sculpher M. *Decision modelling for health economic evaluation*. New York: Oxford University Press; 2006.



4

OBTENCIÓN DEL VALOR DE UTILIDAD DE LAS ALTERNATIVAS EN EVALUACIÓN EN EL MODELO ANALÍTICO DE DECISIÓN

José María Abellán Perpiñán

Catedrático de Economía Aplicada de la Universidad de Murcia



Introducción

La evaluación basada en modelos analíticos de decisión puede definirse como “una comparación cuantificada formal de tecnologías sanitarias, sintetizando las fuentes de evidencia sobre costes y beneficios, con el fin de identificar la mejor opción que pueden adoptar los decisores” (1). En este capítulo vamos a centrar nuestra atención en los beneficios o resultados de las tecnologías sanitarias objeto de comparación en el marco de un determinado modelo. Más concretamente nos centraremos en los resultados cuantificados en forma de Años de Vida Ajustados por Calidad (AVAC), métrica que resume en un único valor numérico cantidad y calidad de vida relacionada con la salud (CVRS) (2). Si bien la cantidad de vida es un dato objetivable, la CVRS es subjetiva, resultado de la valoración que atribuyen los pacientes a su propio estado de salud o la población general a estados de salud hipotéticos. Sin embargo, no cualquier valoración o percepción subjetiva de los resultados de una enfermedad y sus tratamientos —lo que la Agencia Europea del Medicamento (3) define como resultados reportados por el paciente (en inglés, *patient-reported outcomes* o PROs)— es un índice o valor de utilidad, que es el tipo de valoraciones requeridas para poder calcular AVAC en el marco de los modelos coste-efectividad (4,5).

Las utilidades de estados de salud —también llamadas pesos de calidad de vida— son medidas cardinales de las preferencias individuales (6) que describen el valor otorgado a un estado de salud en una escala intervalo, esto es, una escala cuyos intervalos o diferencias entre cada dos estados reflejan la intensidad con que se prefiere uno sobre otro (7). Por esta razón la ganancia en AVAC que proporciona una nueva intervención con respecto al comparador correspondiente (p.ej. la práctica habitual) puede compararse con el incremento en costes que aquella supone, permitiendo así calcular razones incrementales coste-efectividad

(4). La escala en cuestión se define asignando, por convención, el valor 1 a la salud plena o perfecta y el valor 0 a la muerte. Los valores negativos representan, en consecuencia, estados tan intolerables que se juzgan peores que la muerte (8).

En las páginas siguientes explicaremos cómo pueden medirse las preferencias y obtener así utilidades; repasaremos cómo pueden identificarse las utilidades de estados de salud en la literatura y cómo juzgar su calidad metodológica, y concluiremos enumerando las principales limitaciones metodológicas que rodean a la medición de utilidades y a su uso para calcular AVAC.



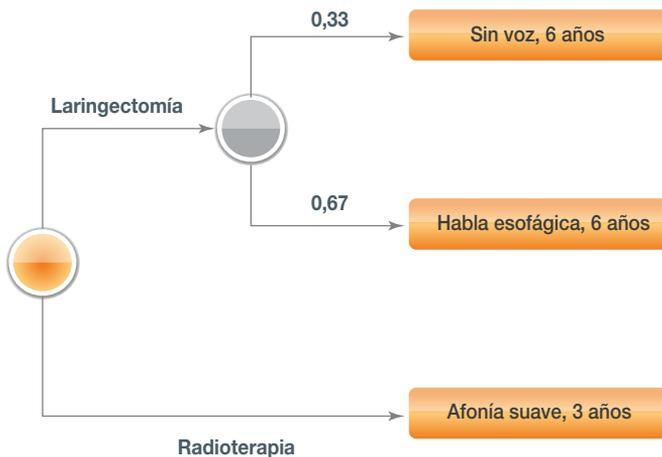
Metodología e instrumentos para su medición

Con el objeto de explicar de forma práctica la medición de utilidades en el marco de un modelo analítico de decisión, utilizaremos en lo sucesivo como ilustración el modelo de cohorte más simple: un árbol de decisión. Los modelos de cohorte caracterizan la experiencia del paciente “medio” en una población que comparte unas mismas características. Reproduciendo el ejemplo adaptado por Abellán (9), el árbol de decisión de la Figura 1 representa el análisis de la decisión entre dos tratamientos para un paciente “medio” de cáncer de laringe (estadio T3, no hay metástasis). Ver Figura 1.

Como puede observarse, los resultados (*outcomes*) del árbol son binarios, compuestos por dos atributos: CVRS y años de vida. Esos dos atributos, como ya señalamos con anterioridad son, precisamente, los requeridos para computar AVAC. La metodología y supuestos implícitos para efectuar dicho cómputo difieren dependiendo de que la CVRS permanezca constante hasta su fallecimiento (perfil o estado de salud crónico) o bien, como es más realista suponer, varíe a lo largo del tiempo (perfil de salud variable).



Figura 1. Dos tratamientos alternativos para el cáncer de laringe



Fuente: elaboración propia a partir de (10)

En el primer caso, los AVAC son el producto de dos factores: los años de vida (en adelante denotados como T) y la utilidad (o bienestar) asociada a la CVRS (denotada como Q) en la que se va a vivir durante esos años. Esta combinación la podemos representar de forma muy simple del siguiente modo:

$$AVAC = U(Q, T) = V(Q) \times T \quad [1]$$

T= años de vida; Q= calidad de vida relacionada con la salud; U(Q, T)= utilidad o bienestar derivado de vivir T años con la calidad de vida Q; V(Q)= utilidad o peso de calidad de vida asociada a Q.

La anterior ecuación puede extenderse a situaciones en las que la CVRS o estado de salud no permanece constante, sumando simplemente tantos productos $V(Q) \times T$ como variaciones experimente Q a lo largo del tiempo. Esto es, si fraccionamos el horizonte vital del paciente en tantos

intervalos temporales disjuntos (no solapados) como ocasiones cambia Q hasta producirse la muerte de aquel, el cálculo de los AVAC se reduce a:

$$AVAC = \sum_{t=1}^n V(Q_t) \times T_t \quad [2]$$

n= número de intervalos temporales disjuntos; t= orden en que se suceden los n intervalos.

Una vez descrito cómo calcular los AVAC vinculados a una determinada intervención sanitaria, podemos concretar en cinco los pasos a seguir para poder comparar los AVAC producidos por dos intervenciones mutuamente excluyentes^a como las representadas en la Figura 1:

1. Describir o caracterizar la CVRS representada por Q.
2. Obtener la utilidad de Q; esto es, medir $V(Q)$.
3. Multiplicar dicha utilidad por los años de vida T .
4. Calcular el número total de AVAC resultante, bien mediante la ecuación 1, bien mediante la ecuación 2, según que el perfil de salud sea constante o variable.
5. Hallar la diferencia entre los AVAC aportados por el nuevo programa y los AVAC aportados por el programa utilizado como comparador (la práctica habitual o *statu quo*).

De estos cinco pasos, los tres últimos son triviales^b, razón por la cual nos centraremos fundamentalmente en las dos primeras fases de la secuencia.

^a Las dos alternativas terapéuticas de la Figura 1 se plantean como mutuamente excluyentes para un mismo paciente: o recibe una o recibe otra. La ratio coste-efectividad incremental que finalmente se obtenga de la comparación de ambas terapias puede luego compararse, siguiendo diferentes reglas de decisión, con las de otras intervenciones independientes (p.ej. vacunas para el virus del papiloma humano, programas de cribado de cáncer de colon, CAR-T, etc.) que, a su vez, tendrán diferentes variantes mutuamente excluyentes. Véase (11) para una descripción en profundidad de dichas reglas de decisión.

^b Los tres últimos pasos son triviales asumiendo, claro está, el modelo AVAC encarnado por las ecuaciones [1] y [2], que trata el tiempo de vida de forma lineal. Por esta razón, este modelo recibe el nombre de modelo AVAC lineal (12) o, simplemente, modelo AVAC (13).

Paso 1: la caracterización de la CVRS o estado de salud

Hay básicamente dos maneras de describir el estado de salud cuya utilidad va a estimarse en el segundo de los cinco pasos antes enunciados. Un método habitual consiste en recurrir a las denominadas PROMs (*patient-reported outcome measures*) o medidas de resultados reportados por los pacientes: cuestionarios estandarizados y validados que describen los resultados en salud de los pacientes mediante una combinación de atributos representativos de su CVRS, o de síntomas específicos, tales como dolor, fatiga, funcionamiento sexual o de sucesos adversos relacionados con los tratamientos o de una combinación de todas las dimensiones referidas (14). En aquellas situaciones en las que no es posible o adecuado recurrir a descripciones estandarizadas de la CVRS de los pacientes se emplean viñetas. Una viñeta puede describir una condición médica, su tratamiento y su impacto en varios dominios de la CVRS. Las viñetas reciben el nombre de “escenarios”, “descripciones del estado de salud”, “viñetas del estado de salud” o simplemente “estados de salud” (15).

Las PROMs pueden clasificarse en genéricas o específicas de una condición o enfermedad determinada (16). Las primeras abarcan dimensiones genéricas de la CVRS, como funcionamiento físico y social, dolor y depresión o ansiedad, lo que las hace adecuadas para efectuar comparaciones entre diferentes enfermedades. Obviamente, esa mayor versatilidad tiene como contrapartida, en comparación a instrumentos específicos, una menor sensibilidad a cambios clínicamente significativos del estado de salud ocasionados por el tratamiento de una enfermedad en particular. Son ejemplos de PROMs genéricas, entre muchas otras existentes, el cuestionario SF-36 (Short-Form 36-Item) en las dos versiones disponibles (17-19) y el autclasificador EQ-5D, también en sus dos versiones (20-22).

En el primer caso, el cuestionario SF-36 recibe ese nombre porque abarca 36 ítems o preguntas agrupados en 8 dimensiones rela-

cionadas con la salud^c. La respuesta dada a cada ítem recibe una puntuación, agregándose con las puntuaciones correspondientes a los restantes ítems de una misma dimensión. La puntuación agregada de cada dimensión se normaliza, de modo que en el SF-36v1 la puntuación normalizada de cada dimensión oscila entre 0 (peor estado de salud) y 100 (mejor estado de salud), mientras que en el SF-36v2 la puntuación normalizada es interpretable directamente en referencia a normas poblacionales, de modo que una puntuación superior a 50 se interpreta como un estado de salud mejor que el de la media de la población de referencia, mientras que una puntuación inferior a 50 se interpreta como un estado de salud peor que el del promedio de la población de referencia^d. Si bien no puede obtenerse una puntuación sintética global de las 8 dimensiones del cuestionario SF-36, sí que pueden calcularse dos puntuaciones resumen, la componente resumen física (PCS) y la componente resumen mental (MCS), nuevamente interpretables en relación a valores poblacionales de referencia.

Si bien, como acabamos de ver, los cuestionarios SF-36 poseen un algoritmo de puntuación que genera medidas representativas del estado de salud de cada una de las 8 dimensiones, así como las dos agregaciones resumen PCS y MCS, no cabe interpretar dichas puntuaciones como medidas de preferencias o utilidades, de modo que no son susceptibles de utilización en la evaluación económica de programas e intervenciones sanitarias, ya que no pueden combinarse con la expectativa de vida y así obtener AVAC. La razón fundamental que impide computar AVAC a partir de las puntuaciones del SF-36 es que los supuestos que subyacen a la interpretación y agregación de

^c Hay otros dos cuestionarios Short-Form originados a partir de la selección de 12 y 8 ítems respectivamente de los 36 que alberga el SF-36. Dichos cuestionarios se denominan SF-12 (23,24) que, al igual que el SF-36 posee dos versiones, y SF-8 (25).

^d Hay normas poblacionales para ambas versiones del cuestionario SF-36, estimadas a partir de diferentes muestras de la población general estadounidense (26). Se disponen, así mismo, valores poblacionales de referencia para la versión española del SF-36v1 (27).

las respuestas otorgadas a cada ítem no están basados en las preferencias de los encuestados^e.

El EQ-5D, por su parte, es una medida estandarizada del estado de salud desarrollada por el grupo EuroQol. Al igual que en el caso del SF-36 hay dos versiones del EQ-5D, la originaria, denominada EQ-5D-3L porque describe el estado de salud del individuo como una combinación de cinco dimensiones y tres niveles de gravedad por dimensión (28), y la versión más reciente, EQ-5D-5L, que amplía a cinco el número de niveles por dimensión (29). Ambos instrumentos constan básicamente de dos páginas ideadas para ser autocompletadas por el entrevistado: el componente descriptivo EQ-5D en sí mismo, el cual consta de las dimensiones movilidad, cuidado personal, actividades cotidianas, dolor/malestar y ansiedad/depresión, admitiendo cada una de ellas tres o cinco niveles, dependiendo de que el instrumento sea el EQ-5D-3L o el EQ-5D-5L, y una escala visual analógica o “termómetro” graduado de 0 (peor estado de salud imaginable) a 100 (mejor estado de salud imaginable), denominada EQ-VAS, en la cual debe situar y puntuar el encuestado su estado de salud, que previamente habrá clasificado de acuerdo al correspondiente componente descriptivo. Por ejemplo, el estado de salud codificado como 21111 significa, en el caso del instrumento EQ-5D-3L, que el sujeto tiene algunos problemas para caminar, mientras que en el caso del instrumento EQ-5D-5L, se interpreta como que tiene ligeros problemas para caminar.

^e Por ejemplo, en el cuestionario SF-36 se asume que cuando un encuestado afirma que el dolor le ha dificultado su trabajo habitual “un poco”, la distancia entre ese nivel de la escala y el nivel “nada”, es el mismo que existe entre los niveles “bastante” y “mucho”. Parece claro que este supuesto dista mucho de ser auto-evidente, ya que dicha igualdad en la ponderación entre niveles adyacentes de respuesta dependerá de cuál sea la verdadera intensidad de preferencias del sujeto. Su preferencia por pasar de tener “un poco” de dolor a no tener “nada” de dolor podría ser, por ejemplo, mayor (más intensa) que su preferencia por pasar de tener “mucho” dolor a tener “bastante” dolor.

Paso 2: la medición de la utilidad del estado de salud

Una vez caracterizada la CVRS, hay que medir su utilidad. En el ejemplo concreto que describe la Figura 1, tres son los estados de salud a valorar (de mayor a menor gravedad): sin voz, habla esofágica y afonía suave.

La utilidad de estos y cualesquiera otros estados de salud puede obtenerse de dos maneras. La primera consiste en medir directamente las preferencias de la población objeto de interés (en este caso los pacientes que comparten las características del paciente “medio” representado en el árbol de decisión) utilizando para ello técnicas o métodos de valoración diseñados al efecto. La segunda pasa por atribuir al estado de salud del paciente “medio” la utilidad predicha por un algoritmo o modelo de utilidad multiatributo estimado a partir de mediciones directas realizadas en amplias muestras de la población general.

Dentro de las medidas directas de preferencias, las predilectas en el ámbito de la economía de la salud son aquellas basadas en elecciones, como es el caso de la lotería estándar y el intercambio temporal (30), en contraposición a procedimientos que podríamos calificar como introspectivos, como es el caso de la escala visual analógica (31)^f. Otra forma de diferenciar los métodos directos de obtención de preferencias es atendiendo al tipo de decisión que evocan durante la medición; así, la lotería estándar estructura la decisión en un entorno de riesgo o incertidumbre, mientras que el intercambio de tiempos lo hace en un entorno de certeza. Hay otros métodos tam-

^f Los métodos basados en elecciones se denominan así porque la preferencia es el resultado de la comparación entre dos alternativas, emulando de esta forma el comportamiento del consumidor en el mercado. Por el contrario, la preferencia obtenida mediante técnicas como la escala visual analógica se fundamenta, en principio, en la introspección. Estos últimos métodos se consideran dotados de una menor validez (esto es, de una menor capacidad para medir realmente preferencias cardinales) ya que son muy sensibles a efectos contextuales (32,33), si bien el debate acerca de la naturaleza y propiedades de la escala visual analógica continúa abierto (34,35).

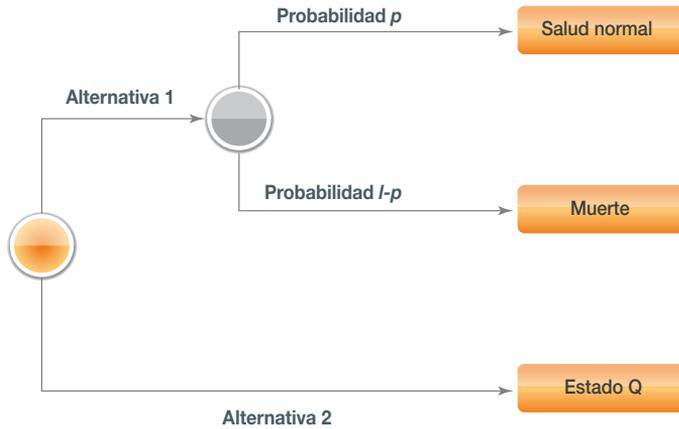
bién empleados para valorar estados de salud como el procedimiento del intercambio de personas (36), diferentes tipos de lotería estándar ‘modificada’, lotería equivalente o doble lotería (37-39), el método del ‘velo de la ignorancia’ (37) o la técnica de escalamiento ‘*best-worst*’ (40). No obstante, por ser las dos técnicas más ampliamente utilizadas, nos detendremos a continuación en explicar sucintamente cómo aplicar el método de la lotería estándar y del intercambio temporal para el caso de estados de salud crónicos, considerados mejores que la muerte.⁹

La lotería estándar compara dos alternativas (Figura 2). La alternativa 1 es un tratamiento con dos posibles resultados: o el paciente recobra su salud normal y vive durante t años más (probabilidad p) o bien muere (probabilidad $1-p$). El tratamiento incierto así descrito, por su semejanza con un juego de azar, recibe el nombre de ‘lotería’. En contraposición, la alternativa 2 ofrece con certeza la seguridad de vivir en el estado de salud crónico Q durante t años. Partiendo de este planteamiento, el analista varía la probabilidad p hasta que el entrevistado se declare indiferente entre ambas alternativas, fijándose la utilidad de Q igual a p .

⁹ La medición de la utilidad de estados de salud particularmente intolerables (considerados peores que la muerte) no está exenta de dificultades, si bien existen diferentes modos de abordarla. Uno de los principales problemas radica en que el ‘*framing*’ o formato de las técnicas tradicionales, como la lotería estándar y el intercambio temporal, para medir utilidades negativas es diferente al utilizado para valorar estados de salud mejores que la muerte (7), lo que da lugar a que la utilidad por debajo de 0 no esté acotada de forma simétrica a la utilidad por encima de ese mismo valor. Una forma de subsanar esta asimetría es reescalando la utilidad “cruda” (41), de forma que el rango de valores esté comprimido entre -1 y +1. Más recientemente, otros autores (42-44) han propuesto dos variantes de intercambio temporal que valora estados de salud de un modo simétrico, ya sean mejores o peores que la muerte. Ambas propuestas tienen, no obstante, sus problemas. En el primer caso, tras el reescalamiento, la utilidad resultante ya no representa fidedignamente las preferencias inicialmente obtenidas de los entrevistados (45). Por su parte, en el segundo caso, se ha constatado empíricamente que con frecuencia los entrevistados no son conscientes de que, según cuáles sean sus elecciones, estas implican que el estado de salud en cuestión lo están considerando como menos deseable que la muerte (46).



Figura 2. El método de la lotería estándar



Fuente: elaboración propia a partir de (7)

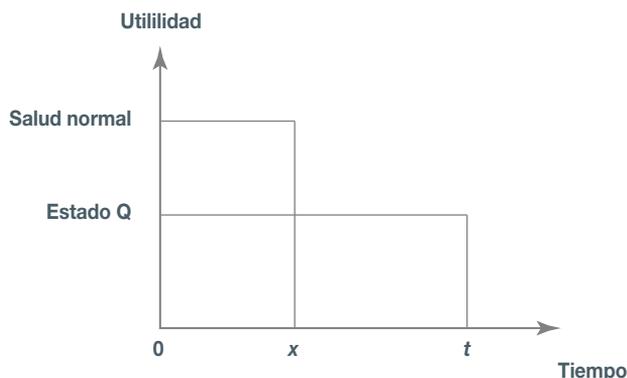
El intercambio temporal también plantea una elección entre dos alternativas (Figura 3). La alternativa 1 representa permanecer en el estado de salud Q (condición crónica considerada preferible a morir) durante t años, tras los cuales fallece, mientras que la alternativa 2 ofrece vivir x años ($x < t$) con salud normal. La duración x se va variando hasta que el individuo se muestra indiferente entre las dos alternativas, fijándose la utilidad de Q igual al cociente x/t .^h

Como alternativa a la medición directa de las preferencias de los pacientes, pueden asignarse a los estados de salud de estos las utilidades previamente estimadas por medio de instrumentos de utilidad multiatributo. Dos de los instrumentos multiatributo más ampliamente utilizados son los ya comentados EQ-5D y SF-6D (48-51). Estos instrumentos poseen dos componentes: un sistema de clasificación de estados de salud multi-

^h En los últimos años se han desarrollado métodos de intercambio temporal alternativos al tradicional expuesto en este capítulo. Así, tenemos los métodos conocidos como *lead-time trade-off* (L-TTO) (43) y *lag-time trade-off* (44) que emplean un mismo formato para medir utilidades positivas y negativas, y el método *composite TTO* (C-TTO) (47), que es un híbrido del intercambio temporal convencional (para los estados mejores que la muerte) y el L-TTO (para los estados peores que la muerte).



Figura 3. El método del intercambio temporal



Fuente: elaboración propia a partir de (7)

dimensional y un algoritmo o regla de puntuación que asocia una utilidad a cada uno de los estados que es capaz de generar el correspondiente sistema de clasificación del instrumento.

El sistema de clasificación EQ-5D-3L genera a partir de la combinación de cinco dimensiones y tres niveles de gravedad por dimensión hasta 243 (= 3^5) estados de salud diferentes. Asimismo, el EQ-5D-5L, combinando cinco dimensiones y cinco niveles, es capaz de describir hasta un total de 3.125 estados de salud. Por su parte, las dos versiones del sistema de clasificación SF-6D están constituidas por seis dimensionesⁱ, cada una de ellas con un número variables de niveles: de 4 a 6 en la primera versión del instrumento (SF-6Dv1), de 5 a 6 en la segunda versión (SF-6Dv2). Ambas versiones provienen de una extracción de ítems del cuestionario SF-36. El SF-6Dv1 es capaz de caracterizar hasta 18.000 estados de salud diferentes, mientras el SF-6Dv2 puede representar hasta 18.750 estados.

ⁱ Las seis dimensiones del SF-6D son las siguientes: funcionamiento físico, limitación de rol, funcionamiento social, dolor, salud mental y vitalidad.

Tanto en el caso del EQ-5D, como del SF-6D, el algoritmo generador de las utilidades asociadas a cada uno de los posibles estados de salud, se infiere estadísticamente de las valoraciones directas de un subconjunto de dichos estados realizadas por amplias muestras de población general. El algoritmo es una fórmula que permite predecir la utilidad de cualquiera de los estados de salud que integran el 'value set' o tarifa social (el conjunto de estados de salud posibles) de cada instrumento.

El Grupo EuroQol desarrolló en 1993 y 2009, respectivamente, dos protocolos estandarizados para estimar la tarifa social del EQ-5D-3L. El primero de ellos, denominado protocolo '*Measurement and Valuation of Health*' (MVH), se creó para valorar un subconjunto de estados de salud EQ-5D-3L por parte de una muestra de población general del Reino Unido (52); a partir de dichas valoraciones se estimaría la tarifa británica (53). El protocolo MVH fue refinado con posterioridad, mejorando el proceso de obtención de valoraciones, dando lugar al segundo de los protocolos a emplear con el sistema EQ-5D-3L, conocido como protocolo de París (54). De acuerdo a ambos protocolos cada entrevistado valora no menos de 10 estados de salud mediante el método del intercambio temporal. A este ejercicio de valoración antecede el de ordenación y valoración del mismo subconjunto de estados con la escala visual analógica. Desde el decenio de 1990 hasta el presente se han publicado 32 'value sets' del EQ-5D-3L empleando ambos protocolos (28), incluida la tarifa española (55).

Más recientemente, el Grupo EuroQol ha desarrollado el llamado protocolo EQ-VT para generar conjuntos de valores estandarizados con el instrumento EQ-5D-5L (46). Este protocolo se basa en la experiencia previa acumulada de la utilización de los dos protocolos precedentes, así como de una serie de estudios piloto multinacionales. El protocolo EQ-VT utiliza el método del intercambio de tiempos compuesto o '*composite TTO*' seguido por un experimento de elección discreta^j. En la actualidad hay alrededor de una veintena de tarifas sociales del EQ-5D-5L

^j Es decir, una serie de elecciones dicotómicas entre dos estados de salud que van cambiando cada vez.

publicadas y otras tantas en proceso de culminación (29), una de ellas para España (56).

Por último, el Grupo EuroQol coordinó un estudio para desarrollar una función de mapeo o ‘*crosswalk*’ entre las utilidades EQ-5D-3L y las utilidades EQ-5D-5L (57), de modo que pueden calcularse utilidades EQ-5D-5L a partir de la descripción de estados de salud de este mismo instrumento y la tarifa social del EQ-5D-3L (58).

Tras la estimación del *value set* del EQ-5D-5L para Reino Unido UK (59), el National Institute for Health and Care Excellence (NICE) hizo público un posicionamiento estableciendo que “en la actualidad no se recomienda el uso del EQ-5D-5L valuation set”^k y continuó recomendando en sus guías la tarifa del EQ-5D-3L como caso de referencia de las evaluaciones económicas que la industria presente a esta institución!. Si se utiliza el sistema descriptivo EQ-5D-5L entonces el NICE establece que debe mapearse al EQ-5D-3L utilizando el ‘*crosswalk*’ (57-58) antes mencionado.

En 2019 el NICE encargó un informe a 4 expertos independientes sobre el estudio que dio lugar a la tarifa EQ-5D-5L británica (59) y concluyeron que se habían producido “efectos entrevistador” (60). A resultados de dicho informe, el NICE actualizó su posicionamiento al respecto^k que sigue desaconsejando el uso de la tarifa EQ-5D-5L, pero anuncia que comisiona un nuevo estudio para reestimarlos, intentando evitar los errores detectados. Dicho estudio tuvo inicio en octubre de 2022 y aún no ha sido concluido^m. El NICE advierte de que “revisará su política sobre el EQ-5D-5L cuando el nuevo estudio esté concluido”^k.

^k <https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/technology-appraisal-guidance/eq-5d-5l>

^l <https://www.nice.org.uk/process/pmg36/chapter/economic-evaluation#the-reference-case-framework>

^m https://euroqol.org/eq-5d-instruments/eq-5d-5l-about/valuation-standard-value-sets/new-uk-5d-5l-valuation-study_blog/

En relación al instrumento SF-6D, desde el estudio inicial conducido en el Reino Unido (49) para estimar la tarifa del SF-6Dv1, se han estimado otros 9 modelos con esta misma versión, entre ellos el 'value set' para España (38). Hay otras 2 estimaciones para esta misma versión del SF-6D, pero abordadas a partir de los ítems del SF-12, no del SF-36. Así mismo, se dispone de 5 tarifas estimadas para el SF-6Dv2. A diferencia del EQ-5D, en el caso de la estimación del algoritmo SF-6D, no hay un protocolo estandarizado para obtener las valoraciones directas de una muestra de población general que sirven luego para inferir la fórmula de generación de utilidades. El método de obtención de preferencias más frecuentemente utilizado en los diferentes estudios ha sido la lotería estándar.

Pasos 3 a 5: cálculo de AVAC de cada alternativa y su diferencia

Ahora que ya sabemos cómo pueden obtenerse las utilidades de la CVRS, volvamos a nuestro ejemplo, y calculemos el número de AVAC asociados a cada uno de los dos tratamientos considerados, laringectomía y radioterapia. Supongamos que, por ejemplo, hemos obtenido la utilidad de cada uno de los tres estados de salud involucrados en la decisión representada en la Figura 1 mediante el método del intercambio temporal. Imaginemos que el paciente a quien hemos valorados sus preferencias, ha declarado que considera equivalente pasar 5 años en buena salud a 10 en el estado 'sin habla', de modo que $V(\text{sin habla}) = 0,5$. Supongamos que de forma semejante, dicho paciente considera equivalente pasar 7,5 años en buena salud a vivir 10 años en el estado 'habla esofágica', y pasar 9,5 años en buena salud a vivir 10 en el estado 'afonía leve'. Las utilidades resultantes son, por tanto, 0,75 y 0,95, respectivamente. Ya tenemos las utilidades que multiplicadas por los años de vida correspondientes, constituyen los resultados de cada una de las ramas del árbol. Tal y como se recoge en la Tabla 1, los AVAC así calculados son 3 (= $U(\text{sin habla}, 6 \text{ años})$), 4,5 (= $U(\text{habla esofágica}, 6 \text{ años})$) y

2,85 (= U(afonía leve, 3 años)). Si bien en el caso de la radioterapia, puesto que no hay riesgo, el número final de AVAC asociado al tratamiento son esos 2,85 AVAC que hemos mencionado, en el caso de la laringectomía, al ser un tratamiento arriesgado, hemos de combinar los resultados del mismo (los AVAC que obtendría el paciente a posteriori) con sus probabilidades de ocurrencia. Dicho de otro modo, para decidir ex ante qué tratamiento es más efectivo (produce más AVAC) hemos de estimar el número de AVAC que cabe esperar que gane el paciente si se somete a laringectomía. Para obtener dicho número esperado de AVAC, aplicamos la regla de evaluación de la teoría de la utilidad esperada (61), esto es:

$$AVAC \text{ esperados} = \sum_{i=1}^n p_i \times U_i(Q_i, T_i) \quad [3]$$

n= número de sucesos; p_i = probabilidad del suceso i; $U_i(Q_i, T_i)$ = utilidad asociada al suceso i.

Aplicando esta ecuación a las probabilidades y utilidades asociadas a la laringectomía, obtenemos que dicho tratamiento produce un total de 4 AVAC (= $0,33 \times 3 + 0,67 \times 4,5 = 4,0$), lo cual representa un incremento de 1,15 AVAC con respecto a la opción radioterapia (= $4,0 - 2,85 = 1,15$).

Tabla 1. AVAC de los tratamientos laringectomía y radioterapia

	Sin habla	Habla esofágica	Afonía leve	Diferencia AVAC
V(Q)	0,5	0,75	0,95	
T	6	6	3	
U(Q, T)	3	4,5	2,85	
p	0,33	0,67	1	
$p \times U(Q, T)$	0,99	3,01	2,85	
AVAC esperados		4,0	2,85	1,15

Fuentes de utilidades de estados de salud y calidad metodológicaⁿ

Por desgracia, el número de ensayos clínicos que incluyen la CVRS como medida de resultado primario o secundario todavía es relativamente reducido, sobre todo en el área oncológica; precisamente la más prolífica en el lanzamiento de nuevas moléculas. Baste decir que el volumen de medicamentos oncológicos en última fase de desarrollo clínico más que se ha doblado entre 2010 y 2020 (62). Sin embargo, la mayoría de los ensayos en este campo emplean medidas de resultado subrogadas o intermedias. Por ejemplo, una revisión reciente de los ensayos clínicos con terapias CAR-T (63) pone de manifiesto que solo 29 de los 424 ensayos incluidos en la revisión (un 6,8%) midió la CVRS de los participantes. Dentro de estos 29 ensayos, la medida más utilizada fue el instrumento EQ-5D (un 65% del total).

Si no se dispone de utilidades generadas partir de datos de ensayos clínicos y no es factible realizar un estudio *ad hoc* para estimarlas, estas deben obtenerse de estudios precedentes que hayan sido publicados. Sin embargo, este recurso puede ser problemático, ya que con frecuencia la calidad de las fuentes originales de las provienen las utilidades utilizadas en evaluaciones previas no ha sido contrastada (p.ej. si la muestra de pacientes de la que proceden es idónea o si el método empleado para medirlas es adecuado). Además, no abundan las revisiones sistemáticas de utilidades de estados de salud y los estándares de calidad seguidos para reportar los valores escogidos son pobres. Todas estas limitaciones pueden conducir a que existan diferentes utilidades, medidas además de forma diversa, para un mismo estado de salud, de modo que la selección discrecional de la evidencia disponible puede conducir a conclusiones sesgadas.

Por las razones apuntadas es importante identificar, combinar y aplicar la evidencia disponible siguiendo unas directrices contrastadas que eviten una selección y utilización arbitraria de las utilidades existentes. En

ⁿ Esta sección se inspira mayormente en las recomendaciones contenidas en (5), transcribiendo diversos pasajes de dicho trabajo. Los lectores interesados en consultar todas las referencias originales citadas en ese artículo son remitidos al mismo, así como para profundizar en los detalles pormenorizados allí, de los cuales solo se reflejan en este capítulo los que se han considerado más relevantes.

este sentido, la *task force* de ISPOR de Buenas Prácticas para la Investigación de Resultados^{o(5)} explicita los factores que tener en cuenta para definir adecuadamente los criterios con los que identificar las utilidades a emplear en el modelo de interés, sugiere una *checklist* de estándares mínimos que cumplir en las revisiones sistemáticas de utilidades de estados de salud abordadas y, finalmente, un conjunto de recomendaciones para la utilización de las mismas en los modelos. Exponemos a continuación todas estas recomendaciones de la citada ISPOR *task force*.

La Tabla 2 recoge los factores que tomar en consideración cuando se definen los criterios de búsqueda para efectuar una revisión sistemática de la literatura, con el fin último de identificar las utilidades de estados de salud a utilizar en un modelo coste-efectividad.

Tabla 2. Factores a considerar en los criterios de búsqueda para identificar utilidades de estados de salud

Factores esenciales

- Las descripciones de los estados de salud proceden del modelo coste-efectividad
- Efectos tratamiento de las intervenciones y comparadores de interés (incluyendo las ganancias de utilidad derivadas de los beneficios de los tratamientos y las pérdidas de utilidad derivadas de sus efectos adversos)
- Efectos tratamiento y gestión de todas las etapas del itinerario clínico incluidas en el modelo
- Utilidades de los estados de salud de los cuidadores
- Sucesos clínicos concurrentes y secuelas
- Normas poblacionales
- Variables contextuales que podrían afectar a las utilidades de los estados de salud (p.ej. método de administración del tratamiento o entorno en el que se aplica)

Factores adicionales que pueden ser relevantes

- Funciones de “mapeo” para estimar utilidades a partir de otras medidas de CVRS no basadas en preferencias^p o de variables clínicas
- El contexto en que será utilizado el modelo (p.ej. localización geográfica o criterios de la agencia de reembolso)

Fuente: elaboración propia a partir de (5)

^o ISPOR Good Practices for Outcomes Research task force.

^p Por ejemplo, el instrumento SF-6D puede asociar una utilidad al estado de salud definido a partir de las respuestas del paciente al cuestionario SF-36. Este cuestionario, por sí solo, como vimos, no puede generar utilidades. En (64) se explica cómo “mapear” las respuestas del cuestionario SF-36 al sistema descriptivo SF-6D.

Más allá de repositorios y bases de datos biomédicas generales como, por ejemplo, MEDLINE^q, hay bases de datos específicas que recopilan una gran cantidad de estudios en los que se han estimado utilidades de estados de salud. Así, por ejemplo, tenemos la base de datos de análisis coste-utilidad del Tufts Medical Center de Boston^r, que contiene más de 10.000 evaluaciones relativas a una gran variedad de enfermedades y tratamientos. O la base de datos del Health Economics Research Centre (HERC) de la Universidad de Oxford^s especializada en recopilar estudios que mapean PROMs o instrumentos clínicos al EQ-5D.

En España no hay una base de datos similar a las anteriores, razón por la cual, en el marco del Plan de Calidad para el SNS, se lanzó hace unos años el proyecto de elaboración de una base de datos, o catálogo de utilidades o preferencias por los estados de salud, obtenidas de población española a partir de una revisión sistemática de la literatura (65,66). La revisión (67) dio como resultado la identificación de 103 artículos (que cumplieren con los criterios de inclusión establecidos) que se corresponden con un total de 742 utilidades, con una media de 7,20 (desviación estándar: $\pm 14,26$) utilidades por artículo. El rango de utilidades incluido en la mayoría de los artículos varía de una a cuatro. En 97 artículos se usó el cuestionario EQ-5D para estimar 538 utilidades (73%), en 10 artículos se usó el SF-6D (185 utilidades) y en tres artículos el HUI-3 (seis utilidades)^t.

La Tabla 3 presenta una lista de comprobación de los criterios mínimos que deben observarse en las revisiones sistemáticas para determinar

^q El *InterTASC Information Specialist Subgroup Search Filter Resource* proporciona filtros para identificar utilidades de estados de salud que aplicar en las búsquedas de repositorios como MEDLINE, por ejemplo.

^r <https://cevr.tuftsmedicalcenter.org/databases/cea-registry>

^s <https://www.herc.ox.ac.uk/downloads/herc-database-of-mapping-studies>

^t Las siglas HUI-3 corresponden a la tercera versión del instrumento de utilidad multiatributo *Health Utility Index*. El HUI3 (68) ofrece un sistema descriptivo compuesto por ocho atributos de 5-6 niveles cada uno: visión (6 niveles), oído (6), habla (5), desplazamiento (6), destreza (6), emociones (5), cognición (6), y dolor (5). Con este instrumento pueden llegar a caracterizarse hasta 972.000 estados de salud diferentes. Los antecedentes del HUI-3 son el HUI-1 (69) y el HUI-2 (70).

si el proceso seguido para seleccionar las utilidades de estados de salud que utilizar en el modelo coste-efectividad es transparente y apropiado⁴.

Tabla 3. Lista de comprobación de los estándares mínimos que reportar en una revisión sistemática de utilidades de estados de salud para modelos coste-efectividad (*ISPOR SpRUCE checklist*)

Crterios	Descripción
<i>Estrategia de búsqueda</i>	
Términos y alcance de la búsqueda	Describa la estrategia final de búsqueda y asegure que se cubre apropiadamente las bases de datos
Criterios para la selección de estudios	Describa los criterios utilizados para identificar y seleccionar estudios para la revisión sistemática (p.ej. muestra del estudio, rango de edad, gravedad de la enfermedad)
<i>Proceso de revisión</i>	
Control de calidad	Describa los criterios de calidad empleados durante la revisión para decidir si incluir o excluir estudios del análisis
Valoración de la relevancia de las utilidades de estados de salud	Describa la relevancia de las utilidades de estados de salud para el modelo coste-efectividad y la agencia de reembolso competente, si procede
<i>Datos extraídos y reportados</i>	
Características de la población o los pacientes	Reporte características relevantes de los pacientes, tales como edad, sexo, comorbilidades, diagnóstico y gravedad de la enfermedad*
Medidas utilizadas	Indique del nombre de la medida utilizada en el estudio
Pesos de preferencias	Describa la técnica utilizada para valorar el estado de salud (p.ej., intercambio de tiempos o lotería estándar) y el país en el que se recogieron los datos
Estadísticas descriptivas sobre utilidades de estados de salud	

*: Verifique la fuente primaria de los datos antes que confiar en datos de un modelo económico semejante en el que no se haya chequeado la relevancia de esos datos.

i: Si esta información es relevante.

⁴ *ISPOR HSU Good Practices Task Force Minimum Reporting Standards of Systematic Review of Utilities for Cost-Effectiveness models* (ISPOR SpRUCE checklist).

Tabla 3 cont. Lista de comprobación de los estándares mínimos que reportar en una revisión sistemática de utilidades de estados de salud para modelos coste-efectividad (*ISPOR SpRUCCE checklist*)

Criterios	Descripción
<i>Datos extraídos y reportados</i>	
Tasa de respuesta para la medida empleada Grado de <i>missing data</i> o datos perdidos para el estudio de seguimiento ⁱ	Indique si es probable que la tasa de respuesta lograda ponga en peligro la validez de la medida Informe de las tasas de pérdidas para el seguimiento y de los datos que faltan, especialmente si la falta de datos puede poner en peligro el grado de representatividad de las utilidades de estados de salud
Referencia original	Cite el estudio original publicado en el que se presentan las utilidades y no un estudio económico previo que haya utilizado esta evidencia
<i>Selección y estimación de utilidades de estados de salud finales para el modelo coste-efectividad</i>	
Base para seleccionar las utilidades	Proporcione las razones para seleccionar las utilidades utilizadas en el modelo
Métodos empleados para combinar las estimaciones	Si se hubieran combinado utilidades de diferentes fuentes, describa el método analítico (p.ej., meta-análisis) utilizado para combinarlas
<i>Métodos utilizados para aplicar utilidades de estados de salud en el modelo</i>	
Utilidades reales utilizadas	Informe de todas las utilidades usadas en el modelo, así como de la medida de las cuales fueron calculadas las utilidades
Ajustes o supuestos	Informe de las utilidades crudas y finales utilizadas con ejemplos, si se requieren, para clarificar el método empleados para ajustar los datos

Fuente: elaboración propia a partir de (5)

Siguiendo el orden establecido en la Tabla 3, en relación a la estrategia de búsqueda, esta idealmente debe describirse en un protocolo antes de iniciar la revisión. La descripción del proceso seguido debe especificar los términos utilizados en las búsquedas y las bases de datos consultadas. El proceso de búsqueda y los criterios de inclusión pueden resumirse en un diagrama de flujo siguiendo el enfoque PRISMA (71).

En cuanto al proceso de revisión, deben describirse los estudios identificados en las búsquedas y su elegibilidad, así como el número de revisores que decidieron la inclusión o no de cada uno de los estudios inicialmente identificados y el procedimiento seguido para resolver los desacuerdos entre ellos. Asimismo, aquellos estudios que inicialmente satisfacían los criterios de inclusión, pero que tras el proceso de revisión no fueron seleccionados, deben ser listados.

Una vez seleccionados los estudios, deben extraerse y resumirse los datos utilizados para evaluar la idoneidad de las utilidades de estados de salud. La revisión debe describir las características de los estudios de los que se extraen estos datos; si son observacionales o ensayos clínicos, y señalar aquellas limitaciones (p.ej., efectos placebo) que pudieran haber conducido a que las utilidades estuvieran distorsionadas. Es fundamental identificar y citar la publicación original y no basarse en referencias secundarias.

Finalmente, si se hubieran identificado múltiples fuentes apropiadas de utilidades, debería de justificarse la selección efectuada de la considerada mejor evidencia o de los métodos (p.ej. meta-análisis) utilizados para sintetizar la evidencia proveniente de diversos estudios. En este último caso deben aportarse pruebas de heterogeneidad que avalen la idoneidad de combinar los diferentes datos disponibles.

La Tabla 4 resume diferentes cuestiones relativas a la utilización de utilidades de estados de salud en los modelos coste-efectividad, así como una serie de recomendaciones en cada caso.

Tabla 4. Recomendaciones metodológicas para utilizar utilidades de estados de salud en modelos coste-efectividad

Criterios	Descripción
Utilidades individuales vs basadas en una función	Las decisiones deben basarse en la relevancia de los datos y la fiabilidad de los análisis reportados
Efectos de utilidad de las comorbilidades y normas de la población general	Si no hubiera utilidades específicas por edad disponibles, estas deberían estimarse empleando normas poblacionales ajustadas por edad. Las utilidades “libres de condición” deberían estimarse de la población diana
Sucesos adversos relacionados con los tratamientos	Debe evaluarse la medida en que los efectos sobre la utilidad de los sucesos adversos importantes se recogen en los datos utilizados para estimar las utilidades de sucesos no adversos del modelo
Sucesos clínicos concurrentes	En ausencia de datos de utilidad recopilados durante o inmediatamente después de un suceso (temporal) agudo, deben considerarse estimaciones plausibles de pérdida de AVAC por suceso en el análisis de sensibilidad
Análisis de sensibilidad	Deben practicarse análisis de sensibilidad uni y multivariados de las utilidades.

Fuente: elaboración propia a partir de (5)

En relación a las utilidades a emplear, estas pueden estimarse, bien analizando los datos de cada estado de salud por separado, bien como una función de una medida relevante del estado clínico. Cuando ambas opciones estén disponibles la alternativa a utilizar en el caso base de la evaluación debe fundamentarse en la relevancia de los datos y la fiabilidad de los análisis (p.ej. la precisión de la utilidad media).

En presencia de comorbilidades, las utilidades deberían reflejar todos los efectos sobre la CVRS no solo de la condición de interés, sino también de dichas comorbilidades. Deben estimarse utilidades específicas

por edad para así reflejar los efectos de las comorbilidades asociados a la edad. Para ello hay que emplear el multiplicador adecuado. Por ejemplo, si se obtiene una utilidad de 0,72 de una muestra con una edad media de 70 años y la norma poblacional para personas de esa edad es de 0,8, el multiplicador será $0,72/0,8 = 0,9$. Las utilidades específicas para otras edades se obtendrán como resultado de multiplicar ese coeficiente por la norma poblacional correspondiente a la edad de interés.

Si las utilidades de los estados de salud del modelo no reflejaran completamente la desutilidad de los efectos adversos debería de recurrirse a evidencia adicional de la cual estimar dicha desutilidad, aun cuando las utilidades originales pudieran capturar parcialmente dichos efectos adversos. La elección de los efectos adversos para los que calcular las utilidades debe justificarse por sus tasas de incidencia en los grupos de tratamiento, su gravedad y duración, así como por la sensibilidad esperada de los resultados de la relación coste-efectividad a las utilidades de lo efectos adversos.

Cuando los participantes en el estudio experimenten sucesos clínicos concurrentes relacionados con la enfermedad de interés (por ejemplo, cuando los pacientes con diabetes experimentan enfermedad cardiovascular y retinopatía), se pueden evaluar los efectos de utilidad de estos acontecimientos clínicos concurrentes: 1) restando la suma de los decrementos estimados de utilidad para los sucesos superpuestos de la utilidad estimada en ausencia de un suceso (método aditivo); 2) multiplicando la utilidad en ausencia de un suceso por el producto de las razones de la utilidad de los individuos con los sucesos clínicos entre la utilidad de los sujetos que no experimentan los sucesos clínicos (método multiplicativo); o 3) utilizando la utilidad más baja para todos los sucesos clínicos (método mínimo).

Los sucesos clínicos agudos, como las exacerbaciones del asma y las fracturas óseas, pueden asociarse a una disminución importante de utilidad. A fin de poder disponer de utilidades para este tipo de sucesos, los participantes en el estudio deberían de completar alguna medida de

CVRS durante o inmediatamente después de haber sufrido los sucesos de interés.

Para evaluar la sensibilidad de los resultados del modelo a los efectos del suceso agudo sobre la utilidad, los analistas deben de generar utilidades plausibles sobre la base de los efectos clínicos esperados del suceso, por ejemplo, determinando las respuestas esperadas a una medida como como el EQ-5D.

En relación al tratamiento de la incertidumbre alrededor de la utilidad promedio, dicha incertidumbre debería de representarse por medio de distribuciones de probabilidad paramétricas. En los análisis de sensibilidad deterministas pueden presentarse intervalos de confianza, mientras que en los análisis de sensibilidad probabilísticos pueden generarse muestras aleatorias a partir de las distribuciones.

Los análisis de sensibilidad univariados deben utilizarse para identificar las utilidades a las que son más sensibles los resultados del modelo. En los análisis de sensibilidad multivariados deben incluirse análisis combinados de todas las utilidades que muevan los resultados coste-efectividad en la misma dirección, esto es, de modo que todas las utilidades seleccionadas incrementen o reduzcan el coste-efectividad.

Limitaciones metodológicas en la medición de utilidades

A pesar del uso generalizado de técnicas de medición de utilidades de estados de salud como las que hemos descrito en la segunda sección de este capítulo, existe una amplia evidencia empírica que cuestiona su validez (4), lo que sugiere que las utilidades, tal y como se obtienen normalmente con métodos como el intercambio temporal (IT), la lotería estándar (LE) y la escala visual analógica (EVA), dependen del contexto. Esto significa que los encuestados no asocian un valor único a cada estado de salud, independientemente de factores “irrelevantes”, como el enmarcado de las preguntas, el método de medición de preferencias empleado

en particular o el procedimiento de búsqueda de la indiferencia entre las dos alternativas confrontadas, sino que las preferencias son contingentes a todos esos elementos (72,73).

La principal implicación que tiene para la evaluación de tecnologías sanitarias la utilización de utilidades de estados de salud sesgadas es que las ganancias de AVAC también lo estarán, lo que dará lugar a RCEI por AVAC distorsionadas. Consideremos, por ejemplo, el caso del EQ-5D, que, como vimos, es el instrumento recomendado por el NICE para medir las utilidades en estudios coste-efectividad^v. Si el ‘value set’ del EQ-5D estuviera sesgado, entonces las recomendaciones del NICE (o de cualquier otra agencia que utilice la tarifa social del EQ-5D) también lo estarán.

No obstante, que las preferencias dependan del contexto no significa que no existan unas “verdaderas” preferencias. Estas pueden “descubrirse” tras las valoraciones instantáneas que efectúan los individuos (74), y pueden representarse (*grosso modo*, al menos) mediante los AVAC. Tal “descubrimiento” de las auténticas preferencias puede realizarse, al menos en teoría, de diferentes maneras, propiciando así utilidades más sólidas y libres de sesgos. Antes de analizar los distintos abordajes que poner en práctica para acceder a preferencias consistentes, veamos a grandes rasgos cuáles son los principales sesgos que pueden contaminar el proceso de medición de utilidades de estados de salud.

Los sesgos

Es ampliamente reconocido que tanto las mediciones del IT como las de la LE se ven afectadas por sesgos sistemáticos (75), es decir, por errores persistentes tanto en dirección como en magnitud. Por el contrario, aunque la EVA no es susceptible de sesgos sistemáticos conocidos, las observaciones individuales están influidas por errores no sistemáticos derivados del

^v <https://www.nice.org.uk/process/pmg9/chapter/the-reference-case#measuring-and-valuing-health-effects>

contexto en que se efectúa la medición (32). Esta distinción entre error sistemático y no sistemático es relevante porque a lo largo de muchas observaciones, el error sistemático perdura y el no sistemático disminuye (76). Además, puede ser útil considerar a la EVA, el IT y la LE más como metodologías que como métodos singulares, ya que pueden aplicarse mediante distintas variantes (77). En este sentido, existen “metodologías” del IT y la LE como el IT compuesto (o C-TTO) propuesto por el grupo EuroQol (47), o la LE modificada propuesta por otros autores (38), ambas variantes presumiblemente menos propensas a algunos de los sesgos que afectan a las versiones tradicionales del IT y la LE.

Las valoraciones del IT tradicional se ven afectadas por tres sesgos diferentes, a saber: la aversión a las pérdidas, la compatibilidad de la escala y la curvatura de la utilidad del tiempo de vida. La aversión a las pérdidas es un factor psicológico que predice que las personas evalúan los resultados como ganancias o pérdidas en relación con un punto de referencia, y que “las pérdidas pesan más que las ganancias” (78,79). La compatibilidad de la escala predice que las personas dan demasiada importancia al atributo utilizado como escala de respuesta para establecer la indiferencia entre las dos opciones confrontadas, en este caso los años de vida (72). Tanto la aversión a las pérdidas como la compatibilidad de la escala tienen como efecto impulsar al alza las utilidades del IT (80,81). Sin embargo, la curvatura de la utilidad, un sesgo causado porque la utilidad para la duración del tiempo de vida no es lineal, tiene como efecto la reducción de los valores del IT (13,77). Dado que la curvatura de la utilidad funciona de forma opuesta a la aversión a las pérdidas y la compatibilidad de escala, algunos investigadores (75,82) han planteado la hipótesis de que los sesgos sistemáticos que afectan al IT quizá pudieran anularse entre sí, si bien hay evidencia reciente que sugiere una relación no lineal entre los años de vida y su utilidad, aun cuando se considera conjuntamente con la aversión a las pérdidas (83).

Por su parte, las utilidades obtenidas por la LE convencional están sesgadas al alza por la aversión a las pérdidas y la transformación de

las probabilidades (84,85), dos sesgos sistemáticos descritos por la denominada teoría prospectiva (86). Según esta teoría, las probabilidades altas tienden a infravalorarse, mientras que las probabilidades pequeñas (probabilidades típicamente inferiores a 0,3) tienden a sobrevalorarse.

Para intentar neutralizar los sesgos que acabamos de describir cabe aplicar, básicamente, dos grandes estrategias: la corrección (*ex post*) de los mismos y la prevención (*ex ante*) de los estos.

La estrategia correctiva de los sesgos

Hay varios autores que han propuesto fórmulas correctivas de sesgos basadas en la teoría prospectiva (87,88). Con la aplicación de dichas fórmulas puede conciliarse el deber ser (la visión normativa acerca de cómo deberían de ser las utilidades) con el ser (la evidencia positiva de que las utilidades están sesgadas). Por ejemplo, y haciendo abstracción del sesgo de aversión a las pérdidas, para simplificar, según la teoría prospectiva el peso otorgado a la probabilidad 0,9 se sitúa en torno a 0,7. Este valor procede de suponer que las probabilidades elevadas, como antes se indicó, tienden a infravalorarse.

Las fórmulas correctivas también pueden aplicarse al IT (89,90). Esto, que los sesgos de ambos métodos, IT y LE, puedan corregirse a posteriori, llevan a algunos autores a defender la necesidad de aplicar este enfoque “correctivo” a la estimación de tarifas sociales con instrumentos como el EQ-5D (91,92).

La estrategia preventiva de los sesgos

Como ya se ha señalado, tanto la LE modificada o doble lotería como el C-TTO se han sugerido como métodos de medición de utilidades que evitan *a priori* ciertos sesgos que distorsionan las valoraciones obtenidas

mediante las versiones tradicionales de la LE y el IT. Así, presumiblemente, la LE modificada evita el denominado efecto certeza o sobrevaloración de la certeza (86), que lleva a “inflar” el valor de la utilidad medida con la LE. Del mismo modo, el C-TTO pretende evitar el efecto enmarcado que afecta a la valoración dispar de estados mejores y peores que la muerte con el IT (46) y el *lead time trade-off*. Por tanto, la utilización de estos métodos alternativos puede evitar ex ante el afloramiento de determinados sesgos.

¿Qué método emplear entonces?

Algunos autores (93) han propuesto un sencillo test para averiguar cuál de los dos métodos —IT o LE— representan más fidedignamente las preferencias de los entrevistados. El test consiste en preguntar a los encuestados, una vez valorados una serie de estados de salud con ambos métodos, que indiquen cuál de los dos tipos de valoraciones refleja mejor sus preferencias, pudiendo ajustarlas, si así lo desean, a la baja o al alza. La principal conclusión de este ejercicio es que la mayoría de los entrevistados considera que el IT es el método cuyas utilidades reflejan más verazmente sus verdaderas preferencias. Este resultado parece respaldar la hipótesis antes enunciada de que los sesgos que afectan al IT tienden a cancelarse entre sí. No obstante, tal y como demuestra este estudio, dicha cancelación no es perfecta, ya que los encuestados ajustaron sus utilidades en la mayoría de las ocasiones. Dicho ajuste, además, fue a la baja, conduciendo, por consiguiente, a utilidades incluso inferiores a las generadas en primera instancia por el IT. La conclusión final de la evidencia registrada es que, si bien, las utilidades medidas con el IT no están exentas completamente de sesgos, sí que parecen reflejar de forma más aproximada que la LE las verdaderas preferencias individuales.



Bibliografía

1. Brennan A, Chick SE, Davies R. A taxonomy of model structures for economic evaluation of health technologies. *Health Econ.* 2006; 15: 1295-1310.
2. Fanshel S, Bush JW. A Health-Status Index and Its Application to Health-Services Outcomes. *Oper Res.* 1970;18: 1021-66.
3. EMA (2005). Reflection paper on the regulatory guidance for the use of health-related quality of life (HRQL) measures in the evaluation of medicinal products. European Medicines Agency, London. [https:// www.ema.europa.eu/en/documents/scientific-guideline/reflection-paperregulatory-guidance-use-health-related-quality-life-hrql-measures-evaluation_en.pdf](https://www.ema.europa.eu/en/documents/scientific-guideline/reflection-paperregulatory-guidance-use-health-related-quality-life-hrql-measures-evaluation_en.pdf) (acceso verificado el 26 de Sept de 2023)
4. Pinto JL, Herrero C, Abellán JM. QALY-Based Cost Effectiveness Analysis. En: Adler MD, Fleurbaey M, editors. *The Oxford Handbook of WellBeing and Public Policy.* New York: Oxford University Press; 2016. p. 160 -192.
5. Brazier J, Ara R, Azzabi I, et al. Identification, Review, and Use of Health State Utilities in Cost-Effectiveness Models: An ISPOR Good Practices for Outcomes Research Task Force Report. *Value Health.* 2019; 22:267-75.
6. Abellán JM, Pinto JL. Quality Adjusted Life Years as Expected Utilities. *Spanish Economic Review.* 2000;2:49-6.
7. Torrance GW. Measurement of health-state utilities for economic appraisal: A review. *J Health Econ.* 1986; 5:1-30.
8. Patrick DK, Starks HE, Cain KC, et al. Measuring preferences for health states worse than death. *Med Decis Making.* 1994;14:9-18.
9. Abellán JM. El AVAC como indicador de la esperanza y la calidad de vida. En: Hidalgo A, Del Llano J, directores. *Medidas de resultados en salud y su incorporación a la evaluación económica.* Cuadernos de Evaluación Sanitaria, nº 5. Instituto Max Weber y Fundación Gaspar Casal. Madrid 2013
10. Sox HC, Blatt MA, Higgins MC, et al. *Medical Decision Making.* Boston, M.A. Butterworth-Heinemann; 1988

11. Abellán Perpiñán JM, Sánchez Martínez FI, Martínez Pérez JE. Evaluación económica de tecnologías sanitarias: ¿valen lo que cuestan? Cuadernos Económicos de ICE. 2008; 75, 189 - 208. Secretaría de Estado de Comercio.
12. Bleichrodt H, Wakker P, Johannesson M. Characterizing QALYs by Risk Neutrality. *J Risk Uncertain.* 1997;15:107–14.
13. Abellán-Perpiñán JM, Pinto-Prades JL, Méndez-Martínez I, et al. Towards a better QALY model. *Health Econ.* 2006;15:665–76.
14. Toumi M, Jaroslowski S, Chouhaid C, et al. Patient-Reported Outcomes in Oncology, Beyond Randomized Controlled Trials. En: Walter E. *Regulatory and Economic Aspects in Oncology. Recent Results in Cancer Research* 213. 2019. p. 57-65
15. Matza LS, Stewart KD, Lloyd AJ, et al. Vignette-Based Utilities: Usefulness, Limitations, and Methodological Recommendations. *Value Health.* 2021; 24:812-21.
16. Kingsley, C. Patel, S. Patient-reported outcome measures and patient-reported experience measures. *Br J Anaesth. Education.* 2017;17:137-44.
17. Ware JE, Sherbourne CD. The MOS 36-Item Short-Form Health Survey 1: conceptual framework and item selection. *Med Care.* 1992; 30:473-83.
18. Ware JE, Snow KK, Kosinski M, et al. *The SF36 Health Survey Manual and Interpretation Guide.* Boston, Massachusetts: The Health Institute, New England Medical Center. 1993.
19. Ware JE Jr, Kosinski M, Dewey JE. *How to score version 2 of the SF-36 Health Survey.* Lincoln RI: Quality Metric Incorporated. 2000.
20. EuroQol Group. EuroQol – a new facility for the measurement of health related quality of life. *Health Policy.* 1990;16:199–208.
21. Brooks R. EuroQol Group: the current state of play. *Health Policy.* 1996;37:53–72.
22. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res.* 2011; 20:1727–36.
23. Ware JE Jr, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care.* 1996; 34: 220-33.

24. Ware JE Jr, Kosinski M, Turner-Bowker DM, et al. How to Score Version 2 of the SF-12 Health Survey (With a Supplement Documenting Version 1). Lincoln, RI: Quality Metric Incorporated. 2002.
25. Ware J. E. Jr., Kosinski M., Dewey J., et al. How to score and interpret single-item health status measures: A manual for users of the SF-8 Health Survey. Boston, MA: QualyMetric. 2001
26. Ware JE, Kosinski M, Keller SD. SF-36 Physical and Mental Health Summary Scales: A User's Manual. Boston, Massachusetts: The Health Institute, New England Medical Center. 1994.
27. Alonso J, Regidor E, Barrio G, et al. Valores poblacionales de referencia de la versión española del Cuestionario de Salud SF-36. Med Clin (Barc). 1998; 111:410-6.
28. EuroQol Research Foundation. EQ-5D-3L User Guide, 2018. Disponible en: <https://euroqol.org/publications/user-guides>. (acceso verificado el 23 de Oct de2023)
29. EuroQol Research Foundation. EQ-5D-5L User Guide, 2019. Disponible en: <https://euroqol.org/publications/user-guides>. (acceso verificado el 23 de Oct de2023)
30. Torrance GW, Thomas WH, Sackett DL. A utility maximization model for evaluation of health care programs. Health services research. 1972;7:118-33.
31. Patrick DL, Bush JW, Chen MM. Methods for measuring levels of wellbeing for a health status index, Health Serv Res. 1973;8: 228-45.
32. Bleichrodt H, Johannesson M. An experimental test of a theoretical foundation for rating-scale valuations. Med Decis Making. 1997;17:208-16.
33. Robinson, A., Loomes, G., Jones-Lee, M. Visual analog scales, standard gambles, and relative risk aversion. Med Decis Making. 2001; 21:17-27.
34. Torrance GW, Feeny D, Furlong W. Visual analogue scales: do they have a role in the measurement of preferences for health states? Med Decis Making. 2001;21:329-34.
35. Parkin D, Devlin N. Is there a case for using visual analogue scale valuations in cost-utility analysis? Health Econ. 2006;15:653-64

36. Nord E. The person-trade-off approach to valuing health care programs. *Med Decis Making*. 1995;15:201-8.
37. Pinto-Prades JL, Abellan-Perpignan JM. Measuring the health of populations: the veil of ignorance approach. *Health Econ*. 2005;14:69-82.
38. Abellán JM, Sánchez FI, Martínez JE, et al. Lowering the 'floor' of the SF-6D scoring algorithm using a lottery equivalent method. *Health Econ*. 2012; 21:1271-85.
39. Sánchez-Martínez FI, Martínez-Pérez JE, Abellán-Perpiñán JM, et al. The Value of Statistical Life in the Context of Road Safety: New Evidence on the Contingent Valuation / Standard Gamble Chained Approach. *J Risk Uncertain*. 2021; 63:203-28.
40. Osman A, Wu J, He X, et al. Eliciting SF-6Dv2 health state utilities using an anchored best-worst scaling technique. *Soc Sci Med*. 2021; 279:114018.
41. Patrick DL, Starks HE, Cain KC, et al. Measuring preferences for health states worse than death. *Med Decis Making*. 1994;14: 9-18.
42. Robinson A, Spencer A. Exploring challenges to TTO utilities: valuing states worse than dead. *Health Econ*. 2006;15:393-402.
43. Devlin NJ, Tsuchiya A, Buckingham K, et al. A uniform time trade off method for states better and worse than dead: feasibility study of the 'lead time' approach. *Health Econ*. 2011; 20:348-61.
44. Devlin N, Buckingham K, Shah K, et al. A comparison of alternative variants of the lead and lag-time TTO. *Health Econ*. 2012;22:517-32.
45. Lamers LM. The transformation of utilities for health states worse than death: consequences for the estimation of EQ-5D value sets. *Med Care*. 2007;45:238-44.
46. Oppe M, Devlin NJ, van Hout B, et al. A program of methodological research to arrive at the new international EQ-5D-5L valuation protocol. *Value Health*. 2014;17:445-53.
47. Janssen BM, Oppe M, Versteegh MM, et al. Introducing the composite time trade-off: a test of feasibility and face validity. *Eur J Health Econ*. 2013;14(Suppl 1):S5-13.
48. Brazier J, Usherwood T, Harper R, et al. Deriving a preference based single index from the UK SF-36 health survey. *J Clin Epidemiol*. 1998;51:1115-28.

49. Brazier J, Roberts J, Deverill M. The estimation of a preference based measure of health from the SF-36. *J Health Econ.* 2002;21:271–92.
50. Brazier JE, Roberts J. The estimation of a preference-based measure of health from the SF-12. *Medical Care.* 2004;42:851–9.
51. Brazier JE, Mulhern BJ, Bjorner JB, et al. Developing a new version of the SF-6D health state classification system from the SF-36v2: SF-6Dv2. *Med Care.* 2020;58:557–65.
52. Williams A. The measurement and valuation of health: a chronicle. Discussion paper 136. York: Center for Health Economics, The University of York; 1995.
53. Dolan P. Modeling valuations for EuroQol health states. *Med Care.* 1997;35:1095–108.
54. Kind P. A revised protocol for the valuation of health states defined by the EQ-5D-3L classification system: learning the lessons from the MVH study. York: Centre for Health Economics, The University of York; 2009.
55. Badía X, Roset M, Herdman M, et al. A Comparison of United Kingdom and Spanish General Population Time Trade-off Values for EQ-5D Health States. *Med Decision Making.* 2001;21:7-16.
56. Ramos-Goñi JM, Craig B, Oppe M, et al. Handling data quality issues to estimate the Spanish EQ-5D-5L value set using a hybrid interval regression approach. *Value Health.* 2018;21:596–604.
57. Van Hout B, Janssen MF, Feng YS, et al. Interim scoring for the EQ-5D-5L: Mapping the EQ-5D-5L to EQ-5D-3L value sets. *Value Health.* 2012;15:708–15.
58. Hernández Alava M, Pudney S. Econometric modelling of multiple self-reports of health states: the switch from EQ-5D-3L to EQ-5D-5L in evaluating drug therapies for rheumatoid arthritis. *J Health Econ.* 2017;55:139–52.
59. Devlin NJ, Shah KK, Feng Y, et al. Valuing health-related quality of life: An EQ5D-5L value set for England. *Health Econ.* 2018;27:7–22.
60. QC report for England. EuroQol Office, Rotterdam, The Netherlands. March 2019. https://eq-5dpublications.euroqol.org/download?id=0_80_133&fileId=81961 (acceso verificado el 23 de Oct de2023)

61. Von Neumann J, Morgenstern O. Theory of Games and Economic Behavior. Princeton NJ: Princeton University Press; 1944.
62. Global Trends in R&D. Overview through 2020, IQVIA Institute, May 2021. <https://www.iqvia.com/-/media/iqvia/pdfs/institute-reports/global-trends-in-r-and-d/iqvia-institute-global-trends-in-rd-0521forweb.pdf> (acceso verificado el 23 de Oct de2023)
63. Raymakers AJN, Regier DA, Peacock SJ, et al. Health-related quality of life data collected in chimeric antigen receptor T-cell (CAR-T) therapy clinical trials. *J Cancer Policy*. 2021;30:100304.
64. Abellán JM. Utilidades SF-6D para España. Guía de uso 2012/8. Sevilla: Cátedra de Economía de la Salud. Universidad Pablo de Olavide. Consejería de Salud de la Junta de Andalucía. 2012
65. García Pérez L, Aguiar Ibáñez R, Linertová R, et al. Plan de Calidad para el Sistema Nacional de Salud del Ministerio de Sanidad, Política Social e Igualdad. Servicio de Evaluación del Servicio Canario de la Salud. 2011. Informes de Evaluación de Tecnologías Sanitarias: SESCO N° 2010/6. https://www3.gobiernodecanarias.org/sanidad/scs/content/5c8f8874-ab93-11e4-9053-9d1690bb437a/SESCS%202010_06_%20RS%20Utilidades.pdf (acceso verificado el 23 de Oct de2023)
66. García Pérez L, Pinilla Domínguez P, Aguiar Ibáñez R, et al. Revisión sistemática de utilidades o preferencias por los estados de salud obtenidas de población española. Plan de Calidad para el Sistema Nacional de Salud del Ministerio de Sanidad, Política Social e Igualdad. Servicio de Evaluación del Servicio Canario de la Salud. 2012. Informes de Evaluación de Tecnologías Sanitarias: SESCO N° 2011/4. https://funcanis.es/wpcontent/uploads/2015/09/SESCS-2011_R.S.-Preferencias-E.S._sin.pdf (acceso verificado el 23 de Oct de2023)
67. García-Pérez L, Aguiar-Ibáñez R, Pinilla-Domínguez P, et al. Revisión sistemática de utilidades relacionadas con la salud en España: el caso de la salud mental. *Gac Sanit*. 2014; 28:77-83.
68. Feeny D, Furlong W, Torrance GW, et al. Multi-attribute and single-attribute utility function for the Health Utility Index Mark 3 system. *Med Care*. 2002;40:113-28.

69. Torrance GW, Boyle MH, Horwood SP. Application of multi-attribute utility theory to measure social preferences for health states. *Oper Res.* 1982;30:1043
70. Torrance GW, Feeny D, Furlong W, et al. Multi attribute preference functions for a comprehensive health status classification system: health utilities index mark II. *Med Care.* 1996;34:702-2.
71. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol.* 2009;62:e1e34.
72. Tversky A, Sattath S, Slovic P. Contingent weighting in judgement and choice. *Psychological Review.* 1988;95:371-84.
73. Tversky A, Simonson I. Context-Dependent preferences. *Management Science.* 1993; 39:1179-89.
74. Plott C. Rational choice in experimental markets en Hogarth M y Reder M (eds). *Rational choice. The constrast between economics and psychology.* Chicago: University of Chicago Press. 1986, p. 117-144
75. Bleichrodt H. A new explanation for the difference between time trade-off utilities and standard gamble utilities. *Health Econ.* 2002;11:447-56.
76. Doctor J, Bleichrodt H, Jill Lin H. Health utility bias: a systematic review and meta-analytic evaluation. *Med Decis Making.* 2010;30:58-67.
77. Attema A, Brouwer W. In search of a preferred preference elicitation method: a test of the internal consistency of choice and matching. *J Econ Psychol.* 2013; 39:126-40.
78. Tversky A, Kahneman D. Loss aversion in riskless choice: a reference-dependent model. *Q J Econ.* 1991;106:1039-61.
79. Tversky A, Kahneman D. Advances in prospect theory: cumulative representation of uncertainty. *J Risk Uncertain.* 1992;5:297-323.
80. Bleichrodt H, Pinto JL. Loss aversion and scale compatibility in two-attribute trade offs. *J Math Psychol.* 2002;46:315-37.
81. Bleichrodt H, Pinto JL, Abellan JM. A consistency test of the time trade-off. *J Health Econ.* 2003;22:1037-52.

82. Wakker PP. Lessons learned by (from?) an economist working in medical decision making. *Med Decis Making*. 2008;28:690-8
83. Lipman SA, Brouwer W, Attema AE. QALYs without bias? Nonparametric correction of time trade-off and standard gamble weights based on prospect theory. *Health Econ*. 2019;28:843-54.
84. Bleichrodt H, Pinto JL. A parameter-free elicitation of the probability weighting function in medical decision analysis. *Manage Sci*. 2000;46: 1485-96.
85. Attema A, Brouwer W, l'Haridon O. Prospect theory in the health domain: a quantitative assessment. *J Health Econ*. 2013;32:1057-65.
86. Kahneman D, Tversky A. Prospect theory: an analysis of decision under risk. *Econometrica*. 1979;47:263-91.
87. Bleichrodt H, Pinto JL, Wakker P. Making descriptive use of prospect theory to improve the prescriptive use of expected utility. *Manage Sci*. 2001;47:1498-1514.
88. Bleichrodt H, Abellan JM, Pinto JL, et al. Resolving inconsistencies in utility measurement under risk: test of generalizations of expected utilities. *Manage Sci*. 2007;53:469-82.
89. Abellan JM, Martinez JE, Sanchez FI, et al. The QALY model which came in from a general population survey, roughlyly multiplicative, broadly nonlinear and sometimes context-dependent. *CENTRA Documento de Trabajo E2009/04*, 2009
90. Attema A, Bleichrodt H, Wakker P. A direct method for measuring discounting and QALYs more easily and reliably. *Med Decis Making*. 2012;32: 583-93.
91. Abellan JM, Martinez JE, Sanchez FI, et al. Debiasing EQ-5D tariffs. New estimation of the Spanish EQ-5D value set under non expected utility. *CENTRA Documento de Trabajo E2009/06*. 2009
92. Lipman SA, Brouwer B, Attema AE. The Corrective Approach: Policy Implications of Recent Developments in QALY Measurement Based on Prospect Theory. *Value Health*. 2019;22:816-21.
93. Lipman SA, Brouwer B, Attema AE. What is it going to be, TTO or SG? A direct test of the validity of health state valuation. *Health Econ*. 2020;29: 1475-81.



5

EL MANEJO DE LA INCERTIDUMBRE EN LOS MODELOS DE ANÁLISIS DE DECISIONES

David Epstein

Catedrático de Economía Aplicada de la Universidad de Granada



Introducción

Cualquier estudio científico debe cuantificar tanto el resultado como el grado de incertidumbre en torno a ese resultado (1). La forma tradicional de expresar la incertidumbre en un sentido estadístico ha sido a través del intervalo de confianza, que proporciona un resumen útil del rango de valores de un parámetro desconocido. De este modo, se estima la probabilidad de que el intervalo de confianza calculado a partir de algún experimento futuro abarque el verdadero valor del parámetro poblacional.

El intervalo de confianza del 95% se basa en un umbral preestablecido de falsos positivos del 5%. La selección de este umbral particular de significación expresa la idea de que la ciencia sólo debe aceptar nuevas teorías (y descartar las existentes) si hay un peso muy fuerte de pruebas a su favor (2). En el contexto de la comparación de tecnologías sanitarias, si no hubiera diferencia en la eficacia de dos tratamientos, entonces por cada 20 experimentos que realicemos, sólo uno mostrará (falsamente) una diferencia.

Si la variable de interés sigue una distribución normal, el umbral del 5% significa que el resultado debe estar a más de 2 desviaciones estándar de la nula para aceptar la hipótesis alternativa (o 2-sigma). La elección de un umbral del 5% (en lugar de, por ejemplo, el 1% o el 10%) es simplemente una convención. Se estableció más o menos a principios del siglo xx en los campos científicos experimentales en los que intervienen seres humanos, como las ciencias de la salud, las ciencias sociales y la economía, porque permite un equilibrio conveniente entre la necesidad de obtener un grado razonable de certeza de que los resultados positivos son válidos y, al mismo tiempo, reclutar tamaños de muestra moderados para evitar estudios innecesariamente costosos (2). Las ciencias físicas experimentales requieren una precisión mucho mayor, de 5 sigmas o

más, para aceptar, por ejemplo, la existencia de una nueva partícula fundamental.

Aunque las evaluaciones económicas de las tecnologías sanitarias suelen informar de los intervalos de confianza, muchos autores reconocen que esta métrica es inadecuada para tomar decisiones sobre las tecnologías sanitarias (en contraposición al esfuerzo científico de desarrollar nuevas teorías). Karl Claxton expresó esta idea en su artículo seminal de tesis doctoral *“The irrelevance of inference”* (3). Claxton argumentó que cuando comparamos dos tecnologías sanitarias para decidir cuál debe adoptarse, no hay ninguna razón a priori para adoptar la nueva tecnología sólo si hay un peso muy fuerte de evidencia a su favor. En cambio, los sistemas sanitarios deberían adoptar una nueva tecnología si su beneficio neto medio es mayor que el de la tecnología existente.

Claxton argumentó que si la evidencia muestra un mayor beneficio neto medio a favor de la nueva intervención, entonces la incertidumbre en ese beneficio neto genera dos tipos de posibles errores para los responsables de la toma de decisiones: 1) los riesgos y pérdidas de aplicar la nueva intervención, cuando la intervención existente era realmente mejor, y 2) los riesgos y pérdidas de *no* aplicar la nueva intervención, cuando la nueva intervención era realmente mejor.

Puede haber circunstancias en las que estos riesgos y pérdidas no sean simétricos. Por ejemplo, si hay importantes costes fijos irrecuperables, como la formación o la inversión en infraestructuras, que deben desplegarse antes de que la nueva tecnología pueda aplicarse (4). No obstante, Claxton manifestó que en la mayoría de las circunstancias, estas pérdidas serán simétricas. Por tanto, no hay razón para exigir más pruebas a favor de la nueva tecnología que de la existente para cambiar de política. Si las pérdidas son simétricas, las decisiones sobre las nuevas tecnologías frente a las existentes pueden tomarse simplemente comparando el beneficio neto medio.

Estos argumentos se han consolidado en la metodología y la práctica de la evaluación de tecnologías sanitarias. Sin embargo, hay varias razones por las que sigue siendo importante que los profesionales evalúen la incertidumbre de las pruebas y la comuniquen eficazmente a los responsables políticos.

La primera razón es que los responsables de la toma de decisiones pueden considerar la posibilidad de posponer la decisión para generar más pruebas (4). Los argumentos de Claxton parten de la premisa de que hay dos opciones: adoptar la nueva intervención o continuar con la existente (3). Puede haber una tercera opción, que es la de esperar a tener más pruebas. Esta tercera opción también tiene, por supuesto, sus ventajas y sus costes. La obtención de más pruebas reduciría la probabilidad de que la decisión sea errónea, pero podría significar que los pacientes se vean privados de los beneficios de la nueva intervención mientras se generan los nuevos datos. Los responsables sanitarios deben conocer el grado de incertidumbre asociado a las pruebas para poder considerar adecuadamente todas estas opciones. El análisis del valor de la información puede ser útil en estas situaciones. Se trata de un conjunto de técnicas que resumen el máximo valor potencial de un nuevo estudio para el responsable de la toma de decisiones y las variables específicas para las que sería más valiosa más información (3).

La segunda razón es que un análisis de la incertidumbre proporciona un medio para comprobar y validar un modelo de decisión. Los modelos económicos de las tecnologías sanitarias, como los árboles de decisión o los modelos de Markov, suelen consistir en una síntesis de datos procedentes de distintas fuentes (probabilidades de estados de salud clínicos, calidad de vida, metaanálisis, recursos sanitarios, etc.). Dado el gran número de variables incluidas, y las interacciones a menudo complejas entre ellas, estos modelos suelen ser difíciles de validar. El análisis de la incertidumbre proporciona un medio para que los analistas puedan depurar los errores, verificar que el modelo se comporta como se espera y comunicar los resultados de forma eficaz a las partes interesadas.

La tercera razón para realizar un análisis de incertidumbre en un modelo es llevar a cabo un análisis de umbral. Esto es especialmente atractivo en situaciones en las que una variable importante es completamente desconocida o indeterminada en la fecha de construcción del modelo (como el precio de la terapia). El análisis de umbral (a veces conocido como análisis de margen o “*headroom*” en inglés) responde a preguntas como “¿cuál sería el precio máximo para que esta intervención sea coste-efectiva?”(5).



Tipos de incertidumbre

Briggs y col. propusieron una tipología de cuatro tipos de incertidumbre: incertidumbre estocástica, incertidumbre de los parámetros, heterogeneidad e incertidumbre estructural (6) (Tabla 1). La incertidumbre estocástica (también llamada variación de Monte-Carlo o incertidumbre de primer orden) se refiere a la variación de los resultados entre individuos similares de una población. En un contexto sanitario, algunos individuos mejorarán tras un determinado tratamiento y otros no. Este tipo de variación no suele ser de interés en el contexto de la evaluación de tecnologías sanitarias. Se supone que el responsable de la toma de decisiones pretende mejorar los resultados medios de la población, y no se espera que cada individuo se beneficie necesariamente. Esto no significa que el conocimiento de la variación entre individuos nunca sea importante en las ciencias de la salud. Los pacientes querrán conocer, en la medida de lo posible, sus riesgos individuales, por ejemplo, su probabilidad de desarrollar una enfermedad cardiovascular o de sobrevivir a una intervención quirúrgica, para poder tomar decisiones personales sobre su atención. Pero la evaluación de tecnologías sanitarias suele adoptar una perspectiva más amplia, al nivel de la población.

Tabla 1. Clasificación de los tipos de incertidumbre

Tipos de incertidumbre	Concepto	Efecto en un modelo de decisión	Enfoques analíticos para comunicar esta incertidumbre a los responsables de la toma de decisiones
Incertidumbre estocástica	Variabilidad aleatoria de los resultados entre individuos idénticos	Los resultados varían dentro de un grupo de individuos idénticos	No suele ser de interés para la toma de decisiones
Incertidumbre de los parámetros	La incertidumbre en la estimación del parámetro de interés	Incertidumbre sobre el valor de cada entrada del modelo de decisión	<p>Análisis de sensibilidad univariante</p> <p>Análisis de sensibilidad multivariante</p> <p>Análisis de umbrales</p> <p>Análisis de sensibilidad probabilística</p>
Heterogeneidad	La variabilidad entre pacientes que puede atribuirse a las características de esos individuos	Los resultados difieren entre los subgrupos	Análisis de subgrupos
Incertidumbre estructural	Los supuestos inherentes al modelo de decisión	Incertidumbre sobre si el modelo representa con precisión el curso de la enfermedad o el tratamiento	<p>Análisis de escenarios</p> <p>Análisis de umbrales</p>

Sin embargo, algunos tipos de modelos de decisión capturan deliberadamente la incertidumbre de primer orden. Son los llamados modelos de microsimulación o simulaciones de eventos discretos. Estos tipos de modelos pretenden seguir a los individuos a lo largo de las distintas etapas posibles de la enfermedad y el tratamiento para captar los resultados a un nivel muy granular. El objetivo de estos modelos es principalmente obtener estimaciones más precisas de los resultados medios a nivel de población, y no porque la variación de los resultados dentro de esa población sea de interés en sí misma.

La incertidumbre de los parámetros (también llamada incertidumbre de segundo orden) es la incertidumbre en la estimación de una variable de interés del modelo. Por ejemplo, supongamos que estamos interesados en la calidad de vida de una población de pacientes con una determinada enfermedad, y realizamos una encuesta a 100 individuos seleccionados al azar de esa población y calculamos la media. Como la muestra es mucho más pequeña que la población, si repitiéramos el experimento con 100 individuos diferentes seleccionados al azar, obtendríamos un resultado diferente. Esta variación aleatoria en la estimación de la media suele resumirse en un único número llamado “error estándar” de la media, y representa nuestra incertidumbre en el valor de un parámetro (en este caso, la calidad de vida media de los pacientes con una determinada enfermedad).

Se denomina heterogeneidad a la variación entre grupos de individuos que puede atribuirse a factores conocidos, como la edad, el sexo o la gravedad de la enfermedad. Por ejemplo, en promedio, los pacientes con una forma más grave de una enfermedad probablemente tendrán peor calidad de vida que los pacientes con una enfermedad moderada. Los factores que causan estas diferencias en los resultados se denominan a veces modificadores o subgrupos.

La incertidumbre estructural, o incertidumbre de modelización, surge porque el analista no está seguro de que el modelo capte

adecuadamente el curso de la enfermedad o el efecto del tratamiento sobre esa enfermedad. Todos los análisis estadísticos, desde los más sencillos hasta los más sofisticados, son simplificaciones de la realidad. Algunos elementos se incluirán y otros se omitirán o, como dijo el estadístico George Box, “todos los modelos son erróneos, pero algunos son útiles”. En el contexto de la evaluación de tecnologías sanitarias, surge una fuente particular de incertidumbre estructural cuando el analista pretende modelar los resultados en un horizonte temporal largo (por ejemplo, la vida de la cohorte de pacientes), pero los principales estudios clínicos son a corto plazo (por ejemplo, unos pocos meses o años). Cualquier suposición sobre los resultados más allá del estudio clínico será necesariamente muy incierta.

La taxonomía de Briggs y col. (6) parece tratar las cuatro fuentes de incertidumbre como totalmente distintas, pero en realidad se solapan en cierto grado. En un sentido puro, la incertidumbre de los parámetros supone que el estudio subyacente no tiene sesgos internos ni externos. En este caso, la incertidumbre de segundo orden está inversamente relacionada con la raíz cuadrada del tamaño de la muestra, y puede reducirse obteniendo más información (es decir, obteniendo una muestra mayor de la población). Sin embargo, en la realidad, los estudios rara vez son perfectos (7). El sesgo interno se produce cuando un estudio carece de rigor. Por ejemplo, es posible que los resultados no estén cegados o que falten muchos datos. El sesgo externo se produce cuando un estudio carece de relevancia. Por ejemplo, puede compararse una intervención con un placebo, en lugar de con el mejor estándar de atención actual, o incluirse pacientes con diversas condiciones clínicas, diferentes a la población diana del modelo de decisión. Por lo tanto, el analista debe ser consciente de que el error estándar de la media notificado puede no representar toda la incertidumbre sobre un parámetro. Si el estudio sufre un sesgo interno o externo considerable, la verdadera incertidumbre será mayor que el error estándar notificado. La evaluación del “riesgo de sesgo” de los estudios que

aportan datos a un modelo de decisión será siempre cualitativa y subjetiva, pero es de vital importancia. Existen listas de comprobación y directrices para evaluar diferentes tipos de estudios (véase <https://www.equator-network.org/>), pero la habilidad y la experiencia de los analistas para interpretar los datos serán decisivas. Esta es una de las razones por las que se recomienda que los modelos de decisión se construyan como un esfuerzo de colaboración entre economistas de la salud, clínicos, epidemiólogos y otros expertos, para que la calidad de las entradas del modelo pueda ser cuestionada y reflejada adecuadamente en los análisis de sensibilidad.



Análisis de modelos de decisión con incertidumbre

La incertidumbre de los parámetros surge en un modelo de decisión porque se desconoce el valor real de un determinado dato de entrada (parámetro). En cambio, disponemos de una estimación sobre la que no estamos seguros, porque se obtiene de un estudio con un tamaño de muestra limitado, y posiblemente también porque el estudio sufre algún grado de sesgo interno o externo.

Un método que permite cuantificar y comunicar eficazmente la incertidumbre a los responsables de la toma de decisiones es el análisis de sensibilidad. La versión del modelo en la que todos los parámetros se fijan en su valor medio (o, a veces, en su valor “más probable”) se denomina el modelo del caso base. A partir de este caso base, los análisis de sensibilidad cambian los datos de entrada seleccionados del modelo para que sean mayores o menores que su valor del caso base y comunican los resultados.

El análisis de sensibilidad univariante cambia un parámetro cada vez e informa del resultado, dejando todos los demás sin modificar. Si el analista cree que el estudio subyacente del que se obtuvo el

parámetro era sin sesgo, entonces la principal fuente de incertidumbre será el tamaño limitado de la muestra utilizada para estimar cada parámetro. En este caso, podría ser apropiado cambiar el parámetro en el análisis de sensibilidad para tomar el límite superior e inferior en su intervalo de confianza del 95%. Por ejemplo, si la variable se distribuye normalmente, el rango inferior y superior sería la media $\pm 1,96$ errores estándar. Sin embargo, si se cree que el estudio subyacente que genera el parámetro estaba sesgado, entonces los valores superior e inferior correspondientes deberían ser más amplios para reflejar la incertidumbre adicional en el valor de la variable. Es importante que los valores inferior y superior de los parámetros reflejen estimaciones razonables del rango probable en el que se cree que existe el valor verdadero.

El análisis de escenarios es un término que suele utilizarse para describir un análisis de sensibilidad que refleja la heterogeneidad o la incertidumbre estructural del modelo. Por ejemplo, el análisis del caso base puede utilizar datos clínicos de un ensayo controlado aleatorio, pero un análisis de escenarios podría utilizar datos de un estudio diferente realizado en un tipo de paciente ligeramente distinto. El análisis de escenarios es especialmente útil para explorar cómo las hipótesis sobre el efecto del tratamiento influyen en los resultados del modelo a largo plazo, cuando hay pocos datos disponibles. Por ejemplo, en ausencia de datos, muchos modelos de terapia génica asumen en el caso base que los pacientes que están respondiendo a la terapia al final de un estudio clínico continuarán libres de enfermedad durante 10 años. Los análisis de escenarios podrían describir un supuesto más pesimista (libre de enfermedad durante 5 años) u optimista (libre de enfermedad durante el resto de la vida del paciente) (8). El análisis de escenarios también puede utilizarse para describir subgrupos, por ejemplo, hombres por separado de mujeres.

El análisis de sensibilidad multivariante cambia dos parámetros simultáneamente, manteniendo otros parámetros sin cambios. La idea es tener en cuenta que algunos parámetros interactúan en el modelo,

por lo que el resultado de la variación del parámetro X puede depender del valor que tome otro parámetro Y. Un ejemplo en el que esto puede ser importante es cuando se expresa un efecto relativo del tratamiento utilizando una medida como la razón de hazard. Un hazard ratio es una medida relativa, lo que significa que la tasa de eventos en el grupo de intervención es el producto del hazard ratio multiplicado por las tasas de eventos en el grupo de control. Por lo tanto, el impacto en los resultados (en términos absolutos) del modelo de la variación del hazard ratio (parámetro X) dependerá de la tasa de eventos en el grupo de control (parámetro Y). Por lo tanto, puede ser útil realizar un análisis de sensibilidad multivariante para mostrar cómo la variación de la razón de hazards afecta a los resultados del modelo para diferentes valores de las probabilidades en el grupo de control. El reto del análisis de sensibilidad multivariante es cómo presentar los resultados a los responsables de la toma de decisiones de forma intuitiva y comprensible (véase la sección posterior “Comunicar la incertidumbre a los responsables de la toma de decisiones”).

El análisis de sensibilidad probabilístico (APS) lleva la idea del análisis de sensibilidad multivariante un paso más allá. Muchos parámetros del modelo son inciertos. En un APS, (preferiblemente) todos los parámetros varían simultáneamente, con múltiples conjuntos de valores de parámetros que se muestrean a partir de distribuciones de probabilidad definidas a priori (véase la sección posterior “Las propiedades estadísticas de los parámetros”). Los resultados de un APS pueden servir de base para varias formas de análisis, como los intervalos de confianza, los planos de coste-efectividad (que muestran las distribuciones de costes y efectos de cada tecnología o servicio evaluado), las curvas de aceptabilidad de coste-efectividad (que muestran la probabilidad de que cada opción tenga el mayor beneficio neto) y los análisis del valor de la información.



Las propiedades estadísticas de los parámetros

Ya se ha señalado que, si una variable se distribuye normalmente, los límites inferior y superior del intervalo de confianza del 95% serían la media $\pm 1,96$ errores estándar. Sin embargo, no todos los parámetros están adecuadamente representados por la distribución normal. La Tabla 2 muestra las propiedades de las distribuciones más utilizadas para los parámetros en los modelos de decisión de la evaluación de tecnologías sanitarias. La forma y la posición de la distribución de probabilidad se rigen por un pequeño número de constantes (normalmente dos). Los nombres de las constantes son los que aparecen en el paquete R. Otros programas informáticos (como Excel) incluyen funciones similares, pero pueden utilizar formas diferentes de especificar las constantes.

Ahora examinaremos las propiedades de estas distribuciones y las diferencias entre ellas. Como ejemplo, imaginemos que estamos interesados en un parámetro concreto “x”, y tenemos datos de que la media de x es 0,6 y su varianza es 0,03 (el cuadrado del error estándar de la media). Podemos utilizar las fórmulas de la Tabla 3 para calcular las constantes que rigen cada una de estas distribuciones. La Figura 1 muestra la densidad de frecuencias del parámetro x bajo las funciones de densidad de probabilidad más utilizadas en la evaluación de tecnologías sanitarias (normal, log-normal, gamma o beta).

La Tabla 3 calcula los constantes para caracterizar las funciones de densidad de probabilidad, para una variable cuya valor medio es 0,6 y cuya varianza es 0,03.

En la distribución normal, x no tiene un límite superior o inferior teórico, y la distribución es simétrica en torno a la media. En las distribuciones log-normal y gamma, el límite inferior de x es cero, pero no hay límite superior, y las distribuciones no son simétricas. En la distribución beta, el límite inferior de x es cero y el límite superior es 1, y la distribución no

Tabla 2. Propiedades de las funciones de densidad de probabilidad

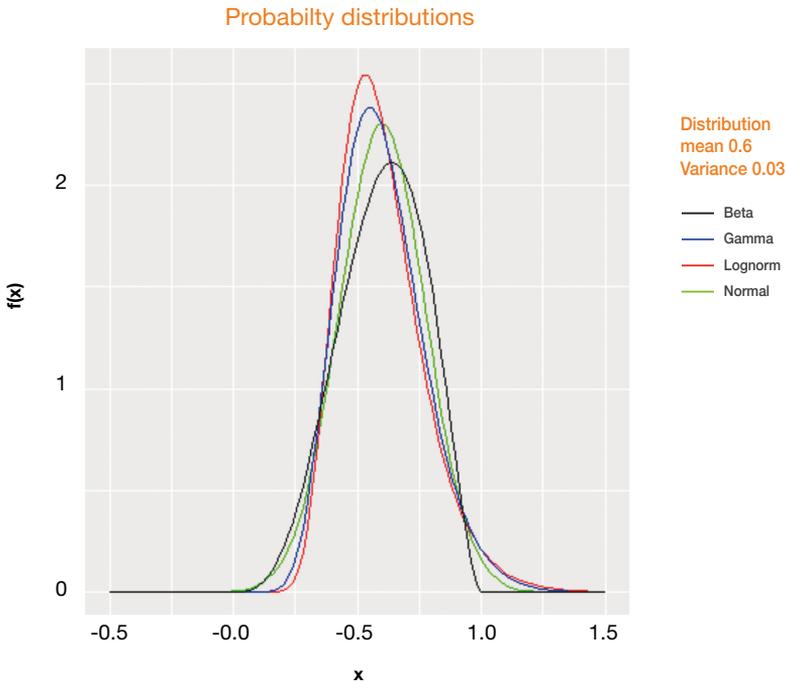
Nombre de la función de densidad de probabilidad	Alcance de mínimo	máximo	y Constantes (como se nombran en R)	Interpretación de las constantes	Parámetros que suelen caracterizarse por la función
Distribución normal	Del infinito negativo al infinito positivo		mean & se		Diferencia de coste, diferencia de utilidad
Distribución logarítmica normal	Estrictamente mayor que cero		mean_log & se_log	La exponencial de una variable con una distribución log-normal tendrá una distribución normal	Odds ratio, Hazard ratio, costes, disutilidad
Distribución Gamma	Estrictamente mayor que cero		scale & rate	El número esperado de eventos por persona en el periodo de tiempo tasa es igual a scale/rate	Tasas de eventos, costes, disutilidad
Distribución Beta	De cero a uno		shape1 & shape2	En un estudio de tamaño N, shape1 es el número esperado de sujetos con un evento y shape2 es el número esperado sin el evento, donde $N = \text{shape1} + \text{shape2}$	Probabilidades de eventos

Tabla 3. Cálculo de las constantes para caracterizar las funciones de densidad de probabilidad a partir de la media y la varianza de una variable

Nombre de la función de densidad de probabilidad	Cálculo de constantes dados los datos de la media m y la varianza v de una variable X	Valor de las constantes, dado que la variable x tiene una media de $0,6$ y una varianza de $0,03$	Intervalo de confianza del 95% para x , si x sigue esa distribución	Código en R para generar 1000 simulaciones de Monte-Carlo	Fórmula en Excel para generar 1 simulación Monte Carlo
Distribución normal	$\text{mean} = m, \text{se} = \sqrt{v}$	$\text{mean} = 0,6, \text{se} = 0,173$	$0,26 - 0,94$	$X \leftarrow \text{rnorm}(1000, \text{mean}, \text{se})$	$=\text{inv.norm}(\text{aleatorio});$ $\text{mean}; \text{se}$
Distribución logarítmica normal	$\text{mean_log} = -\log(m^2/\sqrt{v+m^2})$ $\text{se_log} = \sqrt{(\log(v/m^2 + 1))}$	$\text{mean_log} = -0,55, \text{se_log} = 0,28$	$0,33 - 1,00$	$X \leftarrow \text{rlnorm}(1000, \text{mean_log}, \text{se_log})$	$=\text{inv.lognorm}(\text{aleatorio});$ $\text{mean_log}; \text{se_log}$
Distribución Gamma	$\text{shape} = m^2/v$ $\text{scale} = m/v$	$\text{shape} = 12, \text{scale} = 20$	$0,31 - 0,98$	$X \leftarrow \text{rgamma}(1000, \text{Scale}, \text{rate})$	$=\text{inv.gamma}(\text{aleatorio});$ $\text{Scale}; 1/\text{rate}$
Distribución Beta	$\text{shape1} = m*(1-m)/v - 1$ $\text{shape2} = \text{shape1}*(1-m)/m$	$\text{shape1}=4,2, \text{shape2}=2,8$	$0,24 - 0,90$	$X \leftarrow \text{rbeta}(1000, \text{shape1}, \text{shape2})$	$=\text{inv.beta}(\text{aleatorio});$ $\text{shape1}; \text{shape2}$



Figura 1. Funciones de densidad de probabilidad utilizadas habitualmente en la evaluación de tecnologías sanitarias, para una variable con una media y una varianza determinadas



es simétrica a menos que la media sea 0,5. La distribución beta se utiliza a menudo para modelar probabilidades. Las constantes de la distribución beta tienen una interpretación natural. Por ejemplo, si 70 individuos comienzan un estudio, y 42 experimentan un evento durante el estudio mientras que 28 no lo tienen, entonces la probabilidad del evento puede modelarse como $\text{beta}(42;28)$. También pueden utilizarse constantes que no sean números enteros. La distribución gamma se utiliza a menudo para modelar las tasas de eventos (por ejemplo, el número de eventos por persona-año). Por ejemplo, si un estudio sigue a una cohorte de pacientes durante 20 pacientes-año, y observa 12 eventos, entonces la

tasa media de eventos por paciente-año puede modelarse como gamma(12;20). Nótese que la fórmula que caracteriza la distribución gamma en Excel es diferente a la de R (Tabla 3).

Las distribuciones gamma o log-normal pueden modelar los parámetros de costes o disutilidades. En estos casos, el parámetro no puede ser cero o menor que cero. Las disutilidades son la reducción de la utilidad asociada a un estado de enfermedad más grave, en comparación con un estado de referencia menos grave, como una utilidad igual a uno. La distribución logarítmica normal también se utiliza a menudo para modelar odds ratios o hazard ratios.

El APS se lleva a cabo eligiendo distribuciones de probabilidad para cada uno de los parámetros del modelo de decisión, muestreando aleatoriamente un conjunto de valores de parámetros de cada una de estas distribuciones de parámetros, calculando los resultados del modelo (costes incrementales y AVAC incrementales) para este conjunto de parámetros muestreados, registrando los resultados y repitiendo el procedimiento un gran número de veces. Por convención, 1000 simulaciones iterativas son aceptables. Esto se conoce como muestreo de Monte-Carlo, y requiere un código de programación (por ejemplo, Visual Basic en Excel, o R).



Comunicar la incertidumbre a los responsables de la toma de decisiones

Esta sección describe una serie de herramientas gráficas para comunicar los resultados de los análisis de sensibilidad a los responsables de la toma de decisiones. En Normahani y col. (9) pueden examinar ejemplos de cada una de ellas.

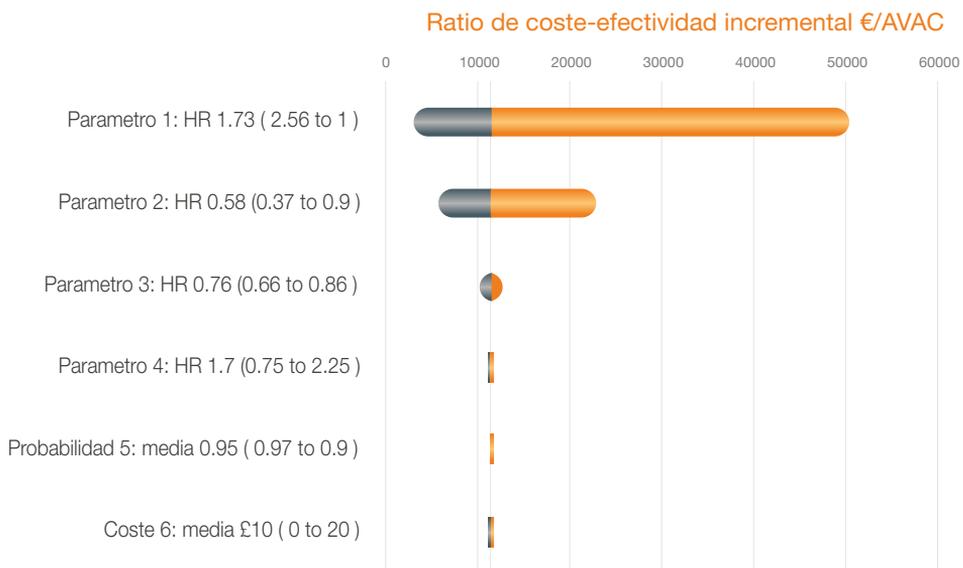
Al informar sobre los análisis de sensibilidad univariantes, es aconsejable mostrar en una tabla o gráfico (como mínimo) la siguiente información: el nombre del parámetro, su valor en el caso base, el resultado (en coste y

AVAC) del caso base, el límite superior e inferior del parámetro y el resultado (en coste y AVAC) de si el parámetro toma su valor superior e inferior. Como se ha señalado anteriormente, es importante que la elección de los límites superior e inferior de cada parámetro utilizado en el análisis de sensibilidad esté cuidadosamente justificada. En principio, los límites deben reflejar estimaciones razonables del rango probable en el que se cree que existe el valor real.

Una herramienta gráfica útil para comunicar el análisis de sensibilidad univariante es el gráfico Tornado. Algunos analistas muestran los resultados por separado para el coste incremental y los AVAC incrementales. La desventaja es que esto no muestra el impacto del parámetro en la decisión, que tiene en cuenta tanto los costes como los AVAC. Algunos analistas muestran los resultados de la relación coste-eficacia incremental (RCEI) (Figura 2). Esto puede ser útil si tanto los costes in-



Figura 2. Gráfico de tornado para mostrar los resultados del análisis de sensibilidad univariante



crementales como los AVAC incrementales son siempre estrictamente positivos en todo el rango de los parámetros analizados, pero si una de estas variables se vuelve cero, o negativa, entonces la RCEI carece de sentido. Otro enfoque consiste en mostrar los resultados del beneficio neto incremental. Esto requiere una elección predefinida sobre el umbral de coste-efectividad que se aplica para esa decisión (10).

Las herramientas gráficas son convenientes para comunicar el análisis de sensibilidad multivariante. Una opción es la que se muestra en la Figura 3. En ella se muestran los posibles valores de un parámetro en el eje horizontal y otro en el vertical. El límite representa el umbral en el que el cambio de un parámetro modificaría el tratamiento que ofrece el mayor beneficio neto medio. Este enfoque requiere una elección predefinida sobre el umbral de coste-efectividad que se aplica para esa decisión.



Figura 3. Tabla de umbrales del análisis de sensibilidad multivariante

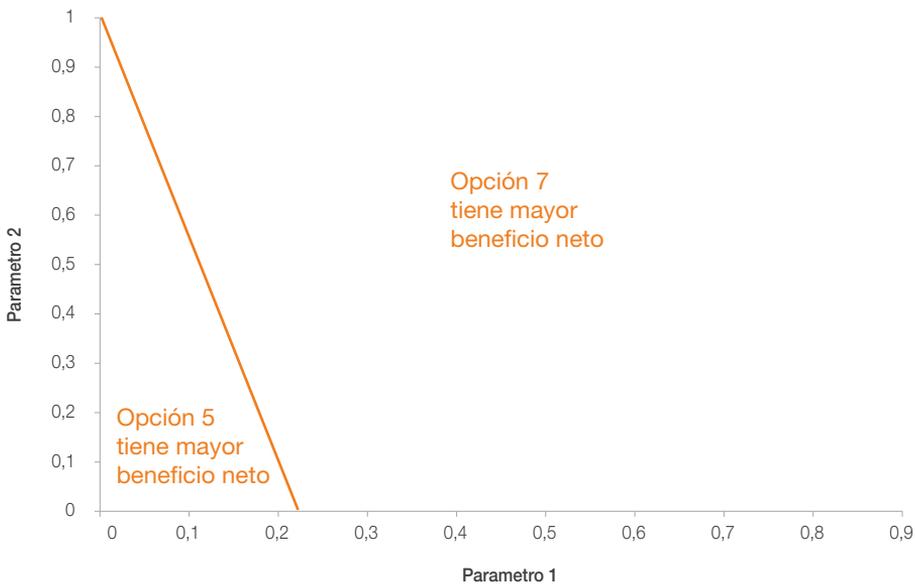
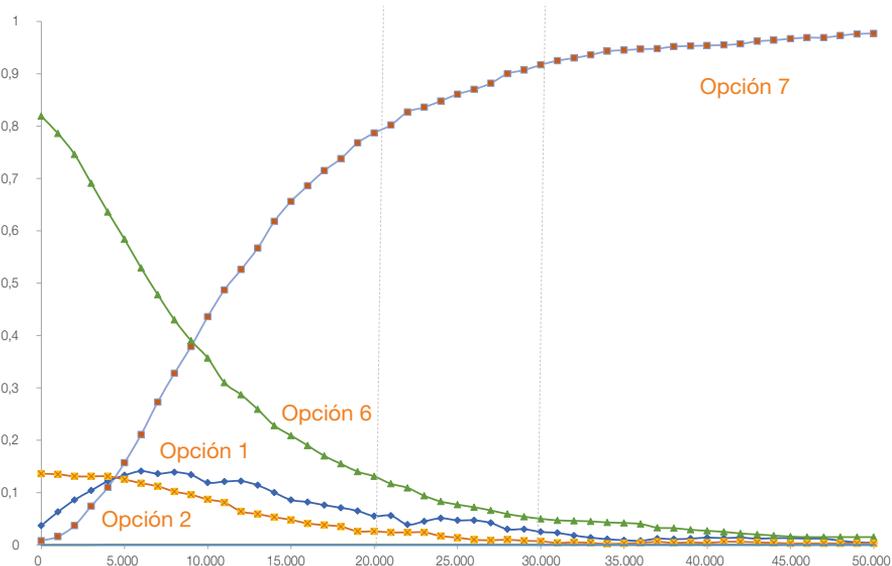


Figura 4. Tabla de umbrales del análisis de sensibilidad multivariante



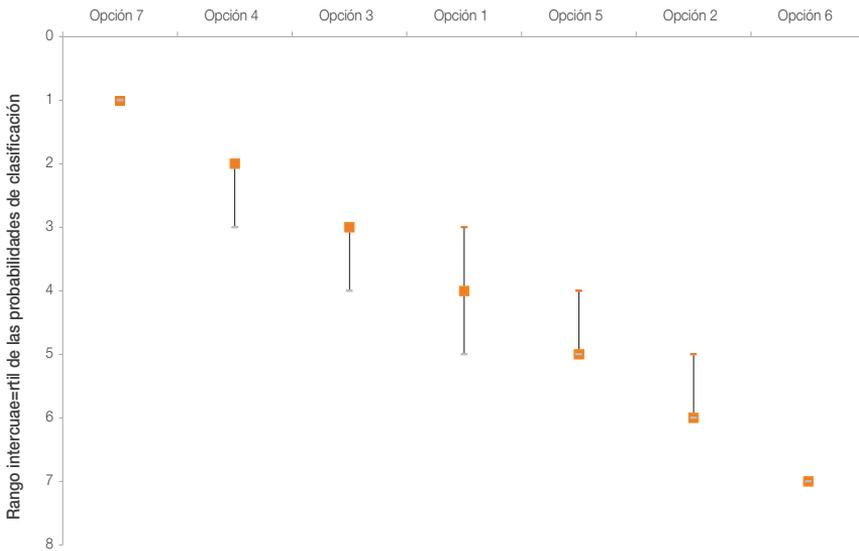
Los resultados del análisis de sensibilidad probabilístico suelen mostrarse mediante una curva de aceptabilidad para cada tratamiento (11). Esto muestra la probabilidad de que un tratamiento tenga el mayor beneficio neto. A menudo se muestra en un rango de umbrales de coste-efectividad, como se muestra en la Figura 4.

Sin embargo, cuando se comparan varios tratamientos, la curva de aceptabilidad no comunica adecuadamente el riesgo (la posibilidad de tomar una decisión incorrecta) al responsable de la toma de decisiones (12). Esto se debe a que la curva de aceptabilidad muestra la probabilidad de que un tratamiento tenga el mayor beneficio neto medio, pero no muestra la probabilidad de que tenga el menor beneficio neto medio.

De ahí que se recomiende que los analistas muestren no sólo la probabilidad de que un tratamiento tenga el mayor beneficio neto, sino



Figura 5. Rango intercuartil de las probabilidades de clasificación, con un umbral de 20.000€ por AVAC



también la probabilidad de que ocupe otros rangos, es decir, el segundo mejor, el tercero mejor y así sucesivamente. Esta información puede calcularse fácilmente a partir del análisis de sensibilidad probabilístico, de forma similar a la curva de aceptabilidad, y puede comunicarse a los responsables de la toma de decisiones mediante un gráfico como el que se muestra en la Figura 5. Los tratamientos se ordenan de izquierda a derecha según el beneficio neto medio. El eje vertical muestra la mediana y el rango intercuartil del rango de esa opción. Por ejemplo, la opción 1 ocupa el tercer lugar o más en el 25% de las simulaciones de Monte-Carlo, y el quinto lugar o más en el 75% de las simulaciones de Monte-Carlo. Este gráfico muestra más claramente que la curva de aceptabilidad que la opción 1 es el tratamiento asociado a la mayor incertidumbre sobre su rango. Existen herramientas gráficas alternativas, como los rank-o-grams, que pueden mostrar información similar sobre el riesgo de tomar una decisión incorrecta.



Conclusión

Los modelos de decisión son herramientas poderosas para mostrar a los responsables de la toma de decisiones cómo los tratamientos podrían alterar el curso de una enfermedad para aportar beneficios a los pacientes y valor a los sistemas sanitarios. Estos modelos son tan exactos y precisos como sus datos y su diseño. Por ello, los analistas deben asegurarse de comunicar adecuadamente a los responsables de la toma de decisiones el grado de incertidumbre asociado a sus resultados. La incertidumbre surge en los modelos de decisión de varias maneras: porque los parámetros se miden con tamaños de muestra limitados, porque los diseños de los estudios utilizados para estimar los parámetros son imperfectos, debido a las diferencias entre las poblaciones, o porque los modelos captan de forma imperfecta el curso de la enfermedad o cómo funciona el tratamiento. En este capítulo se ha descrito una serie de herramientas analíticas y gráficas que pueden ser útiles en este diálogo. Lo más importante es que los analistas deben trabajar en colaboración dentro de equipos multidisciplinares y ser sensibles a los puntos fuertes y las limitaciones de su trabajo.



Bibliografía

1. Diggle PJ, Gowers T, Kelly F, Lawrence N. Decision-making with uncertainty. *Significance*. 2020;17:12.
2. Stigler S. Fisher and the 5% level. *Chance*. 2008;21:12–12.
3. 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.
4. Claxton K, Palmer S, Longworth L, et al. Informing a decision framework for when NICE should recommend the use of health technologies only in the context of an appropriately designed programme of evidence development. *Health Technol Assess*. 2012;16:1–323.
5. Sculpher M, Drummond M, Buxton M. The iterative use of economic evaluation as part of the process of health technology assessment. *J Health Serv Res Policy*. 1997;2:26–30.
6. Briggs A, Weinstein MC, et al. Model Parameter Estimation and Uncertainty: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-6. *Value Health*. 2012;15:835–42.
7. Turner R, Spiegelhalter DJ. Bias modelling in evidence synthesis. *J R Stat Soc Ser A Stat Soc*. 2009;172:21–47.
8. Hettle R, Corbett R, Hinde S, et al. Exploring the assessment and appraisal of regenerative medicines and cell therapy products. York: Centre for Health Economics; 2015.
9. Normahani P, Epstein DM, Gaggero A, et al. Cost-Effectiveness of Diagnostic Tools to Establish the Presence of Peripheral Arterial Disease in People with Diabetes. *Ann Surg*. 2023; 277:e184-e191
10. Vallejo-Torres L, García-Lorenzo B, Serrano-Aguilar P. Estimating a cost-effectiveness threshold for the Spanish NHS. *Health Econ*. 2018;27: 746–61.
11. Barton GR, Briggs AH, Fenwick EAL. Optimal Cost-Effectiveness Decisions: The Role of the Cost-Effectiveness Acceptability Curve (CEAC), the

Cost-Effectiveness Acceptability Frontier (CEAF), and the Expected Value of Perfection Information (EVPI). *Value Health*. 2008;11:886–97.

12. Epstein D. Beyond the cost-effectiveness acceptability curve: The appropriateness of rank probabilities for presenting the results of economic evaluation in multiple technology appraisal. *Health Econ*. 2019;28:801–7.



6

ANÁLISIS, PRESENTACIÓN E INTERPRETACIÓN DE LOS RESULTADOS EN MODELOS ANALÍTICOS DE DECISIÓN

José Manuel Rodríguez Barrios

Early Products, Health Economics & Outcomes Research.

Novartis Farmacéutica



Introducción

La evaluación económica (EE) de tecnologías sanitarias, incluidos medicamentos, programas, servicios y dispositivos sanitarios, se ha convertido en los últimos años en España y en Europa en una herramienta de primer orden para los gestores sanitarios en sus estrategias de asignación de recursos sanitarios y como herramienta de toma de decisiones en la financiación y adopción de nuevas tecnologías (1). El mayor impulso desde el sector público por fortalecer el uso de la EE de intervenciones sanitarias tiene el doble objetivo, en primer lugar, el de conjugar el acceso a sus ciudadanos a aquellos avances terapéuticos que ayuden a mejorar su salud; y por otro lado, conciliar dicho acceso con la sostenibilidad financiera de los sistemas sanitarios públicos. La consideración del criterio de eficiencia convierte a la evaluación económica en herramienta clave en este proceso (2).

La (EE) es una herramienta de investigación evaluativa para informar decisiones complejas. Usando la definición clásica de Drummond podemos definir la EE como el análisis comparativo de cursos de acción alternativos en términos tanto de costes como de consecuencias en salud (3).

Como en cualquier otro ejercicio evaluativo, el análisis correcto de los resultado y la comunicación y reporte de estos, ha de tener como principal objetivo el facilitar la comprensión de este por parte de los usuarios finales del estudio, que van a utilizarlo en la toma de decisiones. En la mayoría de las ocasiones, estos resultados tendrán como objetivo: la elaboración de informes para la presentación de los procesos de financiación, inclusión en guías farmacoterapéuticas o guías clínicas, publicación en un artículo científico, etc.

Por lo tanto, se han de interpretar y comunicar los resultados de forma clara y transparente, con el fin de facilitar su comprensión y uso en la toma de decisiones por parte de los interlocutores finales del proceso de evaluación, a la vez que demostrar de manera transparente y reproducible

como se ha realizado el estudio, reforzando así la credibilidad de este y justificando su utilización en la toma de decisiones.

En diferentes países se han publicado e implementado guías metodológicas que explicitan los criterios acerca del análisis, interpretación y presentación de resultados que describiremos brevemente en las siguientes líneas. Así, podemos destacar guías de diferentes agencias evaluadora, tanto de países de nuestro entorno como Reino Unido, Países Bajos, Suecia (4-6) o de otros con larga tradición en el uso de EE de tecnologías sanitarias como Canadá o Australia (7,8), o incluso en otros de reciente implementación como Japón (9). Cabría también destacar la propuesta de guías realizadas en 2010 por la red de agencias de evaluación de tecnologías (1), y la guía de recomendaciones del CatSalut publicada en 2014 (2).

Por otro lado, y de cara a la publicación de los resultados en revistas científicas, existen también consensos y guías que reúnen los requisitos de calidad para presentar y comunicar las EE (10).

Como habréis visto en capítulos anteriores, la EE, se fundamenta en el *análisis marginal*, es decir en la comparación de resultados incrementales de las opciones evaluadas (en términos de ahorros o costes adicionales producidos al generar una unidad adicional de resultado), con el objetivo fundamental de contribuir a la asignación eficiente de los recursos disponibles (1). En general, el objetivo de una EE es ofrecer información relevante para los tomadores de decisiones, basada en la mejor evidencia disponible, para promover una asignación eficiente de los recursos disponibles que permitan maximizar el bienestar de los usuarios del sistema de salud.

Los resultados de cualquier modelo de análisis de decisión pueden comunicarse a través de dos tipos de formas de presentación de estos: numéricos y gráficos. Es recomendable utilizar tanto la representación numérica como gráfica (11). Los distintos tipos de presentación y análisis de los resultados serán claves para el entendimiento de estos y van a estar presentes en función de cómo hallamos construido y preparado el

modelo de evaluación, el tipo de modelo empleado, etc.; por ejemplo, un modelo sencillo de análisis de decisiones no podrá presentar una representación gráfica mediante curva de aceptabilidad coste efectividad y presentará su análisis de sensibilidad con una gráfica de análisis tornado en el mejor de los casos.

Las formas básicas de presentar y analizar los resultados en un modelo coste efectividad, (también en un modelo coste utilidad, pero por simplificación hablaremos de coste efectividad), son:

- Cálculo coeficiente Coste efectividad incremental.
- Representación gráfica en el plano coste efectividad incremental.
- Cálculo y representación de la curva de aceptabilidad.
- Beneficio neto sanitario.
- Cálculo de la frontera de eficiencia.
- Otros: El valor esperado de la información perfecta.



Cálculo de coeficientes coste efectividad incrementales: La Razón Coste Efectividad Incremental

Partimos de recordar la idea expresada anteriormente de que el análisis coste efectividad (ACE), es una herramienta de investigación evaluativa que se basa en la comparación de estrategias de tratamiento o cursos de acción alternativos, por lo tanto, es una herramienta de comparación. El objetivo real de la evaluación no va a ser el de definir como de efectivas o cuan costosas son las diferentes alternativas de tratamiento evaluadas, sino cuanto de más o menos efectiva y más o menos costosas son las opciones compradas frente a las demás alternativas evaluadas.

A la hora de comparar, por ejemplo, dos tratamientos alternativos en términos de costes y efectos sobre la salud, podemos encontrarnos diferentes escenarios:

1. Que una intervención nueva sea tan efectiva o más que la elegida de referencia en la evaluación y además menos costosa, con lo que aceptaríamos dicha nueva intervención, (la nueva intervención es dominante).
2. Que la intervención nueva a evaluar sea menos efectiva o igual que la de referencia además que tenga un coste mayor, en este caso rechazaríamos la opción nueva, (la nueva intervención es dominada por la intervención con la que se la compara).
3. La nueva intervención es más efectiva que la de referencia, pero más costosa que ella.
4. La nueva intervención es menos efectiva y costosa que la de referencia con la que se la compara.

Es en estos dos últimos escenarios en los que necesitamos de una herramienta como la razón coste efectividad incremental (RCEI) para ayudarnos a la toma de decisiones, ya que nos permitirá obtener el coste adicional que deberíamos pagar por el benéfico adicional que obtenemos con la nueva intervención. Podemos calcular la RCEI mediante el siguiente cálculo:

$$\text{RCEI} = \frac{\text{Unidades Coste alternativa A} - \text{Unidades Coste alternativa B}}{\text{Unidades Efectividad alternativa A} - \text{Unidades Efectividad alternativa B}}$$

Consideremos ahora el siguiente ejemplo: tenemos dos tratamientos farmacológicos para el tratamiento de la diabetes tipo 2: A y B, que deben usarse oralmente de forma crónica. Vamos a considerar para simplificar un horizonte temporal de un año de tratamiento. La diferencia entre ambos tratamientos es la reducción de diferentes complicaciones de la diabetes que han sido medidas en sus ensayos clínicos mediante un resultado agregado, que ha supuesto 1000 complicaciones evitadas para el fármaco A, y 1200 para el fármaco B. Por otro lado, el coste anual

del tratamiento con el fármaco A en una población fija que debemos tratar es de 200.000 euros y con el fármaco B de 300.000 euros.

Para realizar el cálculo de la razón coste efectividad media del tratamiento A, que se realiza dividiendo los costes atribuidos a dicho fármaco por las unidades de efectividad empleadas en la medición de su resultado clínico, es decir, 200.000 € dividido entre 1.000 casos prevenidos nos da como resultado 200 € por complicación evitada. Es decir, con el fármaco A cada complicación de la diabetes tipo 2 de las recogidas en el ensayo clínico evitada cuesta 200 €. Respecto al tratamiento B, la razón coste efectividad media, si realizamos el mismo cálculo, nos da un resultado de 250 € por exacerbación prevenida (resultado de dividir 300.000 € entre 1.200 complicaciones evitadas).

Sí nos quedamos en este análisis de costes medios por resultado conseguido, el tratamiento A parece una opción más atractiva, ya que nos supone un menor coste por complicación evitada que el B (200 €/complicación evitada frente a 250 €/complicación evitada). Sin embargo, para una población a tratar igual en ambos casos debemos tener en cuenta que la elección de la alternativa A la debemos de poner en el contexto poblacional del grupo de pacientes que debemos tratar y para la que queremos maximizar el resultado en salud, (evitación de complicaciones de la diabetes), en función de la inversión económica a realizar.

En este punto, volvamos a recordar que la EE es un análisis comparativo, y por lo tanto en dicho contexto poblacional mencionado anteriormente, la pregunta correcta es ¿merecería la pena pagar 100.000 € adicionales por prevenir en nuestra población cubierta 200 casos adicionales de complicaciones de la diabetes tipo 2?

Dicho de otro modo, el análisis que nos ayudará a tomar la decisión en los términos descritos es el incremental (cálculo de la RCEI), que describiremos a continuación.

Para el cálculo de la RCEI analizamos el balance entre el mayor coste de una intervención respecto a su alternativa y la mayor efectividad de una intervención frente a su alternativa, teniendo en cuenta que nuestra

decisión ha de ser la maximización del resultado de salud con la inversión de la que disponemos.

La RCEI lo caracterizamos por ser la diferencia de costes de las intervenciones comparadas entre la diferencia de efectos de estas, es decir, si llevamos nuestros datos a la razón anteriormente descrita:

$$\text{RCEI} = \frac{\text{Unidades Coste alternativa A} - \text{Unidades Coste alternativa B}}{\text{Unidades Efectividad alternativa A} - \text{Unidades Efectividad alternativa B}}$$

Obtendríamos para nuestro ejemplo un resultado de:

$$\text{RCEI} = \frac{300.000\text{€} - 100.00\text{€}}{1200 \text{ complicaciones evitadas} - 1000 \text{ complicaciones evitadas}} = 500\text{€/complicación evitada}$$

De manera intuitiva, esta razón nos presenta los resultados en términos marginales y su interpretación nos lleva a plantear el coste adicional que hay que invertir para conseguir una unidad adicional de efectividad, en nuestro ejemplo al elegir la intervención A frente a la B habría que pagar 500 euros adicionales por complicación adicional evitada.

Como sin duda, habréis visto en los capítulos de este libro, la EE en salud es una herramienta de ayuda a la toma de decisiones por parte de los responsables de estas, por lo tanto, la decisión de si nuestro resultado adecuad a tener en cuenta por evitar una complicación de la diabetes tipo 2, es algo fuera del ámbito de decisión del equipo que realiza el ACE.

El objetivo de quién realiza la EE es presentar esta información a las personas que han de tomar la decisión de asignar los recursos entre distintas intervenciones o programas, de la forma más clara y cumpliendo los requerimientos metodológicos que apliquen según las guías o recomendaciones de las que hemos hablado anteriormente.

Para realizar dicha toma de decisiones, los responsables de las mismas, en función de los recursos que se está dispuesto a asumir a cambio de cada unidad adicional de mejora en el resultado tomaran dicha decisión. Para ello, deberán tener en cuenta la gravedad del problema analizado y las preferencias sociales frente a dicha situación, e incluso las preferencias de los pacientes, siendo además conscientes de la cantidad de recursos que tienen a su disposición para el tratamiento de la población asignada y conscientes de que sus decisiones deben tener en cuenta la sostenibilidad del sistema, así como otros argumentos complementarios y no sustitutivos al ACE, (por ejemplo la existencia de alternativas de tratamiento para dicha población, la gravedad de los eventos a evitar, el impacto presupuestario, etc.).

Un elemento importante a tener en cuenta a la hora de construir las RCEI es que, al ser un cociente, el denominador del mismo ha de ser positivo. Es decir, cuando en nuestra evaluación estamos hablando de años de vida, lo hacemos como años de vida ganados, (o ganancia en un parámetro fisiológico etc.), sin embargo, cuando hablamos de efectos adversos o complicaciones (riesgo negativo para el paciente), hablamos de efectos “evitados”. Con esto queremos decir, que por ejemplo para los años de vida el denominador será la diferencia de la alternativa con mayor efecto frente a la de menor y en el caso de los efectos adversos será la de mayores al revés. Tener en cuenta esto nos lleva a expresar el análisis de resultados en términos incrementales que es la forma en la que se deben expresar los mismos en EE.

El RCEI es un indicador puntual y unidimensional, tanto si hemos realizado su cálculo a través de simulaciones como si se ha calculado a través de valores medios obtenidos de la literatura o de los resultados de revisiones sistemáticas cuantitativas, (metaanálisis), con el fin de obtener los parámetros que permiten obtener los resultados. Dichos resultados, han de presentarse mostrando la incertidumbre ligada a su cálculo y se deben calcular sus intervalos de confianza, (normalmente al 95%).

La forma tradicional de reportar los resultados que se ha venido utilizando en los ACE, ha venido siendo la combinación de la representación numérica mediante la razón coste efectividad incremental para comparar las razones coste efectividad o coste utilidad (12), sin embargo, existen otras opciones que revisaremos a lo largo de este capítulo.

El tema de la incertidumbre asociada a los resultados de los ACE en los modelos de EE ha sido tratado en un capítulo anterior de este libro, sin embargo y resumiendo mucho, a efectos pedagógicos y para explicar algunas de las formas de análisis y presentación de resultados que vamos a explicar en este capítulo, podemos decir que la incertidumbre se puede analizar utilizando los siguientes tipos de análisis:

1. Mediante métodos estadísticos asociados al cálculo de la RCEI. Métodos tradicionales como el cálculo de intervalo de confianza por el método Fieller, cálculo de cuasi intervalos de confianza, elipses de confianza, etc. (13). Estos métodos pueden aplicarse también en los modelos de simulación, tanto si utilizamos técnicas paramétricas de simulación cuando trabajamos con datos agregados (a través de simulaciones de Montecarlo), como a través de métodos no paramétricos como el bootstrapping.
2. Métodos gráficos mediante el desarrollo y análisis de curvas de aceptabilidad coste efectividad (CACE).
3. Mediante el conteo y presentación del porcentaje de puntos presentes en cada cuadrante del plano coste efectividad incremental.



Representación gráfica del RCEI y reglas de decisión. El plano coste efectividad incremental

Como hemos comentado anteriormente, el análisis de resultados del ACE se basa en la aplicación del análisis marginal entre las alternativas evaluadas, de tal forma que obtenemos una razón que nos informa sobre el coste adicional que debemos invertir en el pago de una unidad adicional de efectividad que proporciona la opción más efectiva, (definición de RCEI).

Para proceder a su cálculo, como hemos visto previamente, en dicho cálculo comparamos dos a dos las alternativas de tratamiento evaluadas, (en caso de tener más de dos alternativas posibles), tanto en términos de los costes como los valores de la variable de efecto usadas en el análisis para el cálculo de los resultados.

Si representamos gráficamente los valores posibles de los costes en un eje vertical, y los efectos, en el eje horizontal, podremos construir el plano coste efectividad incremental, donde representaremos los diferentes resultados obtenidos. En el origen del plano donde se cruzan los ejes horizontal y vertical, se representa los costes y efectividades del comparador tomado como referencia en nuestro análisis (14). Los valores del eje horizontal (X) definen las diferencias en términos de efectividad, mientras los del vertical (Y), definen las diferencias en términos de costes. Veremos a continuación la dinámica de toma de decisiones en función de donde caiga la representación gráfica de los resultados de nuestros análisis coste efectividad, respecto a cada uno de los cuadrantes representados en el plano coste efectividad incremental (PCEI).

Tomando como referencia el punto de origen del PCEI, podemos diferenciar cuatro regiones o cuadrantes diferentes donde se pueden representar, en función de su valor los diferentes RCEI (15). Estos cuatro cuadrantes, suelen representarse por los puntos cardinales y se denominan:

- Cuadrante noroeste (NO), al cuadrante superior izquierdo.
- Cuadrante noreste (NE), al cuadrante superior derecho.
- Cuadrante suroeste (SO), al inferior izquierdo.
- Cuadrante sureste (SE); al inferior derecho.

Este plano, nos ayuda a definir gráficamente si los resultados obtenidos en nuestro RCEI son coste efectivos o no, pero para ello, necesitamos tomar un umbral de aceptabilidad o valor del máximo coste adicional por unidad de efecto que un decisor sanitario está dispuesto a pagar por la innovación aportada por la tecnología más eficiente.

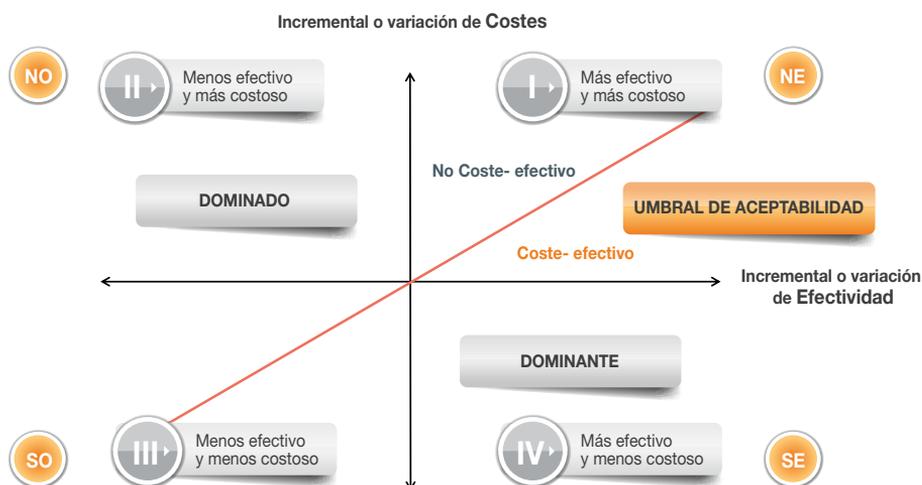
En España no existe un umbral explícito tomado como referencia, si bien Sacristán y col. en el año 2002 tras revisar los estudios de evaluación económica publicados en España hasta la fecha recomendaban la adopción de intervenciones o productos con un coste efectividad incremental por debajo de los 30.000€ por año de vida ganado (AVG) (16). Es por ello, que este es el umbral que se ha venido aceptando mayoritariamente desde entonces, no sólo para ACE que tienen como variable resultada de efectividad el año de vida ganado (17), sino también otras variables como el año de vida ajustado por calidad en el caso de los estudios de coste utilidad (18).

Aunque no existe en España un umbral explícito o disponibilidad máxima a pagar, se ha venido utilizando hasta ahora el umbral de 30.000 € por unidad de efectividad, sin embargo, recientes estudios han planteado diferentes umbrales dinámicos como el estudio de Vallejo-Torres y colaboradores que lo sitúan entre 22.-25.000€ por AVAC (19) o el reciente de Sacristán y colaboradores que lo sitúan entre 25-60.000€ por AVAC (20). Pero, más allá del valor cuantitativo del umbral de eficiencia, tomaremos aquí como referencia, y con fines meramente didácticos, con el afán de facilitar la comprensión de su papel en la toma de decisiones en el PCEI el valor de 30.000€ por unidad de efectividad

Fijémonos en la siguiente figura (Figura 1) donde representamos el PCEI con las cuatro regiones anteriormente mencionadas y en las que la línea roja que pasa por el origen representaría el umbral de aceptabilidad.

Figura 1. Plano de Coste Efectividad Incremental

Plano de Coste-Efectividad Incremental



Los valores situados por encima del umbral de aceptabilidad (línea roja) se considerarán no coste efectivo, ya que superarían dicho umbral, sin embargo, los valores por debajo de dicha línea serían considerados coste efectivo.

La toma de decisiones en cada uno de los cuadrantes donde se sitúe el resultado será diferente. De tal forma, que al comparar una alternativa que sea menos costosa y más efectiva que la de referencia, (punto de corte de los ejes), esta se situaría en el cuadrante SE y es lo que se denomina una alternativa dominante, que debería ser aceptada, ya que proporcionaría un mejor resultado clínico, (mayor efectividad) a un menor coste para el sistema.

En el cuadrante NO se situarían las alternativas menos efectivas y más costosas, por lo que serían alternativas dominadas por la de referencia y no deberían adoptarse.

Sin embargo, las situaciones que más frecuentemente nos encontramos son las de obtener resultados en el cuadrante NE, es decir alternativas más efectivas que la de referencia, pero con un coste mayor. Es en estos casos donde el análisis del RCEI, en función del umbral de aceptabilidad que se ha de tomar de referencia, juegan un papel clave en la toma de decisiones.

Los valores en el cuadrante SO suponen alternativas menos efectivas, y menos costosas que la de referencia, en principio y desde la asunción de que el decisor sanitario priorizaría opciones más efectivas en base a su disponibilidad a pagar, las menos efectivas deberían ser rechazadas.

Resumiendo, a la hora de decidir sobre la caracterización de una alternativa como coste efectivo, deberíamos situar y analizar dicha alternativa en el PCEI y decidir en función de cuál es el cociente coste efectividad aceptable que se está dispuesto a pagar por una unidad más de efectividad, es decir la decisión va a estar supeditada al umbral de aceptabilidad o disponibilidad a pagar por dicha unidad adicional de efecto.



Curva de aceptabilidad de Coste Efectividad

Un problema, respecto a valorar el resultado en el PCEI es que no sabemos qué valor monetario asignara el decisor a cada unidad de efectividad ganada en algunos entornos o cuando no existe un criterio de disponibilidad a pagar explícito. Pero, lo que si conocemos es el valor máximo y mínimo de dicho umbral: entre infinito y cero. Por lo tanto, lo que podemos hacer es determinar la proporción del número de resultados de la nube de puntos que obtenemos de las simulaciones del modelo para diferentes umbrales de aceptabilidad. El porcentaje de dichos resultados representaría la probabilidad en tanto por uno de tener un resultado coste efectivo para cada umbral de decisión dado.

Si hacemos este ejercicio para diferentes umbrales de aceptabilidad posibles en el plano, representamos los diferentes umbrales de aceptabilidad en el plano el eje horizontal y los porcentajes obtenidos para cada umbral en el vertical tendríamos la curva de aceptabilidad de coste efectividad (CACE).

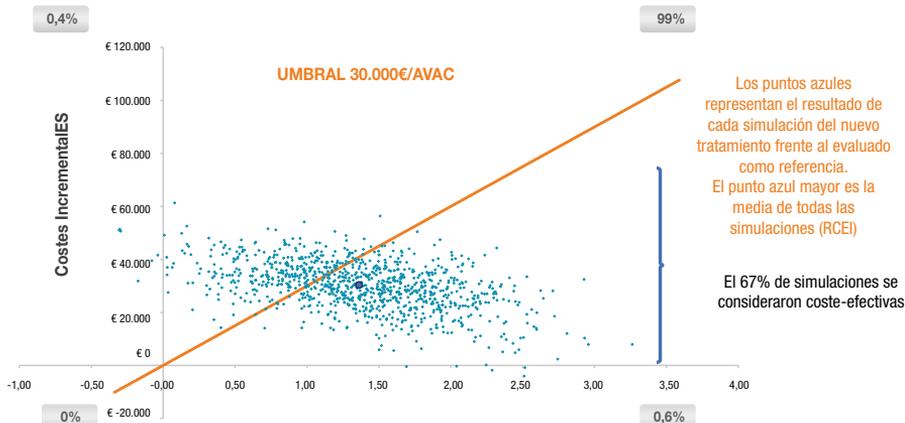
En un modelo de evaluación en el que hayamos obtenido un gran número de resultados de las simulaciones mediante del modelo (nube de puntos), simplemente dividimos el número de simulaciones que corresponden a la parte del plano que son coste-efectivas, entre el número total de replicas (puntos), simuladas en el modelo, y de esta forma obtenemos una estimación de la proporción puntos que corresponden a los resultados la parte coste-efectivos para el umbral dado tenido en cuenta.

Para explicarlo con un ejemplo, supongamos que tenemos un estudio coste utilidad que compara dos terapias diferentes, en el que al representar gráficamente el resultado en el plano coste efectividad incremental de las razones costes utilidad incrementales (RCUI) obtenidos, obtendríamos la representación mostramos en la Figura 2. En dicho



Figura 2. Resultado modelo con 67% de RCUI bajo el umbral

La probabilidad de ser coste-efectivo bajo el umbral de 30.000€ es del 67%



plano para un umbral de aceptabilidad de 30.000 € por unidad de efecto adicional, en este caso años de vida ajustados por calidad, (AVAC), obtenemos que el 67% de los resultados de las simulaciones están bajo el umbral de 30.000€.

Como vemos en la figura, la mayoría de los resultados simulados producirían una RCUI bajo el umbral de aceptabilidad, en este caso el 67%, además existe un porcentaje de puntos (0,6%) en el cuadrante SE, (cuadrante dominante), por lo que podríamos establecer que el resultado del modelo nos presenta un resultado en el que la terapia evaluada frente al tratamiento comparado, (normalmente el estándar de tratamiento), es una opción coste efectiva.

En esta representación gráfica vemos un punto central de mayor tamaño que representaría el valor medio de la RCUI que obtenemos en el modelo.

Sin embargo, los umbrales de aceptabilidad pueden ser diferentes en distintos países, o pueden variar en función de que las preferencias de los decisores (o de la sociedad); también puedan ser distintas en función de elementos como la gravedad de la patología, (enfermedades al final de la vida en el caso de Reino Unido, por ejemplo), o ser distinto para enfermedades raras poco prevalentes, etc., con lo que sería más útil la construcción de una curva que tuviera en cuenta los resultados para diferentes umbrales de aceptabilidad (21).

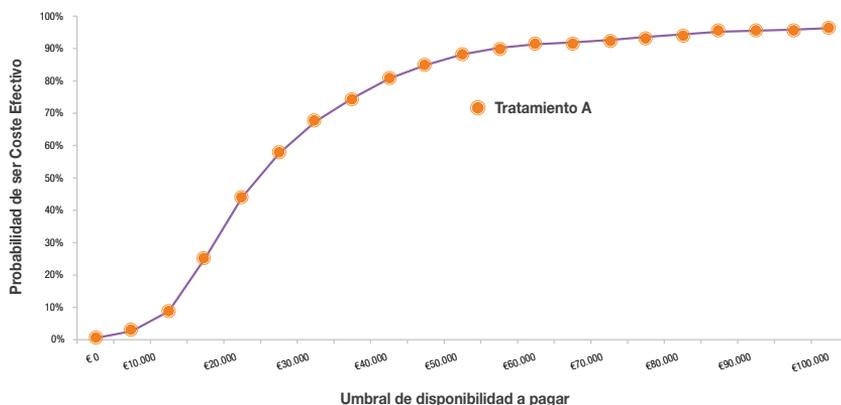
Supongamos que repetimos la gráfica anterior modificando los umbrales de aceptabilidad, es decir, que además del umbral de 30.000€ por año de vida ajustado por calidad, al programar la gráfica, lo hacemos para que tenga en cuenta diferentes umbrales de aceptabilidad, por ejemplo, entre 0€ y 100.000€. Además, representaremos los diferentes umbrales de aceptabilidad en el eje horizontal y en el eje vertical representamos el tanto por ciento entre 0 y 100%, (También podríamos hacerlo en tanto por uno); por último, programaríamos la gráfica para que contase los puntos por debajo de cada uno de los umbrales entre 0 y 100.000. La representación gráfica que obtendríamos es la (CACE), que

vemos en la Figura 3. Es decir, la CACE un gráfico de líneas donde los valores del eje de ordenadas son el porcentaje de simulaciones en que se acepta la intervención analizada, (resultado menor o igual a cada umbral de aceptabilidad) y los valores del eje de abscisas son las etiquetas de los diferentes umbrales. Lo habitual es que al aumentar la disposición a pagar aumente la probabilidad de aceptación de los resultados, esto solo sucede cuando las simulaciones aparecen en el cuadrante de dominancia y en el cuadrante de resultado más efectivo, pero más costoso.



Figura 3. Curva de aceptabilidad del coste efectividad

Curva de aceptabilidad coste efectividad



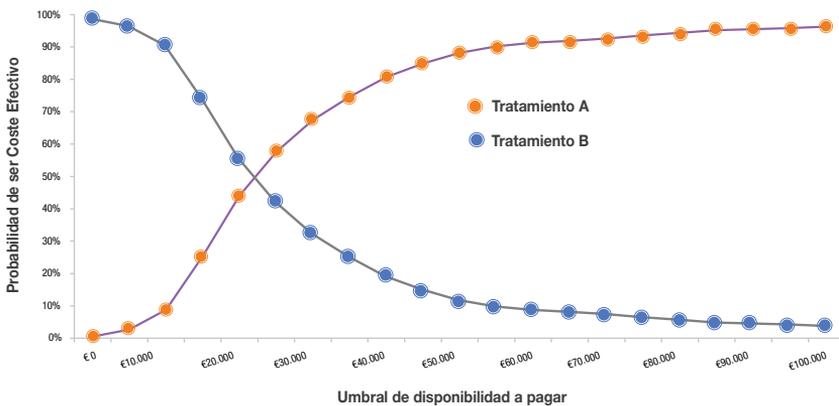
La interpretación de la CACE es tal y como lo hemos descrito anteriormente, es decir nos da la probabilidad de que nuestro resultado de la RCEI sea coste efectivo para diferentes umbrales de las diferentes disponibilidades a pagar; por lo tanto, un resultado con una probabilidad alta (valores claramente por encima de 50%), que sean no sólo costes efectivos para el umbral tipo considerado, sino incluso para umbrales inferiores sería aceptado como coste efectivo.

En la CACE, si hemos evaluado más de una alternativa frente a un comparador común, (estándar de tratamiento, por ejemplo) se pueden representar de manera conjunta las CACE de las alternativas comparadas frente al tratamiento de referencia (Figura 4), lo cual nos sirve para analizar la probabilidad de ser coste efectivo de diferentes tratamientos alternativos en función de la disponibilidad a pagar, pudiendo cambiar la decisión de que tratamiento sería la opción más coste efectiva según el umbral de disponibilidad a pagar.



Figura 4. Representación de la CACE de dos tratamientos

Curva de aceptabilidad coste efectividad (CEAC)



Como observamos en este caso, el tratamiento B sería la opción con mayor probabilidad de ser coste efectivo para disponibilidades a pagar entre 0€ y unos 25.000€ por AVAC, sin embargo a partir de dicha cifra, o dicho de otro modo para disponibilidades a pagar superiores a 25.000€, la alternativa de tratamiento con mayor probabilidad de ser coste efectivo sería el tratamiento A.

La representación de la CACE múltiple es también de utilidad cuando, además de variedad de tratamientos analizados en la EE, tenemos

heterogeneidad de resultados entre grupos de pacientes, y está representación nos ayudaría, al realizar el análisis en cada subgrupo de pacientes, es decir, a ver las diferentes probabilidades de resultado coste efectivo en cada uno de los grupos.



El beneficio neto sanitario

Como hemos visto anteriormente, la regla de decisión habitual para la toma de decisiones en EE se fundamenta en el análisis del RCEI en función de un umbral de aceptabilidad preestablecido.

Este enfoque, sin embargo, puede presentar algunos problemas difíciles de manejar estadísticamente, por ejemplo: a la hora de obtener los intervalos de confianza de la RCEI o a la hora de calcular la varianza de la RCEI, (necesario para el cálculo de los intervalos de confianza de la RCEI), ya que podemos obtener razones coste efectividad con valores negativos de difícil interpretación.

Por ello, el utilizar el beneficio neto sanitario (BNS) como indicador de resultado en la EE, intenta solucionar estos problemas (22), y nos permite calcular y trabajar con un estadístico más fácil de usar que supone una solución a los problemas anteriormente mencionados de los intervalos de confianza de la RCEI, facilitando además una regla de decisión más intuitiva que la valoración en función de valores umbrales de disponibilidad a pagar.

Ahora crearemos la función de beneficio neto sanitario, para ello lo primero que debemos definir es un valor de disponibilidad a pagar por AVAC ganado, (o por años de vida ganados (AVG), si lo hacemos para un modelo coste efectividad cuya variable resultado de efectividad sea el AVG). Normalmente la regla de decisión basada en la RCEI medio nos viene dada por:

$$\widehat{RCEI} = \frac{\bar{C}_1 - \bar{C}_0}{\bar{E}_1 - \bar{E}_0} = \frac{\Delta \bar{C}}{\Delta \bar{E}} = Rc \quad [1]$$

La regla de decisión nos dice que si la RCEI es menor que un parámetro que conocemos como la disponibilidad a pagar por unidad adicional de efectividad ganado (R_c), la tecnología es eficiente, y si está por encima de este umbral, la tecnología que se está evaluando no lo es. El valor de R_c , en los ejemplos que hemos visto hasta ahora sería un valor umbral de 30.000€ por unidad adicional de efectividad.

El problema es que no siempre es sencillo definir lo que llamamos umbral de eficiencia (o de coste-efectividad), la R_c de la ecuación. Además, otro problema todavía mayor, es como hemos comentado anteriormente la dificultad de si cabe, es la dificultad en realizar análisis estadísticos del RCEI, como calcular su varianza.

Para manejar mejor estadísticamente los resultados de los ACE, se ha elaborado un índice llamado BNS. El BNS está directamente relacionado con la RCEI, y se define a partir de la ecuación [1] como:

$$BNS = R_c \cdot \Delta E - \Delta C$$

Por lo que, si la tecnología es eficiente, el BNS debe cumplirse que:

$$BNS = R_c \cdot \Delta E - \Delta C > 0$$

Esto, intuitivamente nos lleva a afirmar que lo que estamos dispuestos a pagar por el beneficio adicional que obtenemos con la terapia que estamos comparando, es mayor al coste incremental al que tengo que incurrir.

La ventaja de la utilización del BNS es que es fácilmente manejable a nivel estadístico, se puede demostrar que la varianza del BNS es, (siendo Var la varianza y Cov la covarianza):

$$Var(BNS) = R_c^2 \cdot Var(\Delta E) + Var(\Delta C) - 2R_c Cov(\Delta E, \Delta C)$$

Por lo que resulta sencillo obtener el intervalo de confianza del BNS:

$$(BNS - z_0 \sqrt{Var(BNS)}, BNS + z_0 \sqrt{Var(BNS)})$$

En la práctica a la hora de programar un modelo de evaluación, los diferentes softwares nos permiten hacerlo de forma sencilla.

Una vez tenemos calculado el BNS y calculamos su intervalo de confianza al 95%, la regla de decisión es clara, es decir, siempre que obtengamos un BNS mayor que 0, y que su intervalo de confianza excluya dicho valor podríamos decir que la intervención evaluada sería coste efectiva. En el caso de más comparadores, calcularemos el beneficio neto medio para cada opción por separado y escogemos la opción con el mayor beneficio neto. Al hacer este cálculo, ya no nos preocupamos sobre la dominancia y tampoco nos preocupamos acerca de la opción con la que comparamos. Como se puede ver, es más fácil de aplicar, ya que únicamente la opción con el mayor beneficio neto tiene un beneficio neto incremental positivo respecto a todas las demás opciones.

En resumen, respecto al beneficio neto, si multiplicamos el efecto incremental en efectividad por el umbral de disponibilidad a pagar tenemos la máxima disponibilidad a pagar por la efectividad ganada. Si a esto le restamos el incremento de costes tenemos el beneficio neto en términos monetarios.

El beneficio neto en términos de salud se obtiene restando al incremento de efectividad el cociente resultante de dividir la variación de costes entre el umbral de aceptabilidad o disposición a pagar. Esto es lo que hemos explicado en los párrafos anteriores.

En ambos casos un resultado positivo significa que la opción es coste efectiva, mientras que si el resultado es negativo la intervención es coste-inefectiva, y a mayor BNS, mejor resultado.



La frontera de eficiencia coste efectividad

Una de las paradojas identificadas en algunos casos es que, aunque una alternativa de tratamiento de las evaluadas tenga la probabilidad más alta de ser coste-efectiva, puede no tener el beneficio neto esperado más alto necesariamente. Por esa razón, la curva de aceptabilidad a pagar no debería ser utilizada para identificar la alternativa óptima. Este efecto depende de la distribución conjunta de los costes y efectos de cada alternativa y la posición relativa en el plano coste-efectividad (23,24). Se podría pensar, en base a la CACE que el tratamiento óptimo es el que tiene el mayor número de simulaciones a su favor, es decir, la opción que tiene el CACE más alto. Sin embargo, cuando el modelo es no lineal, esto no siempre se cumple ya que cuando miramos a las proporciones no consideramos las magnitudes de las simulaciones.

En resumen, la frontera de aceptabilidad coste efectiva no tiene necesariamente por qué coincidir con la CACE más elevada. Esto nos lleva a plantearnos la elaboración de la frontera de decisión por disponibilidad a pagar.

Para su elaboración se deben de seguir los siguientes pasos, (que se programaría si el modelo se hace en Excel o podría realizarse directamente con otros softwares de modelización):

- A. Se calcula la media de coste y efectos mediante la simulación de cada una de ellas; ordenándose las alternativas de menor a mayor efectividad.
- B. Se calcula la RCEI de cada alternativa frente a la alternativa que tiene una efectividad menor. Al estar ordenado, se calcula la ratio de cada alternativa versus la alternativa previa.
- C. Cuando se calcula la RCEI es importante identificar y excluir aquellas opciones que están dominadas (más costosas y menos

- efectivas) y aquellas que tienen dominancia extendida (combinaciones de opciones que pueden mostrar un mayor beneficio con el mismo coste). Una alternativa muestra dominancia extendida cuando la RCEI no aumenta al aumentar la efectividad o, dicho de otro modo, cuando su RCEI es mayor que la RCEI de la siguiente opción más efectiva.
- D. Construir la frontera con las opciones restantes, de tal forma que, para cada umbral, la curva de la frontera de decisión solo nos indica el valor de la alternativa óptima de la curva de aceptabilidad.
 - E. Dibujamos la curva de aceptabilidad, pero sin mostrar las alternativas que hemos identificado como dominadas o que presentan dominancia extendida.
 - F. La opción correcta a elegir como coste efectiva, será aquella que tiene la máxima RCEI siempre que esté por debajo del umbral de decisión.

Estos pasos se realizarían mediante programación; si el modelo se hace en MsExcel© o directamente en otras herramientas de software de modelización

Para poder entender claramente el plano coste-efectividad y la curva de aceptabilidad, se debe tener en cuenta que, en el fondo, la curva de aceptabilidad a pagar nos está dibujando cuántas simulaciones están por debajo de un umbral fijado, de tal forma que para interpretar una CACE con alternativas múltiples de tratamiento, cuando dibujamos la frontera de aceptabilidad, lo que estamos visionando es cuál de las opciones sería la opción seleccionada como más coste efectiva en función de las diferentes disponibilidades a pagar que tenemos en el eje de abscisas.

Dicho de otro modo, la frontera eficiente se define conectando, mediante segmentos, todos los puntos que representan las alternativas

evaluadas situadas más a la derecha y abajo del gráfico (es decir, opciones más razonables desde una perspectiva eficiente: mayor efectividad y menor coste).

Veamos este proceso gráficamente como resultado de un modelo con 3 tratamientos alternativos, en el que se han programado en Excel las gráficas de CACE para la elaboración de la frontera de eficiencia (Figuras 5, 6 y 7).



Figura 5. CACE múltiple para tres tratamientos

Curva de aceptabilidad múltiple: 3 alternativas de tratamiento

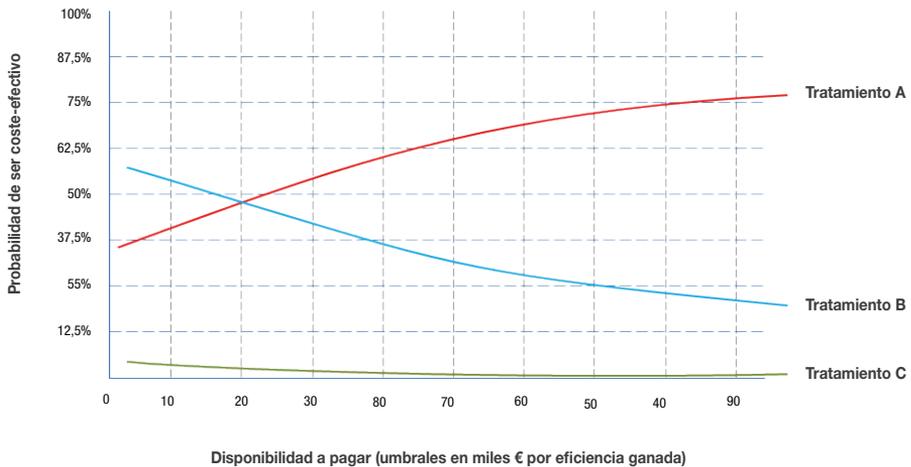


Figura 6. Eliminación de la opción dominada

Curva de aceptabilidad múltiple: Eliminación opción dominada (c)

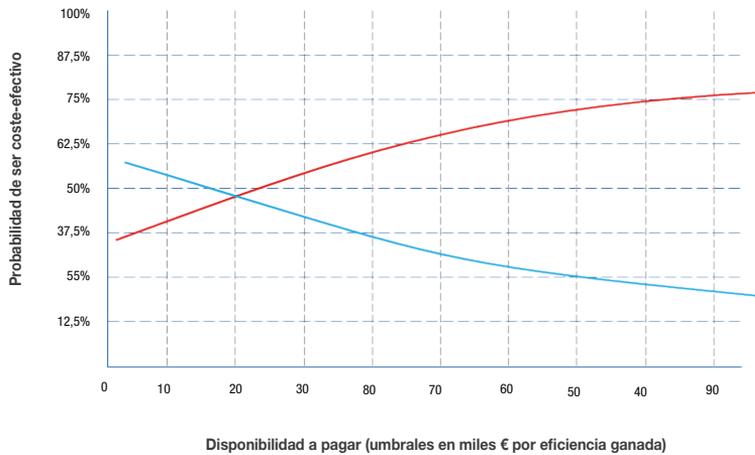
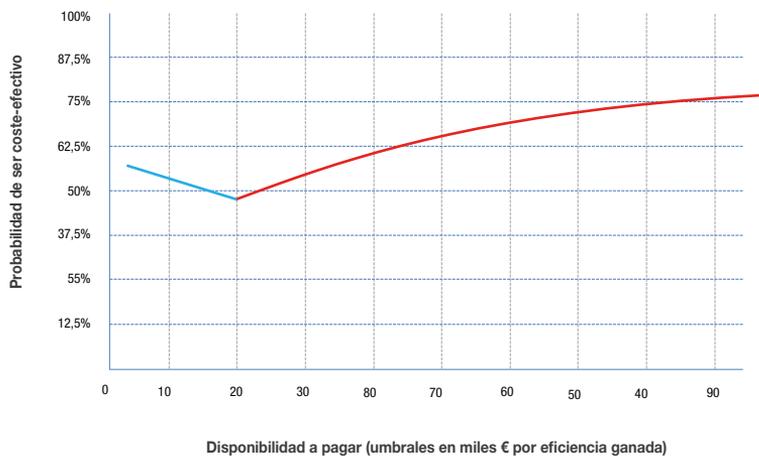


Figura 7. Trazado frontera de eficiencia

Curva de aceptabilidad múltiple: Representación frontera efectividad





El Valor Esperado de la Información Perfecta

La curva de aceptabilidad muestra la probabilidad de cometer un error cuando aceptamos una estrategia (y rechazamos otra). Por tanto, más información reduciría la probabilidad de un error a la hora de asumir un resultado cómo coste efectivo y por lo tanto aumentaría el beneficio a la sociedad, siempre y cuando los costes de obtener dicha información sean menores que los beneficios que obtendríamos por generarla. Ésta es la diferencia entre el máximo beneficio neto alcanzable (con información perfecta) y el beneficio neto asociado a la estrategia favorita.

El valor de la información perfecta es el valor máximo que el decisor querría pagar para eliminar toda la incertidumbre en la decisión. Para poder calcular el valor de la información perfecta, necesitamos tener disponible:

- La estructura del modelo.
- Las distribuciones de los parámetros del modelo, teniendo en cuenta que estamos trabajando con un modelo probabilístico, en el que las variables principales del modelo vienen definidas por distribuciones de probabilidad que conocemos y determinamos, (tal y como se ha explicado en otros capítulos de este libro).
- El valor de la disposición a pagar o umbral.
- El número de personas afectadas por el tratamiento.

El valor esperado de la información perfecta, (VEIP) también da una medida de las oportunidades perdidas por no invertir en la reducción de la desinformación. Si el VEIP es grande, es una señal para que quien toma la decisión, pueda valorar la búsqueda de otra alternativa que no se haya considerado para reducir la incertidumbre, invirtiendo en obtener mediante investigación, más información que disminuye la incertidumbre del modelo.

En el análisis del VEIP; La pérdida de oportunidad se calcula para cada simulación del modelo de manera individual como la diferencia entre el beneficio máximo alcanzable en esa simulación y el beneficio neto de la estrategia favorita (más eficiente). En caso de que el beneficio neto máximo alcanzado en esa simulación coincida con el de la estrategia favorita, no se habrá incurrido en ninguna pérdida de oportunidad.

La pérdida de oportunidad media de todas las simulaciones se calcula posteriormente, y representa el coste esperado de la incertidumbre, que es equivalente al VEIP por paciente. El VEIP total es el VEIP por cada paciente simulado multiplicado por el número de personas que se beneficiarían (prevalencia) de la estrategia (p. ej., tratamiento). El VEIP total equivale al límite máximo del coste de adquisición de más información. Es decir, una vez calculados los valores del VEIP por paciente, se pueden calcular los valores poblacionales multiplicando por la población afectada.

Aunque la decisión final, sería de los políticos sanitarios o planificadores, un VEIP muy bajo en un modelo con probabilidades bajas (por debajo de 50%) en la CACE, podría llevar a pensar que sería socialmente eficiente invertir en más información, (ensayo o estudio), que nos ayudara a reducir dicha incertidumbre, y un VEIP muy alto en modelos en los que los resultados de la CACE respecto a un tratamiento nos proporcionan probabilidades alta, podría no ser valorada la información adicional como una inversión eficiente, aunque habría que tener en cuenta otros elementos para valorarlo adecuadamente, (existencia de alternativas de tratamiento, prioridad social y del propio sistema nacional de salud en cuanto al tratamiento de la patología en cuestión, etc.).

Por lo tanto, para poder calcular el VEIP, se necesita no sólo tener fijado de forma clara un umbral de disponibilidad a pagar, sino delimitada la población total que sería tributaria de ser tratada en la patología evaluada. Además, tendremos que calcular para cada simulación la RCEI y el beneficio neto monetario incremental de las alternativas de tratamiento evaluadas en el modelo.

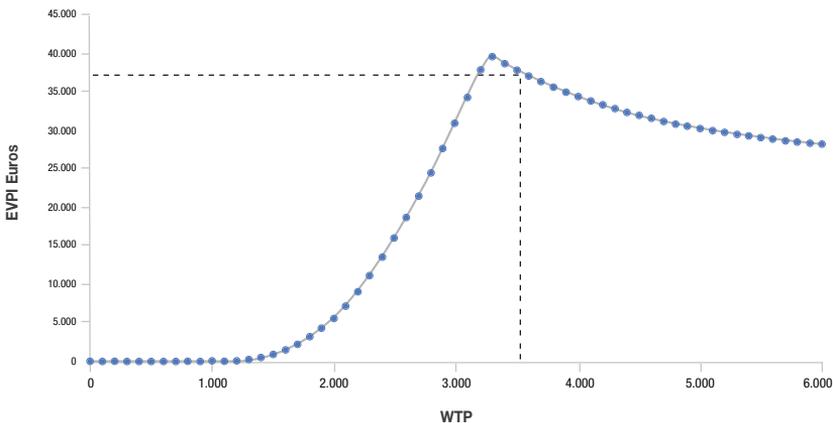
Veamos un ejemplo en el que de las medias de las simulaciones del modelo hemos obtenido un resultado de VEIP de 375 €, para una enfermedad no muy prevalente que podría tener una prevalencia de 1000 pacientes diana y para un umbral de aceptabilidad de 35.000€ por unidad adicional de efectividad adicionalmente ganada. El VEIP, (o EPVI en terminología anglosajona), que deberíamos asegurara para reducir dicha incertidumbre o información imperfecta para esta patología en la población estudiada sería de 375.000€, (resultado de multiplicar los 375€ obtenidos por el número de pacientes diana). La necesidad de dicha inversión debería ser valorada por los planificadores y políticos sanitarios.

La representación gráfica de este ejemplo podemos encontrarla en la Figura 8.



Figura 8. Representación gráfica VEIP

Ejemplo VEIP: Representando diferentes valores (eje vertical), frente a los umbrales de aceptabilidad



Para el umbral 35.000, y un VEIP medio simulación 375 y una población de 1000 pacientes



Bibliografía

1. López Bastida J, Oliva J, Antoñanzas F, et al. Propuesta de guía para la evaluación económica aplicada a las tecnologías sanitarias. *Gac Sanit.* 2010;24:154–70
2. Puig-Junoy J, Oliva-Moreno J, Trapero-Bertrán M, et al. Guía y recomendaciones para la realización y presentación de evaluaciones económicas y análisis de impacto presupuestario de medicamentos en el ámbito del CatSalut. Generalitat de Catalunya. Departament de Salut. Servei Català de la Salut: Barcelona, 2014.
3. Drummond MF, Sculpher MJ, Torrance WG, et al. *Methods for the economic evaluation of health care programmes.* 3rd ed. Oxford: Oxford University Press; 2005.
4. National Institute for Clinical Excellence (NICE). *Guide to the methods of technology appraisals.* April 2013; <http://www.nice.org.uk/process/pmgq> (acceso verificado el 7 Dic 2023)
5. CVZ. *Guidelines for pharmacoeconomic research.* Amstelveen. Health Care Insurance Board, 1999.
6. TVL. *General guidelines for economic evaluations from the Pharmaceutical Benefits Board.* LFNAR 2003; 2; <https://www.tlv.es> (acceso verificado el 7 Dic 2023)
7. Canadian Agency for Drugs and Technologies in Health. *Guidelines for the economic evaluation of health technologies.* 3rd ed. Ottawa, Canada: 2006; <http://www.cadth.ca> (acceso verificado el 7 Dic 2023)
8. Department of Health and Ageing. *Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (PBAC) (version 4.1).* 2007 <http://www.health.gov.au/internet/wcms/publishing.nsf/content/pbacguidelines-index/> (acceso verificado el 7 Dic 2023)
9. Shirowa T, Fukuda T, Ikeda S, et al. Development of an official guideline for the economic evaluation of drugs/medical devices in Japan. *Value Health.* 2017;20:372-8.

10. Huserau D, Drummond M, Petrou S et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS) Statement. *Value Health*. 2013;16:e1-25
11. Tan-Torres T, Baltussen R, Adam T, et al. *Making choices in Health: WHO guide to cost-effectiveness analysis*. WHO Publications. Switzerland 2003.
12. Rubio-Terrés C, Cobo E, Sacristán JA et al. Análisis de la incertidumbre en las evaluaciones económicas de intervenciones sanitarias. *Med Clin*. 2004;122:668-74
13. Petrou S and Gray A. Economic evaluation using decision analytical modelling, design conduct, analyzing and reporting. *BMJ*. 2011; 342: d1766
14. Luengo-Fernández R, Hernández-Quesada C, Rivero-Arias O. Evaluación económica de tecnologías sanitarias: una introducción para médicos. *Rev Neurol*. 2011;53:107-15
15. Klok RM, Postma MJ. Four quadrants of the cost-effectiveness plane: some considerations on the southwest quadrant. *Expert Rev Pharmacoecon Outcomes Res*. 2004; 4:599-601.
16. Sacristán J, Oliva J, Del Llano J et al. ¿Qué es una tecnología sanitaria eficiente en España? *Gac. Sanit*. 2002;16:334-43.
17. Rodríguez-Barrios JM, Pérez F, Crespo C et al. The use of cost per year gained as cost effectiveness result in Spain. *Eur J Health Econ*. 2012;13: 723-40.
18. Rodríguez-Barrios JM, Paz S, Lizán L, et al. The use of Quality Adjusted Life Years (QALYs) in the economic evaluation of health technologies in Spain: a review of the 1990-2009 literature. *Value Health*. 2011;14:458-64.
19. Vallejo-Torres L, García-Lorenzo B, Serrano-Aguilar P. Estimating a cost effectiveness threshold for the Spanish NHS. *Health Econ*. 2018; 27: 746-61.
20. Sacristán JA, Oliva J, Campillo-Artero C et al. ¿Qué es una intervención sanitaria eficiente en España en 2020? *Gac Sanit*. 2020;34:189-93.
21. Fenwick E, Claxton K, Schulpher M. Representing uncertainty: the role of cost-effectiveness acceptability curves. *Health Econ*. 2001; 10:779-87.
22. Zethareus N, Johannesson M, Jönsson B, et al. Advantages of using the Net-Benefit approach for analyzing uncertainty in economic evaluation studies. *PharmacoEconomics*. 2003;21:39-48

23. Fenwick E, O'Brien BJ, Briggs A. Cost-effectiveness acceptability curves facts, fallacies and frequently asked questions" in Health Econ. 2004;13:405-15.
24. Barton GR, Briggs AH, Fenwick EA. Optimal cost-effectiveness decisions: the role of the cost-effectiveness acceptability curve (CEAC), the cost-effectiveness acceptability frontier (CEAF), and the expected value of perfection information (EVPI). Value Health. 2008;11:886-97.



7

EVALUACIÓN DE LA CALIDAD METODOLÓGICA DE UN MODELO ANALÍTICO DE DECISIÓN A TRAVÉS DE UNA LISTA-GUÍA

Javier Soto Álvarez

Médico especialista en Farmacología Clínica

*Consejo de Dirección del Máster de Formación Permanente en Evaluación Sanitaria
y Acceso al Mercado (Fármaco-Economía) de la Universidad Carlos III de Madrid*

En la actualidad, en nuestro país cada vez es más patente la necesidad de incrementar los recursos que se van a destinar al cuidado sanitario de la población, dado que la pirámide poblacional se está derivando a una mayor densidad de población anciana, al haberse incrementado en los últimos lustros la esperanza de vida en nuestro país. Este hecho se acompaña de una demanda creciente y una mayor carga de enfermedad en este grupo de población, lo que se deriva en una mayor necesidad de atención sanitaria por parte del Sistema Nacional de Salud (SNS) de nuestro país.

Dado que los recursos monetarios disponibles en el SNS son limitados y finitos, es crucial destinarlos a financiar con cargo al presupuesto público aquellas opciones terapéuticas que generen el máximo de resultados en salud en los pacientes por cada euro invertido, esto es, que sean lo más eficientes posibles, lo que va a permitir minimizar al máximo el coste de oportunidad de la inversión realizada en el SNS.

En este contexto, las evaluaciones económicas de medicamentos y tecnologías sanitarias (los estudios de eficiencia) cobran un máximo interés, ya que nos van a permitir conocer que alternativas terapéuticas de todas las disponibles van a ser las más coste-efectivas. De esta manera, se va a poder seleccionar las opciones terapéuticas que deberían financiarse y cuáles deberían emplearse de manera rutinaria dentro del SNS, al ser las más eficientes.

A la hora de efectuar una evaluación económica, la metodología más empleada en la actualidad para su realización son los modelos analíticos de decisión (MAD). Los MAD se han utilizado ampliamente para sintetizar la evidencia clínica y económica del uso de distintas opciones terapéuticas en el manejo de una enfermedad, con el fin de orientar a los decisores y profesionales sanitarios a la hora de asignar los recursos disponibles de la manera más eficiente posible (1).

Por la metodología empleada y sus características, los MAD nos van a permitir (2,3):

- 1) extrapolar los datos hallados en un ensayo clínico (normalmente con cortos períodos de seguimiento) a situaciones de uso más crónico, como es el manejo de la mayoría de las enfermedades en la práctica médica rutinaria
- 2) relacionar variables intermedias de eficacia con variables finales de efectividad clínica
- 3) generalizar datos encontrados en un país y un entorno sanitario determinado a otro medio diferente, adaptando los resultados clínicos, económicos y epidemiológicos empleados en el país de referencia
- 4) comparar diferentes opciones terapéuticas entre sí cuando no se hayan realizado estudios clínicos comparativos entre ellas
- 5) trasladar los resultados de los ensayos clínicos (eficacia y seguridad) a los resultados que se obtendrán cuando la nueva opción terapéutica se empiece a emplear en condiciones de uso habitual (efectividad)
- 6) combinar datos de eficacia, seguridad y consumo de recursos provenientes de ensayos clínicos, diseños observacionales (estudios de vida real), revisiones sistemáticas/metaanálisis, comparaciones indirectas, metaanálisis en red, lo que va a permitir incorporar todas las evidencias existentes
- 7) poder obtener rápidamente datos de coste-efectividad cuando se incorporen nuevos tratamientos para tratar la enfermedad diana, o actualizar éstos, cuando aparezcan nuevas opciones terapéuticas para su manejo
- 8) permitir incorporar la calidad de vida y las preferencias de los pacientes en los resultados finales de la evaluación económica

Dado que siempre va a existir un grado de incertidumbre en su desarrollo, es conveniente que en su diseño y elaboración colabore un equipo multidisciplinar, formado por clínicos, farmacéuticos, farmacólogos, estadísticos, economistas de la salud, epidemiólogos, etc., lo que incrementará la calidad y validez de sus resultados, así como su transparencia.

Los MAD pueden ser dinámicos versus estáticos (dependiendo si interviene el tiempo o no), continuos versus discretos (si cambia el estado del sistema continuamente o no) y estocásticos versus determinísticos (si los valores de entrada son aleatorios o no).

De forma adicional y, dependiendo de las características del modelo, existen dos grandes tipos de modelos: a) modelos simples basados en árboles de decisión y b) modelos de transición entre estados, que engloban los modelos de Markov, los modelos de simulación individual de pacientes, los modelos de simulación de eventos discretos y los modelos dinámicos (4,5). Por otra parte, en los últimos años cada vez se emplean más los modelos de supervivencia particional o dividida (partitioned survival models), sobre todo en oncología (6). Además, se puede aplicar la estadística bayesiana a los modelos, lo que podría considerarse como otro tipo de modelo adicional (7).

La elección de un tipo u otro de modelo a la hora de efectuar una evaluación económica va a depender, básicamente, de la enfermedad que se esté evaluando (aguda, sub-aguda o crónica), de las alternativas terapéuticas en estudio, del tipo de resultados demandados por los agentes decisores (corto o medio-largo plazo), de la información de calidad disponible, así como de los conocimientos y experiencia del analista que va a diseñar el modelo.



Características que debería cumplir cualquier modelo analítico de decisión

Para que los resultados de un modelo sean creíbles, válidos y relevantes para los profesionales sanitarios y otros agentes decisores, será necesario que cumplan con unos requisitos mínimos (8):

- su estructura y horizonte temporal debería ser consistente con el objetivo final del modelo
- ser lo más simple posible (mientras represente de manera fidedigna la evolución más probable de la enfermedad y sea válido), lo que ayudará a que sea transparente, fácil de transmitir los resultados y entendible por todos aquéllos que quieran emplear sus resultados
- ser explícitos, definir la perspectiva de este y reflejar las condiciones de uso habitual
- describir la población diana (y sus características) que será evaluada en el modelo
- incluir las opciones terapéuticas más empleadas para el manejo de la enfermedad evaluada en el entorno sanitario donde se efectúa el análisis
- ser transparente y preciso en la medición de los resultados en salud y en la cuantificación de los costes derivados del uso de las alternativas en estudio
- ser lo más realista posible en cuanto a describir el patrón de tratamiento más usual para la enfermedad diana en el medio donde se realiza el estudio
- poder ser reproducible por otro investigador, por lo que será necesario describir en detalle las características y la estructura del modelo
- estar perfectamente documentada la fuente de los datos incluidos en el modelo y estar plenamente justificadas las asunciones incorporadas

- describir el tipo de evaluación económica elegido y argumentar la elección
- analizar los resultados apropiadamente, notificando el cociente coste/efectividad incremental derivado de comparar las diferentes intervenciones entre sí, así como la representación gráfica en el plano coste-efectividad del análisis de sensibilidad probabilístico
- evaluar la incertidumbre del modelo efectuando diferentes análisis de sensibilidad (determinístico, probabilístico y de otro tipo, si se considera necesario)
- validar y auditar el modelo, al ser comparado con otros modelos y/o estudios clínicos, ya realizados o en marcha



Estandarización en el diseño y elaboración de un modelo analítico de decisión: pasos a dar en su elaboración, realización y análisis de los resultados

A la hora de diseñar y realizar un MAD, va a ser muy importante seguir una metodología correcta y tener en cuenta todos aspectos relevantes relacionados con su elaboración, de tal forma que los resultados sean válidos, creíbles y ayuden en la toma de decisión en política sanitaria (9).

En la Tabla 1 se detallan las recomendaciones que deberían seguirse y las preguntas que habría que contestar cuando se vaya a diseñar un MAD de novo, o cuando se quiera valorar uno ya disponible, con el fin de poder evaluar su calidad metodológica, validez y transparencia. Por lo tanto, a continuación, se van a describir los pasos que habría que seguir cuando se vaya a diseñar un MAD de nuevo, y las cuestiones que habría que contestar para evaluar la calidad de un MAD ya existente.

1) Definir el objetivo del modelo y razonar la necesidad de su elaboración

Cuando se diseña un MAD es necesario explicar porque se va a realizar y que pregunta se pretende contestar con su elaboración, así como la relevancia que van a tener sus resultados para la ayuda en la toma de decisiones en política sanitaria. También va a ser necesario especificar porque se va a realizar el MAD para efectuar la evaluación económica y no se utilizan otras metodologías disponibles para su realización, tales como incorporar una evaluación económica en un ensayo clínico o en un estudio observacional (estudio de cohortes, registro de pacientes, análisis de historias clínicas electrónicas, etc.) (10).

La realización de un MAD es recomendable en las siguientes situaciones: a) solamente se dispone de resultados intermedios y no se pueden calcular variables finales de eficacia (por ej. Supervivencia global); b) las opciones evaluadas en los ensayos clínicos (en especial los comparadores) no son los tratamientos más relevantes empleados en el medio donde se realiza la evaluación económica; c) es necesario incorporar en la evaluación económica a efectuar los datos de eficacia y seguridad a más largo plazo que la duración del ensayo clínico, acorde a la historia natural de la enfermedad evaluada.

2) Mostrar en detalle el diseño y la estructura del modelo

La elección del tipo de modelo más adecuado a desarrollar va a depender de la enfermedad a valorar y de la disponibilidad de datos para alimentar el modelo (eficacia, seguridad, efectividad, consumo de recursos, valor de utilidad, etc.).

La estructura del MAD deberá ser lo más simple posible, con tal que refleje bien la evolución más plausible de la enfermedad diana e incluya las opciones terapéuticas que más se van a emplear en el tratamiento de los

pacientes, de una manera fácilmente comprensible (11). Además, deberá ser flexible y tendrá que proporcionar suficiente información del modelo, de tal manera que pueda ser replicado por otro investigador interesado.

Por otra parte, es muy recomendable validar la estructura del modelo con expertos en la enfermedad diana (clínicos, farmacéuticos, farmacólogos, etc.), así como con otros expertos en economía de la salud, epidemiología, estadística, etc., lo que va a incrementar su credibilidad y transparencia a la hora de emplear sus resultados para tomar decisiones.

Si el MAD que se va a diseñar es un modelo de transición entre estados de salud (por ej, un modelo de Markov), será necesario especificar y justificar que los estados de salud que tendrá el modelo, que deberán reflejar la evolución más probable de la enfermedad evaluada (12).

3) Describir con detalle las alternativas terapéuticas a comparar en el modelo

La elección de las opciones terapéuticas a incluir en el MAD con las que se va a comparar el nuevo medicamento (o tecnología sanitaria) es un punto crucial a la hora de obtener el cociente coste-efectividad incremental del nuevo medicamento frente al resto de tratamientos ya disponibles; siempre va a ser necesario razonar porque se han elegido unos comparadores y no se han elegido otros posibles, a la hora de compararse con el nuevo medicamento en la evaluación económica.

Como norma general, en un MAD deben incluirse como comparadores las alternativas terapéuticas más empleadas en la práctica médica rutinaria o aquellas a las que va a sustituir el nuevo medicamento, aunque también se pueden incluir las opciones más baratas o las más eficientes, en muchas ocasiones dentro del análisis de sensibilidad. En ocasiones y, dependiendo de la enfermedad diana del MAD, habría que considerar también la posibilidad de incluir como comparador la opción de no hacer nada y vigilar la evolución de la enfermedad (13).

A la hora de decidir las alternativas a incluir en el MAD como comparadores, siempre va a ser recomendable discutir con profesionales sanitarios y otros decisores este tema, incluyendo las opciones de 2ª línea y líneas posteriores de tratamiento cuando la enfermedad evoluciona y los pacientes necesitan tratamientos adicionales.

Cada vez es más normal que en diferentes patologías (en especial en onco-hematología) los comparadores sean secuencias de tratamiento, incluyendo 1ª, 2ª y líneas posteriores, en vez de tratamientos distintos en cada una de las fases evolutivas de la enfermedad puestos por el clínico lo que supone un hándicap y genera algunas tensiones metodológicas, en especial la disponibilidad de datos de efectividad y seguridad válidos y creíbles de la secuencia de tratamientos (14), aunque en el futuro cada vez será más frecuente esta situación.

4) Especificar la población diana a evaluar en el modelo

Es muy recomendable que se describa muy bien la población incluida en el MAD, incluyendo todas sus características, tales como sexo, edad, co-morbilidades asociadas y factores de riesgo asociados, estado evolutivo de la enfermedad y otras características clínicas de interés. Habrá que asegurarse que los datos que van a alimentar el modelo provienen de estudios en los que las características de la población incluida coincidan con la población diana del MAD, ya que si no los resultados no serán del todo válidos (15).

Por otra parte, será necesario evaluar el grado de heterogeneidad de la población incluida en el MAD ya que, por diferentes motivos habitualmente existen diferencias importantes entre las características de la población sobre la que se evalúa la eficacia o efectividad y seguridad de una nueva alternativa terapéutica y los costes en su implementación. Por este motivo, se recomienda efectuar análisis de resultados en diferentes subgrupos de pacientes homogéneos, siempre

que existan datos válidos. El problema es que para poder realizar este análisis en subgrupos es necesario disponer de datos específicos de dichos subgrupos en los estudios en los que se va a basar el MAD y que estos datos se hayan obtenido siguiendo una metodología correcta y válida (16,17).

5) Exponer y razonar el horizonte temporal elegido en el modelo

Cuando se elabora un MAD se recomienda elegir un horizonte temporal que permita capturar todos los recursos consumidos y los resultados en salud obtenidos con las alternativas en evaluación, por lo que el horizonte temporal tendrá que ser suficientemente extenso para poder evaluar todas las diferencias importantes en consumo de recursos y resultados en salud obtenidos entre el nuevo medicamento (tecnología sanitaria) y los comparadores incluidos (18).

En general, la historia natural de la enfermedad evaluada y la evolución más probable de la misma puede ayudar a elegir el horizonte temporal más adecuado; teniendo en cuenta que la mayoría de las enfermedades son crónicas (a veces con episodios sin síntomas y con recaídas) lo más normal es que los horizontes temporales de las evaluaciones económicas sean largos (normalmente años), salvo las patologías agudas puras, donde pueden adoptarse horizontes más cortos.

En muchas ocasiones se van a adoptar horizontes temporales de toda la vida del paciente en los análisis de resultados del caso base de la evaluación económica, con el fin de estar seguros de que se recogen todo el consumo de recursos y todos los resultados en salud generados mientras los pacientes están vivos y, en este caso, es habitual incluir en el análisis de sensibilidad horizontes temporales menores (por ej., 5, 10 ó 15 años) para poder valorar como se modifican los resultados de la evaluación económica cuando se adoptan diferentes horizontes temporales.

6) Describir en detalle la perspectiva elegida en el modelo

La perspectiva se refiere al punto de vista desde el que se lleva a cabo el MAD y a quien sus resultados les va a servir como ayuda en la toma de decisión. Dependiendo de la perspectiva elegida, habrá que incluir un tipo de coste y no otros, por lo que, según la perspectiva escogida, los resultados podrán variar de manera importante.

La perspectiva más empleada en las evaluaciones económicas y en los MAD son las siguientes: a) el SNS (financiador), que incluye únicamente los costes sanitarios, pero no los no sanitarios, ni los indirectos; b) la asistencia hospitalaria que incluye solo los costes sanitarios que se producen en el hospital (tampoco incluyen costes no sanitarios e indirectos), y c) la sociedad que obliga a incluir todo tipo de costes (sanitarios, no sanitarios e indirectos), siendo la perspectiva más completa y la que la mayoría de guías existentes y las sociedades científicas recomiendan como de elección, ya que se recogen todos los costes relevantes para el paciente y la sociedad (19,20).

Es recomendable incluir en el análisis de sensibilidad otras perspectivas diferentes a la elegida en el caso base del análisis de los resultados, con el fin de valorar como afecta a los resultados del MAD.

7) Identificar, recoger y cuantificar el consumo de recursos y costes asociados

Uno de los principales pasos a la hora de efectuar una evaluación económica a través de un MAD es la determinación de los recursos empleados por las alternativas terapéuticas en evaluación, la recogida del uso de estos recursos y el cálculo del uso de estos recursos en costes asociados para el SNS.

La perspectiva elegida en el MAD va a determinar que tipo de recursos habrá que incluir y cuales no será necesario en la evaluación económica en marcha, tal y como se ha comentado en el apartado anterior.

Es muy recomendable recopilar por separado las unidades de recursos empleados (cantidades) de sus precios o costes unitarios antes de completar su transformación monetaria, en aras de incrementar la transparencia del análisis.

A la hora de asignar un coste a cada recurso consumido, lo ideal es emplear el coste de oportunidad de los recursos empleados, si se dispone de una fuente fiable para obtener este dato. Si no fuera el caso, los costes unitarios a emplear podrían obtenerse de la contabilidad analítica de los centros (por ej., a través del uso de GRDs), de precios de mercado, de publicaciones oficiales o de las tarifas aplicadas a los contratos de prestaciones de servicios del SNS a entidades privadas, siempre justificando de donde se han obtenido los datos.

Siempre que se incluyan costes de cuidados informales (principalmente, coste de los cuidadores), cuando sean relevantes para la evaluación efectuada acorde a la perspectiva escogida, se deberá razonar el método empleado para su cálculo (21). De la misma manera, cuando se incorporen en la evaluación económica costes indirectos (pérdida de productividad laboral medida a través de absentismo o presentismo, o costes asociados a la mortalidad precoz) se recomienda justificar la elección del método escogido para su medición (22).

Es necesario especificar claramente la moneda empleada para cuantificar los costes (euros, dólares, libras, etc.), el año de los costes empleados en el modelo y notificar si se ha efectuado algún ajuste a los costes debido a la inflación, y describir el método empleado. (23). Aunque existe cierta controversia entre los investigadores que elaboran evaluaciones económicas, cada vez más se recomienda incluir los costes futuros en las evaluaciones económicas, en especial aquellos que están relacionados con la enfermedad diana en estudio, y en el caso de que no se contemplen en el análisis será necesario especificar la razón de no hacerlo (24).

En el caso de que en el MAD se estén comparando medicamentos, será necesario especificar como se ha calculado el coste de las diferentes

opciones evaluadas (según dosis diaria, características de los pacientes, duración del tratamiento), siendo preferible emplear la media en vez de la mediana de tratamiento, si se dispone de este dato. Además, es muy deseable que se incluyan dentro de los costes sanitarios el coste de las líneas posteriores de tratamiento cuando falla el tratamiento inicial, el coste de los cuidados paliativos cuando la enfermedad está muy evolucionada (especialmente en onco-hematología) y el coste del periodo peri-muerte, así como el manejo de las reacciones adversas, costes que en ocasiones pueden ser de una elevada cantidad (25).

Es recomendable que los datos de los costes sanitarios, costes no sanitarios (cuidados informales y tiempo perdido) y los costes indirectos (perdidas laborales y mortalidad precoz) se comuniquen de manera desagregada, ya que este hecho facilita su análisis y evaluación.

8) Incorporar en el modelo los resultados en salud más relevantes de las opciones en estudio y las probabilidades de que sucedan

Los resultados en salud a incorporar al MAD se pueden obtener de ensayos clínicos aleatorizados o de diseños observacionales (estudios de cohortes prospectivos y retrospectivos, registros de pacientes, análisis de la historia clínica electrónica, etc.), siempre de estudios con una elevada calidad metodológica (26), por lo que será necesario comprobar su validez y fiabilidad a través de listas-guías empleadas para su evaluación (27,28).

Habitualmente, los MAD que se elaboran para la negociación de precio y financiación se suelen nutrir de datos provenientes de ensayos clínicos, datos que se han empleado para el registro del medicamento y su autorización de comercialización, que reflejan la eficacia del nuevo medicamento frente al estándar de tratamiento ya disponible. Una vez el nuevo medicamento se ha comercializado, es deseable complementar el MAD con datos de efectividad que reflejen los resultados en salud

generados en condiciones de la práctica médica diaria del mundo real, bien con datos de ensayos clínicos pragmáticos (29) o con datos de los estudios observacionales de vida real (30,31).

En el caso de que no se hayan realizado ensayos clínicos que hayan comparado directamente las opciones a evaluar en el MAD entre sí, pero los comparadores empleados en los ensayos clínicos de cada una de esas alternativas haya sido el mismo, se pueden emplear las comparaciones indirectas (*indirect comparisons*) para poder obtener datos comparativos entre estas, opción metodológica cada vez más empleada y cuyos resultados son válidos y fiables, si se ha realizado con rigor científico y siguiendo la metodología adecuada (32). Si, por el contrario, los comparadores de los ensayos clínicos han sido diferentes, también se pueden obtener datos comparativos entre las opciones en evaluación en el MAD a través de las comparaciones mixtas de tratamientos (*mixed-treatment comparisons*), los metaanálisis en red (*network meta-analysis*) y las comparaciones indirectas ajustadas por emparejamiento (*matching-adjusted indirect comparisons*), todas ellas técnicas válidas para obtener datos comparativos entre las opciones en estudio, si la metodología empleada en su diseño es la correcta (33-36).

En la mayoría de los MAD va a ser necesario incorporar información y evidencias que no están disponibles en las publicaciones existentes (por ej. manejo de efectos adversos, tratamiento de 2º o líneas posteriores a emplear, consumo de recursos en el manejo de una determinada enfermedad, etc.), por lo que va a ser necesario preguntárselo a un panel de expertos compuesto por clínicos expertos en el manejo de la enfermedad, farmacéuticos, farmacólogos clínicos, etc., dado que van a ser datos muy relevantes para la obtención de resultados válidos en el MAD. Por este motivo, hay que estar seguro de que a la hora de emplear un panel de expertos se siga una metodología válida y creíble, por lo que es recomendable seguir las guías existentes para su realización (37).

De la misma manera, en muchos MAD es necesario incluir asunciones por parte del investigador que lo está elaborando, con el fin de poder

incorporar datos necesarios para su diseño pero que no se pueden encontrar en la literatura. Teniendo en cuenta que las asunciones se pueden ver como datos de baja transparencia y de fácil manipulación y, dado que pueden ser importantes para obtener resultados válidos, es totalmente mandatorio que las asunciones incorporadas sean coherentes, conservadoras y reflejen lo que sucede en el día a día en el entorno donde se va a desarrollar el MAD. Además, en el informe o publicación de este, se debería presentar un listado con todas las asunciones incluidas debidamente razonada su incorporación.

De la misma manera, es esencial incluir en la discusión de los resultados del MAD una parte destinada a comentar las posibles limitaciones del modelo debido a las asunciones, así como que un grupo de expertos externos valide las asunciones incluidas, en aras de incrementar su credibilidad y transparencia (38).

Si la evaluación económica que se va a realizar es un análisis coste-utilidad, la variable de medida va a ser el año de vida ajustado por calidad (AVAC) debido a su validez, simplicidad y facilidad de uso. A la hora de su cálculo es necesario incorporar el valor de utilidad de los pacientes (preferencia por su estado de salud) que va a multiplicar a los años de vida ganados con las alternativas terapéuticas en evaluación, generando los AVACs ganados con cada una de ellas (39).

Con el fin de que los AVACs ganados y utilizados en el MAD sean creíbles y relevantes para los profesionales sanitarios y otros decisores, es necesario que se especifique el cuestionario empleado para la obtención del valor de utilidad del paciente (EQ-5D-3L, EQ-5D-5L, HUI-3, SF-6D, etc.), el valor obtenido para cada estado de salud del MAD y especificar las fuentes de donde se han obtenido los datos, en aras de su transparencia y validez. A la hora de conseguir los datos de utilidad a incorporar en el MAD para calcular los AVACs ganados con cada opción en estudio, lo más correcto desde el punto de vista metodológico es incluir los datos de utilidad obtenidos en los ensayos clínicos realizados con las alternativas evaluadas y, complementar estos datos, con la realización de

una revisión sistemática de la literatura de estudios donde se haya valorado y medido el valor de utilidad en la enfermedad diana del MAD siguiendo las recomendaciones y directrices existentes para efectuar este tipo de revisiones sistemáticas (40-42). Siempre será necesario valorar de manera cuidadosa que los estudios incluidos presentan una elevada calidad metodológica y que son aplicables y relevantes para el objetivo de la revisión sistemática (43), por lo que es deseable emplear herramientas publicadas que van a ayudar en este cometido (44).

En los últimos años, cada vez se ha prestado más atención a la manera de calcular los valores de utilidad cuando se realizan extrapolaciones en el contexto de un MAD (habitualmente se mantiene el mismo valor de utilidad en los estados definidos a lo largo de todo el horizonte temporal del modelo), ya que se sabe que el valor de utilidad es diferente según la duración de la enfermedad en los pacientes y la evolución de la misma en el tiempo (45), aspecto de gran importancia de cara a obtener valores reales y fiables de utilidad, que nos permita generar cálculos de los AVACs válidos y precisos (46).

Cuando se piense que el valor de utilidad medido no refleja de manera fidedigna la disminución de este valor que pueda producirse por la existencia adicional de co-morbilidades asociadas, o por la aparición de efectos adversos graves, se pueden incorporar disutilidades al valor basal obtenido de utilidad, lo que va a permitir disponer de un valor de utilidad más realista y fiable. Dado que la incorporación de disutilidades puede modificar de manera importante el valor final de utilidad (y, por extensión, del AVAC), siempre que se empleen debe estar plenamente justificado su uso para evitar un doble contaje y una disminución irreal del valor de utilidad, y el valor de las disutilidades empleado tendrá que ser comunicado apropiadamente, lo mismo que la fuente de donde se ha obtenido; por este motivo, lo más recomendable es que el valor de las disutilidades sean obtenidos directamente de una revisión sistemática de estudios publicados donde se haya evaluado el posible valor de disutilidad asociados a diferentes efectos adversos y otros posibles factores (47).

En los últimos años se están realizando estudios para adaptar los instrumentos existentes que miden utilidad, de tal manera que puedan capturar y reflejar de manera más fidedigna como afecta al valor final de utilidad la existencia de efectos adversos y la propia sintomatología de la enfermedad, lo que supondría que no fuera necesario el uso de disutilidades para corregir el valor de utilidad obtenido (48).

En los modelos de transición entre estados (modelos de Markov y otros) los estados de salud definidos deben representar bien la evolución más plausible de la enfermedad diana, de tal manera que el MAD represente bien los estados por los que un paciente va a transitar en el proceso evolutivo de su patología y refleje fielmente su evolución más probable (49). A la hora de calcular las probabilidades de transición entre los diferentes estados de salud a lo largo del horizonte temporal del modelo, su valor se va a obtener primordialmente de las tasas de eventos clínicos que aparecen en los estudios clínicos realizados con el nuevo medicamento y el comparador incluido (pudiéndose complementar con datos de otros estudios publicados), siendo necesario seguir una metodología adecuada para que el cálculo de las probabilidades de transición sea válido y preciso (50,51).

En el momento de calcular las probabilidades de transición entre los diferentes estados de salud se pueden producir dos situaciones diferentes, que se disponga de los datos individuales de los pacientes que han participado en el ensayo, y que solo se disponga del resultado del ensayo clínico a través de la curva de Kaplan-Meier (K-M) publicada en un artículo.

Si se dispone de los datos individuales de los pacientes que han participado en los ensayos clínicos del nuevo medicamento o en otro tipo de diseño prospectivo (registro de pacientes, estudio observacional de cohortes prospectivo, etc.), y en este caso las probabilidades de transición se pueden obtener a través de modelos paramétricos y no paramétricos de supervivencia que va a permitir calcular los datos de supervivencia global (SG) a partir de los datos obtenidos en los ensayos

clínicos. En el caso de las enfermedades onco-hematológicas, a partir de los datos de supervivencia libre de progresión (SLP) de los ensayos clínicos se van a poder extrapolar y calcular los datos de SG cuando los estudios se hayan parado antes de haberse obtenido datos de SG. De esta manera, se va a poder calcular las probabilidades de transición desde el estado libre de progresión a los estados de progresión y de mortalidad (52,53), pudiéndose incorporar evidencia externa a estos datos, en aras de conseguir una extrapolación más fiable y precisa (54).

Por el contrario, si solo se dispone de las curvas de K-M de los estudios publicados en la literatura donde se detallan las curvas y el sumatorio de datos estudiados (pacientes a riesgo y el número total de eventos acaecidos), va a ser necesario estimar y parametrizar las probabilidades de transición de los estados de salud del modelo basándose en las curvas de SLP publicada de los ensayos clínicos siguiendo la metodología más adecuada disponible (55).

En esta situación, en la que solo se dispone de la curva de K-M publicada, lo primero que se hace es extraer los puntos de los datos individuales de los pacientes (tiempo de seguimiento y tasa de supervivencia) a través de esta curva empleando un software que permita realizar el cálculo (el más empleado es GetData Graph Digitizer 2.26) (56). A continuación, se reconstruyen los datos individuales del tiempo hasta el evento de los pacientes a través de diferentes métodos existentes para realizarlo, siendo el más empleado el método estadístico desarrollado por Guyott (57-59), lo que va a permitir la extrapolación paramétrica y la estimación de supervivencia a largo plazo. Al final, se va a evaluar qué modelo de supervivencia es el más apropiado y el que mejor ajusta los datos utilizando el criterio de información de Akaike, el criterio de información Bayesiano, la inspección visual y la validez clínica (60,61), eligiéndose aquel modelo que mejor se ajusta (el valor más bajo del criterio de información de Akaike y el criterio de información Bayesiana) (62). El ajuste y extrapolación de datos de supervivencia se

implementa con el paquete “survHE” y “flexsurv” del programa R (versión 4.2.1) (63).

Dado que los modelos de transición entre estados simulan el tiempo en periodos discretos, conocidos como ciclos, y asume que las transiciones de los pacientes de un estado a otro ocurren a intervalos fijos (al inicio o al final de cada ciclo), se pueden producir errores cuando se calculan los resultados acumulados, con una sobre o infraestimación de los AVACs y los costes calculados, si se compara con los que se hubieran obtenido si las transiciones hubieran ocurrido aleatoriamente a lo largo de la duración del ciclo. La aproximación más empleada para corregir estos errores en los resultados totales es añadir y sustraer la mitad de los resultados (coste y efectividad finales) en cada estado de salud en el primer y último ciclo del MAD, lo que se conoce como la corrección del medio ciclo (64).

En los últimos años cada vez se emplean más en onco-hematología los modelos de supervivencia particional (con particiones) (*partitioned survival models*) que va a constar normalmente de 3 estados de salud: libre de progresión, progresión del tumor y muerte. Este tipo de modelos también se denominan modelos bajo la curva, ya que son más intuitivos y fáciles de diseñar y, además, son más sencillos de comunicar y de entender por parte de los profesionales y decisores sanitarios, ya que se basan en las curvas de SLP y de SG obtenidas en los ensayos clínicos. En estos modelos se calcula directamente el número de pacientes en cada estado de salud en cada ciclo, empleando las curvas de SLP y SG de las alternativas en evaluación en los ensayos clínicos, pudiéndose extrapolar estas curvas más allá de la duración del ensayo según el horizonte temporal fijado en el MAD (65).

La proporción de pacientes que va a estar en cada estado de salud (libre de progresión, en progresión y muerte) se va a obtener directamente de la probabilidad de supervivencia acumulada de las curvas de supervivencia, y se calcula basándose en las áreas bajo la curva de las funciones de supervivencia (SLP y SG). La primera curva describe el tiempo desde que el paciente entra en el modelo (libre de progresión)

hasta que progresa o muere, y la segunda curva describe el tiempo que pasa desde que el paciente entra en el modelo y muere. Esta primera curva, la SLP proporciona el porcentaje de pacientes que se quedan en el primer estado de salud (libre de progresión) a lo largo del tiempo. Los pacientes que van a acabar en el estado de salud de muerte son 1-curva de SG en cada momento. Para el estado de salud de los pacientes que progresan, el número de pacientes viene dado por la diferencia entre las curvas de SG y SLP en cada momento, que nos va a proporcionar lo pacientes que están vivos, pero en progresión. En este tipo de modelo de 3 estados de salud, es necesario que las curvas de SG y SLP sean independientes una de la otra y disponer de los datos individuales de los pacientes de los ensayos clínicos (66).

En ocasiones, va a ser necesario recurrir a otros tipos de MAD para efectuar la evaluación económica, como son los modelos de simulación de eventos discretos y los modelos dinámicos, y siempre habrá que razonar porque se realizan y que ventajas van a tener frente a otro tipo de modelos, así como aportar suficiente información sobre la metodología empleada y los datos incluidos en el modelo, de tal manera que se puede evaluar su calidad metodológica por parte de los profesionales sanitarios y otros decisores del SNS (67,68).

En el mundo de la oncología y, tras la llegada de la inmunoterapia, los MAD más empleados (modelos de transición entre estados y modelos de supervivencia parcial) no van a ser los más precisos a la hora de estimar la SG durante toda la vida de los pacientes, ya que no van a ser capaces de realizar una extrapolación adecuada de los datos de supervivencia de los ensayos clínicos más allá del periodo de seguimiento de los pacientes; este hecho se deriva del retraso del comienzo de la eficacia y el beneficio clínico de la inmunoterapia en muchos pacientes y el potencial hecho de que un porcentaje de los pacientes se conviertan en largos respondedores con supervivencia muy largas, lo que cuestiona el uso de los métodos clásicos para realizar una extrapolación simple de la supervivencia (69).

Por este motivo, es necesario emplear modelos especiales que utilizan técnicas más sofisticadas para efectuar la extrapolación de la supervivencia de los pacientes a partir de los datos obtenidos en los ensayos clínicos a horizontes temporales más extensos, tales como modelos de curación mixta (*mixture cure models*), modelos paramétricos flexibles (*flexible parametric models*), modelos de referencia basada en la respuesta (*response-based landmark models*) y modelos de mezcla paramétrica (*parametric mixture models*), en un intento por evitar una infraestimación de la SG a largo plazo cuando se emplea la inmunoterapia en oncología (70,71).

De igual forma, cada vez es más frecuente que los pacientes sean tratados en onco-hematología con tratamientos secuenciales fijos cuando progresan a las líneas iniciales de tratamiento, dado que las nuevas terapias (incluyendo la inmunoterapia) hacen que la supervivencia de los pacientes cada vez sea mayor. En este caso, cada vez se tiende a emplear más los modelos secuenciales de tratamientos que va a ser capaces de evaluar la efectividad de cada tratamiento empleado en el paciente, en especial la supervivencia global (72). Dado que es una metodología novedosa y, que en ocasiones, no va a existir mucha evidencia de los resultados en salud logrados con los tratamientos secuenciales, va a ser necesario seguir las recomendaciones existentes para su realización (73).

9) Efectuar un ajuste temporal de los costes y de los resultados en salud incluidos en el modelo, si el horizonte temporal es mayor de un año

Dado que la sociedad y los individuos prefieren disfrutar los beneficios clínicos cuanto antes y diferir los costes lo más posible, parece lógico actualizar los flujos futuros de costes y resultados alcanzados al año base, momento en el que se van a tomar las decisiones.

Por este motivo, cuando el horizonte temporal de la evaluación económica es superior a un año es necesario aplicar un ajuste temporal, incorporando una tasa de descuento, tanto a los costes como a los resultados en salud, con el fin de actualizarlos y transformarlos en valor del momento presente (74).

En nuestro país se recomienda adoptar una tasa de descuento anual del 3 % para costes y resultados en salud, y el empleo del 0% y el 5% en el análisis de sensibilidad para poder comparar los resultados de la evaluación económica realizada con otros estudios ya existentes (75).

10) Analizar y presentar los resultados de la evaluación económica realizada con el modelo

Es recomendable notificar los resultados obtenidos en el caso base de referencia, incluyendo una parte dedicada a la perspectiva social del análisis incorporando un inventario del impacto de todos los costes incluidos en la evaluación económica, esto es los costes sanitarios y los costes no sanitarios (informales e indirectos) en los resultados finales de la evaluación.

Es deseable comunicar los valores medios de todos los costes y resultados incluidos de cada alternativa evaluada en el MAD, con descuento y sin haber aplicado la tasa de ajuste correspondiente, tanto de manera agregada como desagregada. También es conveniente comunicar los posibles efectos derivados de la heterogeneidad de los pacientes, notificando costes y resultados medios en los subgrupos de pacientes de máximo interés, lo mismo que para las diferentes perspectivas adoptadas en la evaluación económica, si hubiera más de una en la evaluación (76,77).

A la hora de presentar los resultados, es necesario efectuar un análisis incremental, calculando el cociente coste-efectividad (o utilidad) incremental que representa el coste incremental por unidad de resultado

adicional obtenido (CCEI), separado por las perspectivas del análisis, si se ha escogido más de una en la evaluación económica. Además, es deseable presentar el beneficio neto sanitario/monetario y el valor umbral utilizado, junto con una representación gráfica de los datos del CCEI en el plano coste-efectividad (78,79).

11) Manejar adecuadamente la incertidumbre de los resultados del modelo

Siempre va a ser necesario valorar la incertidumbre de los resultados del MAD, incluyendo la incertidumbre metodológica (se han incluido los aspectos metodológicos correctos en el modelo, según las guías existentes para su elaboración), estructural (la estructura del modelo va a permitir capturar los caracteres más relevantes de la enfermedad diana y de las opciones terapéuticas en evaluación), y de parámetros (el valor incorporado de cada parámetro en el modelo es cierto), lo que va a redundar en una mayor validez y fiabilidad de los resultados del modelo (80,81).

A la hora de explorar la incertidumbre estructural del modelo y su implicación en los resultados, será necesario emplear diferentes aproximaciones, tales como el análisis de escenarios (por ej. horizontes temporales cortos versus más largos, inclusión o no de costes no sanitarios, etc.) y la evolución del modelo empleado (si es el más correcto acorde a la enfermedad evaluada y las alternativas incluidas) (82).

A la hora de evaluar la incertidumbre de parámetros del modelo, es recomendable efectuar un análisis de sensibilidad determinístico (univariante, multivariante, umbral) siempre justificando las variables elegidas y su rango de valores empleado, presentando los resultados del análisis en un diagrama de tornado. De la misma manera, en la actualidad es muy recomendable incluir un análisis de sensibilidad probabilístico, justificando la elección de las distribuciones de probabilidad empleadas con cada

variable, y los resultados deberían presentarse con una gráfica del plano coste-efectividad y a través de la curva de aceptabilidad (83,84).

Teniendo en cuenta la limitación de los recursos disponibles para la financiación de nuevos medicamentos y otras tecnologías sanitarias en nuestro país, y que la asignación de recursos para financiar nuevas prestaciones sanitarias se suele realizar por parte de los decisores bajo condiciones de incertidumbre (tanto en datos de beneficio-riesgo, como de coste-efectividad), podría ser de gran valor para tomar decisiones conocer si merece la pena invertir recursos para disminuir este grado de incertidumbre. Conocer este dato (y la inversión necesaria para su conocimiento) se puede cuantificar empleando el análisis del valor de la información, por lo que siempre habría que valorar si merece la pena realizar un análisis de este tipo, calculando el valor esperado de la información perfecta (85).

12) Evaluar la transparencia, relevancia, credibilidad, consistencia y validación del modelo

En todo MAD es necesario valorar la transparencia de este, para lo que es necesario que los autores proporcionen un resumen de las características del modelo a través de una documentación descriptiva para que cualquier lector pueda conocer las características más importantes del mismo, junto con una documentación técnica del modelo (*technical report*) donde se detalla la estructura del modelo, sus componentes, las ecuaciones y los códigos de computación empleados. Además, va a ser necesario realizar una calibración para estimar los valores de los parámetros desconocidos o inciertos en el modelo, de tal manera que se pueda empezar con estimaciones que vengan de datos clínicos y estudios epidemiológicos disponibles (86,87).

Por otro lado, va a ser necesario evaluar la relevancia y credibilidad del MAD para que el decisor sanitario pueda tener en cuenta sus resultados a

la hora de efectuar recomendaciones en política sanitaria, por lo que habrá que valorar si la población incluida es la adecuada, si se han incluido las alternativas terapéuticas más empleadas, si se han tenido en cuenta los resultados más relevantes para la toma de decisión y si las características basales del modelo son relevantes para el decisor (88). De la misma manera, va a ser necesario realizar una verificación del funcionamiento del modelo, esto es, si la programación del modelo ha sido adecuada, si el modelo funciona acorde a lo esperado, si es consistente internamente, si los resultados generados son técnicamente correctos y si los resultados se pueden verificar y son reproducibles, puntos muy importantes para incrementar la credibilidad de los resultados obtenidos (89).

En otro orden de cosas, siempre va a ser necesario revisar la consistencia del modelo, o sea, la calidad global del mismo, incluyendo su estructura y los datos y asunciones incluidos en el mismo. Siempre habrá que valorar la consistencia interna, que se relaciona con la lógica matemática del modelo (al modificar algún dato, los resultados deben cambiar, y al realizar el modelo con otro software distinto, los resultados deben ser muy parecidos), y la consistencia externa, que valora si los resultados del MAD son consistentes con evidencias ya existentes o con los resultados de otros modelos realizados para explorar el mismo tema (90).

También es necesario desarrollar un proceso de validación del MAD, que va a evaluar la exactitud y aplicabilidad de los resultados para tomar decisiones, revisando la validez aparente (estructura y objetivo del modelo, fuente de los datos incluidos, resultados obtenidos), la validación interna, también llamada validez técnica (que examina si los cálculos matemáticos se han realizado correctamente y son consistentes con la especificación del MAD), la validez externa, también llamada validez convergente o validez cruzada (compara los resultados del modelo con otros modelos realizados), y la validez predictiva (predice los resultados del modelo acorde a evidencias futuras de resultados obtenidos en la práctica médica) (91).

En un futuro próximo, con el avance de la digitalización global y la posibilidad de elaborar y validar los MAD on-line a través de equipos en

red, va a ser importante que los MAD estén disponibles públicamente, de tal manera que se pueda desarrollar un proceso de validación continua de los mismos (92,93).

13) Discutir resultados, elaborar conclusiones y presentar las limitaciones del modelo

La discusión de lo encontrado en el MAD va a poner en contexto los resultados hallados y va a ayudar al lector a interpretar y revisar críticamente los hallazgos de la evaluación económica. La parte de la discusión va a sintetizar los resultados más importantes y va a ayudar a generar las conclusiones de la evaluación.

Los resultados obtenidos deberían ser discutidos en el contexto de la literatura existente y los datos ya disponibles de otros estudios efectuados con anterioridad, aportando las posibles discrepancias encontradas, con el fin de poder resaltar las aportaciones y el potencial impacto de los nuevos hallazgos de la evaluación realizada con el modelo.

Es totalmente recomendable exponer de manera crítica las principales limitaciones del modelo, que pueden ir desde las fuentes de información empleadas y el tipo de modelo empleado, hasta las asunciones incluidas o la validez externa de los resultados obtenidos. El hecho de presentar las posibles limitaciones del modelo incrementa su transparencia y validez, y siempre deberán presentarse de manera clara los aspectos a mejorar en el caso de que se pudiera realizar el modelo en el futuro.

Una vez se ha planteado la discusión y las posibles limitaciones de la evaluación, los autores deberán exponer las conclusiones de esta que deberían responder a la pregunta planteada con el modelo y que deberían ser coherentes con los resultados encontrados. Uno de los puntos a incluir en las conclusiones es valorar si los resultados se pueden transferir a otros contextos, entornos o países, con el fin de que se puedan tener en cuenta para tomar decisiones sanitarias (94).

14) Declarar el conflicto de interés de los autores y de la financiación del modelo

En toda evaluación económica realizada a través de un MAD es necesario especificar quién lo encargo y quién lo financió, independientemente de si el ente que lo ordeno y pago es público o privado, de tal manera que exista una total transparencia en este punto.

Además, es imprescindible que quede claro el papel que ha tenido el financiador en la elección de los participantes que han participado en el MAD, el diseño y la realización de la evaluación y la comunicación de los resultados, con el fin de incrementar la validez y relevancia de los resultados. Es importante que el financiador se mantenga al margen a la hora de presentar los resultados y en su publicación en una revista científica.

Por otra parte, es imprescindible informar sobre la posible existencia de conflictos de interés de los participantes en el diseño y validación del MAD, circunstancias en las que el juicio de los profesionales y expertos que han participado en la evaluación económica y el análisis de sus resultados pueda estar influenciado por diferentes factores, en especial relaciones financieras.



Futuro de los modelos analíticos de decisión

Cada día es más importante disponer de datos de la eficiencia relativa de las nuevas intervenciones terapéuticas con relación a las ya existentes en el mercado, lo que nos permitirá conocer si van a ser o no una inversión eficiente y rentable para el SNS.

Los modelos van a ser metodologías claves para demostrar la eficiencia de los nuevos medicamentos y otras tecnologías sanitarias en el momento de la “peri-comercialización”, cuando es más necesario tomar

decisiones que afectarán a su financiación y a su posible inclusión en los formularios hospitalarios, guías terapéuticas y protocolos asistenciales. Por este motivo, la realización de MAD se va a incrementar en el futuro ya que van a ser necesarios para conocer que alternativas terapéuticas son las más coste-efectivas, opciones que habría que priorizar su uso dentro del SNS, con el fin de incrementar los resultados en salud de los pacientes con costes adicionales asumibles por el SNS. Esta realidad exige que los MAD se diseñen, realicen, analicen y comuniquen de la manera más transparente y válida posible.

Para que esto sea una realidad, se ha desarrollado esta guía que aborda el diseño y realización de un MAD de nuevo, y que también sirve para evaluar la calidad metodológica de uno ya existente. En la misma línea, sería deseable que existiesen unas normas de buena práctica en modelización, que deberían incorporar a los pacientes e incluir una valoración de la transparencia, la consistencia interna, la reproducibilidad y la interpretabilidad del modelo, y que fomentase su consistencia externa y su parsimonia (se debería evitar una complejidad excesiva e introducir sólo aquellas variables y componentes estructurales que sean importantes para el objetivo del modelo) (95). Además, se debería incrementar el papel de los clínicos expertos y las sociedades científicas en el proceso de validación de los MAD elaborados, lo mismo que los pacientes y los ciudadanos en general, de tal manera que los resultados de las evaluaciones económicas efectuadas con modelos sean más creíbles y relevantes para los profesionales sanitarios y los agentes decisores del SNS (96,97).

Dado que en el futuro cada vez se van a realizar más MAD, sería deseable que se crease una red colaborativa a nivel nacional e internacional (que podría estar albergada en alguna página web), donde se podría elaborar un registro de MAD diseñados y realizados que cubriesen diferentes enfermedades y distintos medicamentos y productos sanitarios. De esta manera, se incrementaría la calidad de los MAD realizados, se evitaría la repetición de muchos de ellos y se podría mejorar la eficiencia de su elaboración y realización.



Bibliografía

1. Briggs AH, Claxton K, Sculpher MJ. Decision modelling for health economic evaluation: Handbooks in Health Economics Evaluation. Oxford: Oxford University press, 2006.
2. Buxton MJ, Drummond MF, Van Hout BA, et al. Modelling in economic evaluation: an unavoidable fact of life. *Health Econ.* 1997;6:217-27.
3. Sun X, Faunce T. Decision-analytical modelling in health-care economic evaluations. *Eur J Health Econ.* 2008;9:313-23.
4. Inadomi JM. Decision analysis and economic modeling: a primer. *Eur J Gastroenterol Hepatol.* 2004;16:535-42.
5. Siebert V, Alagoz O, Bayoumi AM, et al. State-transition modeling: a report of the ISPOR-SMDM modelling good research practices Task Force 3. *Value Health.* 2012;15:812-20.
6. Woods BS, Sideris E, Palmer S, et al. Partitioned survival and state transition models for healthcare decision making in oncology: where are we now? *Value Health.* 2020;12:1613-21.
7. Cooper NJ, Sutton AJ, Abrams KR, et al. Comprehensive decision analytical modeling in economic evaluation: a Bayesian approach. *Health Econ.* 2004;13:203-26.
8. Soto J. Health economic evaluations using decision analytic modeling. Principles and practices-Utilization of a checklist to their development and appraisal. *Int J Techn Assess Healthcare.* 2002;18:94-111.
9. Caro JJ, Briggs AH, Siebert V, et al. Modeling good research practices Overview: A report of the ISPOR-SMDM Modeling good research practices Task Force-1. *Value Health.* 2012;15:796-803.
10. 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;6:9-17

11. Dahabred IJ, Troikalinos TA, Balk EM, et al. Recommendations for the conduct and reporting of modelling and simulation studies in health technology assessment. *Ann Intern Med.* 2016;165:575-81.
12. Roberts M, Russell LB, Paltiel AD, et al. Conceptualizing a model: a report of the ISPOR-SMDM modelling good research practices task force-2. *Value Health.* 2012;15:804-11.
13. Petrou S, Gray A. Economic evaluation using decision analytical modelling: design, conduct, analysis, and reporting. *BMJ.* 2011;342:d1766.
14. Huang M, Ramsey S, Xue W, et al. Conceptual framework and methodological challenges for modelling effectiveness in oncology treatment sequence models. *PharmacoEconomics.* 2022;40:257-68.
15. Philips Z, Bojke L, Schulpher M, et al. Good practice guidelines for decision-analytic modelling in health technology assessment: a review and consolidation of quality assessment. *PharmacoEconomics.* 2006;24:355-71.
16. Sculpher M. Subgroups and heterogeneity in cost-effectiveness analysis. *PharmacoEconomics.* 2008;26:799-806.
17. Shields GE, Wilberforce M, Clarkson P, et al. Factors limiting subgroups analysis in cost-effectiveness analysis and a call for transparency. *PharmacoEconomics.* 2022;40:149-56.
18. Kim DD, Wilkinson CL, Pope EF, et al. The influence of time horizon on results of cost-effectiveness analyses. *Exp Rev Pharmacoecon Outcomes Res.* 2019;17:615-23.
19. Kim DD, Silver MC, Kunst N, et al. Perspective and costing in cost-effectiveness analysis, 1974-2018. *PharmacoEconomics.* 2020;38:1135-45.
20. Jönsson B. Ten arguments for a societal perspective in the economic evaluation of medical innovations. *Eur J Health Econ.* 2009;10:357-9.
21. Urwin S, Lau Y-S, Grande G, et al. The challenges of measuring informal care time: a review of the literature. *PharmacoEconomics.* 2021;39:1209-23.
22. Yuasa A, Yonemoto N, Lopresti M, et al. Use of productivity loss/gain in cost-effectiveness analyses for drugs: a systematic review. *PharmacoEconomics.* 2021;39:81-97.

23. Spacirova Z, Epstein D, Garcia-Monchon L, et al. A general framework for classifying costing methods for economic evaluation of health care. *Eur J Health Econ.* 2020;21:529-42.
24. Mokri H, Kvamma I, de Vries L, et al. Future medical and non-medical costs and their impact on the cost-effectiveness of life-prolonging interventions: a comparison of five European countries. *Eur J Health Econ.* 2023;24:701-15.
25. Hay SW, Smeeding J, Carroll NV, et al. Good research practices for measuring drug costs in cost-effectiveness analyses: issues and recommendations: the ISPOR drug cost Task Force-Part 1. *Value Health.* 2010;13:3-7.
26. Braithwaite RS, Roberts MS, Justioce AC. Incorporating quality of evidence into decision analytic modeling. *Ann Intern Med.* 2007;146:133-41.
27. Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ.* 2010;340:C869.
28. Vanderbroucke JP, Von Elm E, Altman DG, et al. Strengthening the reporting of observational studies in epidemiology (STROBE): explanation and elaboration. *Ann Intern Med.* 2007;147:W163-W194.
29. Revicki DA, Frank L. Pharmacoeconomic evaluation in the real word. *Pharmacoeconomics.* 1999;15:423-34.
30. Smith A. The challenge of assessing treatment effectiveness in a real-world setting. *Clin Ther.* 2021;43:2026-9.
31. Conceto J, Corrigan-Curay J. Realworld evidence-where are we now? *N Engl J Med.* 2022;386:1680-2.
32. Hoaglin DC, Hawking N, Jansen JD, et al. Conducting indirect treatment comparison and network-meta-analysis studies: report of the ISPOR task force on indirect treatment comparisons Good Research Practices-part 2. *Value Health.* 2011;14:429-37.
33. Sutton A, Ades AE, Cooper N, et al. Use of indirect and mixed treatment comparisons for technology assessment. *Pharmacoeconomics.* 2008;26:753-67.
34. Tonin FS, Borba HH, Leonard CD, et al. Methodological quality assessment of network meta-analysis of drop interventions: implications from a systematic review. *Int J Epidemiol.* 2019;48:620-32.

35. Signorouitch JE, Sikirica V, Erder M, et al. Matching-adjusted indirect comparisons: a new tool for timely comparative effectiveness research. *Value Health*. 2012;15:940-7.
36. Cooper NJ, Peters J, Lai MCW, et al. How valuable are multiple treatment comparison methods in evidence-based health-care evaluation? *Value Health*. 2011;14:371-80.
37. Iglesias CP, Thompson A, Rogowski WH, et al. Reporting guidelines for the use of expert judgement in model-based economic evaluations. *PharmacoEconomics*. 2016;34:1161-72.
38. Dahabreh IJ, Trikalinos TA, Balk EM, et al. Recommendations for the conduct and reporting of modeling and simulation studies in health technology assessment. *Ann Intern Med*. 2016;165:575-81.
39. McDonough CM, Tosteson ANA. Measuring preferences for cost-utility analysis. *PharmacoEconomics*. 2007;25:93-106.
40. Petrou S, Kwon J, Madan J. A practical guide to conducting a systematic review and meta-analysis of health state utility values. *PharmacoEconomics*. 2018;36:1043-61.
41. Brazier J, Ara R, Azzabi I, et al. Identification, review, and use of health state utilities in cost-effectiveness models: An ISPOR good practices for outcomes research task force report. *Value Health*. 2019;22: 267-75.
42. Pourrahmat MM, Kim A, Kansal AR, et al. Health state utility values by cancer stage: a systematic literature review. *Eur J Health Econ*. 2021;22:1275-88.
43. Zoratti MJ, Pickard AS, Stalmeier PFM, et al. Evaluating the conduct and application of health utility studies: a review of critical appraisal tools and reporting checklists. *Eur J Health Econ*. 2021;22:723-33.
44. Zoratti MJ, Chan KKW, Husereau D, et al. Towards transparency in the selection of published health utility inputs in cost-utility analyses: the health utility application tool (HAT). *PharmacoEconomics*. 2021;22:723-33.
45. Scalone L, Stalmeier PFM, Milani S, et al. Values for health states with different life durations. *Eur J Health Econ*. 2015;16:917-25.

46. Boike L, Manca A, Asaria M, et al. How to appropriately extrapolate costs and utilities in cost-effectiveness analysis. *PharmacoEconomics*. 2017;35:767-76.
47. Pavacha N, Abdulla A, Macgilchrist KS. Systematic review of health state utility values in metastatic non-small cell lung cancer with a focus on previously treated patients. *Health Qual Life Outcomes*. 2018;16:179.
48. Shah KK, Bennett B, Lenny A, et al. Adapting preference-based utility measures to capture the impact of cancer treatment-related symptoms. *Eur J Health Econ*. 2021;22:1301-9.
49. Siebert V, Alagoz O, Bayoumi AM, et al. State-transition modelling: a report of the ISPOR-SMDM modelling good research practices task-force-3. *Value Health*. 2012;15:812-20.
50. Gidwam R, Rusell LB. Estimating transition probabilities from published evidence: a tutorial for decision modelers. *PharmacoEconomics*. 2020;38:1153-64.
51. Srivastava T, Latimer NR, Tappenden P. Estimation of transition probabilities for state-transition models: a review of NICE appraisals. *PharmacoEconomics*. 2021;39:869-78.
52. Olariu E, Cadwell KK, Hancock E, et al. Current recommendations on the estimation of transition probabilities in Markov cohort models for use in healthcare decision-making: a targeted literature review. *Clinicoecon Outcomes Res*. 2017;9:537-46.
53. Latimer NR. Survival analysis for economic evaluations alongside clinical trial-extrapolation with patient-level data: inconsistencies, limitations, and a practical guide. *Med Decis Making*. 2013;33:743-54.
54. Bullement A, Stevenson MD, Baio G, et al. A systematic review of methods to incorporate external evidence into trial-based survival extrapolations for Health Technology Assessment. *Med Decis Making*. 2023; 43: 610-20.
55. Djaby V, Adunlin G, Montero AJ. Survival modeling for the estimation of transition probabilities in model-based economic evaluations in the absence of individual patient data: A tutorial. *PharmacoEconomics*. 2014;32:101-8.

56. Get data-graph-digitizer. (<http://Getdata-graph-digitizer.com>) (Acceso verificado el 20 Dec 2023)
57. Hoyle MW, Henley W. Improved curve fits to summary survival data: application to economic evaluation of health technologies. *BMC Med Res Methodol.* 2011;11: 139.
58. Guyot P, Ades AE, Ovweng MJ, et al. Enhanced secondary analysis of survival data: reconstructing data from published Kaplan-Meier survival curves. *BMC Med Res Methodol.* 2012;1:9.
59. Wan X, Peng L, Li Y. A review and comparison of methods for recreating individual patient data from published Kaplan-Meier survival curves for economic evaluations: a simulation study. *PLoS ONE.* 2015;e0121353.
60. Kuk D, Varadhan R. Model selection in competing risks regression. *Stat Med.* 2013;32:3077-88.
61. Bullement A, Meng Y, Cooper M, et al. A review and validation of overall survival extrapolation in health technology assessments of cancer immunotherapy by the national institute for health and care excellence: how did the initial best estimate compare to trial data subsequently made available. *J Med Econ.* 2019;22:205-14.
62. Baio G. SurvHE: survival analysis for health economic evaluation and cost-effectiveness modeling. *J Stats Software.* 2020;95:1-47.
63. R software version 4.2.1. <http://www.r-project.org>) (Acceso verificado el 20 Dec 2023)
64. Elbasha EH, Chhatwal J. Theoretical foundations and practical applications of within-cycle correction methods. *Med Decis Making.* 2016;36:115-31.
65. Woods BS, Sideris E, Palmer S, et al. Partitioned survival and state transition models for healthcare decision making in oncology: where are we now? *Value Health.* 2020;23:1613-21.
66. Rui M, Wang Y, Fei Z, et al. Will the Markov model and partitioned survival model lead to different results? A review of recent economic evidence of cancer treatments. *Expert Rev Pharmacoecon Outcomes Res.* 2021;21:373-80.

67. Caro JJ, Möller J. Advantages and disadvantages of discrete-event simulation for health economic analyses. *Expert Rev Pharmacoecon Outcomes Res.* 2016;16:327-9.
68. Lochen A, Anderson RM. Dynamic transmission models and economic evaluations of pneumococcal conjugate vaccines: a quality appraisal and limitations. *Clin Microbiol Infect.* 2020;26:60-70
69. Gibson EJ, Begum N, Koblbauer I, et al. Modeling the economic outcomes of immunooncology rugs: alternative model frameworks to capture clinical outcomes. *ClinicoEconomics Outcomes Res.* 2018;10:139-54.
70. Ouwens MJ, Mukhopadhyay P, Zhang Y, et al. Estimating lifetime benefits associated with immune-oncology therapies: challenges and approaches for overall survival extrapolations. *PharmacoEconomics.* 2019; 37:1129-38.
71. Felizzi F, Paracha N, Pöhlmann J, et al. Mixture cure models in oncology: A tutorial and practical guidance. *Pharmacoeconomics Open.* 2021;5: 143-55.
72. Huang M, Ramsey S, Xue, et al. Framework and methodological challenges for modeling effectiveness in oncology treatment sequence models. *PharmacoEconomics.* 2022;40:257-68.
73. Lewis RA, Hughes D, Sutton AJ, et al. Quantitative evidence synthesis methods for the assessment of the effectiveness of treatment sequences for clinical and economic decision making: a review and taxonomy of simplifying assumptions. *PharmacoEconomics.* 2021;39:52-61.
74. O'Mahony JF, Newall AT, van Rosmalen J. Dealing with time in health economic evaluation: methodological issues and recommendations for practice. *PharmacoEconomics.* 2015;33:1255-68.
75. Lopez bastida J, Oliva J, Antoñanzas F, et al. Propuesta de guía para la evaluación económica aplicada a las tecnologías sanitarias. *Gac Sanit.* 2010;24:154-70.
76. Sanders GD, Neumann PJ, Basu A, et al. Recommendations for conduct, methodological practices, and reporting of cost-effectiveness analyses. Second panel on cost-effectiveness in health and medicine. *JAMA.* 2016; 316:1093-103.

77. McQueen R, Slejko JL. Toward modified impact inventory tables to facilitate patient-centered value assessment. *PharmacoEconomics*. 2021;39: 379-82.
78. Husereau D, Drummond M, Augustouski F, et al. Consolidated Health Economic Evaluation reporting standards (CHEERS) 2022 Explanation and elaboration: A report of the ISPOR CHEERS II Good Practices task force. *Value Health*. 2022;25:10-31.
79. Paulden M. Calculating and interpreting ICER, and net benefit. *PharmacoEconomics*. 2020;38:785-807.
80. Bilcke J, Beutels P, Brisson M, et al. Accounting for methodological structural and parameter uncertainty in decision-analytic models: A practical guide. *Med Decis Making*. 2011;31:675-92.
81. Afzali HHA, Bojke L, Karnon J. Model structuring for economic evaluations of new health technologies. *PharmacoEconomics*. 2018;36:1309-19.
82. Bojke L, Claxton K, Sculpher M, et al. Characterizing structural uncertainty in decision analytic models: a review and application of methods. *Value Health*. 2009;12:739-49.
83. Briggs AH, Weinstein MC, Fenwick EAL, et al. Model parameter estimation and uncertainty: A report of the ISPOR-SMDM modelling good research practices task force-6. *Value Health*. 2012;15:835-42.
84. Peñazola Ramos MC, Barton P, Jowett S, et al. A systematic review of research guidelines in decision-analytic modeling. *Value Health*. 2015;18: 512-24.
85. Fenwick E, Stevlen L, Knies S, et al. Value of information analysis for research decisions-An introduction: Report of the ISPOR value of information analysis emerging good practices task force. *Value Health*. 2020;23:139-50.
86. Eddy DM, Hollingworth W, Caro JJ, et al. Model transparency and validation: A report of the ISPOR_SMDM modeling good research practices task force-7. *Value Health*. 2012;15:8423-850.
87. Vanni T, Karnon J, Madena J, et al. Calibrating models in economic evaluation: A seven step approach. *PharmacoEconomics*. 2011;29:35-49.

88. Caro JJ, Eddy DM, Kau H, et al. Questionnaire to assess relevance and credibility of modeling studies for informing healthcare decision making: an ISPOR-AMCP-NPC good practice task force report. *Value Health*. 2014; 17:174-82.
89. Dasbach EJ, Elbasha EH. Verification of decision-analytic models for health economic evaluations: an overview. *PharmacoEconomics*. 2017;35:673-83.
90. Adarkwah CC, van Gils PF, Hiligsmann M, et al. Risk of bias in model-based economic evaluations: the ECOBIAS checklist. *Expert Rev Pharmacoecon Outcomes Res*. 2016;16:513-23.
91. Otten TM, Grimm SE, Ramaekers B, et al. Comprehensive review of methods to assess uncertainty in Health Economic evaluations. *PharmacoEconomics*. 2023; 41: 619-632.
92. Feenstra T, Corro-Remos I, Hamerlijnck D, et al. Four aspects affecting health economic decision models and their validation. *PharmacoEconomics*. 2022;40:241-8.
93. Harvard S, Adibi A, Easterbrook A, et al. Developing an online infrastructure to enhance model accessibility and validation: The peer models' network. *PharmacoEconomics*. 2022;40:1005-9.
94. Lim KK, Kolera-Kolarova R, Fox-Rushby J. A comparison of the content and consistency of methodological quality and transferability checklists for reviewing model-based economic evaluations. *PharmacoEconomics*. 2022; 40:989-1003.
95. Harvard S, Werker GR. Health economists on involving patients in modeling: potential benefits, harms, and variables of interest. *PharmacoEconomics*. 2021;39:823-33.
96. Tremblay G, Humphries B. A practical guide for clinical and key opinion leader validation of health economic models. *J Med Econ*. 2023;26:473-6.
97. Harvard S, Winsberg E. Patient and public involvement in Health Economics Modelling raises the need for normative guidance. *PharmacoEconomics*. 2023;41:733-40.

Tabla 1. Recomendaciones a seguir cuando se vaya a diseñar un modelo analítico de decisión (MAD) de nuevo o cuando se quiera evaluar la calidad metodológica de uno ya existente

1) Definir el objetivo y la finalidad del modelo y justificar la necesidad de su elaboración

- ¿Se razona la necesidad de diseñar un modelo en vez de haber realizado la evaluación económica (EE) a través de un ensayo clínico o de un estudio de vida real?

2) Detallar el diseño y la estructura del modelo

- ¿Se justifica la elección del tipo de modelo que se va a construir y las alternativas que se van a evaluar?
- ¿Se describe en detalle y de forma transparente la estructura del modelo? ¿Es la más apropiada para contestar el objetivo y finalidad de la EE?
- ¿Es correcto conceptualmente el modelo y representa fielmente la evolución de la patología, el patrón de tratamiento y las opciones más empleadas en el entorno dónde se realiza la EE?
- ¿Se ha consensuado con profesionales sanitarios y decisores las características básicas del modelo, la fuente de los datos, las opciones terapéuticas a comparar y las asunciones incorporadas?
- Si se trata de un modelo de transición de estados, ¿los estados de salud definidos y la duración de los mismos son adecuados y reflejan la evolución de la enfermedad diana?

3) Describir con detalle las alternativas terapéuticas que se evalúan en el modelo

- ¿El comparador empleado frente a la nueva opción terapéutica es el más empleado en la práctica médica diaria en el medio donde se van a utilizar los resultados del modelo?
- ¿Se especifica en detalle las opciones de 2ª línea y los tratamientos subsecuentes en el modelo?

Tabla 1 cont.

4) Especificar la población evaluada en el modelo

- ¿Se detalla la población evaluada en el modelo y sus características principales: edad, sexo, co-morbilidades, riesgo basal, estado evolutivo de la enfermedad, etc.?
- ¿Se ha evaluado la heterogeneidad de los pacientes incluidos en el modelo y se han realizado análisis de subgrupos de pacientes?

5) Exponer y razonar el horizonte temporal escogido en el modelo

- ¿Se describe y justifica el horizonte temporal elegido para la EE realizada con el modelo? ¿Es el más adecuado acorde a la evolución de la enfermedad estudiada y las opciones evaluadas?
- ¿Se han incluido también otros horizontes temporales en el análisis de sensibilidad?

6) Describir la perspectiva elegida en el modelo

- ¿Es la perspectiva más adecuada, acorde a la patología y las alternativas terapéuticas evaluadas en el modelo? ¿Se razona su elección?
- ¿Se han incluido otras perspectivas en el análisis de sensibilidad?

7) Identificar, recoger y cuantificar el consumo de recursos y los costes asociados

- ¿Se especifican los recursos y los costes incorporados en el modelo y se describe la metodología para obtenerlos y las fuentes de donde se han obtenido? ¿Se detalla la unidad monetaria utilizada?
- ¿Se notifica y justifica la incorporación de todos los costes relevantes (sanitarios, no sanitarios e indirectos) en función de la perspectiva elegida? ¿Se comunican de manera desagregada?
- ¿Se detalla la metodología seguida para calcular el coste de los medicamentos según las distintas características de los pacientes y la duración de cada tratamiento (media o mediana)

Tabla 1 cont.

- ¿Se han actualizado los costes incorporados en el modelo por la inflación? ¿Se detalla el método empleado?
- ¿Se han incluido los costes de las líneas posteriores de tratamiento, de los cuidados paliativos y del manejo de los efectos adversos generadas por las alternativas evaluadas?
- ¿Se han incorporado en el análisis los costes futuros relacionados con la enfermedad evaluada?

8) Incorporar en el modelo los resultados en salud más relevantes de las opciones en estudio, así como las probabilidades de que éstos sucedan

- ¿Se aportan datos y detalles del diseño de los estudios de donde se han obtenido los resultados en salud de las opciones incluidas en el modelo: ensayos clínicos, estudios de vida real- metaanálisis, comparaciones indirectas?
- ¿Se han obtenido los resultados de fuentes precisas, validas y creíbles?
- Si se han incorporado asunciones en el modelo, ¿son lógicas y conservadoras? ¿Se han contrastado con expertos clínicos?
- Si se ha empleado la opinión de expertos ¿Se describen y justifican los métodos empleados?
- Si se ha medido utilidad, ¿se detallan los cuestionarios empleados para su obtención y se aportan los valores de utilidad de los diferentes estados de salud considerados en el modelo, junto a las fuentes de donde se han obtenido?
- ¿Se especifica si se han empleado disutilidades por efectos adversos y otros posibles factores a la hora de calcular el valor de utilidad de las opciones evaluadas?
- En los modelos de transición entre estados (Markov y otros), ¿los estados de salud representan la evolución de la enfermedad?; ¿se han calculado apropiadamente las probabilidades de transición entre ellos?; ¿se justifica el uso de la corrección del medio ciclo?

Tabla 1 cont.

- ¿Si se ha empleado un modelo de supervivencia particional (Partitioned survival model), ¿se describe la metodología empleada para extrapolar las curvas de supervivencia libre de progresión y supervivencia global al horizonte temporal del modelo?
- Si se ha diseñado un modelo de eventos discretos o un modelo dinámico, ¿se justifica su elección?; ¿se aporta suficiente información para poder evaluar su calidad metodológica?
- En los modelos realizados con terapias de inmuno-oncología, ¿se justifica el tipo de modelo empleado?; ¿se ha incorporado un estado de salud relacionado con los pacientes largos supervivientes?; ¿se describe en detalle la metodología empleada para extrapolar la supervivencia global a largo plazo?; ¿se han incluido los modelos secuenciales de tratamiento?

9) Efectuar un ajuste temporal de los costes y de los resultados en salud incluidos, si el horizonte temporal es mayor de un año

- ¿Se justifica la tasa de descuento empleada, tanto para costes, como para los resultados en salud?

10) Analizar y presentar los resultados de la evaluación económica realizada con el modelo

- ¿Se notifican los resultados del caso de referencia de la evaluación económica realizada con el modelo?
- ¿Se notifican los costes generados y los resultados en salud obtenidos de manera desagregada para cada alternativa en evaluación?
- ¿Se ha efectuado un análisis incremental, calculando el cociente coste/efectividad incremental (CCEI) de las opciones en evaluación?
- ¿Se ha calculado el beneficio neto sanitario de las alternativas estudiadas?

Tabla 1 cont.

- ¿Se ha efectuado análisis de subgrupos de pacientes; ¿se justifica su realización o no haberlos realizado en base a la posible heterogeneidad de los pacientes incluidos en el modelo?

11) Manejar adecuadamente la incertidumbre de los resultados

- ¿Se ha evaluado la incertidumbre del modelo, incluyendo la incertidumbre metodológica, estructural y de parámetros?
- ¿Se describe la incertidumbre estructural del modelo y se discute su posible implicación en los resultados obtenidos?; ¿Se ha evaluado a través de un análisis de escenarios pertinente?; ¿Se justifican los parámetros incluidos?
- ¿Se ha efectuado un análisis de sensibilidad determinístico, modificando el valor de las variables con mayor incertidumbre y de las asunciones incorporadas?; ¿se describen los resultados obtenidos y se detallan a través del diagrama de tornado?
- ¿Se detalla y justifica el análisis de sensibilidad probabilístico y se muestra el resultado con gráficas del plano coste-efectividad y la curva de aceptabilidad?; ¿se justifican las distribuciones de probabilidad empleadas en las variables incluidas?
- ¿Se ha realizado un análisis del valor de la información, calculando el valor esperado de la información perfecta?

12) Evaluar la transparencia, relevancia, credibilidad, consistencia y validación del modelo elaborado

- ¿Se discute la transparencia, relevancia y credibilidad del modelo, incluyendo una calibración de los datos de este?; ¿se ha efectuado una verificación del modelo, incluyendo el software empleado?
- ¿Se ha realizado una evaluación de la consistencia del modelo?
- ¿Se ha efectuado una validación del modelo para valorar la exactitud y aplicabilidad de los resultados?

Tabla 1 cont.

- ¿Se especifica si el modelo esta disponible de manera pública y donde se puede obtener?

13) Discutir resultados, elaborar conclusiones y presentar las limitaciones del modelo

- ¿Se discuten los resultados obtenidos acorde a la evidencia ya existente y como van a ayudar en la toma de decisión?
- ¿Están justificadas las conclusiones/recomendaciones de la evaluación económica en base a los resultados y otros análisis efectuados con anterioridad?
- ¿Se han evaluado las limitaciones del modelo empleado?
- ¿Se ha evaluado la transferibilidad de los resultados del modelo a diferentes medios y países?

14) Declarar el conflicto de interés de los autores y la financiación del modelo

- ¿Se especifica quién encarga y financia la evaluación económica?
¿Se explicita en el contrato la independencia de la investigación a la hora de publicar los resultados?
- ¿Se informa sobre la posible existencia de conflicto de interés en las personas que han participado en la elaboración del modelo?



8

ELABORACIÓN DE UN MODELO ANALÍTICO DE DECISIÓN DE NOVO: PASOS A SEGUIR

María Mareque^a, Itziar Oyagüez^{a,b}

^a *Pharmacoeconomics & Outcomes Research Iberia*

^b *Fundación PORIB*

La evaluación económica es un análisis que implica la identificación, cuantificación y valoración de los parámetros y la comparación de resultados de dos o más intervenciones sanitarias y/o alternativas terapéuticas con el fin de determinar cuál es la eficiencia de las mismas, de forma que se pueda seleccionar la opción más favorable y de esa manera poder distribuir de la manera más eficiente posible los recursos sanitarios disponibles (1,2).

Para el desarrollo de las evaluaciones económicas, uno de los métodos que se pueden emplear son los modelos analíticos de decisión. Este tipo de modelos utilizan una metodología analítica que tiene en cuenta: 1) los resultados a largo plazo, 2) los distintos grupos de población y 3) los datos extraídos de la literatura, con el objetivo de estimar los efectos de una intervención sobre las consecuencias y costes sanitarios considerados (3).

En el presente capítulo se pretende describir o establecer los aspectos fundamentales para el desarrollo desde cero de un modelo analítico de decisión. En este sentido, para la elaboración de este tipo de modelos es recomendable llevar a cabo cada una de las fases, que se muestran gráficamente en la Figura 1:

- 1) Definición y contexto del análisis
- 2) Búsqueda de evidencia
- 3) Conceptualización y diseño del modelo
- 4) Programación del modelo
- 5) Resultados del modelo



Figura 1. Fases principales para la elaboración de un modelo analítico de decisión



Fuente: Elaboración propia



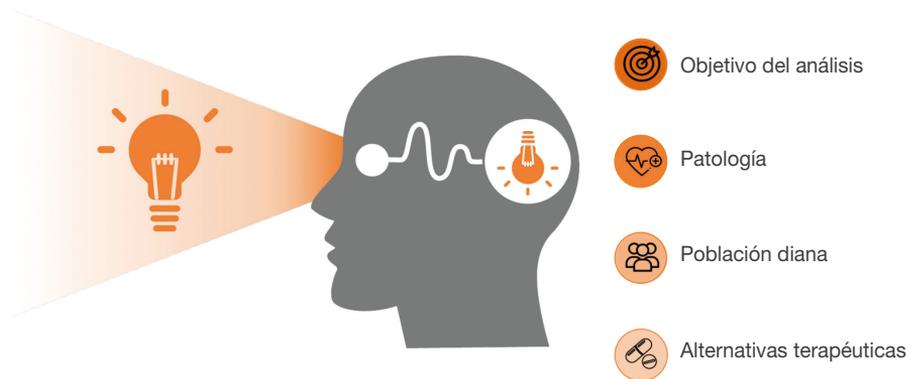
Fase I. Definición y contexto del análisis

Como primer paso para llevar a cabo este tipo de modelos se debe plantear lo que se denomina “pregunta de investigación o evaluación”, la cual determinará cual es el alcance, el objetivo y el contexto del análisis. Es importante que esta pregunta sea definida con claridad, concreción y el suficiente detalle desde un primer momento, ya que será el punto de partida para la elaboración del modelo y establecerá cual es la información requerida para su desarrollo (4).

Para ello, tal y como se muestra en la Figura 2, a la hora de definir la pregunta de investigación se deben tener claros una serie de conceptos comunes a todo tipo de evaluaciones económicas:



Figura 2. Elementos comunes a definir para el desarrollo del modelo



Fuente: Elaboración propia

- **Objetivo del análisis:** Determinar el propósito de la evaluación.
- **Patología de estudio:** Describir con precisión la entidad patológica sobre la que queremos realizar el análisis.

- **Población diana:** Definir la cohorte de pacientes a incluir en el análisis de acuerdo con la intervención objeto de estudio. Adicionalmente, se deberán determinar las características basales de los pacientes, teniendo en cuenta las características sociodemográficas (ej. edad, peso, altura, superficie corporal, proporción de hombres y mujeres, etc.) y clínicas de los mismos.
- **Alternativas terapéuticas o intervenciones sanitarias a comparar:** Incluir aquellas alternativas disponibles que mejor se ajusten a la patología en estudio y a la población considerada y que se empleen con mayor frecuencia en la práctica clínica habitual del entorno geográfico sanitario elegido para el análisis.



En esta primera fase es importante definir detalladamente el objetivo del análisis y la patología a evaluar ya que si estos dos conceptos no están bien establecidos desde el inicio del planteamiento del análisis podrían aparecer dificultades a la hora de diseñar el modelo e identificar los diferentes parámetros necesarios para su elaboración.



Fase II. Búsqueda de evidencia

Una vez definido el contexto y el objetivo del análisis en la fase anterior, lo más recomendable sería realizar una búsqueda en la literatura con el objetivo de identificar si existe algún modelo previo que dé respuesta a la cuestión planteada, o bien, recopilar la evidencia científica disponible en relación con el objetivo y patología en estudio (5).

Para ello, se recomienda llevar a cabo una revisión sistemática o al menos exhaustiva de la literatura, siguiendo las recomendaciones publicadas en la guía PRISMA (*Preferred Reporting Items for Systematic reviews and Meta-Analyses*) (6).

Una vez realizada la búsqueda de la evidencia disponible, se procederá al análisis de las publicaciones identificadas, las cuales servirán de base para determinar el diseño del modelo e identificar posibles parámetros necesarios para el desarrollo del análisis (Figura 3).

Figura 3. Búsqueda de evidencia disponible



Fuente: Elaboración propia

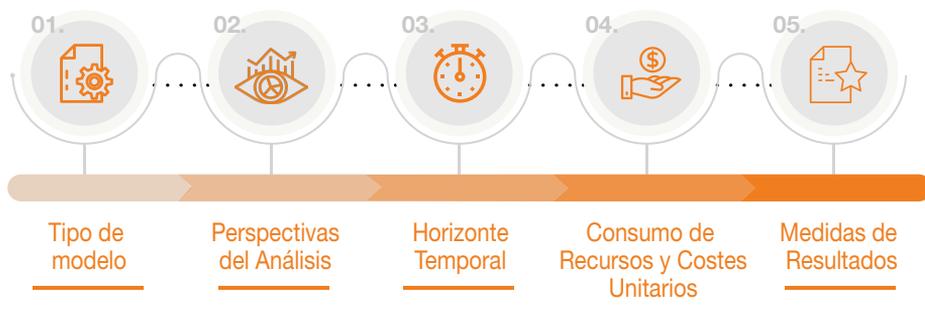


Fase III. Conceptualización y diseño del modelo

Una vez que se ha completado la revisión de la literatura disponible acerca de la patología y las posibles evaluaciones económicas existentes, el siguiente paso sería conceptualizar y diseñar el modelo. En esta fase, deberán determinarse los diferentes parámetros a incluir en el análisis (Figura 4).



Figura 4. Parámetros a determinar para la conceptualización y diseño del modelo



Fuente: Elaboración propia

Tipo de modelo

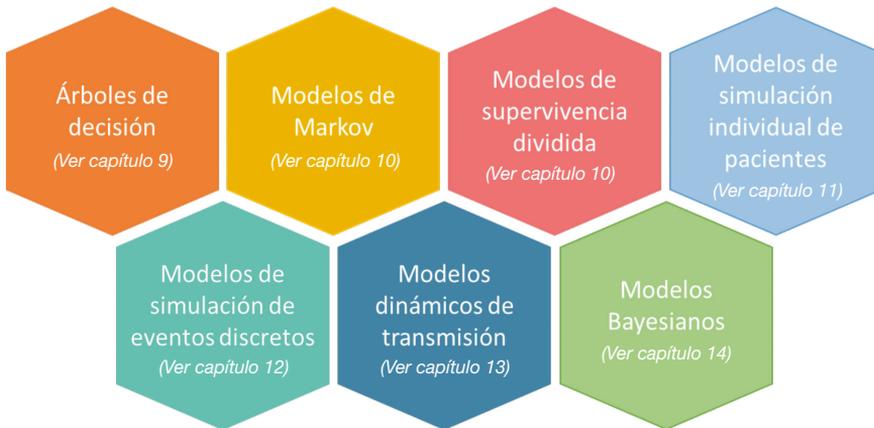
En primer lugar, se deberá definir el tipo de modelo a emplear, el cual tiene como finalidad identificar y comparar los resultados y costes de las distintas alternativas incluidas en el análisis. Hay que tener en cuenta que según el tipo de modelo elegido los resultados se medirán de manera

diferente y, por tanto, variará el tipo de pregunta al que pueden responder (ver capítulo 1 para más información).

A continuación, en la Figura 5 se muestra un resumen de los tipos de modelos más empleados para el desarrollo de modelos analíticos de decisión.



Figura 5. Tipos de modelos



Fuente: Elaboración propia

Perspectiva del análisis

Otro aspecto a tener en cuenta es la perspectiva del análisis, la cual expresa el alcance y punto de vista del modelo, determinando de este modo los costes derivados del consumo de los recursos afectados y los resultados que se deberán considerar a la hora de realizar el análisis (7,8).

Horizonte temporal

El horizonte temporal representa el periodo de tiempo en el que se producen todos los resultados y los costes de las intervenciones evaluadas, el cual debe ser lo suficientemente amplio para recoger todas las diferencias significativas en costes y resultados en salud entre las alternativas a comparar (5,9).

Consumo de recursos y costes

Para la estimación, cuantificación y valoración de los costes y recursos a incluir en el análisis se deberá tener en cuenta la perspectiva considerada, ya que la misma condiciona los recursos que deben incluirse en el análisis (Tabla 1) (*ver capítulo 3 para más información*).

Tabla 1. Clasificación tradicional de los costes a considerar en el modelo

	Sanitarios	No sanitarios
Directos	Cuidados hospitalarios, tratamiento farmacológico, etc.	Gastos de desplazamiento del paciente, cuidados en casa, etc.
Indirectos	Consumo de servicios sanitarios a lo largo de los años de vida ganados como consecuencia de la intervención sanitaria, entre otros	Pérdida de productividad; coste de oportunidad del tiempo invertido en el tratamiento, etc.

Fuente: Prieto L, et al. Med Clin (Baro) 2004;122(11):423-9

Medidas de resultados

La determinación de la medida de resultado en salud, así como la unidad de efectividad seleccionada, es de gran relevancia a la hora de desarrollar el análisis. A continuación, se muestran las medidas de resultados más utilizadas en el desarrollo de modelos analíticos de decisión (*ver capítulo 2 para más información*).

Tabla 2. Medidas de resultados

Dimensión clínica	Mortalidad, morbilidad y años de vida ganados
Valoración de las preferencias de los pacientes	Años de vida ajustados por calidad
Impacto en salud en términos monetarios	Disposición a pagar por las mejoras en salud de una intervención

Estas medidas pueden obtenerse de diversas formas: si bien, lo más habitual es que se obtengan de ensayos clínicos, en ciertas ocasiones se pueden obtener de estudios observacionales epidemiológicos o de información retrospectiva como bases de datos, registros de hospitales, opiniones de expertos, etc.



En esta fase se recomienda realizar un listado con los parámetros necesarios para el desarrollo del análisis, detectando de este modo que parámetros se han identificado en la revisión de la literatura y que parámetros faltarían para poder realizar el análisis



Fase IV. Programación del modelo

La programación del modelo supone traducir la conceptualización y diseño del modelo a un lenguaje informático. Para ello, existen diferentes softwares tales como Microsoft Excel®, Data TreeAge, R, WinBUGS, Open Markov, etc. (ver capítulo 15 para más información).

Microsoft Excel® es una herramienta muy versátil que permite comparar, verificar y analizar datos reduciendo al mínimo los posibles errores de cálculos. Además, Excel cuenta con Visual Basic, el cual permite personalizar sus hojas de cálculo mediante la programación de macros o fórmulas automatizadas.

Para la programación del modelo, se recomienda incluir los apartados señalados en la Figura 6.

A la hora de realizar la programación del modelo se recomienda incluir un apartado específico para abordar la incertidumbre del análisis, ya que la mayoría de los modelos son planteados a través de la combinación de información de varias fuentes como, por ejemplo, datos primarios de ensayos clínicos, datos estimados a partir de la búsqueda bibliográfica realizada, datos estimados por un panel de expertos, etc. Esta incertidumbre se puede testar mediante la realización de análisis de sensibilidad univariantes o multivariantes, análisis de umbral, análisis de escenarios y análisis de sensibilidad probabilístico (ver capítulo 5 para más información).



La representación gráfica del análisis de sensibilidad determinístico (diagrama de tornado) se podría realizar mediante un gráfico de barras y para los análisis de sensibilidad probabilísticos (plano coste-efectividad) mediante un gráfico de dispersión



Figura 6. Apartados recomendados para la programación del modelo



Fuente: Elaboración propia

Por último, con el objetivo de que el modelo sea susceptible de adaptación a otros entornos sanitarios, se recomienda programar los diferentes apartados de tal manera que permita al usuario modificar alguno de los valores de los parámetros elegidos y que, posteriormente, exista

la opción de que el usuario pueda reestablecer los valores originales. Para ello, se recomienda incluir un apartado específico donde se detallen todos los valores inicialmente definidos para el caso base para cada uno de los inputs considerados en el análisis. En caso de utilizar Microsoft Excel® para la programación del modelo, la opción de reestablecer los valores se podría realizar a través de Visual Basic.



Fase V. Resultados del análisis

La presentación de los resultados del análisis es un aspecto fundamental en los modelos de decisión analíticos. Estos resultados deberán de ser lo más detallados posibles, claros y transparentes para que cualquier lector pueda interpretarlos conociendo qué es lo que se ha hecho y por qué.

Para ello, se recomienda desarrollar un informe de resultados. En dicho informe se aconseja incluir los siguientes aspectos: 1) objetivo del análisis; 2) estructura del modelo, así como una breve explicación del mismo; 3) resumen de los parámetros empleados en el análisis; 4) análisis de sensibilidad realizados; 4) resultados del análisis tanto en términos agregados como desagregados; 5) resultados del análisis de sensibilidad; 6) representación gráfica de los resultados obtenidos; 7) un apartado con las limitaciones y asunciones realizadas en el modelo; y, por último, 8) un apartado de conclusiones (*ver capítulo 6 para más información*).



Bibliografía

1. Drummond MF, Sculpher MJ, Claxton K, et al. *Methods for the Economic Evaluation of Health Care Programmes*. 4th Ed. New York Oxford University Press, Oxford. 2015.
2. Sacristán JA, Costi M, Valladares A, et al. Health economics: the start of clinical freedom. *BMC Health Serv Res*. 2010;10:183.
3. 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;6:9-17.
4. Byford S, McDaid D, Sefton T. *Because it is worth it. A practical guide to conducting economic evaluations in the social welfare field*. York: Joseph Rowntree Foundation. 2003.
5. Ortega A, Marín R, Fraga MD, et al. (GENESIS-SEFH). *Guía de evaluación económica e impacto presupuestario en los informes de evaluación de medicamentos. Guía práctica asociada al programa MADRE v 4.0*. Madrid: SEFH (ed.), 2016. <http://gruposdetrabajo.sefh.es/genesis/index.php> (acceso verificado 7 Dic 2023)
6. Page MJ, Moher D, Bossuyt PM, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ*. 2021;372:n160.
7. Oliva J, Brosa M, Espín J, et al. Cuestiones controvertidas en evaluación económica (I): Perspectiva y costes de intervenciones sanitarias. *Rev Esp Salud Publica*. 2015;89:5-14.
8. Drummond M, McGuire A. *Economic Evaluation in Healthcare—Merging Theory With Practice*. Oxford, England: Oxford University Press; 2001.
9. National Institute for Health and Care Excellence (NICE). *Guide to the Methods of Technology Appraisal*. NICE, London 2013. <https://www.nice.org.uk/process/pmg9/resources/guide-to-the-methodsof-technology-appraisal-2013-pdf> (acceso verificado 7 Dic 2023).



9

MODELOS ANALÍTICOS DE DECISIÓN DETERMINÍSTICOS BASADOS EN ÁRBOLES DE DECISIÓN

Javier Soto Alvarez

Médico especialista en Farmacología Clínica

*Consejo de Dirección del Máster de Formación Permanente en Evaluación Sanitaria
y Acceso al Mercado (Fármaco-Economía) de la Universidad Carlos III de Madrid*

Los modelos analíticos de decisión (MAD) se utilizan para sintetizar de manera explícita y sistemática la mejor evidencia disponible sobre la historia natural y la evolución de una enfermedad, así como la efectividad, seguridad, costes y datos calidad de vida relacionada con la salud de diferentes opciones terapéuticas existentes para su tratamiento, con el fin de poder conocer cual de estos tratamientos es más eficiente, esto es, más coste-efectivo. Por lo tanto, se puede definir un modelo analítico de decisión como una aproximación sistemática a la evolución más probable de una enfermedad tras la administración de diferentes alternativas terapéuticas para su tratamiento, proyectando los resultados en salud que se van a obtener con cada una de las opciones administradas y los costes asociados en su consecución. Siempre va a ser necesario representar en el modelo la evolución más probable de la enfermedad diana de la manera más sencilla y comprensible que se pueda, pero siempre deberá ser lo más realista posible, lo que permitirá que el modelo sea fiable y preciso (1).

Existen muchos tipos de MAD que deberán emplearse según la enfermedad evaluada, su más plausible evolución, los tratamientos existentes para su manejo y los datos disponibles y siempre habrá que intentar emplear el MAD que mejor se ajuste y que sea el más sencillo posible, siempre que los resultados que nos aporten sean válidos, creíbles y de calidad (2). Dentro del mundo de los MAD, los modelos determinísticos basados en árboles de decisión (deterministic decision tree models) son los modelos más simples y se emplean de manera rutinaria sobre todo en enfermedades agudas y en enfermedades crónicas que pueden curar con descompensaciones agudas, ya que su elaboración no requiere mucho tiempo y los resultados son relativamente fáciles de obtener y de analizar (3). En este capítulo se van a repasar las características más importantes de este tipo de MAD, los pasos a dar para su elaboración y análisis de los resultados, así como la manera idónea de presentar y

comunicar los resultados de un modelo determinístico basado en árboles de decisión

Un MAD basado en un árbol de decisión es una representación visual de la evolución más probable de una enfermedad tras la administración de distintas opciones terapéuticas empleadas en el tratamiento de esa enfermedad, junto con los resultados clínicos que pueden derivarse de cada una de ellas y el consumo de recursos y costes asociados en su consecución. De cada alternativa terapéutica empleada van a salir diferentes ramas que representan la posible evolución de la enfermedad según una probabilidad determinada, la cual va a depender de las características de los pacientes (ya que van a presentar diferentes factores de riesgo) y de las opciones terapéuticas en sí (su nivel de eficacia y seguridad). El final de cada rama es un evento (curación, complicación, muerte, nivel de utilidad, etc.) y unos costes asociados (4).

Este tipo de modelo es adecuado cuando se evalúan enfermedades agudas (por ej. una neumonía hospitalaria) con un horizonte temporal corto, donde no suelen existir recaídas y recidivas de la enfermedad, y donde la mortalidad de los pacientes suele ser similar en los grupos en comparación (5).

Cuando se elabora un modelo basado en un árbol de decisión, primero hay que decidir qué alternativas terapéuticas se van a evaluar. Después, el árbol de decisión se desarrolla de izquierda a derecha añadiendo los posibles eventos que van a surgir de cada una de las opciones valoradas. A continuación es necesario incorporar la probabilidad de que aparezca cada evento, así como los resultados en salud obtenidos y los costes generados. Los datos necesarios para construir el árbol de decisión (probabilidades, eventos, resultados en salud, valores de utilidad, costes) se van a obtener de revisiones sistemáticas de la literatura y meta-análisis de ensayos clínicos, de comparaciones indirectas, de estudios observacionales y de paneles de expertos. En general, en estos MAD los datos de resultados en salud (curaciones, años de vida ganados, años de vida ajustados por calidad, etc.) generados se incorporan al modelo al final de las ramas del árbol, mientras que el consumo de

recursos y costes asociados se pueden incluir en los eventos que aparecen en las ramas intermedias del árbol, o incorporarlas únicamente al final de las ramas.

Este tipo de modelos son determinísticos, lo que significa que todos los datos incluidos en el modelo (efectividad, efectos adversos, probabilidades, costes, utilidades, etc.) se van a incorporar al modelo a través de fórmulas algebraicas sin emplear técnicas de simulación para conocer la variabilidad de estos datos a lo largo de la modelización. En los modelos determinísticos se utiliza la media de los datos de la población objeto del estudio, ya que se asume que hay certidumbre en estos datos y que no van a fluctuar de manera aleatoria debido al azar (por ejemplo, si se sabe por artículos publicados que la probabilidad de que aparezca un evento determinado es del 85 %, en el modelo se pondrá siempre esta cifra y no se modificará en ningún momento a lo largo del mismo). En un modelo determinístico, tanto a los parámetros como a las variables temporales, se les asignan valores determinados con certeza absoluta que no van a cambiar a lo largo del horizonte temporal del estudio.

A la hora de elaborar un modelo determinístico basado en un árbol de decisión es necesario seguir los siguientes pasos:

1) Establecer el objetivo del modelo y desarrollar su estructura

La pregunta de investigación que intenta contestar el modelo deberá estar claramente especificada, así como los tratamientos que se van a comparar, la población diana del mismo y el tipo de costes y resultados en salud que se van a evaluar. También debería especificarse la perspectiva de la evaluación económica y el horizonte temporal de la misma.

La estructura del modelo deberá estar acorde a la naturaleza de la enfermedad y su evolución clínica más plausible en el mundo real, y siempre debería estar validada por clínicos expertos en el patología y en los tratamientos evaluados

2) Asignar probabilidades a las ramas del árbol de decisión

Las probabilidades incorporadas a las ramas del árbol siempre tienen que provenir de estudios clínicos publicados (ensayos clínicos, estudios observacionales, estudios de vida real, estudios epidemiológicos, etc.), en la mayoría de ocasiones a través de revisiones sistemáticas de la literatura/metaanálisis y comparaciones indirectas, y deberían estar validados por expertos en estas técnicas

3) Incorporar costes y resultados en salud en las ramas finales del árbol de decisión

Los costes deberán ser asignados a los eventos acaecidos en las ramas del árbol de decisión que van a definir los resultados en salud, y deberá incorporarse el tipo de costes acorde a la perspectiva de la evaluación económica (SNS, hospital, sociedad, etc.) y siempre habrá que proporcionar la fuente de donde se han obtenido y año de los mismos.

Los resultados en salud a incluir en el nodo final de las ramas del árbol de decisión deberán ser aquellos que permitan realizar diferentes tipos de evaluaciones económicas: análisis coste-efectividad (años de vida ganados, porcentaje de curaciones, etc.), análisis coste-utilidad (años de vida ajustados por calidad), análisis de minimización de costes (años de vida ganados, porcentaje de curaciones, porcentaje de efectos adversos, eventos evitados, etc.).

4) Calcular los resultados en salud, costes y efectuar el análisis incremental

Los resultados en salud y los costes se calculan por separado en un proceso de análisis hacia atrás (folding back, desde los nodos

finales de las ramas hacia atrás hasta los nudos iniciales) que incluye sumas y multiplicaciones hasta llegar al resultado final (los resultados en salud de cada nodo del modelo son la suma de la efectividad en cada rama multiplicada por la probabilidad de que cada rama suceda).

Es necesario también efectuar un análisis incremental de los costes y los resultados en salud de las alternativas en evaluación, calculando el cociente coste/efectividad (utilidad) incremental.

5) Efectuar el análisis de sensibilidad

Siempre es necesario realizar un análisis de sensibilidad a los resultados del caso base con el fin de examinar la robustez y precisión de los resultados del modelo; este análisis de sensibilidad puede ser determinístico (univariante, multivariante, escenario o umbral) o probabilístico, incorporando los valores de las variables en las distribuciones de probabilidad que mejor se ajustan a cada una de ellas.

Por otra parte, en el desarrollo y diseño de un árbol de decisión será necesario considerar los siguientes apartados:

- a) patrón de tratamiento de la enfermedad diana en el entorno donde se van a emplear los resultados del modelo
- b) progresión y evolución de la enfermedad diana
- c) probabilidad de que aparezcan los diferentes eventos y resultados con cada intervención evaluada (acorde a los datos de la literatura publicada, la experiencia y la práctica clínica habitual en el medio)
- d) consumo de recursos y costes generados que se asocian a cada evento y resultados en salud producidos

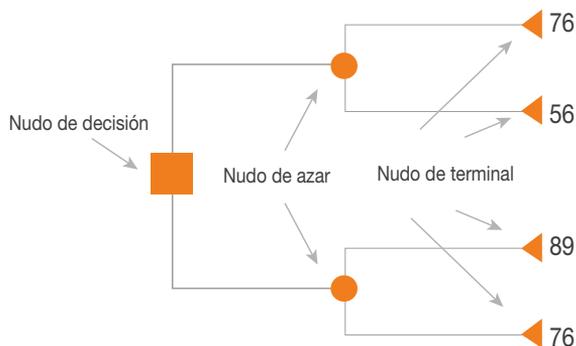
A la hora de su elaboración, los árboles de decisión están compuestos de dos partes diferentes:

- 1) nudos, que se representan mediante ciertos símbolos geométricos
- 2) ramas, que partiendo siempre de un nudo se representan mediante líneas continuas

Todos los árboles se desarrollan de izquierda a derecha, a partir de un punto original o nudo de origen, que se va irradiando de forma ramificada (a partir de ramas más grandes se llega a otras más pequeñas) hasta alcanzar el desenlace final, representado por los nudos terminales.

Por lo tanto, en el diseño de cualquier árbol de decisión, se van a diferenciar tres clases de nudos:

- de decisión, también llamado de elección o de acción
- de acontecimientos o de azar
- finales o terminales



Los nudos de acción (o de decisión) se encuentran casi siempre en el punto que da origen al árbol. Se representan, acorde a un símbolo aceptado de forma universal, a través de un cuadrado, y de éste pueden salir dos o más ramas, según las opciones terapéuticas que se van a evaluar, ya que cada una de ellas va a estar representada por una rama del árbol. El investigador que está realizando el estudio tiene un control total sobre las alternativas que van a emanar de él en forma de ramas, de tal manera que va a decidir cuantas alternativas se van a incluir en el estudio. Estas alternativas incorporadas no llevan asociadas probabilidades, ya que carece de sentido dado que son sucesos excluyentes entre sí (si se opta por dar una opción no se puede administrar la otra/s).

Los nudos de acontecimientos (o de azar) se pueden hallar en cualquier posición del árbol, excepto en aquellos puntos que corresponden a su inicio y final. Siempre se van a representar mediante un círculo o un punto, y de él pueden salir dos o más ramas, tantas como acontecimientos o desenlaces ocurran. Posteriormente, cada desenlace a su vez puede conducir a otros nudos de acontecimientos hasta que desemboquen en un nudo final. El investigador no tiene ningún control sobre los acontecimientos, sucesos o resultados que, en forma de ramas, pueden salir de ellos (ya que van a depender de la evolución de la enfermedad y de los datos de eficacia y seguridad de las opciones en estudio) y, además, cada rama que parte de estos nudos deben ir siempre acompañados de sus respectivas probabilidades de que ocurra, ya que se trata de sucesos que no son excluyentes (como sucedía con los nudos de decisión), sino complementarios entre sí.

Los nudos finales o terminales se encuentran siempre en el extremo situado a la derecha del árbol y en ellos se explicitan los resultados clínicos esperados a los que de lugar el desarrollo del árbol. Se representan mediante un rectángulo, óvalo o un triángulo (el más frecuente), y de ellos no puede partir ninguna otra rama (6).

Para una correcta resolución del árbol de decisión, es recomendable aplicar las siguientes reglas emanadas de la teoría de la probabilidad:

- a) la primera regla especifica que la suma de las probabilidades que están vinculadas a las ramas que salen de un mismo nudo de azar ha de ser igual a la unidad. De esta manera, es posible conocer la probabilidad de cualquier acontecimiento siempre que se conozca el resto de probabilidades de las otras opciones
- b) la segunda regla establece que la probabilidad de que se produzca un resultado concreto, está condicionada por las probabilidades de las ramas que le preceden
- c) la tercera regla señala que el resultado medio esperado de un acontecimiento es igual al sumatorio de los resultados finales por sus respectivas probabilidades. Mediante su aplicación se puede llegar a la resolución del análisis, siempre que se disponga de información precisa y valida en relación con la probabilidad de todos los eventos aparecidos y se hayan cuantificado y medido con fiabilidad los resultados obtenidos

La resolución adecuada de un árbol de decisión pasa por calcular las consecuencias promedio esperadas para cada una de las intervenciones comparadas, multiplicando la probabilidad de cada rama (que lleva a un nudo terminal) por su correspondiente resultado clínico y costes asociados, sumando después todos los valores conseguidos con cada una de las opciones en estudio (7). En la Figura 1 se detalla la representación de un árbol de decisión, su resolución y su análisis.

Los modelos basados en árboles de decisión son muy flexibles y, aunque, se pueden emplear para modelizar prácticamente cualquier proceso, se emplean especialmente cuando los parámetros del modelo no dependen del tiempo y cuando el horizonte temporal no es muy largo. Además, son muy intuitivos y fáciles de entender para aquellas personas no expertas en modelos.

Sin embargo, en enfermedades crónicas y en aquellas enfermedades agudas con múltiples recaídas o con recidivas, si se quiere representar apropiadamente la evolución de la enfermedad es necesario crear modelos muy complejos y difíciles de manejar y analizar. Por estos

motivos, en la actualidad se utilizan habitualmente para modelizar procesos agudos, como ya se ha comentado anteriormente.



Presentación y comunicación de los resultados de los modelos analíticos de decisión determinísticos

A la hora de comunicar los resultados de estos modelos, es necesario seguir unas recomendaciones, lo que facilitará que quienes vayan a emplear sus resultados puedan disponer de la información necesaria para valorar su validez, precisión y relevancia. Además, una presentación adecuada del modelo es clave de cara a incrementar su transparencia, aspecto crucial para que los resultados obtenidos se utilicen en la toma de decisión en política sanitaria.

En la Tabla 1, se especifica el formato que habría que seguir y la información que habría que comunicar a la hora de presentar un MAD determinístico a través del informe de su realización (8). En la parte de introducción del informe habrá que especificar los datos epidemiológicos de la enfermedad diana, el pronóstico de la misma y la progresión en el tiempo de la misma, datos claves para más tarde explicar el tipo de modelo elegido. Los datos de coste de la enfermedad y los resultados de las evaluaciones económicas disponibles de las diferentes opciones terapéuticas existentes para tratar la enfermedad, van a permitir disponer de datos importantes a la hora de elegir el comparador en el modelo y diseñar su estructura final.

En el apartado de la descripción de las intervenciones terapéuticas que van a ser comparadas en el modelo, se proporcionará información sobre la dosis y régimen posológico de las mismas. Es muy aconsejable presentar los datos de costes y resultados en salud obtenidos e incorporados al modelo (incluyendo el valor de utilidad) en tablas separadas, indicando la fuente de donde se han obtenido. De la misma manera, es recomendable que se elabore una tabla con las probabilidades incorpora-

das en el modelo (probabilidades fijas en los modelos determinísticos), indicando la fuente de donde se han sacado. Asimismo, debería elaborarse una tabla con los parámetros empleados para medir los resultados clínicos obtenidos con el uso de las opciones en estudio (RR, odds ratio, etc.).

Si se va a efectuar un análisis de sensibilidad probabilístico, es necesario que se presente en una tabla el tipo de distribución de probabilidad que se va a seguir cada variable del modelo (beta, log-normal, gamma, etc.) y los valores que van a definir cada distribución.

A la hora de comunicar los resultados del modelo, es importante presentar en una tabla el consumo de recursos y los costes asociados y los resultados en salud obtenidos con las opciones terapéuticas en estudio, conjuntamente con el cociente coste/efectividad incremental y el beneficio neto sanitario del caso base de referencia del análisis, si se estima oportuno. De manera adicional, va a ser necesario elaborar una tabla que describa el resultado del análisis de sensibilidad uni y multivariante, junto a la gráfica que detalle los resultados de este tipo de análisis de sensibilidad (diagrama de tornado) y el resultado del análisis de sensibilidad probabilístico (nube de puntos en el plano coste-efectividad y la curva de aceptabilidad).



Selección del tipo de modelo a emplear

¿Cuándo se recomienda emplear un MAD determinístico basado en un árbol de decisión?

Como regla general, cuando se vaya a elegir el tipo de modelo a utilizar, se recomienda usar modelos lo más simples posibles (siempre que reflejen de manera adecuada la evolución de la enfermedad diana y los eventos que van a aparecer y, por lo tanto, sus resultados sean válidos y creíbles), ya que van a ser más entendibles por los profesionales sanitarios y otros agentes sanitarios (9).

Además de esta recomendación, a la hora de elegir el tipo de modelo a emplear será necesario tener en cuenta las siguientes premisas (10-13):

- Quien va a utilizar los resultados del modelo (sólo la persona que lo ha desarrollado o también otras personas: colegas, decisores, etc.)
- En qué contexto se van a emplear sus resultados: reunión interactiva con una persona (por lo que se necesitará que el modelo se pueda revisar de manera más o menos rápida), o en reuniones con grupos para presentar los resultados (se podrá disponer de más tiempo para presentar y discutir los resultados del modelo)
- Cuanto tiempo se tiene para disponer de los resultados del modelo: es necesario tenerlos de manera rápida o se dispone de más tiempo
- Qué tipo de datos están disponibles para incorporar al modelo, en relación a la evolución de la enfermedad o de las opciones en evaluación en el mismo: si la disponibilidad de datos es limitada, no deberían seleccionarse modelos que necesitan muchos datos para su realización (la incorporación de asunciones debería ser la menos posible)
- El tipo de enfermedad a evaluar: aguda o crónica; con recaídas/recaídas o enfermedades infecciosas
- El horizonte temporal que pretende explorar el modelo: corto en el tiempo o más largo (incluso toda la vida de los pacientes)
- Si es importante la historia previa de los pacientes a la hora de obtener resultados en salud derivadas del uso de las alternativas en estudio y si es esperable que puedan existir interacciones entre los pacientes que padecen la enfermedad

En la Figura 1, se describe un algoritmo a seguir a la hora de seleccionar el tipo de modelo a emplear cuando se vaya a efectuar una evaluación económica a través de un modelo analítico de decisión.



Limitaciones de los modelos analíticos de decisión determinísticos

Aunque los modelos se emplean cada vez más para efectuar evaluaciones económicas de medicamentos y productos sanitarios, esta técnicas presentan limitaciones, debidas principalmente a las suposiciones incorporadas (que en ocasiones no están del todo justificadas y son poco conservadoras) y a los datos incluidos en el mismo, ya que en unas ocasiones no se incorporan todos los disponibles y, en otras, se incluyen datos de fuentes que no son las mejores y más validas. Además, la transparencia de los modelos no suele ser la más idónea (sobre todo en los cálculos numéricos), y casi nunca se realiza una validación de sus resultados con lo que acontece después en la práctica médica habitual (14).

Por una parte, los datos de los resultados en salud asociados a las alternativas en evaluación en el modelo, o bien son sacados de ECC publicados, o bien son obtenidos de revisiones sistemáticas de la literatura y meta-análisis. Sin embargo, ambas fuentes presentan importantes lagunas de validez externa y generabilidad de los resultados al conjunto global de la población. La incorporación al modelo de datos naturalísticos (estudios de vida real), bien provenientes de ensayos clínicos con orientación pragmática, o bien obtenidos a través de diseños observacionales, podría evitar en parte esta problemática, pero es necesario emplear una metodología correcta donde los posibles sesgos y factores de confusión se controlen apropiadamente. Además, en la mayoría de los modelos se suelen incorporar datos provenientes de paneles de expertos, sin que se haya seguido la metodología más apropiada para buscar el consenso, lo que limita la obtención de resultados creíbles y válidos.

Por otra parte, en muchas ocasiones las asunciones incluidas en el modelo son muy poco conservadoras y buscan los mejores resultados para una de las opciones evaluadas en el modelo. Este punto es extremadamente importante a la hora de lograr que los resultados de los

modelos sean transparentes y creíbles para los profesionales sanitarios y otros agentes decisores.

En otras ocasiones, el análisis de sensibilidad se realiza incluyendo sólo aquellas variables que se sabe de antemano que no van a modificar sustancialmente los resultados, omitiendo la inclusión de otras que podrían cambiar los resultados en una dirección no deseada. Por este motivo, siempre será necesario exponer las razones de haber elegido unas variables y no otras en el análisis de sensibilidad.

Además, es necesario tener claro que los MAD determinísticos no van a poder representar fielmente la evolución de patologías crónicas, por lo que deberían emplearse mayoritariamente para evaluar enfermedades agudas con evoluciones clínicas cortas en el tiempo, sin que suelen aparecer recaídas o recidivas de manera habitual.

Para que los MAD determinísticos efectuados tengan suficiente credibilidad y calidad, es necesario que los analistas que están desarrollando el modelo sigan los más altos estándares de calidad en su diseño y elaboración, de tal manera que sean totalmente transparentes en su realización y siempre se incorporen los datos que provengan de las fuentes más fiables y de mayor calidad. Es totalmente necesario que la estructura del modelo refleje la evolución más probable de la enfermedad y que el comparador sea siempre la opción terapéutica más relevante de las existentes en el mercado, y que se incorporen al modelo todos los datos disponibles (no sólo los más favorables) de las opciones terapéuticas evaluadas (la de referencia y la nueva en llegar al mercado).



Bibliografía

1. Hay JW. Evaluation and review of pharmacoeconomic models. *Expert Opin Pharmacother.* 2004;5:1867-80.
2. Buxton MJ, Drummond MF, Van Hout BA, et al. Modelling in economic evaluation: an unavoidable fact of life. *Health Econ.* 1997;6:217-27.
3. Siebert U. When should decision-analytic modeling be used in the economic evaluation of health care? *Eur J Health Econ.* 2003;4:143-150.
4. Haji A, Bojke L, Karnon J. Model structuring for economic evaluations of new health technologies. *PharmacoEconomics.* 2018;36:1309-19.
5. Weinstein MC, Toy EL, Sandberg, et al. Modeling for health care and other policy decisions: uses, roles, and validity. *Value Health.* 2001;4:348-61.
6. Inadomi JM. Decision analysis and economic modeling: a primer. *Eur J Gastroenterol Hepatol.* 2004;16:535-42.
7. Lang DL, Lopert R, Hill SR. Use of pharmacoeconomics in prescribing research. Part 5: modeling-beyond clinical trials. *J Clin Pharm Therapeutics.* 2003;28:433-9.
8. Nuijten MJC, Pronk MH, Brosens MJA, et al. Reporting format for economic evaluation. Part II: Focus on modeling studies. *PharmacoEconomics.* 1998;14:259-68.
9. Caro JJ, Möller J. Decision-analytic models; current methodological challenges. *PharmacoEconomics.* 2014;32:943-50.
10. Stahl JE. Modelling methods for pharmacoeconomics and health technology assessment. An overview and guide. *PharmacoEconomics.* 2008;26: 131-48.
11. Karnon J, Brown J. Selecting a decision model for economic evaluation: a case study and review. *Health Care Manag Sc.* 1998;1:133-40.
12. Barton P, Bryan S, Robinson S. Modelling in the economic evaluation of healthcare: selecting the appropriate approach. *J Health Serv Res Policy.* 2004;9:110-8.

13. Scholte M, Rovers MM, Grutters JPC. The use of decision analytical modeling in the evaluation of surgical innovations: A scoping review. *Value Health*. 2021;24:884-900.
14. Petrou S, Gray A. Economic evaluation using decision analytical modeling: design, conduct, analysis, and reporting. *BMJ*. 2011;342:d1766.

Tabla 1. Formato y datos que debería aportar el informe donde se comunican los resultados de un modelo analítico de decisión determinístico

Introducción del informe

- Datos epidemiológicos de la enfermedad diana (incidencia, prevalencia, edad presentación, factores de riesgo asociados, etc.)
- Pronóstico de la enfermedad (en general y en distintos subgrupos de pacientes)
- Progresión de la enfermedad
- Patrón de tratamiento de la enfermedad en el entorno donde se desarrolla el modelo
- Costes del manejo de la patología (estudios de coste y carga de la enfermedad)
- Evaluaciones económicas de las alternativas terapéuticas más empleados en el manejo de la enfermedad diana

Opciones terapéuticas que se comparan en el modelo

- Definir que opciones terapéuticas se van a comparar en el modelo: la nueva intervención terapéutica frente a alternativas ya disponibles en el mercado
- Presentar las dosis recomendadas y los regímenes posológicos más adecuados de cada una de ellas
- Aportar datos de las evidencias existentes sobre los resultados en salud que producen las opciones en evaluación

Hipótesis y objetivo del modelo

- Describir el porque se realiza el modelo y que como van a ayudar sus resultados en la toma de decisiones

Tabla 1. cont.

Diseño y metodos empleados en el modelo

- Describir del tipo de modelo realizado
- Presentar la estructura del modelo
- Definir la población que va a ser evaluada en el modelo
- Especificar el horizonte temporal y la perspectiva del modelo
- Notificar los datos clínicos y de resultados en salud que se han incorporado en el modelo, junto con la fuente de donde se han obtenido
- Detallar las probabilidades incorporadas a cada rama del árbol de decisión y las fuentes de donde se han extraído
- Describir los recursos consumidos y los costes asociados secundarios al empleo de las intervenciones en estudio, conjuntamente con las fuentes de donde se han obtenido
- Especificar los valores de utilidad que se han empleado en el modelo, si se ha efectuado un análisis coste-utilidad, y la fuente de los mismos
- Describir el análisis de sensibilidad empleado en el modelo, justificando su elección
- Si se ha realizado un análisis de sensibilidad probabilístico, describir la distribución de probabilidad elegida para cada variable incluida en el análisis
- Especificar el programa empleado para la realización del modelo

Resultados obtenidos

- Presentar los costes y los resultados en salud obtenidos (AVG, AVAC, evento evitado, etc.) de manera desagregada (costes, costes incrementales, resultados en salud, resultados incrementales)
- Comunicar el cociente coste/efectividad incremental obtenido y el beneficio neto sanitario del caso de referencia
- Describir los resultados de los análisis de sensibilidad efectuados
- Si se ha efectuado un análisis de sensibilidad probabilístico, presentar los resultados en el plano coste-efectividad y a través de la curva de aceptabilidad

Tabla 1. cont.

Discusión de los resultados

- Presentar la validación de los resultados obtenidos al compararlos con otros estudio ya efectuados, o bien a través de opiniones de expertos
- Discutir las discrepancias encontradas y recomendar la realización de futuros modelos
- Valorar el uso de los resultados para la toma de decisiones en el medio donde se ha realizado el modelo

Conflicto de interes

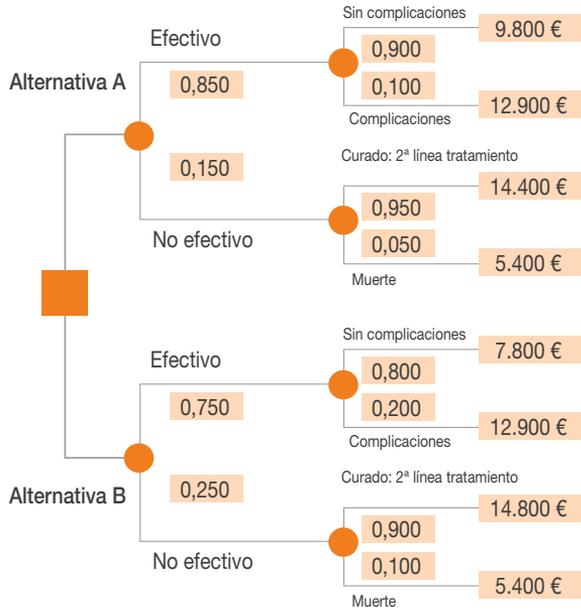
- Especificar si alguno de los autores del modelo presenta algún conflicto de interés
- Detallar la relación y el contrato entre el investigadores que ha desarrollado el modelo y el ente que ha financiado el mismo

Apendices y referencias

- Detallar la siguiente información del modelo: fuentes de los datos incluidos, lista de asunciones incorporadas, lista de formulas incluidas en el modelo, nombre y filiación labora de expertos que han participado en su desarrollo, etc.
- Especificar todas las referencias en las que se basa el modelo



Figura 1. Modelo determinístico basado en un árbol de decisión y su análisis (horizonte temporal de 1 año y se asume que todas las complicaciones viven)



Alternativa terapéutica A

Resultado	Costes €	Probabilidad	Costes x Probabilidad
Efectivo sin complic.	9.800	0,7650	7.497
Efectivo con complic.	12.900	0,0850	1.096,5
No efectivo con curación	14.400	0,1425	2.052
No efectivo con muerte	5.400	0,0075	40,5

Costes esperados

➔ **10.686**



Figura 1. cont. Modelo determinístico basado en un árbol de decisión y su análisis (horizonte temporal de 1 año y se asume que todas las complicaciones viven)

Alternativa terapéutica B	Resultado	Costes €	Probabilidad	Costes x Probabilidad
	Efectivo sin complic.	7.800	0,6000	4.680
	Efectivo con complic.	12.900	0,1500	1.935
	No efectivo con curación	14.800	0,2250	3.330
	No efectivo con muerte	5.400	0,0250	135
Costes esperados				10.080

Alternativas	Costes esperados	Probabilidad paciente vivo (AVG)	Cociente coste/efectividad medio	Cociente coste/efectividad incremental
A	10.686	0,9925 (*)	10.766	34.628 €/AVG
B	10.080	0,9750 (**)	10.338	

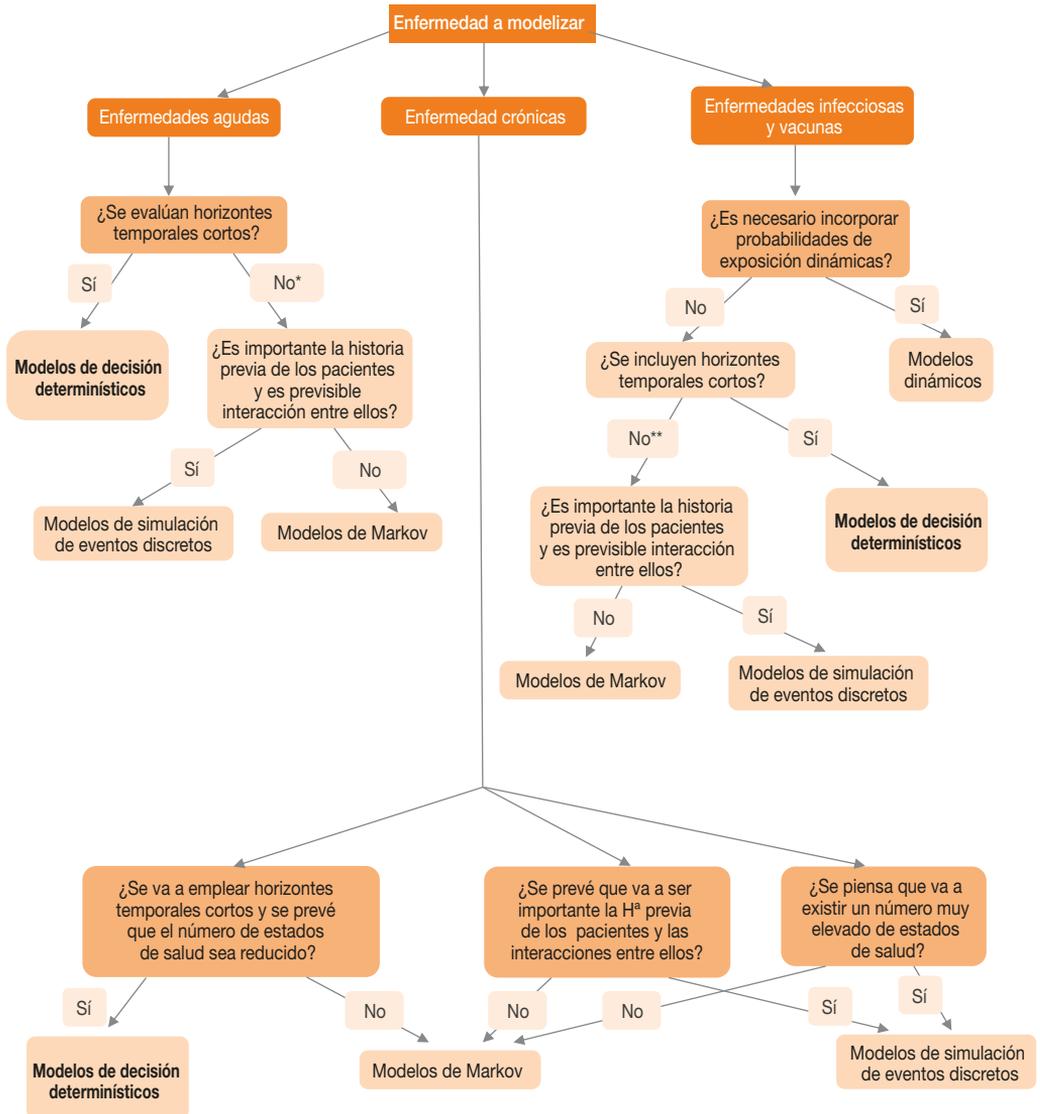
AVG = Año de vida ganado

(*) = $0,765 + 0,085 + 0,1425$

(**) = $0,600 + 0,1500 + 0,225$

Conseguir un año de vida adicional supone un coste adicional de 34.628 €, cuando se emplea la alternativa terapéutica A en vez de utilizar la alternativa terapéutica B

Figura 1. Algoritmo a seguir para elegir cuando emplear un MAD determinístico





10

MODELOS DE MARKOV Y MODELOS DE SUPERVIVENCIA DIVIDIDA

Antonio Olry de Labry Lima^a, Zuzana Špacírová^b

^a *Centro Andaluz de Documentación e Información de Medicamentos (CADIME).
Escuela Andaluza de Salud Pública. Granada, España*

*Centro de Investigación Biomédica en Red Epidemiología y Salud Pública
(CIBERESP), España*

Instituto de Investigación Biosanitaria ibs.GRANADA, Granada, España

^b *PhD (Técnica de Proyectos. Escuela Andaluza de Salud Pública)*

La evaluación económica es un enfoque sistemático que compara los costes y los resultados clínicos obtenidos tras la aplicación de dos o más tecnologías sanitarias (tratamiento, intervención o estrategia de prevención) alternativas. Así, el objetivo de una evaluación económica es determinar qué tecnología sanitaria representa el uso más eficaz de los recursos sanitarios, lo que suele denominarse como la alternativa más coste-efectiva. La limitación de recursos en la atención sanitaria, hace que, para la aprobación de una tecnología, junto con la calidad, la seguridad y la eficacia, se hace necesario demostrar que una intervención es rentable, y para ello algunos países usan como regla para decidir si financian o no una nueva tecnología la razón de coste-efectividad (1,2).

El análisis de evaluación económica proporciona a los decisores sanitarios, una base objetiva para la ayuda en la toma de decisiones. Así, el análisis de coste-efectividad es el cociente entre la diferencia de los costes y de la efectividad entre las alternativas. Cuando esta efectividad se mide en Años de Vida Ajustados por Calidad (AVACs), hablamos de análisis de coste-utilidad. Las personas encargadas de la toma de decisiones pueden considerar si los costes adicionales están justificados por la mejora de resultados sanitarios. Hay que tener en cuenta que existe una gran demanda de incluir entre las prestaciones nuevos tratamientos eficaces y seguros, sin perder de vista el limitado presupuesto sanitario, y en este sentido, la modelización desempeña un papel importante en la estimación del valor de los nuevos tratamientos.

La modelización solventa ciertos problemas como en aquellas situaciones de incertidumbre con información preliminar de variables. Debido a que los ensayos clínicos no suelen tener un seguimiento o

duración adecuada, se hace necesaria la realización de una extrapolación de resultados de los ensayos clínicos a horizontes temporales más largos. Hay que tener en cuenta que en los estudios no se suelen comparar todas las alternativas relevantes, por ejemplo, se compara la intervención con un grupo de control o con placebo, y no con la práctica clínica más habitual. Además, en los ensayos clínicos suelen faltar variables que resultan imprescindibles en la toma de decisiones (uso de recursos sanitarios, calidad de vida, duración de vida, etc.). Por último, hay que ser cuidadoso con su manipulación y con el riesgo de introducir datos o planteamientos sesgados (3).

Cuando se utilizan modelos de decisión para generar estimaciones tanto de los costes, como de la efectividad de las intervenciones, se pueden aplicar diferentes enfoques de modelización. Los factores que determinan la selección de un modelo específico son el objetivo del estudio, la naturaleza de la condición (su dinamismo) y la disponibilidad de datos (4).

En el capítulo 9 de este libro (modelos determinísticos basados en árboles de decisión (*deterministic decision tree models*) se han abordado las características básicas y la elaboración de un árbol de decisión. Un enfoque alternativo a los árboles de decisión son los modelos de Markov, que permiten la representación de determinados problemas de salud mediante un número recurrente de estados de salud (estados de Markov) a lo largo del tiempo. Los modelos de Markov fueron inicialmente teorizados por matemático ruso Andréi Markov (1856-1922), a principios del siglo xx (5).

El modelo de Markov debe diseñarse teniendo en cuenta la historia natural de la enfermedad. El modelo se estructura en torno a los estados de salud y los movimientos (o “transiciones”) entre ellos. El uso de los modelos de Markov está justificado cuando se trata de enfermedades que a) se repiten; b) son irreversibles; c) son de larga duración; d) cuyos estados de salud cambian en el tiempo; e) cuando algunos parámetros del modelo son tiempo-dependientes e importa el momento en el que

se inicia el evento; y, f) en enfermedades cuyos eventos se deben a la exposición a riesgos (e.g., una persona físicamente inactiva e hipertensa (factor de riesgo) sufre un evento cardiovascular) (6,7).

El modelo comprende un número finito de estados de salud que son mutuamente excluyentes (un paciente no puede estar en dos estados al mismo tiempo durante el mismo ciclo) y exhaustivos (tienen que incluir todos los estados posibles de la enfermedad en cuestión) y proporcionan una base para estimar diferentes trayectorias a través de estos estados para los pacientes que reciben diferentes intervenciones sanitarias. Hay dos tipos de estados de salud: a) absorbentes (el paciente tiene que permanecer en el mismo estado, e.g., la muerte), y b) no absorbentes (el paciente puede pasar a otro estado al final del ciclo). Los estados de salud deben abarcar a todos los individuos de la cohorte inicial, es decir, todas las personas deben estar encajadas en un estado de salud en cualquier momento. Conforme va evolucionando la enfermedad, los pacientes se mueven de un estado a otro en períodos discretos, fijos y uniformes (que se nominan ciclos de Markov y que pueden ser semanas, meses, años) (6) en función de las probabilidades de transición que son específicas para cada estado de salud y cada ciclo. La duración de los ciclos debe ser el intervalo de tiempo más corto clínicamente significativo, y debe ser constante a lo largo de la simulación. Al final de cada ciclo, el paciente puede permanecer en el mismo estado de salud o pasar a otro (8). Cabe destacar que sólo se permiten unas determinadas transiciones entre estados previamente establecidas. Cada paciente sólo puede hacer una transición en cada ciclo. Todos los pacientes en un estado de salud determinado y en un ciclo concreto están sometidos a las mismas probabilidades de transición (7).

Aunque, la representación mediante los diagramas de Markov no se utiliza para hacer los cálculos, visualmente resultan muy útiles, ya que nos ayuda a comprender la trayectoria y cómo las personas o pacientes transitan entre los diferentes estados de salud. La transición entre los estados de salud se representa mediante las flechas que conectan los

estados y de esta manera podremos comprender los posibles recorridos de los pacientes. Cuando pensamos en el recorrido del paciente, tenemos que pensar en el tiempo que necesitamos para seguir un proceso, el tiempo que las personas permanecen en cada estado, la frecuencia con la que las personas permanecen en cada estado y los costes y resultados asociados a estar en cada uno de ellos (9).

En la Figura 1 se representa un caso hipotético de un proceso que tiene tres posibles estados de salud, “Sin evento”, “Con evento” y “Muerte”. En la Tabla 1, se muestran los valores de las probabilidades de transición que representan cómo se realiza el paso de los pacientes entre los diferentes estados de salud. Hay que tener en cuenta que, por tratarse de probabilidades, deben tener valores que oscilan entre cero y uno, y que los valores de probabilidad de cada fila deben sumar uno.

En primer lugar, se asume que todos los pacientes comienzan en el mismo estado de salud al inicio del seguimiento, es decir, en el periodo de tiempo cero (t_0) y en el estado de salud “Sin evento”. Posteriormente, en el primer periodo de tiempo (t_1), los pacientes pueden sufrir un evento o pueden fallecer, sin pasar por el estado evento (directamente). Lógicamente, una vez que el paciente ha llegado al estado de muerte, ya no puede salir de ahí (estado absorbente). Supongamos que tenemos un total de 100 pacientes. En el primer punto temporal 1 (t_1), vemos que, de los 100 pacientes iniciales, el 65% de los pacientes permanecen en el estado ‘sin evento’, el 20% de los pacientes pasan al estado ‘con evento’ tras sufrir un evento y el 15% de pacientes han muerto. El segundo periodo (t_2), en el estado de salud ‘sin evento’ estará el 65% de los pacientes del ciclo t_1 . En el estado de salud ‘con evento’ estará el 70% de los pacientes del estado ‘con evento’ del ciclo t_1 más el 20% de los pacientes del estado ‘sin evento’ del ciclo t_1 . El estado de salud ‘muerte’ tendrá el 15% de pacientes del estado ‘sin evento’ del ciclo t_1 , el 30% de pacientes del estado ‘con evento’ del ciclo t_1 más los pacientes muertos del ciclo t_1 . De esta manera, en cada periodo de tiempo (t_i),

hay una nueva distribución de los pacientes en los distintos estados de salud (10).



Figura 1. Ejemplo de diagrama de Markov con tres estados de salud



Tabla 1. Tabla de probabilidades de transición entre los diferentes estados de salud durante cada ciclo

Desde	a	Sin evento	Con evento	Muerte
Sin evento		0,65	0,20	0,15
Con evento		0	0,70	0,30
Muerte		0	0	0

En la Tabla 2, se muestra un ejemplo del cálculo del modelo de Markov para los pacientes con el tratamiento habitual (es decir, sin la intervención nueva). Comenzando en el t_1 , los 65 pacientes en el estado “sin evento” tienen una probabilidad de permanecer en el estado sin evento del 0,65, lo que significa que 42 pacientes permanecen en este estado ($65 \times 0,65 = 42,25$). Igualmente, un total de 27 pacientes están en el estado de “con evento” que provienen de aquellos que pasan desde “sin evento” a “con evento” ($65 \times 0,20 = 13$) más aquellos que permanecen en el estado de “con evento” ($20 \times 0,70 = 14$), lo que nos da el total de 27 pacientes ($13 + 14 = 27$). Por último, en el segundo ciclo hay un total de 31 personas fallecidas, que provienen de las 10 personas que desde el estado “sin evento” fallecen ($65 \times 0,15 = 9,75$), más las 6 personas que provienen del estado “con evento” ($20 \times 0,3 = 6$), más las 15 personas que estaban muertas en el ciclo anterior t_1 (15 personas), es decir ($9,75 + 6 + 15 = 30,75$).

Tabla 2. Ejemplo del cálculo del modelo de Markov para un grupo de tratamiento dado

Tiempo (t)	Sin evento	Con evento	Muerte	Total pacientes
0	100	0	0	0
1	65	20,00	15	100
2	42,25	27,00	30,75	100

Llegados a este punto debemos tener en cuenta dos aspectos relacionados con el tiempo. En primer lugar, el tiempo utilizado para el periodo debe ser lo suficientemente pequeño para poder reflejar con detalle todos los cambios que se vayan produciendo, pero el uso de periodos pequeños requiere de un mayor uso de recursos para realizar sus cálculos.

los, pudiendo no ser eficiente utilizar periodos cortos (días o semanas). Por ejemplo, la utilización de periodos de meses cuando los pacientes viven años o de años para los pacientes que viven sólo unos meses tampoco sería sensato. Así, la elección del periodo dependerá de la enfermedad analizada.

Como puede verse en el ejemplo anterior, los modelos de Markov pueden realizarse perfectamente en una hoja de cálculo (Excel o LibreOffice), aunque también hay programas gratuitos que pueden resultar de gran utilidad como OpenMarkov (disponible en: <http://www.openmarkov.org/>; y su tutorial en <http://www.openmarkov.org/docs/tutorial/openmarkov-tutorial.pdf>).

Cabe destacar que, en el ejemplo anterior, todos los pacientes comenzaron en el estado de “sin evento”, pero una variante sería aquel donde una proporción de pacientes pueden comenzar desde el estado “con evento” u otros estados, diferentes de la muerte, que tenga el modelo. De esta manera, los pacientes que comiencen desde el estado “con evento” seguirán su trayectoria de transición por el modelo desde el momento t_0 hacia adelante.

Hay distintos métodos para la estimación del modelo para los pacientes con la nueva intervención. Lo más común es utilizar un modelo con los mismos estados de salud, pero con probabilidades de transición diferentes, que reflejan la efectividad de la intervención para cambiar la evolución de la enfermedad, paliar los síntomas, o efectuar una cura. Así, el ejemplo de la enfermedad descrita en la Figura 1 y la Tabla 1, desde el estado de “sin evento”, la probabilidad de un evento no fatal durante un ciclo podría ser 0,10 y la probabilidad de morir durante un ciclo podría ser 0,05, así que la probabilidad de permanecer en el estado de “sin evento” sería $1 - (0,1 + 0,05) = 0,85$. En este ejemplo, la nueva intervención no cambiaría la probabilidad de morir de una persona con un evento, que sigue siendo 0,3.

Las probabilidades de transición pueden extraerse de estudios de intervención (ensayos clínicos o cuasiexperimentales), estudios

observacionales o de registros. Igualmente, se pueden utilizar aquellos datos que procedan de aquellos documentos que sinteticen la información de diferentes estudios, pudiéndose utilizar la literatura publicada y/o la opinión de los expertos. Es importante tener en cuenta que las probabilidades de transición condicionan la forma en la que los pacientes van transitando por los diferentes estados del modelo. Por ello, la forma de estimar las probabilidades de transición es un elemento crucial en el desarrollo de la evaluación económica y condicionará enormemente los resultados (11).

Los modelos de Markov tienen de base una suposición bastante restrictiva, la denominada suposición markoviana. En esta suposición, el modelo carece de «memoria» y se asume que, conociendo únicamente el estado de salud presente de un paciente, es posible simular la evolución futura del mismo; es decir, se asume que todos los pacientes que se encuentran en un estado de salud, en un momento concreto, tienen el mismo pronóstico, independientemente de su estado previo. Por tanto, un paciente que pasa del estado “sano” al estado de “enfermo” tendría el mismo pronóstico que otro paciente “enfermo” que, transcurrido el ciclo, siguiese en el mismo estado de “enfermo”. Esta suposición del modelo no se ajusta estrictamente a la realidad de numerosas enfermedades, por lo que supone un inconveniente de los modelos de Markov (12).

Un ejemplo, podría verse cuando estamos evaluando una enfermedad que puede estar condicionada por la aparición previa de otra. Para ello sería recomendable incluir diferentes estados de salud que permitiesen capturar combinaciones importantes de enfermedades. Esto conlleva un problema, ya que al incluir demasiados estados de salud en el modelo de Markov, puede hacerlo inmanejable y estar sometido a una mayor incertidumbre. Un ejemplo puede ser el estimar probabilidades de transición para las dos enfermedades que están condicionales a la presencia o no de la otra enfermedad (13).

En segundo lugar, cuando las entradas del modelo (como las probabilidades de transición, las utilidades y los costes) evolucionan con el tiempo. Así, cuando la edad de los sujetos del modelo aumenta con los ciclos repetidos y, por consiguiente, sus riesgos de enfermedades aumentarán. Para ello, si hay datos disponibles, se pueden incluir entradas específicas del ciclo para los individuos en el primer estado de salud (“sin eventos” en este ejemplo), siendo necesario calcular los cambios relacionados con la edad en los riesgos de enfermedad. Sin embargo, debida a la “suposición markoviana” las probabilidades de transición para el segundo y siguientes estados tienen que ser constantes en el tiempo.

El efecto de la suposición markoviana se puede minimizar introduciendo estados de túnel temporales, que no son estados en sí mismos, pero sí sirven para contemplar el cambio temporal en el estado de los pacientes (por ejemplo, los pacientes que sufren un infarto agudo de miocardio tienen peor calidad de vida en los doce meses siguientes) (14).



Asignación de la utilidad y de los costes a los estados descritos

Supongamos que la calidad de vida de todos los pacientes es igual a 1 (salud perfecta). En este caso, el resultado en salud del modelo sería los años de vida ganados. Si, por el contrario, suponemos que la utilidad de los pacientes depende del estado de salud en el que se encuentran, incluso varía con el tiempo, el resultado en salud del modelo serán los años de vida ajustado por calidad.

Para calcular los costes del modelo, primero se calculan los costes de cada estado, luego se suman los costes de todos los estados del ciclo 1, ciclo 2, etc. para al final sumar los costes de todos los ciclos.

No obstante, hay que hacer dos ajustes a los costes y resultados en salud. El primer tipo de ajuste es el ajuste temporal, y que consiste en aplicarle, tanto a los costes como a los resultados en salud, una tasa de descuento oficial del 3%, cuyo rango oscila entre 0% y 5% para los análisis de sensibilidad (15).

El segundo tipo de ajuste se denomina 'la corrección de la mitad del ciclo'. Este tipo de ajuste es necesario realizarlo porque los modelos de Markov asumen que todos los pacientes se mueven de un estado a otro al comienzo o al final de cada ciclo. No obstante, en realidad los pacientes se pueden mover de un estado a otro en cualquier momento del ciclo. Con esta corrección se asume que, de media, los pacientes se moverán de un estado a otro a mitad del ciclo y se minimiza el sesgo de los resultados en salud. Es importante realizarla si el objetivo de los analistas es reportar los costes y los resultados en salud medios. En cambio, si el objetivo es calcular los costes y resultados en salud incrementales, éstos no deben verse afectados por no realizar la corrección de la mitad del ciclo (14).



Consideraciones con la estimación de las probabilidades de transición

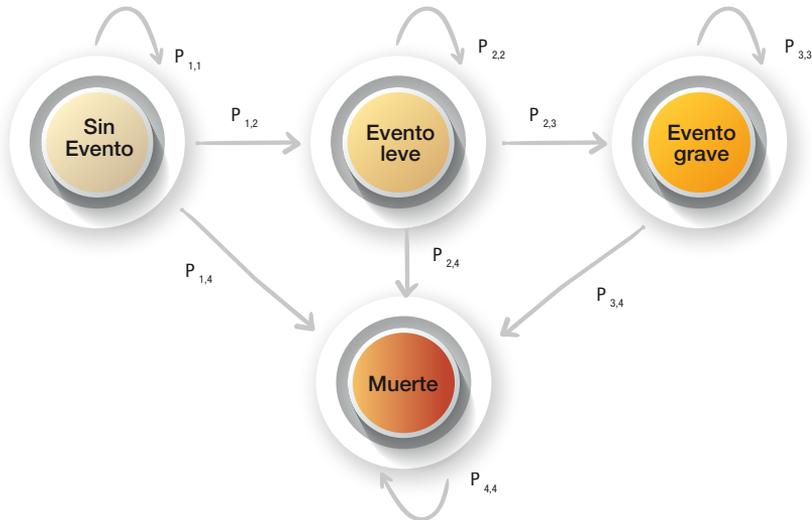
Los movimientos entre estados de salud se denominan transiciones, y la velocidad a la que se producen estas transiciones, ocurren como probabilidades de transición. Los valores de los estados se utilizan para reflejar los valores tanto de los costes, como de la medida de efectividad o de la calidad de vida relacionada con la salud. Ambos valores dependerán de estas probabilidades de transición, ya que condicionan el número de pacientes en cada estado de salud.

Para ilustrar las consideraciones con las probabilidades de transición entre estados, usaremos un ejemplo (representado mediante la Figura 2) con un modelo de Markov con cuatro estados de salud: "Sin evento"

y “con evento” se divide en “con evento leve” y “con evento grave” y por último el estado “muerte”. Sólo son posibles las transiciones hacia adelante, es decir, no es posible una mejoría en el estado de salud (11).



Figura 2. Ejemplo de modelo de transición de estado y probabilidades de transición permitidas



$P_{m,n}(t)$ es una probabilidad de transición, que describe la probabilidad de pasar del estado m al estado n (estado diferente a m , en el tiempo t).
 $P_{m,m}(t)$ es la probabilidad de permanecer en el estado m en el tiempo t .

- **Fuentes de datos:** en la estimación de las probabilidades de transición puede que sea necesario utilizar fuentes de datos, que pueden tener cierta heterogeneidad en diversos aspectos como: seguimiento, diseño del estudio (por ejemplo, ensayo clínico controlado y aleatorizado, estudios observacionales y registros). La utilización de diferentes fuentes de datos puede suponer un reto, al tener que combinar estadísticamente estos datos o usarlos directamente. Por ejemplo,

imaginemos que la probabilidad de transición $P_{2,3}$ la obtenemos que un estudio observacional. Sin embargo, el resto de las probabilidades de transición, que alimenta el modelo, proviene de un ensayo clínico que presenta unas características basales de los pacientes incluidos diferentes (por ejemplo, en edad).

- **Falta de transiciones:** Imaginemos que tenemos información sobre los valores de unas probabilidades de transición, pero no de otras, sea por la exclusión de pacientes con una determinada gravedad, censura, pérdida de seguimiento o porque ciertas transiciones son raras. Por ejemplo, puede haber datos disponibles sobre $p_{1,1}$ y $p_{2,3}$, pero debido a que los pacientes del ensayo fueron censurados, no está disponible información sobre las probabilidades $p_{1,3}$ y $p_{3,4}$. En estos casos, podría ser necesario realizar una síntesis de múltiples fuentes para estimar estas probabilidades. Por otro lado, es posible que no siempre se observen transiciones factibles en todos los conjuntos de datos, lo que implica que la probabilidad de que ocurra una determinada transición sea del 0%. La literatura recomienda no excluir posibles transiciones porque no se observaron en los datos (16,17).
- **Ausencia de datos:** un caso particular de la consideración anterior sería cuando, aun teniendo información para todas las probabilidades de transición ($P_{1,n}$ a $P_{3,n}$), el número de sujetos a utilizar para calcular una determinada probabilidad de transición puede ser muy escaso, haciendo que estas estimaciones pueden no ser generalizables a toda la población. En estas situaciones, es necesario hacer suposiciones sobre la ausencia de datos para evitar sesgos; teniendo en cuenta que diferentes suposiciones conducen a diferentes estimaciones de las probabilidades de transición.
- **Datos sobre subgrupos no disponibles:** imaginemos que teniendo información sobre todas las probabilidades de transición ($P_{1,n}$ a $P_{3,n}$), no tenemos información para determinados subgrupos de sujetos. Asumir que las probabilidades de transición son equivalentes entre los diferentes subgrupos y la población general puede suponer un sesgo.

Igualmente, de forma inversa, usar la información disponible de determinados subgrupos, puede generar unas estimaciones sesgadas por utilizar tamaños de muestra pequeños, transiciones faltantes, etc. Una estrategia puede ser realizar una imputación múltiple de datos (18).

- **Necesidad de extrapolación:** supongamos que las probabilidades de transición ($p_{1,n}$ a $p_{3,n}$) fueron obtenidas de un ensayo clínico que, por diferentes motivos, se detuvo antes de que todos los pacientes experimentaran el evento de interés. En este caso, se hace necesario extrapolar las probabilidades de transición más allá del periodo observado, lo que implica hacer suposiciones sobre los efectos relativos de las alternativas en momentos del tiempo posteriores. La recomendación habitual es que el horizonte temporal debe ser lo suficientemente largo para tener en cuenta todos los resultados de salud significativos, lo que a menudo requiere modelar largos horizontes temporales (19).
- **Intervalos largos entre evaluaciones:** es posible que haya cierto desajuste entre el tiempo entre las mediciones a los pacientes y el tiempo de las probabilidades de transición del modelo. Por ejemplo, el estudio tiene planificado que se realizan mediciones de seguimiento a los pacientes de hasta 6 meses. Sin embargo, podemos tener una falta de precisión en las estimaciones del modelo al utilizar ciclos de 6 meses, pudiendo ser necesario la adopción de ciclos más cortos.
- **Incongruencia de los datos:** supongamos que tenemos información que nos permite calcular las probabilidades en el tiempo hasta un determinado evento (por ejemplo, tiempo hasta el estado de muerte $p_{1,4}$, $p_{2,4}$ y $p_{3,4}$), pero no tenemos información sobre las probabilidades de transición para el resto de estado de salud, es decir todas las probabilidades de transición $p_{1,n}$ a $p_{4,n}$). Para estimar las tasas de transición subyacentes, tendríamos que utilizar la información que tenemos disponible (u observada) para estimar esa información, dada la estructura del modelo, por ejemplo, mediante la calibración del modelo.



Medidas de frecuencia

Antes de nada, es importante señalar la diferencia entre tasa y probabilidad. Así, una tasa es una medida de “velocidad” y describe el número de veces que sucede un evento para un número determinado de pacientes por unidad de tiempo, y oscila entre cero e infinito. Por el contrario, la probabilidad describe la posibilidad de que se produzca un acontecimiento para un solo individuo en un periodo de tiempo determinado, oscilando entre 0 y 1. La diferencia entre ambas medidas es el rol del tiempo.

La tasa de incidencia es la velocidad de aparición de casos nuevos en la población en riesgo. Se calcula dividiendo los casos nuevos de una enfermedad por la población en riesgo (“población expuesta”), teniendo en cuenta cuánto tiempo se ha seguido a esta población. Así, el denominador es la suma de cada tiempo a riesgo de los diferentes individuos. Por el contrario, para el cálculo de la probabilidad no se incluye el tiempo.

Para la conversión entre tasa y probabilidad, si el evento ocurre siguiendo una tasa constante r por unidad de tiempo t , entonces la probabilidad de que un evento ocurra durante el tiempo t viene dada por (nótese que la unidad de tiempo utilizada en r y t debe ser la misma) (20).

$$p = 1 - e^{-rt}$$

Por el contrario, si se tiene una probabilidad y se quiere obtener la tasa se haría siguiendo la siguiente fórmula, donde p es la probabilidad, r es la tasa y t es la unidad de tiempo.

$$r = \frac{1}{t} \ln(1 - p)$$

Por ejemplo, asumamos una cohorte con 100 pacientes con un seguimiento de 3 años. Durante el seguimiento un total de 70 pacientes experimentaron el evento de interés. Así, usando la ecuación 2, la

probabilidad de 1 año se puede determinar directamente a partir de la probabilidad del 70% de 3 años.

$$r = \frac{1}{3} \ln(1 - 0,7) = 0,4013$$

Al traducir esta tasa a una probabilidad de transición usando la Ecuación 1, se obtiene el resultado de 0,3306, que nos arroja un total de 70 casos incidentes durante los 3 años.

$$p = 1 - e^{-0,4013 \times 1} = 0,3306$$

Las fórmulas y procedimientos anteriores no se ajustan adecuadamente cuando, en un ciclo, pueden ocurrir tres o más transiciones entre estados; ya que se puede sobreestimar e infraestimar las probabilidades de transición, debido al riesgo competitivo entre los estados (21). Una de las posibles soluciones sería revisar la estructura del modelo, para que cada nodo tenga sólo dos transiciones. Cuando esta descomposición no sea posible, otra opción sería la utilización de la descomposición de la probabilidad de transición en vectores y valores de Eigen (22). Para ello, se debe tener una única fuente de datos para informar las transiciones desde un estado de salud inicial, o, aunque es raro, tener datos de múltiples estudios que tengan el mismo tiempo de seguimiento. Por ejemplo, un modelo de una patología concreta donde la eficacia del tratamiento y los efectos secundarios se obtienen de un único ensayo clínico, mientras que los datos de mortalidad por todas las causas se obtienen de otra fuente, ya que el ensayo tuvo un periodo de seguimiento corto. Cuando el periodo de seguimiento es diferente, no puede aplicarse este enfoque de descomposición propia. Igualmente, este método de descomposición no es aplicable cuando en el modelo se realizan análisis de subgrupos, con múltiples fuentes sobre las probabilidades de transición. Por último, en aquellas situaciones donde: 1) Sólo hay tres transiciones de estado de salud posibles, 2) Dos de las probabilidades publicadas

son muy pequeñas, y 3) La duración del ciclo del modelo es más corta que la probabilidad publicada, el error al usar la fórmula para dos estados es pequeño. Para más información recomendamos la lectura del artículo de Jones E (23).



Cálculos de las probabilidades de transición (24)

Una de las medidas de asociación más utilizadas es el riesgo relativo (RR) que es el cociente entre la probabilidad del evento entre las personas expuestas a una intervención o condición (p_1) y la probabilidad en aquellas personas no expuestas (p_0) y su rango oscila entre 0 e infinito.

$$RR = \frac{p_1}{p_0} \quad (\text{Ecuación n}^\circ 1)$$

Teniendo en cuenta la fórmula del riesgo relativo, si multiplicamos el valor del riesgo relativo por la probabilidad del evento entre aquellas personas no expuestas (p_0), entonces se anula el denominador, teniendo como resultado la probabilidad del evento entre los expuestos (p_1).

$$p_1 = RR \times p_0 = \left(\frac{p_1}{p_0}\right) \times p_0 \quad (\text{Ecuación n}^\circ 2)$$



Uso del riesgo relativo para determinar transiciones en subgrupos

En determinadas ocasiones, cuando la probabilidad del evento entre los no expuestos no está disponible es necesario acudir a la literatura. En este caso, es necesario valorar que ambas probabilidades provengan de poblaciones similares y sean válidas y aplicables a la población

modelizada. Por otro lado, cuando el riesgo relativo disponible esté ajustado por diferentes variables, la probabilidad del evento en el grupo, en el denominador de la fórmula no se elimina. En esta situación, por falta de datos, se puede utilizar la ecuación 3, teniendo en cuenta que existe cierto grado de error en la estimación de la probabilidad p_1 . Es recomendable realizar un análisis de sensibilidad para probar un rango plausible de valores de la p_1 y determinar la importancia de este error.

$$p_1 \approx \left(\frac{p_1 \text{ ajustada}}{p_0 \text{ ajustada}} \right) \times p_0 \text{ No ajustada} \quad (\text{Ecuación n}^\circ 3)$$

En general, las probabilidades disponibles suelen hacer referencia a la población general, pero no para los subgrupos. Los riesgos relativos informan sobre la probabilidad de población que es un promedio ponderado de las probabilidades de los diferentes subgrupos. Supongamos una situación, donde queremos saber la mortalidad según diferentes grupos de edad por la infección por COVID-19.

$$pm_{\text{total}} = (p_0 \times pm_0) + (p_1 \times pm_1) + (p_{23} \times pm_{23})$$

pm_{total} es la probabilidad conocida de morir por la COVID-19 para el grupo de edad en su conjunto. Así, las probabilidades de morir por la COVID-19 para el grupo de edad 1 que no recibieron la vacuna (p_0), una dosis (p_1) o dos o tres dosis de vacuna (p_{23}). Multiplicando el lado derecho por pm_0/pm_0 permite reformular la ecuación en términos de los RR de muerte por estado de vacunación:

$$pm_{\text{total}} = \left(p_0 \times pm_0 + \left(p_1 \times pm_0 \times \left(\frac{pm_1}{pm_0} \right) \right) + \left(p_{23} \times pm_0 \times \left(\frac{pm_{23}}{pm_0} \right) \right) \right)$$

o

$$pm_{\text{total}} = pm_0 \times (p_0 + (p_1 \times RR_1) + (p_{23} \times RR_{23}))$$

Los riesgos relativos RR_1 y RR_{23} y las probabilidades p_0 , p_1 y p_{23} se obtienen de la literatura y utilizando la ecuación 2 se pueden obtener las probabilidades pm_1 y pm_{23} .



Riesgo relativo y Odds Ratio

La odds es la probabilidad de que ocurra un evento (p) dividido por la probabilidad de que no ocurra ese evento y tiene un rango entre 0 e infinito. La fórmula de la odds es

$$Odds = \frac{p}{(1 - p)}$$

Para obtener las probabilidades a partir de la odds ratio (OR), hay que tener en cuenta dos situaciones, En la primera de ellas, si el evento de interés es raro ($<10\%$) y/o la OR tiene un valor cercano a 1. En estos casos, la OR es una aproximación razonable al RR, y puede utilizarse directamente la ecuación nº 2. En el resto de caso se puede utilizar la siguiente ecuación, donde la p_0 es la probabilidad del evento en el grupo no expuesto.

$$RR = \frac{OR}{(1 - p_0 + (p_0 \times OR))}$$

Mediante esta fórmula, a medida que aumenta la probabilidad del evento en el grupo no expuesto, el valor de la OR se aleja del RR. A menudo, la estimación de la OR proviene de una regresión logística multivariable, que variará su valor dependiendo de las variables incluidas en el modelo multivariante, y lógicamente de la base de datos utilizada. Por ello, es recomendable realizar un análisis de sensibilidad con el rango de valores del RR, para determinar su influencia en los resultados del modelo.

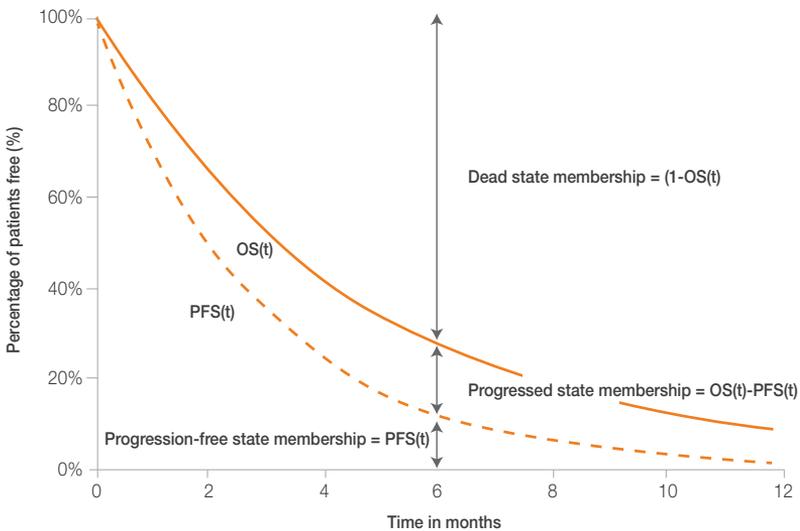


El modelo de supervivencia dividido (Partitioned-survival analysis)

En los últimos años, los modelos de supervivencia dividido (*Partitioned-survival model*) se han convertido en el método de modelización más utilizados para evaluar medicamentos en oncología (25). El modelo de supervivencia dividido simula la probabilidad de que un paciente se encuentre, en un momento determinado, en cada uno de los diferentes estados de salud cuando es tratado con una terapia concreta. La pertenencia a un estado se estima a partir de un conjunto de curvas de supervivencia no mutuamente excluyentes. Así, los modelos de supervivencia dividido no utilizan transiciones entre estados para determinar la proporción de



Figura 3. Determinación de la membresía estatal en un modelo de supervivencia dividido en 3 estados. OS(t) denota la curva OS en el tiempo t, y PFS(t) la curva de supervivencia libre de progresión en el tiempo t



OS indicates overall survival; PFS, progression-free survival

Tomada de Woods 2020 (26)

pacientes en cada estado de salud en cada punto (pertenencia estado). Destacar que, para un modelo que tenga N estados de salud, se necesitan $N-1$ curvas de supervivencia.

La supervivencia de supervivencia a lo largo del tiempo se suele representar como una curva de Kaplan-Meier que traza la proporción de pacientes vivos, muertos o que han progresado (según el resultado) a lo largo del tiempo (26,27) (Figura 3). En este sentido, si nos fijamos en la Figura 1, tenemos dos curvas de supervivencia:

La primera de ellas es la supervivencia global que se calcula desde el momento de la aleatorización (o primera dosis de tratamiento) hasta la muerte (28). Así, la función de supervivencia $OS(t)$ es la probabilidad de que un individuo sobreviva desde la fecha de entrada en el estudio hasta un momento determinado en el tiempo t y determina la proporción de pacientes vivos (y muertos) en cada momento del tiempo.

La otra curva de supervivencia es la supervivencia libre de progresión que se calcula desde el momento de la aleatorización (o la primera dosis recibida) hasta la progresión o la muerte. Es decir, la curva de supervivencia libre de progresión describe el tiempo desde la entrada en el estudio hasta la salida del estado libre de progresión a través de la progresión o la muerte y proporciona la pertenencia al estado libre de progresión a lo largo del tiempo. El cálculo de la pertenencia al estado de salud con progresión, es la diferencia entre la Supervivencia Global y la curva de Supervivencia Libre de Progresión en cada punto, porque representa la proporción de pacientes que están vivos y no han progresado.

El área bajo la curva de la supervivencia libre de progresión representa la proporción de pacientes que aún no han progresado, y el área entre las curvas de la supervivencia global y la supervivencia libre de progresión representa la proporción de pacientes que han progresado y siguen vivos. La proporción restante de pacientes (es decir, los que no han progresado ni van a progresar) representa a los que han muerto (25).

En los modelos de supervivencia dividido, la proporción de pacientes en los tres estados en cada punto de tiempo se determina mediante cur-

vas de supervivencia, que trazan el número de eventos predefinidos que han ocurrido a lo largo del seguimiento. El modelo de supervivencia dividido utiliza el área bajo la curva para representar el número de pacientes en cada estado (29). Esto implica una asunción inherente de la estructura del modelo, ya que la progresión de la enfermedad es irreversible; lo pacientes pueden progresar desde el estado de libre de progresión al estado con progresión, pero no al revés. Así, las curvas de supervivencia que informan las estimaciones de pertenencia a un estado (por ejemplo, la Supervivencia Libre de Progresión y la Supervivencia Global) se modelan de forma completamente independiente; siendo esta la diferencia fundamental con el modelo de Markov. El hecho de modelar ambas curvas de manera independiente puede hacer que se obtengan resultados que sean imposibles, como por ejemplo que se crucen las curvas de supervivencia global y supervivencia libre de progresión. Posteriormente, los AVAC y los costes totales asociados a un determinado estado de salud de una estrategia de tratamiento se calculan integrando la probabilidad de supervivencia.

En la extrapolación de las curvas de supervivencia a largo plazo, es decir, estimar la trayectoria de supervivencia una vez finalizado el periodo de seguimiento del ensayo, se utilizará la evidencia externa, junto a los parámetros de ajuste estadístico y la opinión de expertos para evaluar la plausibilidad de las extrapolaciones. Destacar que, si todos los pacientes fallecen durante el seguimiento del ensayo, no será necesaria la extrapolación.



Diferencias entre el modelo de Markov y los modelos de supervivencia dividida

En general, los resultados obtenidos mediante ambos abordajes son similares (29,31), aunque diferentes asunciones en ambos abordajes pueden llevar a diferencias en el Ratio coste-efectividad incremental (RCEI) que puede afectar a la toma de decisiones (30,32). La estructura de tres

estados ha sido el más utilizado, tanto para el modelo de supervivencia dividida, como para el modelo Markov. En cuanto a los datos, el modelo de supervivencia dividida necesita los datos individuales de los pacientes o su reconstrucción, mientras que el modelo de Markov requiere de las probabilidades de transición. Destacar que, en aquellas situaciones más complicadas, como recidiva y la metástasis, el modelo de Markov se muestra como una aproximación más apropiada.

El modelo de Markov es un modelo de transición de estado discreto que se utiliza para simular y extrapolar el progreso de eventos aleatorios que cambian con el tiempo. Por el contrario, el modelo de supervivencia dividida utiliza el área bajo la curva para representar el número de pacientes en cada estado. Se utiliza principalmente para evaluar el impacto de las intervenciones que pueden prolongar la vida y la calidad de vida del paciente (29).

El hecho de que el modelo de Markov sea más flexible, es la principal ventaja, ya que se pueden añadir más estados de salud según necesidad. Por el contrario, la desventaja que presenta el modelo de Markov es que, requiere más asunciones y supuestos para la construcción del modelo, cuando los datos son limitados. Igualmente, añadir muchos estados de salud puede hacer el modelo muy complejo. La mayor ventaja del modelo el modelo de supervivencia dividida es su sencillez, y que no es necesario hacer suposiciones. Por el contrario, la desventaja es que se suelen construir con menos estados de salud. Además, el modelo utiliza las curvas de supervivencia libre de progresión y supervivencia global que proporcionarían directamente la información de los datos, y por ello no resulta tan flexible como el modelo de Markov. Por último, en la literatura científica se sugiere que el modelo de supervivencia dividida presenta sesgos inherentes a favor de la intervención que tiene impacto sobre la estimación de la progresión de la enfermedad.



Bibliografía

1. McManus E, Sach TH, Levell NJ. An introduction to the methods of decision-analytic modelling used in economic evaluations for Dermatologists. *J Eur Acad Dermatol Venereol*. 2019;33:1829-36.
2. Campillo-Artero C, Ortún V. El análisis coste-efectividad: por qué y cómo. *Rev Esp Cardiol*. 2016;69:370-3.
3. Soto Álvarez J. Estudios de farmacoeconomía: ¿Por qué, cómo, cuándo y para qué? *Medifam*. 2001;11:147-55.
4. Gold MR, Siegel JE, Russell LB, et al. (Eds). Assessing the effectiveness of Health Interventions. In *Cost-Effectiveness in Health and Medicine*. New York Oxford University Press, 1996; p. 151-163
5. Basharin GP, Langville AN, Naumov VA. The life and work of A.A. Markov. *Linear Algebra Appl*. 2004;386:3-26.
6. Mar J, Antoñanzas F, Pradas R, et al. A. Los modelos de Markov probabilísticos en la evaluación económica de tecnologías sanitarias: una guía práctica. *Gac Sanit*. 2010;24:204-19.
7. Rubio-Terrés C, Echevarría A. Modelos de Markov: una herramienta útil para el análisis farmacoeconómico. *Pharmacoeconomics – Spanish Research Articles*. 2006;3:71-8.
8. Castillo-Riquelme M. El uso de modelos matemáticos en evaluación económica de intervenciones de salud. *Rev Med Chile*. 2010;138 (Supl 2): 98-102.
9. Edlin R, McCabe C, Hulme C, et al. *Cost Effectiveness Modelling for Health Technology Assessment*. London: Springer International Publishing; 2019.
10. Wilson ECF. Methodological Note: Reporting Deterministic versus Probabilistic Results of Markov, Partitioned Survival and Other Non-Linear Models. *Appl Health Econ Health Policy*. 2021;19:789-95.
11. Srivastava T, Latimer NR, Tappenden P. Estimation of Transition Probabilities for State-Transition Models: A Review of NICE Appraisals. *Pharmacoeconomics*. 2021;39:869-78.

12. Beck JR, Panker SG. The Markov process in medical prognosis. *Med Decis Making*. 1983;3:419-58.
13. Ademi Z, Kim H, Zomer E, et al. Overview of pharmacoeconomic modelling methods. *Br J Clin Pharmacol*. 2013;75:944-50.
14. Briggs A, Sculpher M. An Introduction to Markov Modelling for Economic Evaluation. *PharmacoEconomics*. 1998;13:397-409.
15. Ortega Eslava A, Marin Gil R, Fraga Fuentes MD, et al. (GENESIS-SEFH). Guía de evaluación económica e impacto presupuestario en los informes de evaluación de medicamentos. Guía práctica asociada al programa MADRE v 4.0 . Madrid: SEFH (ed). 2016. Disponible en: https://gruposedetrabajo.sefh.es/genesis/genesis/Documents/GUIA_EE_IP_GENESIS-SEFH_19_01_2017.pdf (acceso verificado el 25 Sept 2023)
16. 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.
17. Olariu E, Cadwell KK, Hancock E, et al. Current recommendations on the estimation of transition probabilities in Markov cohort models for use in health care decision-making: a targeted literature review. *Clinicoecon Outcomes Res*. 2017;9:537-46.
18. Gomes M, Díaz-Ordaz K, Grieve R, et al. Multiple imputation methods for handling missing data in cost-effectiveness analyses that use data from hierarchical studies an application to cluster randomized trials. *Med Decis Making*. 2013;33:1051-63.
19. Siebert U, Alagoz O, Bayoumi AM, et al. State-transition modeling: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-3. *Value Health*. 2012;15:812-20.
20. Fleurence RL, Hollenbeak CS. Rates and probabilities in economic modelling: transformation, translation and appropriate application. *PharmacoEconomics*. 2007;25:3-6.
21. Chhatwal J, Jayasuriya S, Elbasha EH. Changing cycle lengths in state-transition models: challenges and solutions. *Med Decis Making*. 2016;36:952-64.

22. Jung J. Estimating Markov Transition Probabilities Between Health States Using U.S. Longitudinal Survey Data. [consultado 17 febrero 2022]. https://web.archive.org/web/20210428052356id_/http://webapps.towson.edu/cbe/economics/workingpapers/2020-06.pdf (acceso verificado el 25 Sept 2023)
23. Jones E, Epstein D, Garcia-Mochon L. A procedure for deriving formulas to convert transition rates to probabilities for multistate Markov models. *Med Decis Making*. 2017;37:779–89.
24. Gidwani R, Russell LB. Estimating Transition Probabilities from Published Evidence: A Tutorial for Decision Modelers. *PharmacoEconomics*. 2020; 38:1153-64.
25. Bullement A, Cranmer HL, Shields GE. A Review of Recent Decision-Analytic Models Used to Evaluate the Economic Value of Cancer Treatments. *Appl Health Econ Health Policy*. 2019;17:771-80.
26. Woods BS, Sideris E, Palmer S, et al. Partitioned Survival and State Transition Models for Healthcare Decision Making in Oncology: Where Are We Now? *Value Health*. 2020;23:1613-21.
27. Ranstam J, Cook JA. Kaplan-Meier curve. *Br J Surg*. 2017;104:442.
28. Rebasea P. Conceptos básicos del análisis de supervivencia. *Cir Esp*. 2005; 78:222-30.
29. Rui M, Wang Y, Fei Z, et al. Will the Markov model and partitioned survival model lead to different results? A review of recent economic evidence of cancer treatments. *Expert Rev Pharmacoecon Outcomes Res*. 2021;21: 373-80.
30. Smare C, Lakhdari K, Doan J, et al. Evaluating Partitioned Survival and Markov Decision-Analytic Modeling Approaches for Use in Cost-Effectiveness Analysis: Estimating and Comparing Survival Outcomes. *PharmacoEconomics*. 2020;38:97-108.
31. Goeree R, Villeneuve J, Goeree J, et al. Economic evaluation of nivolumab for the treatment of second-line advanced squamous NSCLC in Canada: a comparison of modeling approaches to estimate and extrapolate survival outcomes. *J Med Econ*. 2016;19:630–44.

32. Williams C, Lewsey JD, Mackay DF, et al. Estimation of survival probabilities for use in cost-effectiveness analyses: a comparison of a multi-state modeling survival analysis approach with partitioned survival and Markov decision-analytic modeling. *Med Decis Making*. 2017;37:427–39.
33. Coyle D, Coyle K. The inherent bias from using partitioned survival models in economic evaluation. *Value Health*. 2014;17:A194.



11

MODELOS DE TRANSICIÓN ENTRE ESTADOS (STATE-TRANSITION MODELS)

Carlos Crespo Palomo

Profesor asociado Universidad de Barcelona,

Profesor asociado Universitat Pompeu Fabra,

CEO de Axentiva Solutions



Modelos de simulación individual de pacientes (microsimulation models)

Características básicas

El uso de modelos matemáticos se ha venido usando de una manera creciente en el ámbito de las evaluaciones económicas en el sector sanitario (1–3). A través del uso de modelos, es posible estructurar y armonizar de una forma clara tanto la evidencia clínica, como las consecuencias económicas del uso de una tecnología sanitaria (fármacos, dispositivos médicos, apps, ...), de forma que puede ser de gran provecho a la hora de apoyar decisiones relativas a la práctica clínica, o la gestión de los presupuestos y los recursos vinculados.

Dependiendo de la pregunta de investigación, el enfoque de modelización óptimo debe seleccionarse en función de las diferencias esperadas entre los comparadores, el número de covariables, el número de subgrupos de pacientes, las interacciones entre las covariables y el tiempo de simulación. Es por ello que Brennan y col. (4) propusieron una taxonomía de estructuras de modelos para la evaluación económica de tecnologías sanitarias. Dentro de la taxonomía, distinguieron los modelos a nivel de cohorte y de paciente. Los elementos claves utilizados para clasificar las estructuras de los modelos de evaluación económica fueron la interacción de los pacientes, el tiempo de los eventos, los valores esperados frente a los estocásticos y el supuesto markoviano. Las cuestiones subyacentes incluían la necesidad de describir la variabilidad; el ciclo de los estados de salud; la heterogeneidad de la población; la dimensionalidad; la interacción entre covariables y las asociaciones no lineales entre los factores de riesgo individuales y el resultado. Todo ello, sin menoscabar la necesidad de realizar el análisis probabilístico de sensibilidad y el análisis de subgrupos. Tras determinar qué tipo de modelo era el más adecuado, indicaron que hay factores como la disponibilidad de programas informáticos de modelización,

la experiencia del investigador con el tipo de modelo y las limitaciones de tiempo del proyecto que se debían tener en cuenta (4). Por tanto, el modelo más adecuado siempre será el modelo más sencillo que por sus características pueda reflejar mejor la patología que estamos evaluando.

Los modelos de simulación individual, llamados también modelos de microsimulación o simulación de primer orden, permiten estudiar la evolución de cada individuo en un sistema complejo (por ejemplo, en los síndromes de disfunción de múltiples órganos o cuando se desea controlar los efectos de la osteoporosis por cada tipo de fractura) en lugar de cohortes hipotéticas (5). Los modelos de microsimulación son modelos matemáticos que operan desde el nivel del individuo hacia la población, sin centrarse en un individuo representativo o promedio. Estos modelos permiten seguir la progresión de los individuos potencialmente heterogéneos acumulando al historial sus transiciones individuales, los costes y los beneficios clínicos (Tabla 1). Esto implica que la población se estratifica a través de los estados de salud y los atributos identificados como relevantes para el problema analizado (por ejemplo, edad, gravedad de la enfermedad, exposición al riesgo). Una ventaja de utilizar modelos a nivel individual es la capacidad de modelar las características individuales como variables continuas y de evaluar estrategias de intervención dinámicas, en las que las decisiones futuras dependen de las características actuales y pasadas del paciente.

Al igual que los modelos de Markov requieren ciclos fijos de igual longitud, aunque pueden simular múltiples eventos que ocurren en paralelo sobre el mismo individuo (Tabla 1). Esto último les hace realmente útiles para poder evitar el problema subyacente de la pérdida de memoria de los modelos de Markov, ya que se simula cada individuo y sus características a lo largo del tiempo, pudiendo disponer de variables auxiliares vinculadas al individuo para controlar cómo se modifican sus características. La microsimulación permite por tanto condicionar las probabilidades de transición a partir de características del paciente o incluso de cuánto tiempo lleva en un estado de salud (por ejemplo, reducir el riesgo de mortalidad tras varios años como paciente estable).

Tabla 1. Resumen de las características diferenciales entre los modelos de Markov y los modelos de microsimulación

Modelo de Markov	Modelo de simulación individual de pacientes
Cohorte	Individuos
Pocos estados de salud (<12)	Muchos estados de salud
Eventos independientes	Eventos en paralelo, ya que se trabaja a nivel de paciente
Falta de memoria	Con memoria a partir de variables indicadoras
Sin gran necesidad computacional	Complejidad elevada computacionalmente

Mientras que los modelos de cohortes se analizan como cohortes individuales que progresan a través de los estados simultáneamente, lo que no permite distinguir a un individuo de otro salvo por las descripciones de los estados, la microsimulación nos permite emular la evolución individual de cada paciente desde el inicio de tratamiento hasta su fallecimiento a través de una cadena de estados de salud descriptivos de la enfermedad con ciclos fijos (Tabla 1). Incorporar el comportamiento de cada individuo permite mostrar la variabilidad biológica de las respuestas que pueden producirse, lo cual supone una ventaja sobre los enfoques de modelos agregados típicos como son los modelos de Markov en los que las variables agregadas representan únicamente el comportamiento de la media poblacional.

No obstante, cabe destacar que los modelos de simulación son unos modelos matemáticos complejos que necesitan de una alta capacidad de cálculo para su ejecución. Es por ello, que su uso debe ser cuando los modelos de Markov se descartan debido a que el número de estados de salud son tan elevados que la ejecución del mismo es demasiado compleja o tenemos la imperiosa necesidad de soslayar la

pérdida de memoria de los modelos Markovianos por la patología de estudio. Seguramente, aunque depende según el investigador, cuando estamos hablando de más de 12-15 estados de salud, entonces es mejor utilizar el modelo de estados de salud de microsimulación, ya que su implementación será más realista y con una menor necesidad de tiempo para su elaboración. Como los otros modelos de estados de transición, su utilización es razonable cuando el problema de decisión puede enmarcarse en términos de estados de salud, la interacción entre individuos no es relevante y la población de interés es una cohorte específica (6). Una de las ventajas de los modelos de microsimulación frente a los modelos de Markov es que se pueden hacer análisis para subgrupos sin necesidad de generar más estados de salud o estados túneles para controlar la ejecución, ya que se puede mantener la estructura con pequeños cambios en las probabilidades para incorporar las nuevas interacciones.

Ejemplos de modelos de microsimulación

Teniendo en cuenta las características de los modelos de microsimulación, en la literatura se pueden localizar estudio vinculados a la enfermedad pulmonar obstructiva crónica (7), a colitis ulcerosa (8), a hepatitis C (9), al cribado del cáncer de próstata por imagen (10), en evaluar el impacto a largo plazo de la diabetes tipo 2 (11), en la enfermedad de injerto contra huésped crónica a (12), en pacientes con esquizofrenia (13), en osteoporosis (14) o el virus de la inmunodeficiencia humana (15) entre otros.

Para poder mostrar la idoneidad de la aplicación de los modelos de microsimulación se va a ahondar un poco más en alguno de estos casos:

1. La enfermedad de injerto contra huésped crónica es una complicación que puede ocurrir después de un trasplante de médula

ósea o de células madre, en el cual el material recientemente trasplantado ataca el cuerpo del receptor del trasplante (16). Este ataque principalmente incide sobre el hígado, piel, pulmones, ojos y mucosas con diferente nivel de severidad. Por tanto, como el pronóstico de los pacientes depende de los órganos afectados y de la severidad de cada uno de ellos, no es viable llevar a cabo un modelo de Markov aunque tenga eventos repetitivos (12). En este ejemplo se pone de manifiesto que no se puede simular una cohorte de pacientes, sino que es necesario simular qué le sucede a cada órgano en el paciente y evaluar su evolución individual.

2. Otro ejemplo donde se ha aplicado la microsimulación es en la osteoporosis. Es conocido que en la epidemiología de la osteoporosis, las características de los pacientes tienen un rol clave dentro de la transición de la enfermedad. Tal es así, que el riesgo relativo de padecer una fractura tras haber padecido una fractura previa se eleva y además varía en función del sexo, el lugar de la fractura y el nivel de densidad mineral ósea (17). De modelarse como un modelo de Markov, obligaría tener múltiples estados para poder evitar la pérdida de memoria y poder identificar en qué momento temporal está cada paciente para poder modificar su riesgo. Además, también es conocido que el riesgo de mortalidad aumenta tras la fractura para complicarlo aún más (18). Por tanto, es necesario simular los individuos para incorporar todas las características de los pacientes. Siendo necesario en este caso tener identificado el lugar de la fractura, el número de fracturas, el tiempo hasta la primera fractura y el tiempo desde la última fractura, así como el tiempo tras el inicio del tratamiento para poder simular correctamente la patología (14).

Posicionamiento de los modelos de microsimulación

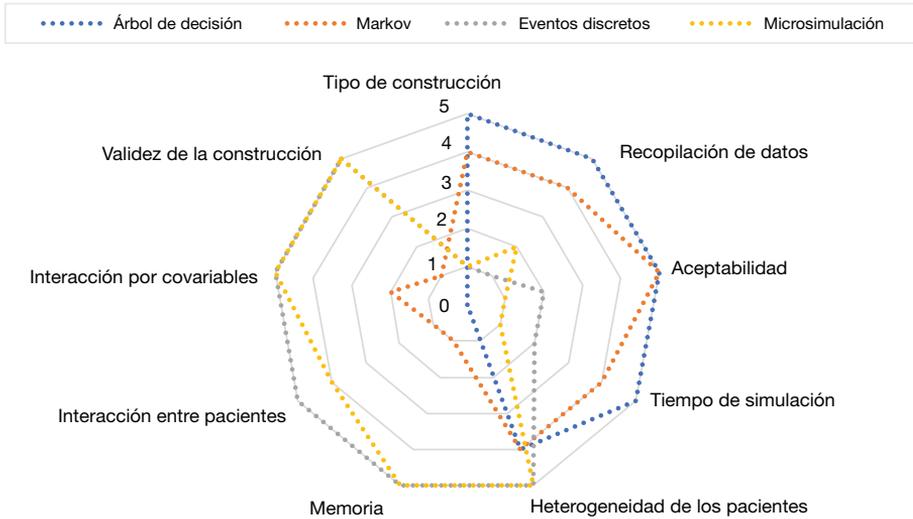
Teniendo más claro cuándo son útiles los modelos de simulación, es importante poder identificar sus fortalezas y debilidades en el contexto de los modelos de evaluación económica (árboles de decisión, modelos de Markov, simulación de eventos discretos y modelos de microsimulación). En concreto, desde el punto de vista del proceso global para la creación de modelos de simulación podemos encontrar una serie de factores claves:

- El tiempo necesario para su construcción
- La facilidad con la que se pueden recopilar la información en la literatura para poder incorporarla al modelo.
- La aceptabilidad por las autoridades sanitarias o las agencias de evaluación económica, así como por las revistas científicas.
- El tiempo necesario para poder generar las simulaciones.
- La facilidad para incorporar la heterogeneidad de los pacientes en el modelo.
- La capacidad para recordar (memoria) las características relevantes de los pacientes o de la enfermedad que pueden condicionar el riesgo.
- La facilidad para incorporar la interacción entre pacientes, como la que sucede por la lista de espera condicionando el poder disponer de una intervención.
- La facilidad para incorporar la interacción por covariables.
- La facilidad para poder validar el modelo con la realidad.

En este sentido claramente los árboles de decisión y los modelos de markov son más fuertes en los tiempos de construcción, la recopilación de datos, la aceptabilidad, el tiempo de simulación y algo menos en incorporar la heterogeneidad de los pacientes (Figura 1). En cambio, los modelos de simulación de eventos discretos y los modelos de microsimulación son más fuertes en la capacidad de memoria, la interacción entre pacientes, la interacción por covariables y la validez de la construcción (Figura 1).



Figura 1. Fortalezas (puntuación 5) y debilidades (puntuación 0) de los modelos de evaluación económica: árbol de decisión, Markov, simulación de eventos discretos y microsimulación



Una de las razones de que los modelos de microsimulación tengan un perfil tan parecido a los modelos de simulación de eventos discretos, aunque su uso en evaluación económica está muy diferenciado, es debido a que realmente son complementarios. En la simulación de eventos discretos lo que estimamos es cuánto tiempo es necesario fijado un evento, en cambio los modelos de simulación estiman cuántos eventos suceden fijando el tiempo. Realmente, es muy parecido a la relación estadística que existe entre la distribución exponencial, utilizada para estimar el tiempo de una cola, por ejemplo, para ser atendido en la consulta médica, y la distribución Poisson, utilizada para estimar el número de individuos en una cola.

Existen otros métodos de simulación como son los sistemas dinámicos (19) o la simulación de modelos basado en agentes (20,21) que no suelen utilizarse tanto en el ámbito de la evaluación económica dejándolos fuera de este capítulo.



El arte de la simulación

La generación de números aleatorios tiene como objetivo el poder construir modelos que describan la parte esencial del comportamiento de un sistema de interés, así como en diseñar y realizar experimentos con el modelo y extraer conclusiones de sus resultados para apoyar a la toma de decisiones.

En muchas aplicaciones los números aleatorios suelen ser generados directamente por el ordenador manteniendo las características de una distribución deseada. Pero los números que son generados por el ordenador trabajan en un punto estrictamente determinista, los valores no son realmente aleatorios, ya que estos pueden ser predichos de forma escrita, por eso estos números son llamados números pseudoaleatorios. Salvando esta consideración en la generación de números aleatorios, el método más utilizado es el método de inversión.

Supongamos que la variable aleatoria X tiene la función de distribución F continua y estrictamente creciente, siempre que $0 < F(x) < 1$. Sea Y una variable aleatoria con distribución uniforme en $(0, 1)$. Entonces, la variable aleatoria $F^{-1}(U)$ tiene función de distribución F .

Esta proposición sugiere que para hacer un muestreo de una variable aleatoria X de la que se conoce $F^{-1}(x)$, se pueden generar números U uniformes entre $(0, 1)$ y hacer $X = F^{-1}(U)$.

Tenemos entonces el siguiente algoritmo de inversión

```

Generar  $U \sim U(0, 1)$ 
Hacer  $X = F^{-1}(U)$ 
Retornar  $X$ 
    
```

El método en si mismo lo que hace es simular la probabilidad que es un valor entre 0 y 1, para después una vez fijada una distribución con sus parámetros poder indicar qué valor corresponde a la probabilidad. Por ejemplo, si estuviéramos ante una distribución normal de media 50 y

desviación 23, entonces si tenemos una probabilidad de 20% el valor simulado sería 30,64.

Por tanto, una condición mínima para la aplicación del método de inversión es conocer la forma explícita de F^{-1} . Esta condición es muy común en múltiples distribuciones, siendo algunas de ellas la distribución normal, la beta, la log-normal, la gamma o la exponencial.

Para poder iniciar la generación de números aleatorios tal como hemos comentado es necesario incluir un valor inicial, denominado semilla. Es a partir del valor semilla y de unos cálculos internos se generan estos valores aleatorios. Por tanto, cuando estemos haciendo simulaciones es recomendable en la fase de desarrollo fijar el valor de la semilla para replicar siempre los mismos valores y de esa forma poder verificar y validar el modelo. Una vez ya se dispone del modelo de microsimulación óptimo entonces se permitirá al sistema que utilice el valor de semilla que tenga por defecto, que muchas veces se basan en el día hora de la ejecución.

A nivel práctico fijar la semilla va a permitir evaluar cómo se modifica el resultado para cada pequeño cambio que hagamos. Por ejemplo, si se modifica la probabilidad de curarse de un tratamiento del 10% al 14% y hacemos simulaciones sin semilla podemos ver modificaciones del resultado que están más vinculados al hecho de simular que al cambio que hemos realizado. Se debe tener en cuenta que cada vez que simulemos estaremos generando valores aleatorios distintos y que, por tanto, el resultado se va a modificar si el número de simulaciones es reducido.

Convergencia

Como sucede siempre que hagamos simulaciones de primer orden, como este caso, o de segundo orden se debe alcanzar la convergencia de las estimaciones. Los criterios de convergencia describen el nivel mínimo de precisión de los resultados del modelo en comparación con los valores observados de los objetivos de calibración (22,23). Dicho de una

forma más profana debemos hacer tantas simulaciones hasta que las distribuciones y probabilidades que estamos simulando se estabilicen.

No hay consenso sobre la especificación de los umbrales de convergencia, pero un enfoque pragmático consiste en especificar un intervalo de confianza o una tolerabilidad a partir de la cual indicamos que no hay diferencia en el resultado. Es recomendable identificar la tolerabilidad para la diferencia de costes y la diferencia de efectos por separado, ya que si se hace sobre la ratio coste-efectividad incremental, equivalentemente para la ratio coste-utilidad incremental, puede estar influenciado por el hecho que haya pocas diferencias en el denominador y por tanto su volatilidad sea algo mayor. No obstante, siempre que se encuentre ante esta tesitura deberá tener en cuenta la interpretación de la ratio coste-efectividad incremental basado en el umbral de eficiencia. Dicho de otro modo, cuando los valores de la ratio coste-efectividad no modifiquen la decisión, podremos decir que se ha llegado a la convergencia al no ser relevantes.

En el caso en concreto de las microsimulaciones se debe tener en cuenta que estamos haciendo simulaciones de individuos. Por tanto, normalmente se simula un conjunto de individuos, muestra de simulación, y se realizan varias muestras. Por tanto, si generamos 100 muestras de tamaño 20 entonces se generarán en total 2.000 individuos. No obstante, se recomienda simular como mínimo 1.000.000 individuos ya sea por muestras de simulación grandes o por muchas muestras para evitar potenciales errores de simulación. El error de simulación se comete al aproximar la distribución teórica mediante una distribución empírica a partir de réplicas y se puede controlar al elevar el número de simulaciones. Esto conlleva que estos modelos necesiten de sistemas de cálculo potentes para poder evitar grandes tiempos de espera para obtener los resultados. Tema que se magnifica cuando se realizan análisis de sensibilidad probabilísticos o análisis del valor de la información perfecta (24). Hay que tener en cuenta que la incertidumbre de segundo orden debe considerarse por separado de la incertidumbre de primer orden, siendo esta última la que utilizamos para los modelos de microsimulación (25).



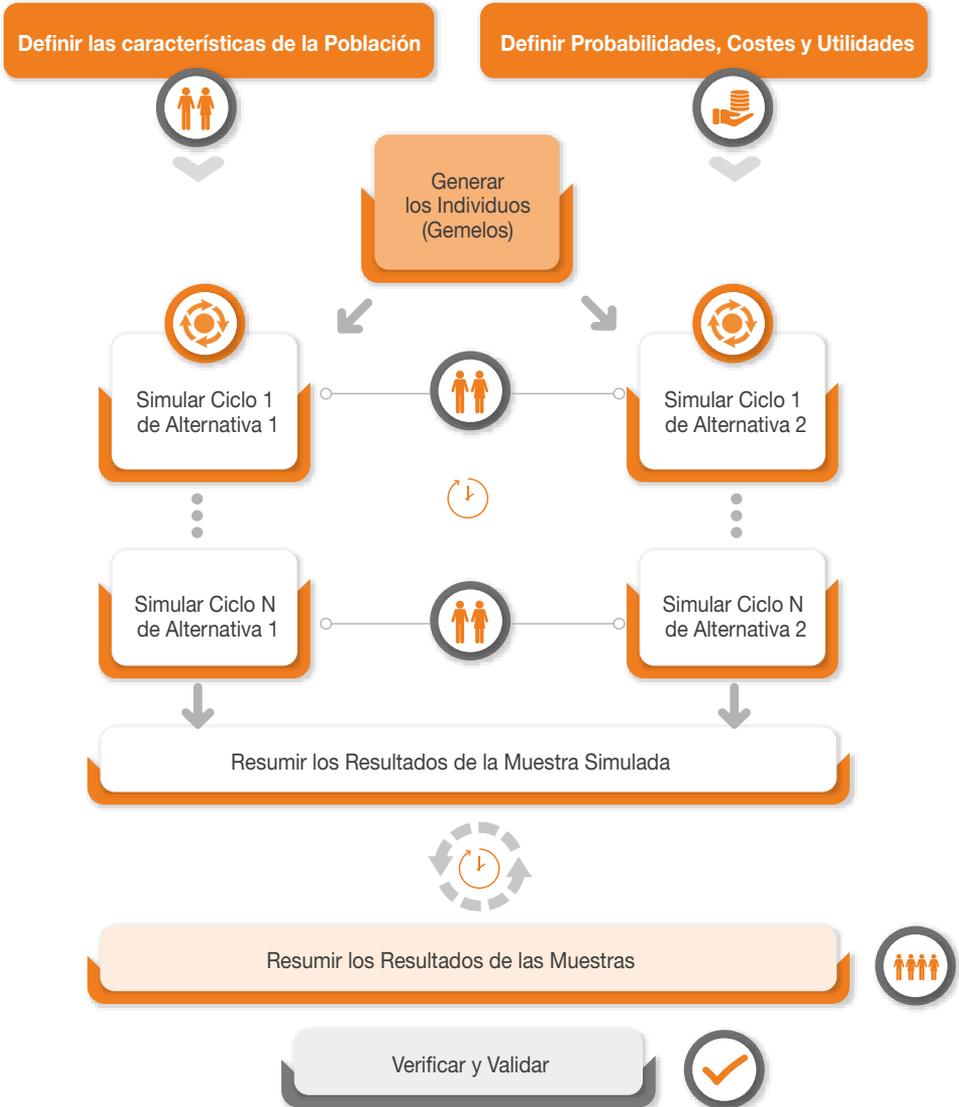
Pasos para construir un modelo de microsimulación

Como se ha comentado, los modelos de microsimulación no están limitados por la propiedad markoviana ya que se simula un individuo cada vez. Estas microsimulaciones se evalúan mediante la simulación Monte Carlo de primer orden. Dicho de otra forma, se evaluará la variabilidad mediante la desviación estándar, a diferencia de en las simulaciones de Monte Carlo de segundo orden donde la variabilidad se realiza sobre la media y, por tanto, se hace mediante el error estándar.

Como en todo modelo debemos disponer de la información de la literatura. En este caso, no solo se necesitan las probabilidades de transición, los costes o las utilidades, sino que es necesario definir las características o variables de control que se deben simular a nivel de individuo antes de poder empezar (Figura 2). A partir de ese punto, el investigador debe determinar el número de muestras simuladas y el tamaño de la muestra, teniendo en cuenta la elevada necesidad de que las distribuciones converjan. En ese punto ya se estará en disposición de generar la muestra simulada de pacientes a partir de las características fijadas para hacer evolucionar los pacientes por las dos alternativas de tratamiento. Para cada ciclo se debe estimar los costes y las medidas de efectos, como puede ser los años de vida o los años de vida ajustados por calidad. Al alcanzar el horizonte temporal, se debe resumir la información de los ciclos previos y calcular los resultados acumulados. Este proceso de generación y evolución de los pacientes se debe repetir para tantas muestras simuladas como el número que se haya previsto, para obtener el resultado final del modelo agregando la información de cada muestra simulada. Finalmente, y como paso previo a divulgar los resultados, se debe verificar y validar el modelo. En el caso que la verificación o la validación no sea la adecuada entonces se deberá revisar los parámetros y las relaciones lógicas que se han realizado en el modelo hasta alcanzar la verificación y la validación.



Figura 2. Pasos para generar un modelo de microsimulación



Para poder ilustrar mejor los pasos para la construcción de un modelo de microsimulación nos basaremos en un ejemplo ficticio, pero con características reales para ilustrar cada paso. En ese sentido basaremos nuestro ejemplo de microsimulación en una patología con afectación multiorgánica. En concreto asumiremos que la patología puede afectar a 4 órganos con niveles de severidad distintos (leve, moderado y grave). En ese sentido la población son pacientes con una edad media de 50 años (DE: 15), el 60% mujeres donde el 10% de los pacientes tienen afectación del órgano 1, el 30% del órgano 2, el 80% del órgano 3 y el 50% del órgano 4. Para estimar la evolución pronóstica de los pacientes es necesario calcular un estado de salud global a partir de los órganos afectados y el nivel de severidad (leve, moderado y severo). Se ha previsto un horizonte temporal de 3 años con ciclos semestrales, siendo necesario descontar costes y efectos a partir del primer año tal como indican las guías (26,27).

El ejemplo se ha desarrollado sobre Excel y se han dispuesto unas hojas para el desarrollo del mismo: '0- Características pacientes', '1-Generación población', '1-Generación población', '2- Probabilidades', '3- Costes y Utilidades', una hoja para emular la evolución de los ciclos en el tratamiento actual '4- Ciclo 1' a '4- Ciclo 6', y el tratamiento nuevo '5- Ciclo 1 Nuevo' a '5- Ciclo 6 Nuevo', para finalizar '6-Resumen población' y una auxiliar '7-Mortalidad'.

Paso 1: Definir las características de la población y generar los individuos

La gran diferencia de los modelos de microsimulación es que al inicio de la simulación se crean pares de individuos «gemelos», es decir, se generan individuos con las mismas características, que acceden a cada una de las estrategias de tratamiento comparadas en el modelo. Por tanto, que la diferencia en el tratamiento que se les aplica a cada gemelo es la única distinción entre ellos a la entrada del modelo. Esto es importante, ya que de otra forma podemos crear un sesgo en las poblaciones evaluadas por cada tratamiento, siendo, por tanto, los resultados erróneos en este caso. Este proceso es equiparable a la aleatorización que se realiza en los ensayos clínicos previo el inicio del estudio.

Para poder generar los pacientes se debe representar las características basales de los individuos en la hoja '0- Características pacientes' (Figura 3). En ese sentido, se debe definir las distribuciones necesarias para hacer las simulaciones. En este caso en las celdas de color verde se puede ver cómo se utiliza la distribución normal para la edad y la distribución Bernoulli para el resto de las variables. En las celdas azules se han indicado los parámetros necesarios para cada distribución.

Para el caso de la edad se debe identificar tanto la media como la desviación estándar y calcularemos el valor específico del paciente en la celda E13 a partir de la función inversa de la distribución normal. Para generar el valor aleatorio, se utiliza una función volátil denominada ALEATORIO() que mostrará un número nuevo cada vez que se haga un cálculo en Excel, equivalentemente cuando pulsemos la tecla F9.

CELDA	FÓRMULA
E13	=INV.NORM(ALEATORIO();G13;I13)

Figura 3. Características de los pacientes a simular

CARACTERÍSTICAS DE LOS PACIENTES

	Valor simulado	Distribución		Parámetro 1		Parámetro 2		Parámetro 3	
EDAD	20,40	Normal	20,40	Media	50	Desviación estándar	15		
SEXO (Mujer=1)	Hombre	Bernoulli	0	% Mujeres	60%				
ÓRGANOS AFECTADOS (Max:4)	1								
Afectación órgano 1	No	Bernoulli	0	% Afectación	10%				
Severidad. Órgano 1		Bernoulli	Leve	Aleatorio	14%	% Leve	25%	% Moderado	50%
Afectación órgano 2	No	Bernoulli	0	% Afectación	30%				
Severidad. Órgano 2		Bernoulli	Leve	Aleatorio	4%	% Leve	5%	% Moderado	30%
Afectación órgano 3	Sí	Bernoulli	1	% Afectación	80%				
Severidad. Órgano 3	Leve	Bernoulli	Leve	Aleatorio	29%	% Leve	30%	% Moderado	30%
Afectación órgano 4	No	Bernoulli	0	% Afectación	50%				
Severidad. Órgano 4		Bernoulli	Severo	Aleatorio	75%	% Leve	60%	% Moderado	20%
ESTADO DE SALUD GLOBAL*: Score	1								
ESTADO DE SALUD GLOBAL*: Severidad	Leve								

*Valor estimado a partir de los órganos afectados y el nivel de severidad. Cada órgano puntúa como 1 y el nivel de severidad multiplica la puntuación por 1 en leve, 2 en moderado y 3 en grave. En el caso de tener afectado el órgano 1, automáticamente la puntuación total es 12. El estado de salud global es leve por debajo de 3, moderado por debajo de 7 y el resto severo.

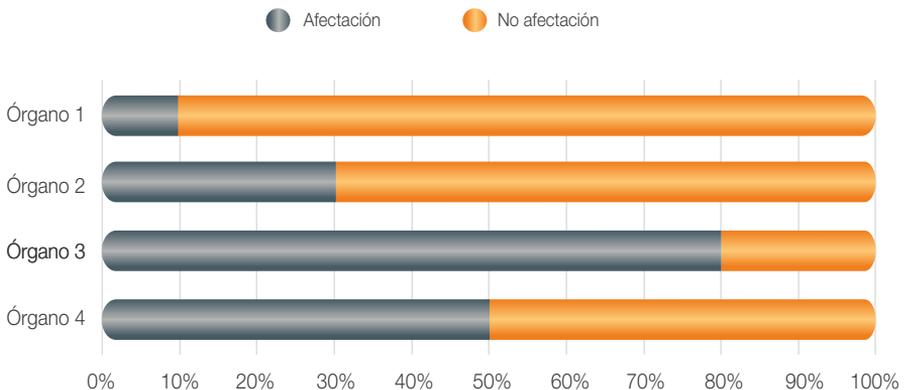
Para asignar el sexo del paciente simulado como se ha comentado se debe asignar una distribución de Bernoulli, pero esta función como tal no está definida dentro de las funciones de Excel. No obstante, lo que se necesita es utilizar la función volátil ALEATORIO() condicionando el resultado en función de la probabilidad de éxito del caso en concreto. En este caso necesitamos indicar que el paciente es “Mujer” cuando el valor aleatorio sea igual o inferior a 60%, para el resto de los valores será “Hombre”. Tal como se ha generado en el Excel, en la columna E se genera el valor aleatorio con resultado 1 u 0 para luego hacer las transformaciones en la celda C14.

CELDA	FÓRMULA
E14	=SI(ALEATORIO())<=G14;1;0)
C14	=SI(E14=1;"Mujer";"Hombre")

Para el caso de cada órgano, primero deberemos simular si el órgano está afectado para luego ver en qué nivel de severidad está (Figura 4). Como en el caso anterior utilizamos la distribución de Bernoulli, que para el órgano 1 es a partir del valor 10%.



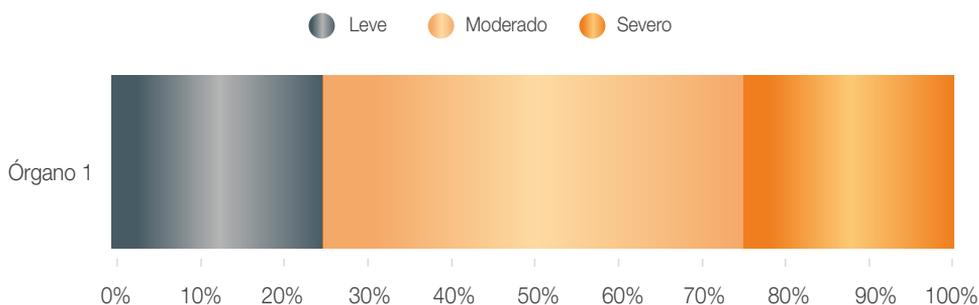
Figura 4. Probabilidad de que esté afectado cada órgano



Para simular el nivel de severidad se deben hacer 2 puntos de corte al haber 3 niveles de severidad. Eso quiere decir que si el valor aleatorio está por debajo del 25% entonces el nivel de severidad es Leve, si está entre 25% y 50% el nivel de severidad es Moderado y por encima el nivel de severidad es Severo (Figura 5).



Figura 5. Niveles de severidad del órgano 1



En este caso en la celda C17 controlamos si el órgano está afectado para mostrar el nivel de severidad o en caso contrario dejarlo vacío.

CELDA	FÓRMULA
E16	=SI(ALEATORIO())<=G16;1;0)
C16	=SI(E16=1;"Sí";"No")
E17	=SI(G17<=I17;"Leve";SI(G17<=K17;"Moderado";"Severo"))
C17	=SI(C16="Sí";E17;"")

Para el resto de los órganos, se deberá realizar de la misma manera.

CELDA	FÓRMULA
E18	=SI(ALEATORIO())<=G18;1;0)
C18	=SI(E18=1;"Sí";"No")
E19	=SI(G19<=I19;"Leve";SI(G19<=K19;"Moderado";"Severo"))
C19	=SI(C18="Sí";E19;"")
E20	=SI(ALEATORIO())<=G20;1;0)
C20	=SI(E20=1;"Sí";"No")
E21	=SI(G21<=I21;"Leve";SI(G21<=K21;"Moderado";"Severo"))
C21	=SI(C20="Sí";E21;"")
E22	=SI(ALEATORIO())<=G22;1;0)
C22	=SI(E22=1;"Sí";"No")
E23	=SI(G23<=I23;"Leve";SI(G23<=K23;"Moderado";"Severo"))
C23	=SI(C22="Sí";E23;"")

Por último, se calculará el número de órganos afectados en la celda C15 (=E16+E18+E20+E22). De esta manera podremos controlar posteriormente si hay algún órgano afectado, repitiendo la simulación en el caso de que el número de órganos sea cero. Hay que tener en cuenta que se está simulando por separado cada órgano y fácilmente pueden simularse paciente sin afectación (probabilidad del 6% para nuestro ejemplo).

Asimismo, restará calcular el indicador global vinculado al estado de salud a partir de los órganos afectados y el nivel de severidad. En ese sentido, cada órgano puntúa como 1; y el nivel de severidad multiplica la puntuación por 1 en leve, 2 en moderado y 3 en grave. En el caso de tener afectado el órgano 1, se ha asumido que es un órgano vital y, por tanto, automáticamente la puntuación total es 12. El estado de salud global finalmente se clasifica como leve por debajo de 3, moderado por debajo de 7 y el para puntuaciones superiores como severo.

CELDA	FÓRMULA
C24	=SI(E16=1;12;E18*SI(E19="Leve";1;SI(E19="Moderado";2;3))+E20*SI(E21="Leve";1;SI(E21="Moderado";2;3))+E22*SI(E23="Leve";1;SI(E23="Moderado";2;3)))
C25	=SI(C24<3;"Leve";SI(C24<7;"Moderado";"Severo"))

Una vez ya se tiene la estructura de cálculo para la generación de la muestra simulada, se puede acceder a la hoja '1-Generación población' (Figura 6). Esta hoja se ha dividido en 3 secciones:

- Sección inicial (filas 8 a 10): donde se genera a los individuos a partir de la información de la hoja '0- Características pacientes'. Se dispone de una celda de control de errores (P10) para eliminar dicha simulación, si hace falta.
- Sección acumuladora de la muestra (filas 18 a 370): donde se van acumulando los pacientes que se van simulando.
- Sección resumen de la muestra (filas 12 a 17): donde se hace una descriptiva de las variables de la muestra generada. En principio, si el tamaño muestral es suficiente, la descripción debería ser similar a las características poblacionales que hemos planteado en la hoja '0- Características pacientes'.

Como se ha comentado en la celda P10 se han planteado varios controles que de no cumplirse elimina la simulación actual. Los controles son que la edad del paciente esté comprendida en la horquilla de 20 y 90 años y que el número de órganos afectados sea mayor que 0. Si se cumple el control, entonces en la celda P10 aparecerá "No".

ÓRGANO 3		ÓRGANO 4		ESTADO DE SALUD GLOBAL		ELIMINAR
Afectación	Severidad	Afectación	Severidad	Score	Severidad	
Sí	Leve	No		1	Leve	No

Figura 6. Generación de la población de base



GENERACIÓN DE LA POBLACIÓN DE BASE



Tamaño muestra: 350

Paciente simulado	EDAD	SEXO	ORGANOS AFECTADOS (#)		ORGANO 1		ORGANO 2		ORGANO 3		ORGANO 4		ESTADO DE SALUD GLOBAL		
			Afectación	No	Severidad	Afectación	Severidad	Afectación	Severo	Afectación	No	Severidad	Score	Severidad	Eliminar
	54,63	Hombre	1	No									3	Moderado	No

RESUMEN CARACTERÍSTICAS SIMULADAS

EDAD	SEXO (Mujer)	ORGANOS AFECTADOS (#)	ORGANO 1		ORGANO 2		ORGANO 3		ORGANO 4		ESTADO DE SALUD GLOBAL	
			Afectación	Severidad	Afectación	Severidad	Afectación	Severo	Afectación	No	Severidad	Score
50,26	58,6%	1,80	10,0%	Leve: 22,86%	32,3%	Leve: 3,54%	88,0%	Leve: 30,19%	49,7%	Leve: 60,92%	4,73	Leve: 16,57%
			Moderado: 25,71%		Moderado: 0%		Moderado: 0%		Moderado: 0%			Moderado: 66,86%
			Severo: 51,43%		Severo: 77,88%		Severo: 69,81%		Severo: 39,08%			Severo: 16,57%

POBLACIÓN SIMULADA

ID	EDAD	SEXO	ORGANOS AFECTADOS (#)		ORGANO 1		ORGANO 2		ORGANO 3		ORGANO 4		ESTADO DE SALUD GLOBAL	
			Afectación	No	Severidad	Afectación	Severidad	Afectación	Severo	Afectación	No	Severidad	Score	Severidad
1	34,46	Mujer	1	No									3	Moderado
2	37,37	Hombre	2	No									4	Moderado
3	73,52	Hombre	2	No									4	Moderado
4	34,11	Mujer	2	No			Severo						6	Moderado
5	25,34	Hombre	2	No									4	Moderado
6	76,04	Hombre	2	No									4	Moderado
7	58,00	Mujer	1	No									3	Moderado

CELDA	FÓRMULA
P10	=SI(O(C10<20;C10>90;E10=0);"Sí";"No")

El número de pacientes a simular se identifica en la celda H3 que se ha nombrado como 'TamañoMuestra' para facilitar su uso en las fórmulas y hacerlas más inteligibles. En ese sentido, para el caso actual se ha fijado siempre el tamaño muestral en 350. No obstante, si el lector quiere trabajar con otro tamaño muestral deberá modificarlo y ser consecuente con algunos cálculos que están fijados solo para las 350 primeras filas.

Una vez se dispone de la estructura se deberá ejecutar una macro de Visual Basic que copiará la fila 10 y la pegará a partir de la fila 21 repetidamente mediante un bucle, siempre que la celda P10 tenga el valor "No" y no se ha alcanzado el número máximo de pacientes (n: 350).

```

Sub GeneracionPoblacion()
  Sheets("1-Generación población").Select
  Index = 0
  Trials = Range("TamañoMuestra").Value
  Do
    Range("C10:O10").Select
    Selection.Copy
    If Range("P10") = "No" Then
      ActiveCell.Offset(Index + 11, 0).Range("A1").Select
      Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, Skip-
Blanks:= _
        False, Transpose:=False
      Index = Index + 1
    Else
      Calculate
    End If
  Loop While Index < Trials
End Sub

```

Paso 2: Definir probabilidades, costes y utilidades

Antes de proseguir con las simulaciones se deben definir las probabilidades de transición en la hoja ‘2- Probabilidades’, así como los costes y utilidades en la hoja ‘3- Costes y Utilidades’. Esto se podría hacer antes tal como se indicaba en la Figura 2, pero realmente no es necesario hasta que tenemos generado los pacientes, y de ahí que en esta sección se ha modificado el orden.

Para las probabilidades, realmente se han definido las probabilidades de transición entre niveles de severidad para cada órgano (Tabla 2), la mortalidad vinculada al nivel de severidad del estado de salud global (Tabla 3) y la referencia de tiempo del ciclo (celda C130= 6 meses). En este caso los datos de cada parámetro se han fijado directamente. No obstante, se deberá estimar o inferir a partir de la literatura haciendo las transformaciones adecuadas para poder ajustar las probabilidades al marco temporal del modelo de simulación, 6 meses en este ejemplo (28,29).

Tabla 2. Probabilidades de transición por nivel de severidad de cada órgano por tratamiento

TRATAMIENTO ACTUAL				TRATAMIENTO NUEVO		
ÓRGANO 1	a Leve	a Moderado	a Severo	a Leve	a Moderado	a Severo
De Leve	70,0%	20,0%	10,0%	91,0%	7,0%	2,0%
De Moderado	10,0%	70,0%	20,0%	46,0%	39,0%	15,0%
De Severo	15,0%	55,0%	30,0%	48,0%	32,0%	20,0%

ÓRGANO 2	a Leve	a Moderado	a Severo	a Leve	a Moderado	a Severo
De Leve	91,0%	7,0%	2,0%	94,0%	5,0%	1,0%
De Moderado	25,0%	60,0%	15,0%	42,0%	48,0%	10,0%
De Severo	20,0%	55,0%	25,0%	35,0%	40,0%	25,0%

Tabla 2 cont. Probabilidades de transición por nivel de severidad de cada órgano por tratamiento

TRATAMIENTO ACTUAL				TRATAMIENTO NUEVO		
ÓRGANO 3	a Leve	a Moderado	a Severo	a Leve	a Moderado	a Severo
De Leve	92,0%	3,0%	5,0%	95,0%	3,0%	2,0%
De Moderado	35,0%	55,0%	10,0%	45,0%	53,0%	2,0%
De Severo	60,0%	30,0%	10,0%	65,0%	30,0%	5,0%

ÓRGANO 4	a Leve	a Moderado	a Severo	a Leve	a Moderado	a Severo
De Leve	85,0%	10,0%	5,0%	90,5%	7,5%	2,0%
De Moderado	7,5%	85,0%	7,5%	55,0%	42,0%	3,0%
De Severo	5,0%	10,0%	85,0%	35,0%	25,0%	40,0%

Tabla 3. Probabilidad de mortalidad

MORTALIDAD	VALOR
Leve	1 % DE: 0,5%
Moderado	10% DE: 2%
Severo	30% DE: 5%

DE: Desviación Estándar

Como se puede observar en la Figura 7, correspondiente a la hoja '2- Probabilidades', tenemos una celda (C11) para identificar si queremos utilizar el análisis determinístico, medias de los parámetros, o queremos aplicar la simulación de Montecarlo de 2º orden para cada muestra simulada de Montecarlo de 1er orden. La información de los parámetros se ha dividido por colores, siendo en color azul las distribuciones (columna F-G) y los parámetros de estas (columnas (H-K)); y en color verde las celdas con la media (columna D) y error estándar (columna E) de cada variable. Para facilitar la interpretación de las fórmulas la celda C10 se ha nombrado como 'AnDeterministico' y en la columna 'A' aparece el nombre de cada parámetro del modelo.

Figura 7. Definición de probabilidades

Análisis determinístico		1	0 No; 1 Sí			
PROBABILIDADES DE TRANSICIÓN						
	Valor modelo	Media	Error estándar	Distribución	Parámetro 1	Parámetro 2
TRATAMIENTO ACTUAL						
ÓRGANO 1						
DE LEVE						
pLeveLeveOrg1 Actual	a Leve	70%				
pLeveModOrg1 Actual	a Moderado	20%	20,0%	5,0% Beta	0,30631766 alpha	12,6 beta
pLeveSevOrg1 Actual	a Severo	10%	10,0%	2,0% Beta	0,09389649 alpha	22,4 beta
DE MODERADO						
pModLeveOrg1 Actual	a Leve	10%	10,0%	2,0% Beta	0,06575126 alpha	22,4 beta
pModModOrg1 Actual	a Moderado	70%				
pModSevOrg1 Actual	a Severo	20%	20,0%	5,0% Beta	0,13966699 alpha	12,6 beta
DE SEVERO						
pSevLeveOrg1 Actual	a Leve	15%	15,0%	2,0% Beta	0,11232041 alpha	47,6625 beta
pSevModOrg1 Actual	a Moderado	55%	55,0%	5,0% Beta	0,56954364 alpha	53,9 beta
pSevSevOrg1 Actual	a Severo	30%				
ÓRGANO 2						
DE LEVE						
pLeveLeveOrg2 Actual	a Leve	91%				
pLeveModOrg2 Actual	a Moderado	7%	7,0%	2,0% Beta	0,0596922 alpha	11,3225 beta
						150,4275

A nivel de fórmulas, una vez hemos incorporado la media y el error estándar de cada parámetro, y se ha fijado la distribución Beta entonces podemos calcular los parámetros alfa y beta de la distribución a partir del método de los momentos (30,31). Como en la generación de los individuos de la muestra simulada, tendremos un parámetro volátil que nos indicará la probabilidad y que aplicando la función inversa de la beta nos mostrará un valor de la probabilidad. Por último, se indicará el valor del modelo en la columna C. Como ejemplo se incluye las fórmulas de la probabilidad de transición de leve a moderado del órgano 1 del paciente tratado con el tratamiento actual (fila 19). Para el resto de los casos, se debe aplicar las mismas fórmulas con su referencia relativa a la fila correspondiente.

CELDA	FÓRMULA
D19	Valor de media
E19	Valor del error estándar
I19	$= (D19^2 - D19^3) / (E19^2) - D19$
K19	$= I19 / D19 - I19$
G19	$= \text{INV.BETA.N}(\text{ALEATORIO}); I19; K19$
C19	$= \text{SI}(\text{AnDeterministico}=1; D19; G19)$

De la misma forma, se dispone de los costes semestrales para cada órgano en función del nivel de severidad (Tabla 4), los costes de los tratamientos semestrales (Tabla 5), las utilidades por nivel de severidad (Tabla 6) y el descuento de costes ($\text{desCostes}=3\%$) y efectos ($\text{desEfectos}=3\%$). En todos los casos se han nombrado los costes por estado con un 'c' al inicio, el nivel de severidad y órgano, los costes de los tratamientos como 'cActual' / 'cNuevo' y las utilidades con una 'u' antes del nivel de severidad.

Tabla 4. Costes semestrales según nivel de severidad y órgano

COSTE POR CICLO	LEVE	MODERADO	SEVERO
Órgano 1	350 DE: 50	6.500 DE: 900	12.000 DE: 3000
Órgano 2	125 DE: 50	454 DE: 90	2.835 DE: 600
Órgano 3	50 DE: 2	75 DE: 5	200 DE: 10
Órgano 4	300 DE: 90	1.500 DE: 150	3.000 DE: 300

Tabla 5. Costes semestrales según nivel de severidad y órgano

COSTE POR CICLO	VALOR
Tratamiento actual	250 DE: 20
Tratamiento nuevo	1.750 DE: 150

Tabla 6. Utilidades por estado de salud global

UTILIDADES	VALOR
Leve	0,93 DE: 0,02
Moderado	0,7 DE: 0,02
Severo	0,54 DE: 0,02

Como se puede observar en la Figura 8 la estructura es la misma que se había hecho para las probabilidades. La diferencia es que en este caso la distribución que se utiliza para estimar de forma estocástica los costes es la gamma y para las utilidades la beta.

A nivel de fórmulas se actúa de forma similar, pero aplicando el cálculo apropiado para la estimación por el método de los momentos para la gamma (30,31). Como ejemplo se incluye las fórmulas del coste de seguimiento, visitas médicas y pruebas diagnósticas de los pacientes que tienen afectado el órgano 1 de forma leve independientemente del tratamiento (fila 15). Para el resto de los casos, se debe aplicar las

Figura 8. Definición de costes y utilidades

	Valor modelo	Media	Error estándar	Distribución	Parámetro 1	Parámetro 2
COSTES						
ÓRGANO 1						
cLeveOrg1	350.00	350.00	50.00	Gamma	49.00	beta
cModOrg1	6.500.00	6.500.00	900.00	Gamma	52.16	beta
cSevOrg1	12.000.00	12.000.00	3.000.00	Gamma	16.00	beta
ÓRGANO 2						
cLeveOrg2	125.00	125.00	50.00	Gamma	6.25	beta
cModOrg2	454.00	454.00	90.00	Gamma	25.45	beta
cSevOrg2	2.835.00	2.835.00	600.00	Gamma	22.33	beta
ÓRGANO 3						
cLeveOrg3	50.00	50.00	2.00	Gamma	625.00	beta
cModOrg3	75.00	75.00	5.00	Gamma	225.00	beta
cSevOrg3	200.00	200.00	10.00	Gamma	400.00	beta
ÓRGANO 4						
cLeveOrg4	300.00	300.00	90.00	Gamma	11.11	beta
cModOrg4	1.500.00	1.500.00	150.00	Gamma	100.00	beta
cSevOrg4	3.000.00	3.000.00	300.00	Gamma	100.00	beta



mismas fórmulas con su referencia relativa a la fila correspondiente para los costes con la distribución gamma y para las utilidades se debe aplicar las mismas fórmulas vistas para probabilidades con la distribución beta.

CELDA	FÓRMULA
D15	Valor de media
E15	Valor del error estándar
I15	=D15/K15
K15	=E15^2/D15
G15	=INV.GAMMA(ALEATORIO();I15;K15)
C15	=SI(AnDeterministico=1;D15;G15)

Paso 3: Simular ciclos de cada alternativa

Una vez ya se tienen todos los parámetros para la generación de la evolución de la muestra simulada, se puede acceder a la hoja '4- Ciclo 1' (Figura 9). Esta hoja se ha dividido en 4 secciones:

- Sección inicial (filas 7 a 10): donde se recupera los datos del individuo generado la primera vez o del ciclo anterior. Se dispone de una celda de control para ver si el paciente se ha muerto en el ciclo (R10) y se guarda cuándo se ha muerto (Q10). En el caso que el paciente estuviera muerto entonces se muestra el dato de muerto 'Sí' y el ciclo en que se ha muerto.
- Sección de evolución (filas 12 a 16): donde vamos a simular cómo se modifica el nivel de severidad de cada órgano en función de las probabilidades de transición del tratamiento y calcularemos el indicador de estado de salud global, así como los datos de supervivencia.

Figura 9. Generación de la transición por ciclo

GENERACIÓN DEL CICLO
TRATAMIENTO Actual



CICLO INICIAL	Paciente simulado	CICLO	EDAD	SEXO	ORGANOS AFECTADOS (#)	ORGANO 1		ORGANO 2		ORGANO 3		ORGANO 4		ESTADO DE SALUD GLOBAL		CICLO DE LA MUERTE	Vivo en el ciclo
						Afectación	Severidad	Afectación	Severidad	Afectación	Severidad	Afectación	Severidad	Score	Severidad		
1			34,46	Mujer	1	No		No		No		SI		3	Moderado	SI	

CICLO SIGUIENTE: TRATAMIENTO Actual

CICLO	Paciente simulado	EDAD	SEXO	ORGANOS AFECTADOS	ORGANO 1		ORGANO 2		ORGANO 3		ORGANO 4		ESTADO DE SALUD GLOBAL		CICLO DE LA MUERTE	VIVO
					Afectación	Severidad	Afectación	Severidad	Afectación	Severidad	Afectación	Severidad	Score	Severidad		
2		34,96	Mujer	1	No	No	No	No	SI	Severo	0,372951401	3	Moderado	SI		

RESUMEN CARACTERÍSTICAS SIMULADAS

EDAD	SEXO (Mujer)	ORGANOS AFECTADOS	ORGANO 1 Afectación	ORGANO 1 Severidad	ORGANO 2 Afectación	ORGANO 2 Severidad	ORGANO 3 Afectación	ORGANO 3 Severidad	ORGANO 4 Afectación	ORGANO 4 Severidad	ESTADO DE SALUD GLOBAL Score	ESTADO DE SALUD GLOBAL Severidad	CICLO DE LA MUERTE	VIVO
50,73	56,6%	1,80	10,0%	Leve: 20% Moderado: 57,14% Severo: 22,86%	32,3%	Leve: 23,01% Moderado: 51,39% Severo: 25,66%	88,0%	Leve: 67,86% Moderado: 21,75% Severo: 10,39%	49,7%	Leve: 51,72% Moderado: 12,64% Severo: 35,63%	3,77	Leve: 41,14% Moderado: 46,29% Severo: 12,57%	1,00	89,7%

- Sección acumuladora de la muestra (filas 25 a 377): donde se van acumulando los pacientes que se van simulando una vez evolucionan en el ciclo.
- Sección resumen de la muestra (filas 19 a 24): donde se hace una descriptiva de las variables de la muestra para ese ciclo.
- Sección estimaciones (columnas T a AB): donde se calcula los años de vida, años de vida ajustados por calidad y costes de cada ciclo sin descontar y aplicando el descuento, si es pertinente. Adicionalmente, en la fila 22 se muestra el resumen de los resultados para el total de la muestra simulada en ese ciclo.

Por tanto, ya estamos en disposición de recuperar la información de los pacientes generados al inicio con la función BUSCARV. Para ello, en la celda B10 indicaremos el código identificador del individuo de la muestra simulada y seleccionaremos la columna que haga referencia a la variable del paciente. El indicador de individuo de la muestra simulada se generará en la celda B10 a partir de una función bucle programa da en Visual Basic. Por ejemplo, `BUSCARV(B10;'1-Generación población'!B21:O370;2;0)` busca el identificador del individuo (`B10`) en la primera columna del rango `B21:O370` de la hoja '1-Generación población' identificando de esta forma la fila del rango `B21:O370` donde aparece y mostrará el valor de la columna 2 que corresponde con la Edad. Así se ha hecho con todas las variables de caracterización disponibles.

CELDA	FÓRMULA
D10	=BUSCARV(\$B\$10;'1-Generación población'!\$B\$21:\$O\$370;2;0)
E10	=BUSCARV(\$B\$10;'1-Generación población'!\$B\$21:\$O\$370;3;0)
F10	=BUSCARV(\$B\$10;'1-Generación población'!\$B\$21:\$O\$370;4;0)
G10	=BUSCARV(\$B\$10;'1-Generación población'!\$B\$21:\$O\$370;5;0)
H10	=BUSCARV(\$B\$10;'1-Generación población'!\$B\$21:\$O\$370;6;0)

CELDA	FÓRMULA
I10	=BUSCARV(\$B\$10;'1-Generación población'!\$B\$21:\$O\$370;7;0)
J10	=BUSCARV(\$B\$10;'1-Generación población'!\$B\$21:\$O\$370;8;0)
K10	=BUSCARV(\$B\$10;'1-Generación población'!\$B\$21:\$O\$370;9;0)
L10	=BUSCARV(\$B\$10;'1-Generación población'!\$B\$21:\$O\$370;10;0)
M10	=BUSCARV(\$B\$10;'1-Generación población'!\$B\$21:\$O\$370;11;0)
N10	=BUSCARV(\$B\$10;'1-Generación población'!\$B\$21:\$O\$370;12;0)
O10	=BUSCARV(\$B\$10;'1-Generación población'!\$B\$21:\$O\$370;13;0)
P10	=BUSCARV(\$B\$10;'1-Generación población'!\$B\$21:\$O\$370;14;0)

Una vez tenemos las características de los pacientes debemos valorar si del ciclo del que provienen al nuevo ciclo siguen vivos. Para ello, debemos tener en cuenta que hay dos fuentes de mortalidad.

1. Mortalidad específica según el nivel de severidad ('pMuerteLeve', 'pMuerteModerado' y 'pMuerteSevero').
2. Mortalidad general por edad y sexo que provienen de las tablas de mortalidad del instituto nacional de estadística (32). En ese sentido se debe obtener el riesgo de muerte en tanto por 1 y de ahí que se tenga que transformar en la columna E y F (Figura 10). Esta sección se ha denominado 'tMortalidad' para su utilización posterior.

Para poder saber si el paciente muere (celda R10), se hace un control de que está vivo para posteriormente estimar a partir de una variable aleatoria si se muere en el periodo. Como en este caso tenemos dos potenciales fuentes de mortalidad, entonces el marcador de mortalidad será la mayor de las dos mortalidades posibles. En función del nivel de severidad del estado de salud global (celda P10) se aplicará de probabilidad de muerte específica. Para la mortalidad general, se controla si es

Figura 10. Mortalidad según edad y sexo

Resultados nacionales, por comunidades autónomas y provincias					
Series desde 1991					
Tablas de mortalidad por año, sexo, edad y funciones.					
Unidades: Personas, %1000, Años, Años-persona					
	2020		Hombre	Mujer	
	Hombres	Mujeres			
0 años	2,723213	2,293295	0	0,002723	0,002293
1 año	0,225523	0,171995	1	0,000226	0,000172
2 años	0,123444	0,083722	2	0,000123	8,37E-05
3 años	0,120790	0,093537	3	0,000121	9,35E-05
4 años	0,080290	0,018902	4	8,03E-05	1,89E-05
5 años	0,078429	0,050983	5	7,84E-05	5,1E-05
6 años	0,060818	0,060102	6	6,08E-05	6,01E-05
7 años	0,046813	0,076895	7	4,68E-05	7,69E-05
8 años	0,040675	0,082275	8	4,07E-05	8,23E-05

hombre o mujer para mostrar la columna 2 de hombres o la columna 3 de mujer (SI(E10="Mujer";1;0)) y la fila de la tabla de mortalidad se basa en la edad del individuo en ese periodo (Figura 10).

La celda Q10 para el ciclo 1 es nula, ya que es el primero. Para el resto de los ciclos deberá mostrar en qué ciclo se murió el paciente. De esa forma, podemos controlar qué imputación de costes o de efectos se le debe hacer en el ciclo que muere, al ser coste de transición o la corrección de medio ciclo, en el caso que se incluyan.

CELDA	FÓRMULA
Q10	Nulo para el ciclo 1. Para el resto tiene estructura similar =BUSCARV(\$B\$10;'4- Ciclo 1'!\$B\$28:\$R\$377;16;0) =SI(Q10=""; SI(ALEATORIO)<=MAX(ELEGIR(COINCIDIR(P10;'2- Probabilidades'!\$B\$125:\$B\$127); pMuerteLeve; pMuerteModerado; pMuerteSevero); BUSCARV(D10; tMortalidad;2+ SI(E10="Mujer";1;0); VERDADERO));"No"; "Sí";"No")

Ahora ya estamos en disposición de hacer que los individuos evolucionen (filas 12 a 16). Para ello, deberemos modificar la edad del paciente (D15) en función de si sigue vivo (Edad del ciclo previo más tiempo en el ciclo, $D10 + \text{ciclo}/12$) o, por el contrario, en el caso de que haya fallecido en el ciclo se le añade la corrección de medio ciclo (33) y si ha fallecido con anterioridad ($Q15 \neq C10$) entonces se le mantiene fija la edad. Por otro lado, mantenemos el sexo del paciente (E15) y llevamos un control del número de órganos afectados (F15).

CELDA	FÓRMULA
D15	=SI(\$R\$10="Sí"; D10+ciclo/12;SI(Q15=C10;D10+ciclo/12/2;D10))
E15	=E10
F15	=CONTAR.SI(G15:N15;"Sí")

Para estimar la evolución en el nivel de severidad por órgano se debe actuar de la misma forma que cuando generamos la población con la distribución Bernoulli. Debemos basarnos en una variable aleatorio que nos indicará si, por ejemplo, del estado de salud leve se transita a moderado o a severo, o por si lo contrario se mantiene en el estado leve. En este caso se debe definir en una celda la variable volátil ALEATORIO(), ya que vamos a tener que referenciarla más de 1 vez para poder saber el nivel de severidad del paciente (H15). Para ello, se ha incluido la variable en la fila 16, generando una variable aleatoria para cada órgano. Por tanto, en el caso de tener afectado el órgano, a partir del nivel de severidad del ciclo previo (H10) sabemos qué probabilidad de transición se debe aplicar para cada tratamiento (de leve, de moderado o de severo). Tal como se ha creado la lógica de la afectación de órganos (G15), se mantienen los órganos afectados (G10) aunque pudieran llevar mucho tiempo como leves.

CELDA	FÓRMULA
G15	=G10
H15	=SI(G10="Sí"; ELEGIR(COINCIDIR(H10;NivelSeveridad;0); SI(H16<=pLeveLeveOrg1Actual; "Leve"; SI(H16<=pLeveLeveOrg1 Actual+pLeveModOrg1Actual; "Moderado"; "Severo")); SI(H16<=pModLeveOrg1Actual; "Leve"; SI(H16<=pModLeveOrg1 Actual+pModModOrg1Actual; "Moderado"; "Severo")); SI(H16<=pSevLeveOrg1Actual; "Leve" ; SI(H16<=pSevLeveOrg1A ctual+pSevModOrg1Actual; "Moderado"; "Severo")));H10)

Estas fórmulas se repiten para los 4 órganos (columnas I a N), pudiéndose calcular de nuevo el indicador del estado de salud global (O15) y el nivel de severidad (P15). Si el paciente se muere en ese periodo (R10="No") y no tiene el ciclo en el que se ha muerto (Q10="") entonces se vincula el ciclo actual como el de la muerte (C10), sino se mantiene la información de la celda como vacía (Q15 = ""). Por último, se indica el estado de mortalidad (R15=R10).

CELDA	FÓRMULA
O15	=SI(G15="Sí";12; SI(I15="Sí"; SI(J15="Leve";1;SI(J15="Moderado";2;3));0)+ SI(K15="Sí"; SI(L15="Leve";1;SI(L15="Moderado";2;3));0)+ SI(M15="Sí"; SI(N15="Leve";1;SI(N15="Moderado";2;3));0))
P15	=SI(O15<3;"Leve"; SI(O15<7;"Moderado";"Severo"))
Q15	=SI(R10="No"; SI(Q10="";C10;Q10);"")
R15	=R10

Una vez se dispone de la estructura para la evolución del ciclo se deberá ejecutar una macro de Visual Basic que identificará al paciente de la muestra simulada (B10); y copiará el rango B15:R15 y la pegará a

```

Sub GeneracionCiclo()
  ActiveSheet.Select
  Trials = Range("TamañoMuestra").Value
  For i = 1 To Trials
    Range("B10").Value = i
    Range("B15:R15").Select
    Selection.Copy
    ActiveCell.Offset(i + 12, 0).Range("A1").Select
    Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, Skip-
Blanks:= _
      False, Transpose:=False
  Next i
End Sub
  
```



partir de la fila 28 repetidamente mediante un bucle hasta que se alcance el número máximo de pacientes (n: 350).

Es el momento de preparar el modelo para cálculos los años de vida en el ciclo en la columna T, los años de vida ajustados por calidad en el ciclo en la columna U y los costes del ciclo en la columna V tal como se muestra en la Figura 11.

Figura 11. Cálculo de años de vida, años de vida ajustados por calidad y costes por individuo simulado en un ciclo

SIN DESCONTAR			DESCONTADO		
AÑOS DE VIDA	AVACs	COSTE	AÑOS DE VIDA	AVACs	COSTE
0,47	0,37	1.782,38	0,47	0,37	1.782,38

SIN DESCONTAR			DESCONTADO				
AÑOS DE VIDA	AVACs	COSTE	DESCUENTO	AÑOS DE VIDA	AVACs	DESCUENTO	COSTE
0,50	0,47	1.750,00	1	0,50	0,47	1	1750,00
0,50	0,47	600,00	1	0,50	0,47	1	600,00
0,50	0,35	1.800,00	1	0,50	0,35	1	1800,00
0,50	0,35	3.135,00	1	0,50	0,35	1	3135,00
0,50	0,47	600,00	1	0,50	0,47	1	600,00
0,50	0,47	600,00	1	0,50	0,47	1	600,00
0,25	0,23	150,00	1	0,25	0,23	1	150,00
0,50	0,35	754,00	1	0,50	0,35	1	754,00
0,50	0,27	9.800,00	1	0,50	0,27	1	9800,00

De ese modo para el primer individuo que está en la fila 28 para calcular el tiempo del ciclo nos basamos en la edad en el ciclo (D28) menos la edad en el periodo previo ('1-Generación población'!C21 para el primer ciclo y '4- Ciclo 1'!D28 para los siguientes). Para el cálculo de los años deberemos identificar el nivel de severidad para el estado de salud global del individuo (P28) asociarlo a su utilidad ('uLeve', 'uModerado' o 'uSevero') por el tiempo. Para los costes del ciclo de un individuo deberemos identificar qué órganos tiene afectado y asociarle el coste por nivel de severidad y órgano (por ejemplo, para el segundo órgano, 'cLeveOrg2', 'cModOrg2' y 'cSevOrg2'), así como imputar el coste del tratamiento. Todo el coste se multiplica por tiempo del ciclo que está vivo el paciente, y de esa forma, queda igual a 0 cuando el paciente ha muerto en ciclos previos.

CELDA	FÓRMULA
T28	=D28-'1-Generación población'!C21
U28	=ELEGIR(COINCIDIR(P28; NivelSeveridad;0); uLeve; uModerado; uSevero) * (D28-'1-Generación población'!C21)
V28	=(SI(G28="Si"; ELEGIR(COINCIDIR(H28; NivelSeveridad;0); cLeveOrg1; cModOrg1; cSevOrg1);0)+ SI(I28="Si";ELEGIR(COINCIDIR(J28; NivelSeveridad;0); cLeveOrg2; cModOrg2; cSevOrg2);0)+ SI(K28="Si";ELEGIR(COINCIDIR(L28; NivelSeveridad;0); cLeveOrg3; cModOrg3; cSevOrg3);0)+ SI(M28="Si";ELEGIR(COINCIDIR(N28; NivelSeveridad;0); cLeveOrg4; cModOrg4; cSevOrg4);0)+ cActual) *(D28-'1-Generación población'!C21)/(ciclo/12)

Ahora ya estamos en disposición de estimar el descuento para el marco temporal del ciclo para los efectos (columna X) y para los costes (columna AA), para aplicarlo a los años de vida, los años de vida ajustados por calidad y los costes descontados. Tenga en cuenta que el

descuento solo se aplica a partir del año (3 ciclo) y que se ha ajustado en función del tiempo del ciclo (REDONDEAR(C28*ciclo/12;0)-1).

CELDA	FÓRMULA
X28	=1/(1+desEfectos)^(REDONDEAR(C28*ciclo/12;0)-1)
Y28	=T28*X28
Z28	=U28*X28
AA28	=1/(1+desCostes)^(REDONDEAR(C28*ciclo/12;0)-1)
AB28	=V28*AA28

Una vez se ha propagado las fórmulas para los 350 individuos, entonces podemos resumir los resultados de los costes y efectos en la fila 22, haciendo el promedio de los resultados individuales de ese ciclo.

CELDA	FÓRMULA
T22	=PROMEDIO(T28:T377)
U22	=PROMEDIO(U28:U377)
V22	=PROMEDIO(V28:V377)
Y22	=PROMEDIO(Y28:Y377)
Z22	=PROMEDIO(Z28:Z377)
AB22	=PROMEDIO(AB28:AB377)

Este proceso de evolución es el mismo para todos los ciclos (1 a 6) siendo las probabilidades de transición diferentes en función si estamos en tratamiento actual (de '4- Ciclo 1' a '4- Ciclo 6') o en tratamiento nuevo (de '5- Ciclo 1 Nuevo' a '5- Ciclo 6 Nuevo'). Se debe tener en cuenta que para el primer ciclo los datos previos son los de la generación y para el resto siempre es el ciclo previo.

Paso 4: Resumir los resultados

La hoja ‘6-Resumen población’ permite hacer un resumen de los años de vida, los años de vida ajustados por calidad (AVACs) y los costes para ciclo y cada tratamiento (Figura 12).

Figura 12. Resultados por ciclo del tratamiento actual y nuevo por ciclo con su diferencial

TRATAMIENTO ACTUAL

CICLO	SIN DESCONTAR			ACUMULADO			DESCONTADO			ACUMULADO		
	AÑOS DE VIDA	AVACs	COSTE	AÑOS DE VIDA	AVACs	COSTE	AÑOS DE VIDA	AVACs	COSTE	AÑOS DE VIDA	AVACs	COSTE
1	0,47	0,37	1.782,38	0,47	0,37	1.782,38	0,47	0,37	1.782,38	0,47	0,37	1.782,38
2	0,43	0,34	1.436,91	0,90	0,71	3.219,28	0,42	0,33	1.395,06	0,89	0,70	3.177,43
3	0,39	0,32	1.089,05	1,30	1,03	4.308,33	0,38	0,31	1.057,33	1,27	1,01	4.234,76
4	0,37	0,30	1.041,72	1,67	1,34	5.350,05	0,35	0,29	981,92	1,62	1,30	5.216,68
5	0,35	0,29	909,20	2,02	1,63	6.259,25	0,33	0,27	857,01	1,95	1,57	6.073,69
6	0,33	0,27	857,25	2,34	1,90	7.116,51	0,30	0,25	784,51	2,25	1,82	6.858,20

TRATAMIENTO NUEVO

CICLO	SIN DESCONTAR			ACUMULADO			DESCONTADO			ACUMULADO		
	AÑOS DE VIDA	AVACs	COSTE	AÑOS DE VIDA	AVACs	COSTE	AÑOS DE VIDA	AVACs	COSTE	AÑOS DE VIDA	AVACs	COSTE
1	0,47	0,38	678,34	0,47	0,38	678,34	0,47	0,38	678,34	0,47	0,38	678,34
2	0,43	0,36	2.167,27	0,91	0,74	2.845,60	0,42	0,35	2.104,14	0,89	0,73	2.782,48
3	0,41	0,34	1.918,91	1,31	1,09	4.764,52	0,39	0,33	1.863,02	1,29	1,07	4.645,50
4	0,38	0,33	1.634,35	1,69	1,42	6.398,87	0,36	0,31	1.540,53	1,65	1,38	6.186,04
5	0,36	0,31	1.501,44	2,05	1,73	7.900,31	0,34	0,30	1.415,25	1,98	1,67	7.601,29
6	0,35	0,30	1.484,24	2,40	2,04	9.384,55	0,32	0,28	1.358,29	2,30	1,95	8.959,58

DIFERENCIA (Nuevo-Actual)

CICLO	SIN DESCONTAR			ACUMULADO			DESCONTADO			ACUMULADO					
	AÑOS DE VIDA	AVACs	COSTE	AÑOS DE VIDA	AVACs	COSTE	AÑOS DE VIDA	AVACs	COSTE	AÑOS DE VIDA	AVACs	COSTE			
1	-	0,00	0,01	-	1.104,04	-	0,00	0,01	-	1.104,04	-	0,00	0,01	-	1.104,04
2	0,01	0,02	730,36	0,00	0,03	-	373,68	0,01	0,02	709,09	0,00	0,03	-	394,95	
3	0,01	0,02	829,87	0,02	0,05	456,19	0,01	0,02	805,69	0,02	0,05	410,74			
4	0,01	0,03	592,63	0,03	0,08	1.048,82	0,01	0,03	558,61	0,03	0,08	969,36			
5	0,01	0,02	592,24	0,03	0,11	1.641,06	0,01	0,02	558,24	0,03	0,10	1.527,60			
6	0,02	0,03	626,99	0,05	0,14	2.268,04	0,02	0,03	573,78	0,05	0,13	2.101,38			

Al final se muestra la ratio coste-efectividad incremental y la ratio coste-utilidad incremental (Figura 13), así como se han modificado las características de los pacientes del momento basal hasta el último ciclo en función del tratamiento (Figura 14).

Figura 13. Ratio coste-efectividad y coste-utilidad incremental por ciclo

RATIO COSTE-EFECTIVIDAD INCREMENTAL

CICLO	SIN DESCONTAR		DESCONTADO	
	AÑOS DE VIDA	AVACs	AÑOS DE VIDA	AVACs
1	1.545.654,00	Dominante	1.545.654,00	Dominante
2	Dominante	Dominante	Dominante	Dominante
3	24.563,92	8.436,75	22.806,79	7.767,47
4	37.649,90	12.872,32	36.220,82	12.314,94
5	46.887,31	15.603,63	45.606,40	15.117,74
6	42.908,89	16.740,96	42.164,97	16.317,32

Paso 5: Agregar por muestra simulada

El proceso descrito hasta ahora era para la transición a través de los ciclos de solo 1 muestra simulada tanto para la terapia actual como para la terapia nueva. Esto no es suficiente para evitar el potencial error de simulación como se ha comentado en la sección de convergencia. Por tanto, se debe repetir los pasos 1 a 4 con muchas muestras hasta haber simulado 1.000.000 individuos, en el mejor de los casos. Como en el paso 4 de resumen, deberemos acumular la información de cada muestra para poder posteriormente estimar la media de todas las muestras simuladas disponiendo finalmente el resultado global del modelo de microsimulación. Queda para el lector intentar este último paso a partir de lo ya explicado.

Figura 14. Resultados por ciclo entre el tratamiento nuevo y el actual

CARACTERÍSTICAS PACIENTES

INICIAL	ÓRGANOS AFECTADOS (#)	ÓRGANO 1		ÓRGANO 2		ÓRGANO 3		ÓRGANO 4		ESTADO DE SALUD	
		Afectación	Severidad	Afectación	Severidad	Afectación	Severidad	Afectación	Severidad	Score	Severidad
50,26	1,80	10,0%	Leve: 22,86% Moderado: 25,71% Severo: 51,43%	32,3%	Leve: 3,54% Moderado: 18,58% Severo: 77,88%	0%	Leve: 30,19% Moderado: 0% Severo: 69,81%	49,7%	Leve: 60,92% Moderado: 0% Severo: 39,08%	4,73	Leve: 16,57% Moderado: 66,86% Severo: 16,57%

FINAL: TRATAMIENTO ACTUAL

FINAL: TRATAMIENTO ACTUAL	ÓRGANOS AFECTADOS (#)	ÓRGANO 1		ÓRGANO 2		ÓRGANO 3		ÓRGANO 4		ESTADO DE SALUD		CICLO DE LA MUERTE	
		Afectación	Severidad	Afectación	Severidad	Afectación	Severidad	Afectación	Severidad	Score	Severidad		
52,60	1,80	10,0%	Leve: 17,14% Moderado: 57,14% Severo: 25,71%	32,3%	Leve: 70,8% Moderado: 22,12% Severo: 7,08%	83,44%	Leve: 88,0% Moderado: 11,69% Severo: 4,87%	49,7%	Leve: 35,06% Moderado: 32,76% Severo: 32,18%	3,46	Leve: 50% Moderado: 39,71% Severo: 10,29%	2,86	64,0%

FINAL: TRATAMIENTO NUEVO

FINAL: TRATAMIENTO NUEVO	ÓRGANOS AFECTADOS (#)	ÓRGANO 1		ÓRGANO 2		ÓRGANO 3		ÓRGANO 4		ESTADO DE SALUD		CICLO DE LA MUERTE	
		Afectación	Severidad	Afectación	Severidad	Afectación	Severidad	Afectación	Severidad	Score	Severidad		
52,66	1,80	10,0%	Leve: 91,43% Moderado: 8,57% Severo: 0%	32,3%	Leve: 80,53% Moderado: 15,93% Severo: 3,54%	91,88%	Leve: 88,0% Moderado: 5,84% Severo: 2,27%	49,7%	Leve: 85,63% Moderado: 11,49% Severo: 2,87%	2,98	Leve: 68,29% Moderado: 21,71% Severo: 10%	2,59	69,1%

Paso 6: Verificar y validar

Asimismo, antes de finalizar el modelo de simulación individual de pacientes y analizar los resultados se debe verificar y validar.

- **Verificación:** es la comprobación de la consistencia de la estructura lógica del modelo. Por ejemplo, que los pacientes que fallecen no computen a partir del momento del fallecimiento ni años de vida, ni años ajustados por calidad ni costes. Hasta que no se verifique por completo el modelo, se deberá revisar la implementación del mismo.
- **Validación:** es el proceso que sirve para determinar el grado de semejanza entre el modelo de simulación coincide con la casuística observada en la patología a modelar (23). En definitiva, es la comparación de los resultados del modelo respecto el sistema real. En este caso, la simulación de la supervivencia del caso base debería ajustarse a la observada en la literatura. En el caso de no alcanzar dicha validación, se deberá revisar los supuestos sobre las distribuciones o sobre los parámetros de los mismos.

Este proceso es muy importante, ya que de otra forma no se estarán controlando potenciales sesgos asociadas a la implementación del modelo y por tanto podemos estar obteniendo resultados erróneos. Para aquellos que lo deseen pueden ver un proceso completo de verificación y validación en el estudio de Bongers y col. (34) vinculado al cáncer de pulmón.

Una vez ya esté todo confirmado es el momento de poder analizar el caso base y también realizar los análisis de sensibilidad pertinentes. En el caso del querer realizar el análisis probabilístico de Monte Carlo de 2º orden sobre la microsimulación, simplemente será necesario modificar la variable 'AnDeterministico' de la hoja '2- Probabilidades' por el valor 0 y realizar los pasos de 1 a 5.



Más allá del Excel

Aunque el ejemplo de modelo de microsimulación se ha desarrollado sobre Excel, existen otro tipo de programas con mayor capacidad de cálculo para realizar los modelos de simulación como es R. Realmente, si bien Microsoft Excel, incluso otro como TreeAge Pro son buenos programas a nivel didácticos y para producir los tipos de análisis que suelen requerir las agencias de evaluación de tecnologías sanitarias, las ventajas de eficiencia y transparencia de los lenguajes de programación como MATLAB y R se vuelven cada vez más valiosas cuando se realizan análisis más complejos (35,36).

R es el lenguaje de programación estadística más poderoso y flexible del mundo. Tiene la dificultad que la curva de aprendizaje al inicio es bastante elevada, ya que es sobre todo un cambio de mentalidad a diferencia del uso de otros paquetes estadístico. No obstante, pasado ese esfuerzo inicial, el proceso de aprendizaje es rápido. En el fondo, una vez aprendida la sintaxis y comprendida la orientación a objetos, aprender a aplicar una nueva técnica es absolutamente inmediato. Además, la comunidad de usuarios de R es ya la más activa en el mundo de los paquetes estadísticos, siendo la primera herramienta de desarrollo de nuevos modelos y técnicas. Dicho de otra forma, cualquier nueva metodología se encuentra disponible en R al poco de existir.

Todo procedimiento en R es, por su naturaleza de lenguaje de programación, fácilmente reusable y compatible, especialmente dentro de las organizaciones. R puede absorber grandes volúmenes de datos de fuentes como Hadoop, siendo capaz de leer prácticamente cualquier tipo de datos (.txt, .dat, .csv...), así como leer JSON, SPSS, SAS, Excel, STATA... Todo además de una forma absolutamente gratuita, lo cual puede llevar a las organizaciones a ahorrar cientos de miles de euros en licencias a lo largo de los años.

R no es un lenguaje de programación que se haya desarrollado centrándose en el rendimiento computacional. Sin embargo, hay algunas

características del lenguaje de programación R que pueden aprovecharse para que los programas codificados en R tengan un rendimiento eficiente. Un atributo clave de R es su capacidad para procesar problemas de forma más eficiente cuando se presentan en un vector en lugar de en un formato escalar. La vectorización puede reducir la necesidad de recurrir a los bucles iterativos como es la función “for” o la función “Do-While” que hemos utilizado en nuestro ejemplo de código de visual basic y, que, en general, aumentan considerablemente el tiempo de cálculo. La vectorización es especialmente importante en el análisis de decisiones, ya que los procesos suelen ser repetitivos y relativamente sencillos de vectorizar.

A nivel de programación de modelos de simulación hay algunas experiencias exitosas (37) y algunos paquetes específicos para evaluación económica que pueden ser de gran ayuda para generar estos modelos como el HESIM (38) o RDECISION(39), aunque este último más pensado para cohortes que para pacientes individuales. Aquel lector con conocimientos previos en R y que quiera profundizar le recomiendo al artículo de Krijkamp EM y col. (37).



Bibliografía

1. Rodríguez Barrios JM, Pérez Alcántara F, Crespo Palomo C, et al. The use of cost per life year gained as a measurement of cost-effectiveness in Spain: a systematic review of recent publications. *Eur J Health Econ.* 2012;13:723–40.
2. Catalá-López F, Ridao M, Alonso-Arroyo A, et al. The quality of reporting methods and results of cost-effectiveness analyses in Spain: a methodological systematic review. *Syst Rev.* 2016;5:6.
3. Salleh S, Thokala P, Brennan A, et al. Simulation Modelling in Healthcare: An Umbrella Review of Systematic Literature Reviews. *Pharmacoeconomics.* 2017;35:937–49.
4. Brennan A, Chick SE, Davies R. A taxonomy of model structures for economic evaluation of health technologies. *Health Econ.* 2006;15:1295–310.
5. Siebert U, Alagoz O, Bayoumi AM, et al. State Transition Modeling: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-3. *Value Health.* 2012;15:812–20.
6. Barton P, Bryan S, Robinson S. Modelling in the economic evaluation of health care: selecting the appropriate approach. *J Health Serv Res Policy.* 2004;9:110–8.
7. Lin F, Thomas S, Baldwin M. Microsimulation or Cohort Modelling? A Case Study in Chronic Obstructive Pulmonary Disease (COPD). *Value Health.* 2013;16:A586.
8. Dulai PS, Sandborn WJ, Murphy J. Microsimulation Model to Determine the Cost-Effectiveness of Treat-to-Target Strategies for Ulcerative Colitis. *Clin Gastroenterol Hepatol.* 2021;19:1170-1179.
9. Zhou H, Zhao Y, Wojciech A, et al. Differences in Cost-Effectiveness Estimates for Chronic Hepatitis C Treatment Among Cohort Markov Model, Markov Microsimulation and Discrete Event Simulation. *Value Health.* 2016;19:A853.
10. Getaneh AM, Heijnsdijk EA, Koning HJ. Cost-effectiveness of multiparametric magnetic resonance imaging and MRI-guided biopsy in a popu-

- lation-based prostate cancer screening setting using a micro-simulation model. *Cancer Med.* 2021;10:4046–53.
11. Willis M, Fridhammar A, Gundgaard J, et al.. Comparing the Cohort and Micro-Simulation Modeling Approaches in Cost-Effectiveness Modeling of Type 2 Diabetes Mellitus: A Case Study of the IHE Diabetes Cohort Model and the Economics and Health Outcomes Model of T2DM. *Pharmacoeconomics.* 2020;38:953–69.
 12. Crespo C, Pérez-Simón JA, Rodríguez JM, et al. Development of a population-based cost-effectiveness model of chronic graft-versus host disease in Spain. *Clin Ther.* 2012;34:1774–87.
 13. Furiak NM, Ascher-Svanum H, Klein RW, et al. Cost-effectiveness of olanzapine long-acting injection in the treatment of patients with schizophrenia in the United States: a micro-simulation economic decision model. *Curr Med Res Opin.* 2011;27:713–30.
 14. Hiligsmann M, Ethgen O, Bruyère O, et al. Development and Validation of a Markov Microsimulation Model for the Economic Evaluation of Treatments in Osteoporosis. *Value Health.* 2009;12:687–96.
 15. Kuehne F, Chancellor J, Mollon P, et al. M. Microsimulation or cohort modelling? A comparative case study in HIV infection. *Value Health.* 2007;10:A170.
 16. Saidu NEB, Bonini C, Dickinson A, et al. New Approaches for the Treatment of Chronic Graft-Versus-Host Disease: Current Status and Future Directions. *Frontiers in Immunology.* 2020; 11: 5783-4.
 17. Bliuc D, Alarkawi D, Nguyen T V, et al. Risk of Subsequent Fractures and Mortality in Elderly Women and Men with Fragility Fractures with and without Osteoporotic Bone Density: The Dubbo Osteoporosis Epidemiology Study. *J Bone Miner Res.* 2015;30:637–46.
 18. Bliuc D. Mortality Risk Associated With Low-Trauma Osteoporotic Fracture and Subsequent Fracture in Men and Women. *JAMA.* 2009;301:513.
 19. Davahli MR, Karwowski W, Taiar R. A System Dynamics Simulation Applied to Healthcare: A Systematic Review. *Int J Environ Res Public Health.* 2020;17:5741.

20. B. Mielczarek. Review of modelling approaches for healthcare simulation. *Oper Res Decis.* 2016;1:55–72.
21. Yaghoub M, Mansell K, Hassanali V, et al. Patient-level micro-simulation model for evaluating the future potential cost–effectiveness of pharmacy-based interventions in the control and management of diabetes-related complications in Canada. *medRxiv.* 2020; [https://www.medrxiv.org/content/10.1101/2020.03.10.20033597v1/](https://www.medrxiv.org/content/10.1101/2020.03.10.20033597v1) (acceso verificado el 14 Oct 2023)
22. Hatswell AJ, Bullement A, Briggs A, et al. Probabilistic Sensitivity Analysis in Cost-Effectiveness Models: Determining Model Convergence in Cohort Models. *PharmacoEconomics.* 2018;36:1421–6.
23. Karnon J, Vanni T. Calibrating Models in Economic Evaluation. *PharmacoEconomics.* 2011;29:51–62.
24. Griffin S, Claxton K, Hawkins N, et al. Probabilistic Analysis and Computationally Expensive Models: Necessary and Required? *Value Health.* 2006;9:244–52.
25. Lipscomb J. Time Preference for Health in Cost-Effectiveness Analysis. *Med Care.* 1989;27(Suppl.):S233–53.
26. López Bastida J, Oliva J, Antoñanzas F, et al. A proposed guideline for economic evaluation of health technologies. *Gac Sanit.* 2010;24:154–70.
27. Husereau D, Drummond M, Augustovski F, et al. Consolidated Health Economic Evaluation Reporting Standards 2022 (CHEERS 2022) Statement: Updated Reporting Guidance for Health Economic Evaluations. *Value Health.* 2022;25:3–9.
28. Fleurence RL, Hollenbeak CS. Rates and Probabilities in Economic Modelling. *PharmacoEconomics.* 2007;25:3–6.
29. Gidwani R, Russell LB. Estimating Transition Probabilities from Published Evidence: A Tutorial for Decision Modelers. *PharmacoEconomics.* 2020;38: 1153–64.
30. Briggs A, Sculpher M, Claxton K. Decision modelling for health economic evaluation. Oxford: Oxford University Press; 2006.
31. 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.

32. Instituto Nacional de Estadística (INE). Tablas de mortalidad por año, sexo, edad y funciones (2020). Resultados nacionales, por comunidades autónomas y provincias. Instituto Nacional de Estadística. 2021. <https://www.ine.es> (acceso verificado el 14 Oct 2023)
33. Naimark DMJ, Bott M, Krahn M. The half-cycle correction explained: two alternative pedagogical approaches. *Med Decis Making*. 2009;28:706–12.
34. Bongers ML, de Ruyscher D, Oberije C, et al. Multistate Statistical Modeling. *Med Decis Making*. 2016;36:86–100.
35. Hollman C, Paulden M, Pechlivanoglou P, et al. A Comparison of Four Software Programs for Implementing Decision Analytic Cost-Effectiveness Models. *PharmacoEconomics*. 2017;35:817–30.
36. Incerti D, Thom H, Baio G, et al. R You Still Using Excel? The Advantages of Modern Software Tools for Health Technology Assessment. *Value Health*. 2019;22:575–9.
37. Krijkamp EM, Alarid-Escudero F, Enns EA, et al. Microsimulation Modeling for Health Decision Sciences Using R: A Tutorial. *Med Decis Making*. 2018;38:400–22.
38. Incerti D, Janssen JP, Team RDC. Health Economic Simulation Modeling and Decision Analysis (Hesim). 2021. p. 141. <https://cran.r-project.org/web/packages/hesim/hesim.pdf> (acceso verificado el 14 Oct 2023)
39. Sims A, Fairbairn K, Team RDC. Decision Analytic Modelling in Health Economics (rdecision). 2021. p. 72. <https://cran.r-project.org/web/packages/rdecision/rdecision.pdf> (acceso verificado el 14 Oct 2023).



12

MODELOS DE SIMULACIÓN DE EVENTOS DISCRETOS (DISCRETE EVENT SIMULATION MODELS)

Javier Mar

*Osakidetza-Servicio Vasco de Salud
Organización Sanitaria Integral Debagoiena*



Características básicas

La evaluación de tecnologías sanitarias (ETS) como herramienta para la toma de decisiones implica una síntesis de diferentes fuentes de información en la que la aportación de los modelos de representación de la historia natural de la enfermedad es clave para el resultado final (1). Un modelo captura las relaciones de causa y efecto relevantes de un problema de salud y las proyecta en el tiempo en función de las alternativas en estudio. Para ello es necesario que esté validado lo que implica documentar que reproduce la realidad (2). En este proceso resulta de gran ayuda disponer de herramientas o técnicas para convertir el modelo conceptual en un conjunto de módulos matemáticos vinculados y capaces de realizar los cálculos y producir un resultado en un formato estándar. De esta manera, se facilita la transparencia del proceso y la comunicación de los resultados a los responsables de la toma de decisiones (3).

La simulación de eventos discretos (SED) es una herramienta de construcción de modelos de ETS junto con las otras técnicas utilizadas en la ETS que se describen en este libro como los árboles de decisión, los modelos de transición de estado (modelos de Markov) y los modelos de transmisión dinámica. La SED aporta mayor flexibilidad para reproducir la historia natural de la enfermedad, pero adolece de una curva de aprendizaje con mayor pendiente. La consecuencia práctica ha sido que se haya utilizado menos que los modelos de Markov que se han convertido en la herramienta estándar. Sin embargo, las fortalezas de la SED deberían ser tenidas en cuenta ya que permiten analizar problemas de ETS que son inabordables por los otros tipos de modelos como los modelos poblacionales con múltiples cohortes y los modelos de optimización del uso de recursos (4,5). En este capítulo mostraremos qué es la SED y que ventajas aporta y que retos requiere para

su uso en la ETS (5). Para profundizar en su conocimiento se recomienda el libro de Caro y col. “*Discrete Event Simulation for Health Technology Assessment*” que aborda específicamente el uso de la SED en ETS (1).

Origen

DES fue desarrollado originalmente en la década de 1960 en los campos de la ingeniería industrial y de investigación operativa para representar procesos industriales y empresariales. Sus primeras aplicaciones de atención médica aparecen en los años 80. En la literatura de ETS en español, la primera referencia es el artículo de Rodríguez-Barrios y col. (6).

Definición

La SED se define como la técnica de modelización que reproduce el curso de los individuos en función de los eventos que experimentan y del efecto que tienen en su estado de salud actual y futuro y del uso de los recursos y otros componentes a lo largo del tiempo (1,5). La participación del azar es un aspecto básico de la simulación que plantea dificultades de comprensión pero que sin embargo es una de sus fortalezas (7). Como consecuencia de su carácter estocástico, los resultados del modelo pueden ser diferentes para conjuntos de individuos con las mismas características. Dado que las funciones que determinan las probabilidades con las que ocurren los posibles eventos contienen un parámetro aleatorio, entidades con las mismas características tienen un resultado distinto. Sin embargo, cuando el número de entidades es suficientemente grande los resultados estadísticos del modelo o sistema se estabilizan. De hecho, el cálculo del

número de entidades necesario para que los resultados sean válidos depende de la variabilidad estocástica del sistema a representar. Si bien la naturaleza estocástica pueda parecer un mero aspecto técnico, permite que la SED pueda reproducir la naturaleza variable del mundo real (5). Este elemento aporta una gran flexibilidad a la hora de representar las trayectorias de los individuos en función de sus características, exposición a riesgos, ocurrencia de eventos, consumo de recursos y resultados en salud. Por ello la SED ha sido usada en una amplia gama de aplicaciones sanitarias. Inicialmente se utilizó para el análisis de sistemas con recursos limitados, con el propósito de mejorar la organización de los servicios prestados. Posteriormente la SED ha entrado en el grupo de las técnicas de modelización de aplicación en la ETS. El objetivo de este artículo va más dirigido a este segundo grupo de aplicaciones, aunque la descripción de la técnica sea también útil para entender los modelos de optimización de los recursos limitados (8).



Aplicaciones

La gama de problemas de atención médica que se pueden abordar con la SED es enorme. En este libro abordamos principalmente aquellos que son pertinentes a la ETS dirigida a evaluar una tecnología sanitaria para decidir sobre su uso en el sistema sanitario. Hay dos categorías de aplicaciones de la SED en la atención sanitaria: modelos de recursos no limitados dirigidos a la ETS o al análisis epidemiológico (4,9,10) diagnosis and treatment y modelos de recursos limitados. La SED se desarrolló en el mundo de la investigación operativa para analizar estos últimos mediante la optimización del uso de recursos (8,11,12) first-out (FIFO). Estos modelos incorporan explícitamente las limitaciones de capacidad de los recursos existente por los que compiten las

entidades o individuos y esperan en colas desde las que son atendidos en función de su estado de prioridad. Sin embargo, su uso en el mundo sanitario ha sido limitado. En la mayoría de los modelos utilizados para la evaluación económica de intervenciones sanitarias se parte del principio de que todos los recursos necesarios están disponibles sin limitaciones de capacidad. Otra aplicación de la SED que queda fuera del ámbito de este trabajo es la dirigida a construir modelos dinámicos de representación de enfermedades infecciosas en poblaciones mediante modelos basado en agentes (agent-based models) (13,14).

Los últimos años han vivido la publicación de múltiples artículos mostrando aplicaciones de la SED para la evaluación económica tanto en formato análisis coste-efectividad (11,15–18) como análisis del impacto presupuestario (19) o ambos (4,20). Estos trabajos aplican los principios de la evaluación económica en general calculando los años de vida ajustados por calidad y costes incrementales para medir la razón coste-efectividad incremental (7,21) e incluyen herramientas como las simulaciones de Monte Carlo o el valor esperado de la información perfecta (7,15,22). Pero la SED se ha aplicado también en el análisis epidemiológico de poblaciones permitiendo estimar prevalencia (19) y obtener parámetros inobservables mediante procedimientos de calibración (10). A partir de los mismos modelos se pueden evaluar el impacto epidemiológico de las políticas de salud pública como los cribados de cáncer o los impuestos sobre bebidas azucaradas (23). En estas evaluaciones, los sistemas objeto de la simulación pueden ser poblaciones constituidas por una cohorte o por múltiples cohortes. Precisamente estas últimas han sido más frecuentemente aplicadas por representar poblaciones dinámicas. Especialmente en salud pública al evaluar el impacto de políticas de salud a lo largo del tiempo se requiere construir sistemas dinámicos para los cuales los modelos de Markov resultan ineficientes. Por “población dinámica”, nos referimos a aquellas en las que los miembros que la componen varían con el tiempo (24).

A lo largo del tiempo nuevas cohortes entran en el sistema, se producen transiciones del estado sano al enfermo y los individuos de la población pueden morir.

Ventajas

En los artículos comparando diferentes herramientas de modelización de recomienda usar la SED para representar estructuras complejas cuya estructura mediante modelos de Markov resultaría muy costosa de programar por la cantidad de estados requeridos para para capturar las diferencias relevantes entre las diferentes alternativas en cuanto a costes y efectividad (1,3). La revisión de Karnon y Haji Ali Afzali describe con mayor detalle cuando usar la SED en ETS (25). No solamente el número de estados es una barrera para la construcción de los modelos de Markov, también lo es la característica de requerir que la población en cada estado sea homogénea, para poder cumplir con la sunción markoviana de que todos los individuos en un estado tienen las mismas probabilidades de transición (26). Sin embargo, en la ETS no se cumple ese principio y la SED resulta más apropiada. Pocos artículos han comparado ambas técnicas para evaluar una misma intervención (27).

Otro punto señalado en la revisión previamente citada señala al importancia de la progresión de la enfermedad como un proceso continuo medido en una escala continua en la selección de la técnica de modelización (25). La SED permite proyectar en el tiempo la progresión de la enfermedad a partir de funciones matemáticas como los modelos mixtos o la supervivencia paramétrica (18,23) y evita la obligación de los modelos de estados de transición de representar escalas continuas mediante la división en pequeñas categorías discretas.

Una tercera razón para usar la SED se relaciona con la representación de tasas de eventos que varían en el tiempo, donde la probabilidad de eventos posteriores varía según el tiempo pasado en el estado de

salud actual. La implementación de tasas de eventos variables en el tiempo es sencilla en un DES, ya que el tiempo se trata de forma continua y los riesgos pueden cambiar en consecuencia (1).

En resumen, la SED es más adecuada para la ETS en modelos complejos y cuando se requiera flexibilidad en la adaptación a la lógica de la representación de la historia natural de la enfermedad (25). Además, resulta obligatoria cuando la población a representar esté compuesta por poblaciones dinámicas con múltiples cohortes. Esto suele ser habitual en los modelos de evaluación de políticas de salud pública (9,23).

Programas de software para la SED

La utilización de la SED en ETS se ha llevado a cabo mediante tres tipos de programas: hojas de cálculo, lenguajes de programación general y paquetes de software especializados en SED. Excel, con programación en VBA ha sido utilizado aunque con limitaciones para la computación (18). Los lenguajes generales de programación informática como Java, Visual Basic(.net), Fortran. C++ también permiten construir modelos de simulación. Sin embargo, la mayor parte de los trabajos publicados han sido implementados en software específico para DES desarrollado de forma estándar como Arena de Rockwell Software, Simio, AnyLogic from XJ Technologies, ProModel, Witness, Simulation Studio from SAS y Simul8 (1). En los últimos años se ha propuesto una nueva herramienta llamada DICE que se apoya en Excel y que integra tanto el enfoque de los modelos de Markov como la SED (28). Los programas estadísticos como R también están siendo cada más usados en la programación de modelos de simulación tanto de micro-simulación como de SED (20,29,30).

Componentes de los modelos de SED

Los eventos son el principal componente de la SED. La base que utiliza la SED para construir matemáticamente el recorrido que va a seguir un individuo en el sistema analizado son los eventos que le pueden ocurrir. Lo relevante en ETS, es que en función de su ocurrencia, una entidad puede cambiar en un instante de tiempo alguna sus características o atributos y su relación con otras entidades o el sistema en su conjunto. Par el propósito de la ETS, un evento es algo que le sucede a una entidad durante una simulación. Una entidad es generalmente un individuo o un paciente en ETS, pero puede ser también un objeto, que experimenta eventos dentro del modelo. Los atributos contienen las características de una entidad y son la información que va a determinar la probabilidad de los eventos. Generalmente se distingue entre atributos que son características específicas de las entidades de las variables que son valores que afectan al conjunto de individuos y que varían en cada simulación. En la jerga de la modelización los atributos corresponden a la incertidumbre de primer orden y las variables a la de segundo orden (31).

El tiempo es continuo en el sentido de que los eventos pueden ocurrir en cualquier momento durante la simulación y avanza a medida que ocurren los eventos en el sistema. Por el contrario, en los modelos de Markov el tiempo es discreto y avanza ciclo a ciclo. Sin embargo, la SED permite programar bucles que pueden reproducir ciclos si la información disponible está en este formato. Así en un modelo de hepatitis construido con SED, la etapa de progresión de la fibrosis se programó de forma similar a la llevada a cabo en los modelos de Markov ya que la información disponible estaba en formato de probabilidades de transición (20).

Los recursos llevan a cabo servicios a una entidad. Tienen la capacidad de servir a un número particular de entidades simultáneamente pero como en la vida real su capacidad de atender simultáneamente a las entidades es limitada. Cuando esa capacidad se sobrepasa se

generan colas. Eso significa que las entidades tienen que esperar a que el recurso tenga disponibilidad. Proporcionar un servicio requiere tiempo. Definidos en un tiempo determinado, su consumo lleva asociado el gasto de una serie de unidades de dichos recursos en momentos de tiempo concretos. En el ámbito sanitario estos recursos se pueden dividir en personas (médicos, enfermeras, cuidadores), bienes (medicamentos, pruebas de laboratorio), espacios (camas hospitalarias, quirófanos) y otros.

Una cola es cualquier módulo en la que una entidad espera para el acceso a un recurso. Su análisis ha desarrollado la teoría de colas que permite el análisis matemático y gestión de las colas para conseguir un uso óptimo de los recursos. La gestión de las colas puede seguir diferentes criterios. Por ejemplo, en un cajero de un supermercado se aplica el criterio de que el primero que llega es el primero que es atendido First In/First Out o lógica (FIFO). Por el contrario, al bajar de un avión se suele seguir el orden contrario al de entrada Last In /First Out o lógica (LIFO) y, por último, la sala de espera en un servicio de urgencias atiende en primer lugar a los pacientes con mayor riesgo Highest Value First o lógica (HVF).



Elaboración *de novo* de un modelo de estas características: pasos a seguir

Al igual que en cualquier otra ETS, el primer paso consiste en la definición y justificación de la pregunta de investigación. Aclarar por qué se requiere llevar a cabo la ETS objeto del estudio es clave ya que el desarrollo, programación y validación modelos exigen un uso intensivo de recursos humanos. Por ello, para justificar la necesidad del modelo es necesario previamente formular el objetivo en términos de la población de referencia, establecer la intervención bajo evaluación, la alternativa a comparar y las medidas de efectividad que se van a utilizar para medir el

impacto en salud. Al mismo tiempo se necesita seleccionar el diseño del estudio que se va a seguir en la evaluación y que va a ser modelado mediante SED para llevar a cabo el estudio: estudio coste-efectividad, análisis del impacto presupuestario o ambos. Deberá incluir las unidades de medición de los resultados finales.

El segundo paso que también es común con estudios basados en otro tipo de modelos es la conversión de la pregunta de investigación en un modelo conceptual representando la historia natural de la enfermedad objeto del estudio. Este deberá ser lo suficientemente complejo como para identificar las incertidumbres clave en la estructura del modelo y justificar los análisis de sensibilidad necesarios para tenerla en cuenta (31,32). El modelo conceptual se suele representar mediante un diagrama que incluya las especificaciones necesarias para responder a la pregunta de investigación. Estas incluyen las características clínicas de la población y la posible necesidad de llevar a cabo análisis de subgrupos. Debe ser lo suficientemente complejo como para incluir los elementos críticos del modelo, pero no más y debería mantener la validez aparente que deberá ser aportada por expertos clínicos en el tema bajo estudio. En el modelo conceptual se deben incluir los riesgos asociados a la pregunta de investigación, los riesgos competitivos relevantes como el de muerte por otras causas y los cambios asociados a la intervención a evaluar.

El paso siguiente es la obtención de los parámetros necesarios para la programación del modelo mediante SED. Ello implica la observación del sistema a representar como historia natural de la enfermedad y la definición de los sujetos, los eventos, los atributos, los parámetros o variables, el horizonte temporal, nivel de detalle y la perspectiva del estudio. La estimación de los parámetros suele depender de la disponibilidad de la información y del plazo estipulado para el final del estudio. La situación ideal es disponer de tiempo para poder recoger los datos de entrada del modelo a partir de datos propios de la población a estudio. Sin embargo, la realidad es que en general la información utilizada suele proceder de

varias fuentes que incluyen la literatura. En el estudio de hepatitis C citado la evolución durante las etapas de fibrosis procedía de la literatura ya que ese periodo puede durar hasta 20 o 30 años y es imposible esperar ese tiempo. Sin embargo, los tiempos en cirrosis y cirrosis descompensada se obtuvieron en un estudio retrospectivo de la propia población navarra(20). Los parámetros necesarios para el cálculo de los costes deberán ser obtenidos a ser posible de la propia población. Tanto el uso de recursos como los costes unitarios suelen ser muy específicos.

El siguiente paso es la obtención de las funciones matemáticas que van a convertir las características de las entidades en eventos y estados. Para esta tarea el análisis de supervivencia paramétrico suele ser la herramienta estadística más habitual (33). Sin embargo, el tipo de funciones incluye otros enfoques como distribuciones empíricas para las cuales los programas específicos disponen de módulos preparados (34,35).

Con las funciones y los parámetros se programa el modelo en el software seleccionado. El informe del grupo de buenas prácticas de ISPOR-SMDM describe con detalle el procedimiento a seguir para que el modelo matemático sea consistente con el modelo conceptual (5). Una vez programado hay que proceder a la validación del modelo. Hay referencias que detallan los procedimientos de validación (36). Estos procedimientos son comunes a las diferentes técnicas que se utilicen. Como en el proverbio chino acerca del color de los gatos y su habilidad para cazar ratones, lo importante no es la técnica que se utilice, sino que el modelo reproduzca fielmente el sistema que debe representar y que el informe técnico documente adecuadamente los diferentes pasos de la validación (36,37). En este sentido la ventaja de los modelos programados mediante SED es que permiten separar la obtención de los resultados de su análisis estadístico para poder llevar a cabo la validación a partir de indicadores específicos.

La última etapa es la obtención de los resultados de investigación. Con ella se responde a la pregunta de investigación que es el objetivo final de la ETS. El formato dependerá del objetivo y del diseño aplicado.



Figura 1. Pasos a seguir en la construcción de un modelo de simulación de eventos discretos





Bibliografía

1. Caro JJ, Möller J, Karnon J, et al. Discrete Event Simulation for Health Technology Assessment. 1st Ed. Chapman and Hall/CRC; 2015. 374 p.
2. Eddy DM, Hollingworth W, Caro JJ, et al. Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-7. *Med Decis Making*. 2012;32:733-43.
3. Caro JJ, Briggs AH, Siebert U, et al. ISPOR-SMDM Modeling Good Research Practices Task Force. Modeling good research practices--overview: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-1. *Med Decis Making*. 2012;32:667-77.
4. Arrospe A, Rue M, van Ravesteyn NT, et al. Economic evaluation of the breast cancer screening programme in the Basque Country: retrospective cost-effectiveness and budget impact analysis. *BMC Cancer*. 2016;16:344.
5. Karnon J, Stahl J, Brennan A, et al. Modeling using discrete event simulation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-4. *Med Decis Making*. 2012;32:701-11.
6. Rodríguez Barrios JM, Serrano D, Monleón T, et al. Discrete-event simulation models in the economic evaluation of health technologies and health products. *Gac Sanit*. 2008;22:151-61.
7. Briggs A, Sculpher M, Claxton K. *Decision Modelling for Health Economic Evaluation*. London. Oxford University Press; 2006. 256 p.
8. Comas M, Castells X, Hoffmeister L, et al. Discrete-event simulation applied to analysis of waiting lists. Evaluation of a prioritization system for cataract surgery. *Value Health*. 2008;11:1203-13.
9. Soto-Gordoa M, Arrospe A, Moreno-Izco F, et al. Projecting Burden of Dementia in Spain, 2010-2050: Impact of Modifying Risk Factors. *J Alzheimers Dis*. 2015;48:721-30

10. Mar J, Soto-Gordoa M, Arrospide A, et al. Fitting the epidemiology and neuropathology of the early stages of Alzheimer's disease to prevent dementia. *Alzheimers Res Ther.* 2015;7:2.
11. Cooper K, Davies R, Roderick P, et al. The development of a simulation model of the treatment of coronary heart disease. *Health Care Manag Sci.* 2002;5:259-67.
12. Jahn B, Pfeiffer KP, Theurl E, et al. Capacity constraints and cost-effectiveness: a discrete event simulation for drug-eluting stents. *Med Decis Making.* 2010;30:16-28.
13. Auchincloss AH, Diez Roux AV. A new tool for epidemiology: the usefulness of dynamic-agent models in understanding place effects on health. *Am J Epidemiol.* 2008;168:1-8.
14. Bicher M, Rippinger C, Urach C, et al. Evaluation of Contact-Tracing Policies against the Spread of SARS-CoV-2 in Austria: An Agent-Based Simulation. *Med Decis Making.* 2021; 41: 1017-32
15. Ibarrondo O, Álvarez-López I, Freundlich F, et al. Probabilistic cost-utility analysis and expected value of perfect information for the Oncotype multi-genic test: a discrete event simulation model. *Gac Sanit.* 2020;34:61-8.
16. Arrospide A, Mar J, Vivancos-Mora J, et al. [Cost-effectiveness analysis of using high doses of atorvastatin for the secondary stroke prevention in Spain]. *Rev Neurol.* 2010;51:1-11.
17. Karnon J, Czoski-Murray C, Smith KJ, et al. A hybrid cohort individual sampling natural history model of age-related macular degeneration: assessing the cost-effectiveness of screening using probabilistic calibration. *Med Decis Making.* 2009;29:304-16.
18. Mar J, Gorostiza A, Ibarrondo O, et al. Economic evaluation of supplementing the diet with Souvenaid in patients with prodromal Alzheimer's disease. *Alzheimers Res Ther.* 2020;12:166.
19. Mar J, Arrospide A, Comas M. Budget impact analysis of thrombolysis for stroke in Spain: a discrete event simulation model. *Value Health.* 2010;13: 69-76.

20. Mar J, Ibarrondo O, Martínez-Baz I, et al. Economic evaluation of a population strategy for the treatment of chronic hepatitis C with direct-acting antivirals. *Rev Esp Enferm Dig.* 2018;110:621-8.
21. Drummond MF, Sculpher MJ, Claxton K, et al. *Methods for the Economic Evaluation of Health Care Programmes.* 4th Ed. New York. Oxford University Press; 2015.
22. Neumann PJ, Ganiats TG, Russell LB, et al. *Cost-Effectiveness in Health and Medicine.* 2nd Ed. New York: Oxford University Press; 2016
23. Arrospe A, Rue M, van Ravesteyn NT, et al. Evaluation of health benefits and harms of the breast cancer screening programme in the Basque Country using discrete event simulation. *BMC Cancer.* 2015;15:671.
24. Vandenbroucke JP, Pearce N. Incidence rates in dynamic populations. *Int J Epidemiol.* 2012;41:1472-9.
25. Karnon J, Haji Ali Afzali H. When to use discrete event simulation (DES) for the economic evaluation of health technologies? A review and critique of the costs and benefits of DES. *PharmacoEconomics.* 2014;32:547-58.
26. Sonnenberg FA, Beck JR. Markov models in medical decision making: a practical guide. *Med Decis Making.* 1993;13:322-38.
27. Larrañaga I, Stafylas P, Fullaondo A, et al. Economic Evaluation of an Integrated Health and Social Care Program for Heart Failure Through 2 Different Modeling Techniques. *Health Serv Res Manag Epidemiol.* 2018;5: 1-6
28. Caro JJ. Discretely Integrated Condition Event (DICE) Simulation for Pharmacoeconomics. *PharmacoEconomics.* 2016;34:665-72.
29. Krijkamp EM, Alarid-Escudero F, Enns EA, et al. Microsimulation Modeling for Health Decision Sciences Using R: A Tutorial. *Med Decis Making.* 2018;38:400-22.
30. Incerti D, Thom H, Baio G, et al. R You Still Using Excel? The Advantages of Modern Software Tools for Health Technology Assessment. *Value Health.* 2019;22:575-9.
31. Briggs AH, Weinstein MC, Fenwick EAL, et al. Model parameter estimation and uncertainty analysis: a report of the ISPOR-SMDM Modeling Good

- Research Practices Task Force Working Group-6. *Med Decis Making*. 2012;32:722-32.
32. Roberts M, Russell LB, Paltiel AD, et al. Conceptualizing a model: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-2. *Med Decis Making*. 2012;32:678-89.
 33. Latimer NR. Survival analysis for economic evaluations alongside clinical trials--extrapolation with patient-level data: inconsistencies, limitations, and a practical guide. *Med Decis Making*. 2013;33:743-54.
 34. Kelton WD, Sadowski RP, Zupick BM. *Simulation with Arena*. 6th Ed. New York. Mc Graw-Hill Education 2014.
 35. Law A, Kelton WD. *Simulation Modeling and Analysis*. 3rd Boston. Mc Graw-Hill Education 2000.
 36. Vemer P, Corro Ramos I, van Voorn G a. K, et al. AdViSHE: A Validation-Assessment Tool of Health-Economic Models for Decision Makers and Model Users. *Pharmacoeconomics*. 2016;34:349-61.
 37. Kopec JA, Finès P, Manuel DG, et al. Validation of population-based disease simulation models: a review of concepts and methods. *BMC Public Health*. 2010;10:710.



13

MODELOS DINÁMICOS DE TRANSMISIÓN

M^a Reyes Lorente Antoñanzas, Fernando Antoñanzas Villar

Departamento de Economía y Empresa. Universidad de la Rioja

La evaluación económica en el ámbito sanitario constituye una herramienta de decisión que relaciona los efectos en salud y los económicos, para una intervención determinada sobre un colectivo de sujetos concreto. Existen distintos tipos de modelos para llevar a cabo la evaluación económica, como ya se han visto en los capítulos anteriores: desde los árboles de decisión que tienen restricción en el número de retornos, ya que un excesivo número de nudos de repetición disminuye su operatividad, hasta modelos más avanzados, bien sean de carácter determinista o estocástico.

Una de las principales características de los escenarios habituales objeto de evaluación es la existencia de recurrencia de determinadas estados o situaciones. Si bien los modelos Markov permiten recoger esta característica de recurrencia y la de proceso continuo con temporización de los eventos -aunque habitualmente son en un contexto de tiempo discreto-, su concepción general emplea grupos de sujetos con ciclos de tiempo en la transición de duración fija, a los que en modelos más avanzados se pueden integrar 'variables rastreadoras', que pueden afectar modificando las nuevas probabilidades de transición (1-3).

Con el objeto de superar el inconveniente de no individualizar a los sujetos surgen los modelos de microsimulación, en los que las probabilidades de transición son acordes a los atributos asignados respecto a los estados de salud o al tiempo de permanencia en los mismos, lo que genera un error cuasi-estándar independiente de la incertidumbre en los parámetros de entrada (4,5).

En un paso más para la abstracción más completa de la realidad, se trabaja con los modelos de simulación con eventos discretos, donde se pueden incorporar fenómenos emergentes y los individuos se desplazan a través del sistema pudiendo interactuar con otros a lo largo de distintas formas y duraciones de la temporalización, obteniendo o

modificando sus atributos y reformulando la secuencia de eventos de cada individuo. Estas características conllevan que se requiera un mayor tiempo de ejecución según aumente el tamaño estructural del modelo y el número de individuos objeto de estudio (6-9).

Una vez repasados brevemente los diferentes tipos de modelos descritos en capítulos anteriores, nos centraremos en los modelos dinámicos de transmisión. Este tipo de modelos se refiere, principalmente, a los que emplean ecuaciones en diferencias finitas y ecuaciones diferenciales. Los modelos dinámicos de transmisión se diferencian de los de transición en el sentido de que recogen la influencia de unos estados y de unos individuos en otros.

Ahora bien, la distinción entre ecuaciones en diferencias finitas y las ecuaciones diferenciales se sustenta en cómo se considera el tiempo. Las primeras hacen referencia al tiempo como variable discreta y las segundas asumen que el tiempo es continuo. Si bien, el tratamiento del tiempo como continuo no impide obtener resultados para cortes de tiempo concreto (los resultados a uno, tres y cinco años, por ejemplo).

Los modelos dinámicos de transmisión se aplican como solución al tiempo requerido de ejecución, que pueden precisar otros modelos que realizan sus análisis por individuos, por lo que se utilizan habitualmente en problemas que afectan a un tamaño de población extenso (10), trabajando con el supuesto de fracciones de cohortes como el de los modelos de cohorte de Markov. En este sentido, cabe señalar que los modelos de cohorte de Markov deterministas son, en casos especiales, resueltos con sistemas dinámicos —concretamente, ecuaciones con diferencia finita—, pero sin la posibilidad de interacción (11).

Además, hay que tener presente la distinción entre modelos deterministas y estocásticos, debido a que los modelos dinámicos de transmisión pueden ser tratados de una forma u otra. Los modelos deterministas no consideran el efecto del azar —el error o desviación en la

estimación—. Ahora bien, la estimación de los parámetros busca la fiabilidad y validez de los resultados, lo cual conlleva a la minimización del error, por lo que se podría asumir una proximidad entre el estocástico y el determinista. Es decir, en términos matemáticos, de manera introductoria, y como un simple ejemplo lineal con regresores parciales: si partimos de $Y = B_1X_1 + B_2X_2 + \varepsilon$, donde Y es la variable explicada, B_i es el regresor o parámetro que indica la parte de X_i —variable predictora— que explica Y , y ε es el error en la estimación, la estimación será $\hat{Y} = b_1X_1 + b_2X_2$, donde \hat{Y} es la variable estimada, b_i es el parámetro correspondiente a X_i que estima \hat{Y} . En el tema siguiente se tratará la estimación de los modelos bayesianos que abordan las probabilidades condicionadas.

Algunas de las características de los modelos dinámicos de transmisión —denominados por algunos autores como ‘sistemas dinámicos’— incorporan mecanismos de realimentación interna, relaciones no lineales y procesos acumulativos (7,12); además, se considera el flujo de individuos de forma agregada y determinista con interacción entre los individuos (1,3,11). Los modelos dinámicos de transmisión emplean una perspectiva holística, ya que capturan explícitamente sistemas de realimentación complejos; de tal modo que puede dar lugar a que el sistema funcione de manera distinta, cuando la capacidad de tener un número de entidades en un estado se completa o se sobrepasa (13).

Téngase en cuenta que los modelos dinámicos de transmisión también recogen relaciones no lineales, es decir, en sus ecuaciones quedaría reflejada, mediante la multiplicación de las variables, dichas relaciones.

El diseño de un modelo dinámico de transmisión se basa habitualmente en un sistema de ecuaciones, que parte de su representación mediante un diagrama causal que recoge todos los factores y vínculos relevantes para los individuos; posteriormente, se representa el número de individuos por estado en el momento inicial —que permite limitar la capacidad de transición, es decir, la población entera de cada momen-

to estará repartida entre los distintos estados, no pudiendo ser superada— y los flujos o ratios con los que aumenta o disminuye el número de individuos en un estado. La estructura fundamental de los modelos dinámicos de transmisión puede influir en el comportamiento progresivo del sistema —por ejemplo, las variaciones temporales de los distintos aspectos considerados— (10).

Como ya se ha señalado, los modelos de sistemas dinámicos corresponden a la aplicación de ecuaciones diferenciales o de ecuaciones en diferencia, dependiendo de que el tratamiento del tiempo sea continuo o discreto, respectivamente. Habitualmente, se aplican ecuaciones ordinarias —consideran una única variable explicativa—, aunque también se pueden utilizar ecuaciones parciales que trabajan con varias variables independientes explicativas para complementar a las ecuaciones ordinarias y mejorar la riqueza del modelo (13).

Al iniciar la introducción matemática, las ecuaciones en diferencia finita ordinarias permiten describir la ratio de cambio de estado como, $x_{i+1} = x_i + f(x_i, t_i) \Delta t$, donde x_i es el estado en el momento $t_i = i \Delta t$; mientras que las ecuaciones diferenciales ordinarias permiten describir la ratio de cambio del estado del siguiente modo $\frac{\partial x}{\partial t} = f(x, t)$, donde x es el estado objeto de análisis y t es el tiempo. De este modo, nos podemos encontrar con un planteamiento análogo a las cadenas Markov a tiempo continuo; por ejemplo,

$$\frac{\partial I}{\partial t} = \frac{c\beta I(N - I)}{N} - \frac{I}{d}$$

Donde I es el número de individuos en el estado de origen —infectados, por ejemplo—, N es el total de individuos en la población objeto de estudio —por lo que $(N-I)$ son los susceptibles de infectarse—, c es la tasa de contactos, β es la probabilidad de contagio por contacto y d es la duración media de la infección —entonces, $(1/d)$ es la tasa promedio de recuperación uniforme temporalmente—. En consecuencia, la

variación del número de infectados a lo largo del tiempo es el resultado de los nuevos posibles infectados menos los que han dejado de estar infectados.

Como ejemplo numérico, si tenemos una población de estudio de 1.000 individuos (N) y un total de infectados de 50 (I), los susceptibles a infectarse, en el primer instante de tiempo, son 950 individuos ($N-I$); si suponemos una tasa de contactos del 20% (c) y una probabilidad del contagio del 4% (β), obtenemos que la posibilidad de que un infectado contagie es 0,000008 individuos de la población ($c\beta/N$). En consecuencia, los contagios en los susceptibles, en ese primer instante de tiempo, son 0,38 ($=0,000008*50*950$). Supongamos que la unidad de medida temporal es la hora. Además, si consideramos que la duración media de la infección es de 7 días ($d=168$ horas) y teniendo en cuenta un reparto uniforme a lo largo del tiempo en la recuperación, tal y como presenta la fórmula expuesta más arriba, obtenemos que 0,297 ($=50/168$) de los infectados iniciales se recuperan en esa primera hora. De ahí que, en la primera hora, la variación de los infectados es un aumento de 0,083 ($=0,38-0,297$ individuos). Como se puede intuir fácilmente, los resultados van cambiando en cada momento al variar los valores de individuos en cada estado. Téngase en cuenta que la explicación numérica, dada en este párrafo, es una aproximación inicial a la interpretación de la resolución de este tipo de modelos dinámicos de transmisión.

Antes de presentar los diferentes tipos de modelos dinámicos de transmisión aplicables en economía de la salud y de los métodos de resolución de este tipo de modelos, cabe señalar algunas de las limitaciones consecuencia de las propias características de estos sistemas de ecuaciones. En concreto, el grado de complejidad del modelo afecta a la factibilidad de su resolución, ya que 1) no pueden describir interacciones entre procesos concurrentes, 2) sólo pueden considerar un limitado registro de la historia de la cohorte de individuos, y 3) sin analizar individualmente las características propias de cada sujeto.



Tipos de modelos dinámicos de transmisión

Desde finales del siglo XVIII, se han empleado modelos de transmisión dinámica para conocer matemáticamente el impacto epidemiológico de las infecciones. A partir de entonces se han desarrollado diferentes modelos que permiten estudiar el impacto de la transmisión y su repercusión en diferentes estados. Centrándonos en los sistemas de ecuaciones diferenciales, los más conocidos son el modelo SIR, creado por Kermack y McKendrick en 1927, y las posteriores adaptaciones que incorporan parámetros añadidos al modelo (13). Todavía hoy, se está profundizando en la integración de los modelos matemáticos y los métodos estadísticos para una composición más avanzada de la realidad (14). En este capítulo de libro se realiza una presentación de carácter introductoria.

Como punto de partida hay que conocer el proceso de transmisión de la enfermedad objeto de estudio y así seleccionar y diseñar el tipo de modelo más adecuado. Su representación mediante una figura de flujos ayuda a plasmar visualmente de manera sencilla lo que posteriormente se querrá modelizar matemáticamente. Se puede utilizar una representación sencilla del estilo $S \rightarrow I \rightarrow R$ o más visual, como la siguiente, que permite incorporar los diferentes parámetros de transmisión.



Figura 1. Representación del modelo SIR básico con parámetros para contagio y recuperación



La selección de los colectivos y la secuenciación entre ellos dan lugar a diferentes modelos. En este sentido, podemos asumir que los individuos que superan la infección se recuperan y permanecen en ese nuevo estado [modelo SIR], que todos los individuos recuperados vuelven a ser susceptibles a contagiarse nuevamente [modelo SIS] o que parte de esos individuos recuperados permanecen en ese estado [modelo SIRS]; también podría asumirse que puede haber individuos contagiados pero en estado latente o de incubación con posibilidad de infectar o no a otros [modelos SEIS, SEIR o SEIRS]; incluso incorporar infantes con inmunidad pasiva que irán perdiendo tras cierto tiempo [todos los modelos presentados anteriormente, pero con M delante para identificar a este colectivo con inmunidad pasiva].

Estos modelos admiten complementos según los detalles que se añadan para acercarlos más a la realidad objeto de análisis, lo cual suele conllevar una mayor complejidad matemática. Se puede considerar distintos tipos de infectados conforme a su progreso en diferentes estadios —siguiendo la secuenciación en cadena tradicional— o distinguir entre infectados inmunodeprimidos o no —en este último caso, se recomienda hacer dos modelos en paralelo para cada colectivo de infectados—. También, se puede incorporar determinados parámetros complementarios en el modelo matemático; por ejemplo, nacimientos, muertes naturales y muertes por la enfermedad, vacunados, etc.

El siguiente paso es elaborar el constructo matemático de las ecuaciones —consideremos ecuaciones diferenciales—. A continuación, se presentan cuatro ejemplos: el primero es el modelo SIR básico representado en la Figura 1, el segundo es un SEIR, el tercero un modelo SIRS y el cuarto una propuesta de un modelo SIR bastante más complejo en su parametrización. De este modo, se verán, a partir del modelo más clásico, las diferentes vertientes, ya sea por integración de otro colectivo —individuos contagiados,

pero en proceso de incubación (E)—, por retorno de parte de un colectivo a un estado ya transitado, o por inclusión de más parámetros en el modelo.

El sistema de ecuaciones diferenciales en un modelo SIR básico, con tasa de contagio (β) y tasa de recuperación (γ), viene representado de la siguiente manera:

$\frac{\partial S}{\partial t} = -\beta SI$	La variación en el número de susceptibles a lo largo del tiempo es resultado de la disminución de los mismos a consecuencia del contagio entre los infectados y susceptibles.
$\frac{\partial I}{\partial t} = \beta SI - \gamma I$	La variación en el número de infectados a lo largo del tiempo es resultado del aumento de los mismos a consecuencia del contagio entre los infectados y susceptibles, así como de la disminución por los infectados que se recuperan.
$\frac{\partial R}{\partial t} = \gamma I$	La variación en el número de recuperados a lo largo del tiempo es resultado del aumento debido a los infectados que se recuperan.

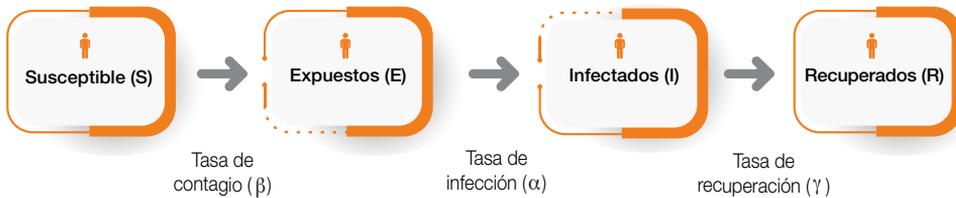
En este planteamiento del modelo SIR básico, a diferencia de la variación de infectados presentado previamente en este capítulo, vemos que β antes era contagio por contacto y c era tasa de contactos; sin embargo, ahora β es la tasa de contagio —es decir, el resultado de multiplicar el contagio por contacto por la tasa de contactos—. Esta pequeña advertencia puede ser extrapolable a cualquier otro parámetro y modelo; en consecuencia, se ha de tener presente y bien claro lo que realmente representa cada parámetro en cada ecuación. Además, la estimación de la tasa de contagio como indicador global es más inmediato de obtener mediante modelos de regresión sobre series epidemiológicas, pero la segregación de contagio por contacto y tasa de contactos requiere métodos más apropiados para su estimación, tales como los bayesianos que se verán en el siguiente capítulo. Otra diferencia entre la variación de los infectados vista en la ecuación inicial del capítulo y la presentada aquí es la parte que resta, donde antes teníamos la recuperación representada

como uno dividido entre el tiempo de recuperación ($1/d$) y ahora como una tasa (γ); sin embargo, el significado es el mismo, pero presentado de manera diferente. En ambos la recuperación es un elemento modelado de manera constante, pero podría ser variable en el tiempo mediante componentes estacionarios en modelos más sofisticados.

En el siguiente ejemplo, se incluye el modelo SIR previo distinguiendo a los individuos contagiados en proceso de incubación (E). La representación gráfica de este modelo SEIR es similar al anterior, pero con un nuevo colectivo que denominamos como expuestos.



Figura 2. Representación del modelo SEIR básico con parámetros para contagio, infección y recuperación



Este modelo SEIR tiene como sistema de ecuaciones diferenciales los siguientes elementos:

$$\frac{\partial S}{\partial t} = -\beta SI$$

La variación en el número de susceptibles a lo largo del tiempo es resultado de la disminución de los mismos a consecuencia del contagio entre los infectados y susceptibles.

$$\frac{\partial E}{\partial t} = \beta SI - \alpha E$$

La variación en el número de expuestos a lo largo del tiempo es resultado del aumento de los mismos a consecuencia del contagio entre los infectados y susceptibles, así como de la disminución por los expuestos que presentan infección no latente.

$$\frac{\partial I}{\partial t} = \alpha E - \gamma I$$

La variación en el número de infectados a lo largo del tiempo es resultado del aumento de los mismos a consecuencia de los expuestos que pasan a considerarse infectados, así como de la disminución por los infectados que se recuperan.

$$\frac{\partial R}{\partial t} = \gamma I$$

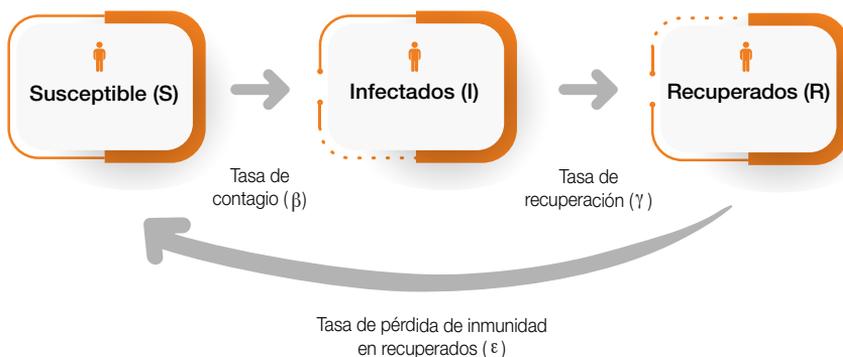
La variación en el número de recuperados a lo largo del tiempo es resultado del aumento debido a los infectados que se recuperan.

Téngase en cuenta que aquí estamos considerando que el estado de expuesto o con infección latente supone la posibilidad de no transmisión de la infección a los susceptibles, cuestión que no necesariamente es así. Sin embargo, la distinción entre expuestos e infectados conlleva el impacto posterior en la cuantificación de carga económica asignable a cada uno de estos grupos, de ahí su relevancia en determinados estudios.

A continuación, el ejemplo de un modelo con retorno de parte de un colectivo a un estado ya transitado como puede ser un SIRS; aunque, como ya se indicó, hay otras posibilidades de modelos con retorno a estados previos.



Figura 3. Representación del modelo SIRS básico con parámetros para contagio, recuperación y pérdida de inmunidad



El modelo SIRS expuesto en la figura presenta las siguientes ecuaciones diferenciales:

$$\frac{\partial S}{\partial t} = -\beta SI + \varepsilon R$$

La variación en el número de susceptibles a lo largo del tiempo es resultado de la disminución de los mismos a consecuencia del contagio entre los infectados y susceptibles, así como el aumento por los recuperados que pierden su inmunidad volviendo a ser susceptibles.

$$\frac{\partial I}{\partial t} = \beta SI - \gamma I$$

La variación en el número de infectados a lo largo del tiempo es resultado del aumento de los mismos a consecuencia del contagio entre los infectados y susceptibles, así como de la disminución por los infectados que se recuperan.

$$\frac{\partial R}{\partial t} = \gamma I - \varepsilon R$$

La variación en el número de recuperados a lo largo del tiempo es resultado del aumento debido a los infectados que se recuperan, así como a la disminución por los recuperados que pierden su inmunidad.

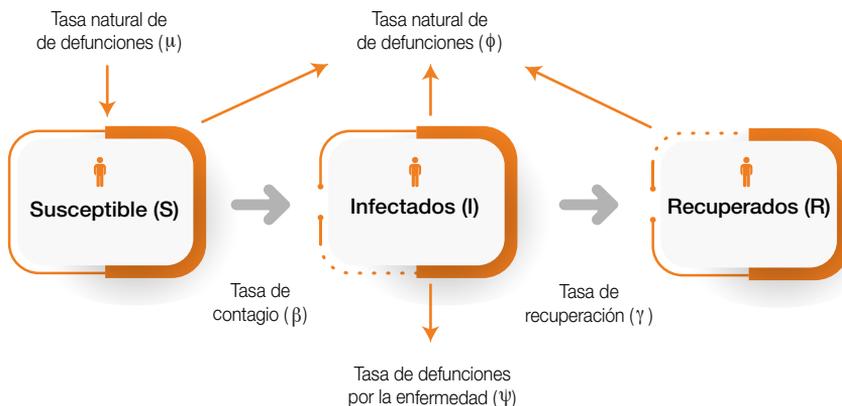
En el último ejemplo se presenta la inclusión de otros parámetros en el modelo, mediante un modelo SIR sencillo (como el visto en la Figura 1). En concreto, se va a considerar nacimientos y defunciones, como elementos que permiten la adaptación demográfica continua, quedando su representación como aparece en la Figura 4.

A partir de una población inicial que se reparte entre susceptibles, infectados y recuperados, en esta propuesta de modelo se incorporan nuevos individuos mediante una tasa promedio de nacimientos y se excluyen los fallecimientos naturales que suceden por cualquier motivo ajeno a la infección —estos parámetros suelen ser obtenidos a partir de algún modelo que utilice series temporales, de manera similar a la de otros posibles parámetros—, así como los fallecimientos derivados de la infección.

El sistema de ecuaciones diferenciales del modelo SIR básico, presentado en el primer ejemplo, se ve adaptado matemáticamente con la inclusión del comportamiento demográfico previsto durante el horizonte temporal del estudio.



Figura 4. Representación del modelo SIR básico con parámetros para contagio, recuperación, nacimientos y defunciones



$$\frac{\partial S}{\partial t} = N\mu - \beta SI - \phi S$$

La variación en el número de susceptibles a lo largo del tiempo incorpora a los nuevos nacimientos que, en el modelo planteado, se asignan íntegramente a este colectivo y resta los susceptibles anteriores que se contagian y los que fallecen por motivo natural.

$$\frac{\partial I}{\partial t} = \beta SI - \gamma I - (\phi + \psi) I$$

La variación en el número de infectados a lo largo del tiempo incorpora a los nuevos contagiados y resta los que se recuperan y los que fallecen bien por motivo natural o como consecuencia de la infección.

$$\frac{\partial R}{\partial t} = \gamma I - \phi R$$

La variación en el número de recuperados a lo largo del tiempo incorpora los infectados que se recuperan y resta los recuperados que fallecen por motivo natural.

Como se ha indicado anteriormente, existen modelos más complejos que siguen la misma mecánica de modelización y que pueden incluir múltiples parámetros en el sistema de ecuaciones. Algunos de gran interés como el análisis del efecto del distanciamiento social, bien voluntario por la percepción de riesgo de los individuos o bien el establecido por la autoridad pública (15).

En modelos más avanzados estos parámetros fijos presentados como promedios podrían ser adaptados a situaciones con estacionariedad en el comportamiento de los mismos; por ejemplo, por motivos ambientales. Sin embargo, en los modelos de transmisión dinámica publicados es más habitual considerar sus valores promedio estimados para la fragmentación temporal de los cambios de tiempo que rigen el modelo. Es decir, si por ejemplo el horizonte temporal de la evaluación económica es 3 años, no significa que el salto temporal del modelo sea tal; sino que se considera el proceso evolutivo, ya sea con ecuaciones diferenciales o con ecuaciones en diferencia, acorde con las particularidades del problema que podría requerir análisis evolutivo de semanas, días, horas, etc. De ahí que los parámetros de este modelo deben estar valorados con la misma medida temporal que damos a la variación del tiempo (t).

La forma en la que se trata la estacionalidad en los modelos de transmisión dinámica puede ser, también, de diferente grado de complejidad. A continuación, se van a poner dos nuevos ejemplos que permitan ver este particular de manera más clara e introductoria. El primero considera la vacunación recurrente en un periodo de tiempo determinado y repartida de manera proporcional durante dicho periodo (16):

$$V(t) = \begin{cases} 0 & t_{i-p} \leq t \leq c_i \\ \frac{N * v * \varepsilon}{c_f - c_i} & c_i < t \leq c_f \\ 0 & c_f < t \leq t_{f-p} \end{cases}$$

Donde $V(t)$ son los vacunados en el momento t , N es la población objeto de estudio, v es la tasa de cobertura de la vacunación, ε es la efectividad o eficacia de la vacuna, c_f es el último momento del periodo de vacunación, c_i es el último momento previo al de arranque de la vacunación, $(c_f - c_i)$ es el periodo de vacunación, t_{i-p} es el primer momento del fragmento temporal (por ejemplo, la primera semana del año), t_{f-p} es el último momento del fragmento temporal.

La relación entre las variables en el modelo SIR sencillo descrito con cambios en la población debidos a nacimientos y fallecimientos, con la incorporación de la vacunación descrita, se puede expresar mediante el siguiente sistema no lineal de ecuaciones diferenciales ordinarias:

$$\left\{ \begin{array}{l} \frac{\partial S}{\partial t} = N(t)\mu - \beta S(t)I(t) - \phi S(t) - V(t) \\ \frac{\partial I}{\partial t} = \beta S(t)I(t) - \gamma I(t) - (\phi + \varphi)I(t) \\ \frac{\partial R}{\partial t} = \gamma I(t) - \phi R(t) + V(t) \end{array} \right.$$

El segundo ejemplo de cómo considerar la estacionalidad de alguna variable, en un modelo SIR, consiste en distinguir entre vacunación efectiva al nacer (ρ) y vacunación efectiva de refuerzo posterior ($\theta \cong \nu * \varepsilon$) que en este caso será sobre los susceptibles, en vez de sobre toda la población (17):

$$\left\{ \begin{array}{l} \frac{\partial S}{\partial t} = N(t)\mu(1-\rho) - \beta S(t)I(t) - \phi S(t) \\ \frac{\partial I}{\partial t} = \beta S(t)I(t) - \gamma I(t) - (\phi + \varphi)I(t) \\ \frac{\partial R}{\partial t} = N(t)\mu\rho + \gamma I(t) - \phi R(t) \end{array} \right\} t \neq \kappa\tau$$

$$\left. \begin{array}{l} S(t^+) = (1 - \theta) S(t) \\ I(t^+) = I(t) \\ R(t^+) = R(t) + \theta S(t) \end{array} \right\} t = \kappa\tau$$

En este último modelo planteado se diferencia entre la evolución de las variables en el tiempo y el efecto en el momento concreto de la vacunación de refuerzo; por lo que τ es el tiempo entre dos momentos de vacunación de refuerzo consecutivas.



Tipos de métodos de resolución de ecuaciones ordinales

En este apartado nos centramos en las ecuaciones ordinales para simplificar la presentación y por ser las más habitualmente utilizadas en el ámbito de la economía de la salud. Si bien existen diferentes métodos analíticos para la resolución exacta de ecuaciones diferenciales ordinales, en función de las características del modelo matemático específico analizado, los métodos de aproximación numérica, presentados a continuación, son los más frecuentes a la hora de resolver los sistemas de ecuaciones diferenciales ordinales que evolucionan con el tiempo (18).

En primer lugar, téngase presente que la nomenclatura matemática para la expresión y resolución de ecuaciones es $\frac{dy}{dx} \equiv f(X_t, Y_t)$; es decir, se representa la variación de Y respecto a variaciones en X mediante una función del valor de ambas variables en ese momento t, o lo que es lo mismo la pendiente o tangente en el punto (X_t, Y_t) . Dicha variación necesariamente no se ha de mantener constante en el tiempo, lo que nos sugiere que la evolución de la variable Y respecto a la variable X tiene una función curvilínea que viene expresada por la propia función $f(X, Y)$.

Comencemos con el más sencillo que es el método Euler, donde una variable explicativa X y una variable explicada Y cambian a lo largo del tiempo y consideremos que X evoluciona de manera constante a lo largo del tiempo (es decir, $X_{t+1} = X_t + \Delta X$, donde $\forall \Delta X = h$) por lo que $Y_{t+1} = Y_t + hf(X_t, Y_t)$, siendo $f(X_t, Y_t)$ la derivada de primer orden de Y res-

pecto al tiempo. Sin embargo, una adaptación de ese método produce un mejor ajuste: el de Euler mejorado que consiste en adaptar $Y_{t+1} = Y_t + \frac{h}{2} [f(X_t, Y_t) + f(X_{t+1}, (Y_{t+1})^*)]$, donde $(Y_{t+1})^* = Y_t + hf(X_t, Y_t)$.

Otro de los métodos de aproximación en la resolución de ecuaciones diferenciales es el de Runge-Kutta, que puede ser de diferente orden. El modelo general de Runge-Kutta es $Y_{t+1} = Y_t + h(a_1k_1 + a_2k_2 + \dots + a_nk_n)$, donde n es el orden del método, y previamente se han de obtener los valores de a_i y de k_i . El desarrollo matemático para su obtención es irrelevante para el objetivo de este libro; en tanto en cuanto, podemos aplicar las valoraciones usuales en matemáticas para dichos parámetros —diferentes según se aplique el método Runge-Kutta de un orden u otro—, dado que la condición para dichos parámetros es que $\sum_1^n a_i = 1$. Evidentemente, cuanto mayor es el orden del método mayor es su complejidad, pero también mayor es la precisión en los resultados (mejor se ajusta a la solución exacta de la ecuación, por lo que menor será el error cometido en la estimación del resultado). En todos los métodos de Runge-Kutta, sea cual sea su orden, $k_1 = f(X_t, Y_t)$; en consecuencia, este método de orden 1 es equiparable al método de Euler.

Entonces, lo relevante son métodos de Runge-Kutta de orden superior a uno. En el caso de un método de orden 2, se puede estimar considerando que $a_1 = a_2 = \frac{1}{2}$ y que $k_2 = f(X_t + h, Y_t + k_1h)$, siendo este k_2 el resultado de aplicar a la función original los cambios señalados; es decir, X_t pasa a ser $(X_t + h)$ y Y_t pasa a ser $(Y_t + k_1h)$. En consecuencia, en Runge-Kutta de orden 2 tenemos $Y_{t+1} = Y_t + h(\frac{1}{2}f(X_t, Y_t) + \frac{1}{2}f(X_t + h, Y_t + k_1h)) = Y_t + h\frac{1}{2}(f(X_t, Y_t) + f(X_t + h, Y_t + k_1h))$.

En Runge-Kutta de orden 3, se suele considerar $a_1 = a_3 = \frac{1}{6}$, $a_2 = \frac{4}{6}$, $k_2 = f(X_t + \frac{1}{2}h, Y_t + \frac{1}{2}k_1h)$ y $k_3 = f(X_t + h, Y_t - k_1h + 2k_2h)$, por lo que tenemos $Y_{t+1} = Y_t + h\frac{1}{6}(f(X_t, Y_t) + 4f(X_t + \frac{1}{2}h, Y_t + \frac{1}{2}k_1h) + f(X_t + h, Y_t - k_1h + 2k_2h))$.

Habitualmente, en economía de la salud se suele utilizar para resolver este tipo de ecuaciones diferenciales ordinales el Runge-Kutta de orden 4, que es el que se presenta a continuación. Sin embargo, los de orden inferior presentados previamente también podrían ser válidos según el grado de error dispuesto a asumir en las estimaciones.

El método Runge-Kutta de orden 4 suele considerar $a_1 = a_4 = \frac{1}{6}$, $a_2 = a_3 = \frac{2}{6}$, $k_2 = f(X_t + \frac{1}{2}h, Y_t + \frac{1}{2}k_1h)$, $k_3 = f(X_t + \frac{1}{2}h, Y_t + \frac{1}{2}k_2h)$ y $k_4 = f(X_t + h, Y_t + k_3h)$, lo que da lugar a $Y_{t+1} = Y_t + h\frac{1}{6}(f(X_t, Y_t) + 2f(X_t + \frac{1}{2}h, Y_t + \frac{1}{2}k_1h) + 2f(X_t + \frac{1}{2}h, Y_t + \frac{1}{2}k_2h) + f(X_t + h, Y_t + k_3h))$.

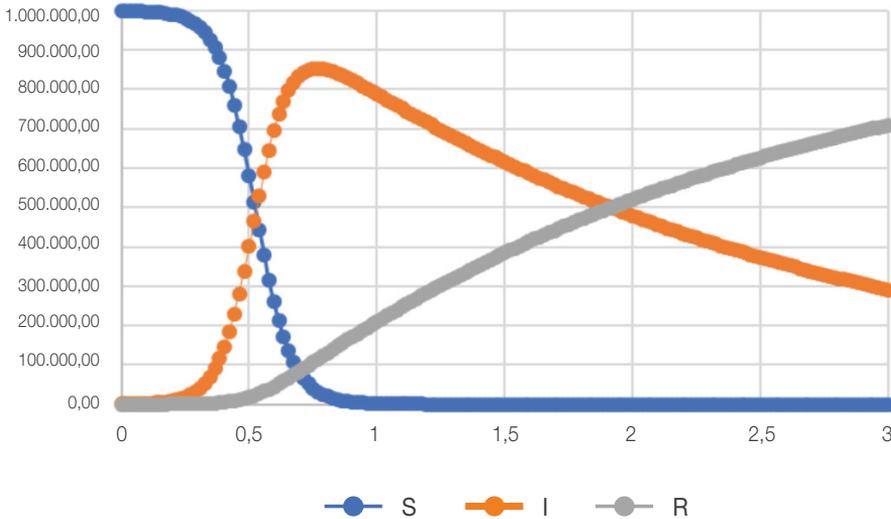
Las funciones presentadas previamente en este apartado son directamente aplicables para un modelo dinámico de transmisión sin configuración de sistema de ecuaciones. Al programar estas ecuaciones, se ha de tener presente que los modelos dinámicos de transmisión presentados son sistemas de ecuaciones y que, por lo tanto, hay vinculación entre las ecuaciones. En consecuencia, cada k_i debe fijarse para la variable explicativa y la explicada, en cada fórmula programada.

A continuación, se pone un ejemplo ilustrativo de los resultados numéricos que podríamos obtener. Imagine una población de 100.000 habitantes con 50 individuos infectados inicialmente y una tasa de contagio de 15/100.000 y una tasa de recuperación de 0,5 semanal por cada infectado, supongamos que queremos conocer el efecto de esta infección dado los parámetros señalados en un horizonte temporal de 3 años, distinguiendo los resultados de manera anual, y con intervalos evolutivos semanales (h). Consideremos un modelo SIR básico como el presentado en el primer ejemplo detallado de modelos dinámicos de transmisión con ecuaciones diferenciales. Si programamos en Excel las fórmulas indicadas anteriormente, podemos obtener los siguientes resultados:

	Euler			Runge-Kutta 4 orden		
	Año 1	Año 2	Año 3	Año 1	Año 2	Año 3
Susceptibles	184,14	0,004	0,00001	182,96	0,016	0,00001
Infectados	81.160,56	49.223,77	29.783,58	78.806,01	47.941,36	29.061,54
Recuperados	18.655,303	50.776,22	70.216,42	21.011,03	52.085,62	70.938,46



Figura 5. Resultados del modelo SIR para cada uno de los estados con Runge-Kutta de 4° orden



El método Runge-Kutta de 4 orden es el que mejor se ajusta a lo solución exacta del modelo. Evidentemente, los resultados serán diferentes según los valores de los parámetros considerados. Con los datos del ejemplo anterior, vemos que el método de Euler está sobreestimando el número de infectados e infraestimando el de los recuperados.



Reflexiones sobre estos modelos para evaluaciones económicas

Ante el surgimiento de nuevas infecciones con alto riesgo de transmisión y con efectos negativos inmediatos en la salud de la población general, es habitual que se elaboren modelos de ecuaciones diferenciales ordinarias

cuyos resultados no están exentos de cuestionamiento debido a diversos aspectos: 1) las características del propio modelo —los individuos se consideran homogéneamente distribuidos y conectados entre sí—, 2) la modelización explícita realizada, 3) el método de resolución, 4) el grado de conocimiento del problema objeto de estudio y 5) la calidad de los datos disponibles (19). Sin embargo, el único inconveniente que cabe abordar para sustentar una defensa específica de este tipo de modelos es su diseño y modelización resolutiva concreta, ya que las otras limitaciones son comunes a cualquier tipo de modelo de los presentados en este libro.

Como se ha querido transmitir anteriormente, hay que prestar atención a la hora de diseñar un modelo. Los modelos dinámicos de transmisión con ecuaciones diferenciales o en diferencia tratan a grupos de población y no realizan un análisis individualizado, para ello existen otros tipos de modelos. Sin embargo, se puede segmentar la población objeto de estudio en subpoblaciones, por ejemplo, personas hospitalizadas o no, ya que el valor de los parámetros diferirá, quizá sustancialmente, e incluso requerirá adaptación en el diseño del modelo base. De este modo, el mismo tipo de modelo puede ser utilizado para diferentes contextos de caracterización o segmentación de la población de índole sanitario, social o geográfico.

En este sentido, es condición necesaria que los gestores sanitarios tengan presente el constructo del modelo y la parametrización del mismo; además de recordar que las situaciones cambian y que, por ejemplo, un parámetro como la tasa de contagio puede ser modelizado como fijo para el horizonte temporal de estudio o con variabilidad estacional mediante una formulación específica —apareciendo en el modelo como $B(t)$ —.

Hasta ahora, se han presentado los modelos para el análisis epidemiológico, pero los componentes necesarios para hacer una evaluación económica, evidentemente, conllevan la incorporación de los recursos consumidos y los costes vinculados. Este aspecto suele realizarse paralelamente, mediante la aplicación de un promedio de recursos consu-

midos para cada estado en toda o en parte de la población que incluye en cada momento. Como ejemplo de asignación a toda la población de un estado en cada momento puede ser el consumo de un fármaco para todos los individuos que están en un estado mientras permanezcan en él; sin embargo, si es para parte de la cohorte, por ejemplo, solo el 2% de los infectados requieren hospitalización, como las ecuaciones van generando la evolución del total acumulado, la base para ese porcentaje es la diferencia entre el momento actual y el anterior. A partir de ahí, se aplicarán los costes unitarios para los recursos consumidos y se obtendrá la parte económica del estudio.



Bibliografía

1. Stahl JE. Modelling methods for pharmacoeconomic and health technology assessment: an overview and guide. *PharmacoEconomics*. 2008; 26:131-48.
2. Cooper K, Brailsford SC, Davies R. Choice of modelling technique for evaluating health care interventions. *J Oper Res Soc*. 2007;58:168-176.
3. Barton P, Bryan S, Robinson S. Modelling in the economic evaluation of health care. Selecting the appropriate approach. *J Oper Res Soc*. 2004;9:110-8.
4. Katsaliaki K, Mustaffe N. Applications of simulation within the healthcare context. *J Oper Res Soc*. 2011;62:1431-51.
5. Briggs AH. Handling uncertainty in cost-effectiveness models. *Pharmacoeconomis*. 2000;17:479-500.
6. Karnon J, Stahl J, Brennan A, et al. Modeling Using Discrete Event Simulation: A Report of the IPOR-SMDM Modeling Good Research Practices Task Force-4. *Med Decis Making*. 2012;32:701-11.
7. Mielczarek B, Uzialko-Mydlikowska J. Application of computer simulation modelling in the health care sector: a survey. *Simulation*, 2012; 88: 197-216.
8. Caro JJ. *PharmacoEconomics*. analyses using discrete event simulation. *PharmacoEconomics*. 2005;23:323-32.
9. Davies R, Roderick P, Raftery J. The evaluation of disease prevention and treatment using simulation models. *Eur J Oper Res*. 2003;150:53-66.
10. Taylor K, Lane D. Simulation applied to health services: opportunities for applying the system dynamics approach. *J Health Serv Res Policy*. 1998; 3:226-32.
11. Brennan A, Chick SE, Davies R. A taxonomy of model structures for economic evaluation of health technologies. *Health Econ*. 2006;15:1295-310.
12. Mustafee N, Katsaliaki K, Taylor SJE. Profiling literature in healthcare simulation. *Simulation*, 2010; 86:543-58.

13. Diekmann O, Heesterbeek JAP. *Mathematical Epidemiology of Infectious Diseases: Model Building Analysis and Interpretation*. New York: Wiley, 2000.
14. Bekiros S y Koulompou D. SBDiEM: A new mathematical model infectious disease. *Chaos, Solitons and Fractals*. 2020;136:109828.
15. Huang H, Chen Y, Yan Z. Impacts of social distancing on the spread of infectious diseases with asymptomatic infection: a mathematical model. *Appl Math Comput*. 2021;398:125983.
16. Pradas R, Gil de Miguel A, Álvaro A, et al. Budget impact analysis of a pneumococcal vaccination programme in the 65-year-old Spanish cohort using a dynamic model. *BMC Infect Dis*. 2013;13:175.
17. Gao S, Ouyang H, Nieto JJ. Mixed vaccination strategy in SIRS epidemic model with seasonal variability on infection. *Int J Biomaths*. 2011;4:473-91.
18. Arrieta JM, Ferreira R, Pardo RS, et al. Análisis numérico de ecuaciones diferenciales ordinarias. *Paraninfo*, 2020.
19. Chen S, Robinson P, Janies D, et al. Four challenges associated with current mathematical modelling paradigm of infectious diseases and call for a shift. *Open Forum Infect Dis*. 2020;7:ofaa333.



14

MODELOS BAYESIANOS

Miguel Ángel Negrín, Francisco José Vázquez Polo

Departamento de Métodos Cuantitativos en Economía y Gestión

Universidad de Las Palmas de Gran Canaria



Introducción

Los modelos analíticos de decisión son una herramienta ampliamente utilizada en la evaluación económica de tecnologías sanitarias. Estos modelos permiten simplificar procesos complejos y, de esta forma, poder tomar decisiones sobre la preferencia entre tecnologías sanitarias. La gran mayoría de estos modelos han sido analizados desde el punto de vista de la estadística clásica o frecuentista. En este capítulo mostraremos qué puede ofrecer la estadística bayesiana a la estimación de estos modelos.

La estadística bayesiana ha vivido su gran expansión a partir de los años 90. Aunque la formulación matemática de los modelos bayesianos es muy anterior, su crecimiento sólo ha sido posible a raíz del desarrollo de métodos de computación avanzados, como los métodos Markov Chain Monte Carlo (MCMC) (1), que han permitido aplicar métodos bayesianos a problemas más complejos. Alguna de las primeras aplicaciones de esta metodología en el contexto sanitario se encuentran en Eddy y col. (2,3), donde se muestran las bondades de la metodología bayesiana en el tratamiento de la incertidumbre en la estimación de los parámetros de un meta-análisis. Posteriormente, Spiegelhalter y col. (4) y Jones (5) discuten la aproximación bayesiana para la inferencia estadística en la comparación de tecnologías sanitarias con datos individuales provenientes de ensayos clínicos. Los trabajos posteriores de O'Hagan y col. (6-8), permitieron formalizar la estimación bayesiana del análisis coste-efectividad con datos individuales. En estos trabajos se destacan como principales argumentos para proponer la estadística bayesiana frente a la frecuentista la posibilidad de incorporar información a priori sobre los parámetros de interés; una estimación más robusta de la incertidumbre asociada a la ratio de coste efectividad incremental (RCEI); y la posibilidad de interpretar la curva

de aceptabilidad coste-efectividad (CEAC) como una probabilidad, en concreto la probabilidad de preferencia de un tratamiento frente a otro. Aunque ésta es la interpretación habitualmente empleada para la curva CEAC, la interpretación en términos de probabilidad solo tiene sentido desde un punto de vista bayesiano.

Palmer y col. (9) fueron los primeros autores en proponer la estimación bayesiana en modelos de decisión. En su artículo los autores comparan la estimación bayesiana y por remuestreo (bootstrapping) de un análisis coste-efectividad basado en cadenas de Markov, mostrando cómo los métodos bayesianos ofrecen un marco adecuado para evaluar la incertidumbre en la toma de decisiones en salud. Las probabilidades de transición del modelo y el impacto de la intervención son estimados de forma conjunta por métodos bayesianos. En Parmigiani (10) se puede consultar una completa revisión de los métodos bayesianos aplicados a modelos de decisión como árboles de decisión o meta-análisis.

Durante los primeros años del siglo XXI crece muchísimo el uso de la estadística bayesiana en modelos de decisión analíticos como son los ejemplos de Fryback y col. (11) y Cooper y col. (12,13). En este capítulo revisaremos qué pueden aportar los métodos bayesianos a través de tres secciones. En la primera de ellas revisaremos los conceptos básicos de la metodología bayesiana. En la segunda parte se mostrarán distintas áreas de los modelos analíticos de decisión donde la estadística bayesiana puede ofrecer soluciones óptimas. En concreto nos centraremos en el meta-análisis, la estimación de las probabilidades de transición, y el análisis de la sensibilidad. En el último capítulo mostramos una aplicación práctica real.



Métodos bayesianos

Los métodos bayesianos son una alternativa a los métodos clásicos de inferencia estadística. El argumento más utilizado para diferenciar ambas aproximaciones es que la estadística bayesiana permite la incorporación de información externa, adicional a los datos, en el modelo. Esta información se denomina “información a priori”, $p(\theta)$, y es combinada con la información que proviene de los datos, denominada “verosimilitud”, $p(x|\theta)$. La obtención del resultado final, o “distribución a posteriori”, se realiza gracias a la aplicación del Teorema de Bayes (14).

$$p(\theta|x) = \frac{p(\theta) \cdot p(x|\theta)}{p(x)} \quad (1)$$

Sin embargo, la incorporación de información a priori no es el único argumento a favor de la estadística bayesiana, ni siquiera es un requerimiento obligatorio para realizar modelos bayesianos. Para evitar la principal crítica recibida por la estadística bayesiana, en relación a la subjetividad de sus resultados y su dependencia de la distribución a priori, se ha desarrollado el denominado análisis bayesiano objetivo (15). Dicho análisis consiste en utilizar distribuciones a priori “vagas” o “no informativas” de tal forma que la distribución a posteriori dependa principalmente de la información suministrada por los datos. Un ejemplo paradigmático de distribución a priori no informativa es la distribución uniforme, que asigna la misma probabilidad a priori a un rango de posibles valores de un determinado parámetro. Una revisión de las distintas distribuciones a priori no informativas propuestas en la literatura puede encontrarse en Consonni y col. (16). En la práctica la gran mayoría de los análisis bayesianos realizados son objetivos.

Una segunda diferencia fundamental entre la perspectiva bayesiana y la frecuentista es cómo resuelve la estadística bayesiana el problema

de la inferencia sobre los parámetros del modelo. Desde la perspectiva bayesiana, la incertidumbre sobre los parámetros, por el hecho ser desconocidos, puede venir descrita a través de distribuciones de probabilidad. De esta forma se describe la distribución a priori (información inicial sobre los parámetros) y la distribución a posteriori (información sobre los parámetros tras incorporar los datos). Recordemos que en la estadística frecuentista los parámetros son considerados desconocidos pero fijos, y por lo tanto no es posible asignarles una distribución de probabilidad. Desde una perspectiva bayesiana la incertidumbre sobre los resultados puede ser descrita de una forma natural como probabilidades, y podemos hablar de que la probabilidad de que un tratamiento tenga un coste superior a una cierta cantidad es el 90%, o que el ICER se sitúa entre dos valores con una probabilidad del 95%. Esta lectura natural de los resultados no es posible desde la perspectiva frecuentista donde no se puede hablar de probabilidades asociadas a los parámetros de interés, sustituyendo este concepto por otros términos como “intervalo de confianza”, “niveles de significación” o “valor-p”, que en muchos casos no se interpretan adecuadamente.

A pesar de las ventajas del análisis bayesiano, su expansión no ha sido posible hasta los años 90 del siglo pasado, debido principalmente a sus dificultades computacionales. Durante años, el análisis bayesiano se vio limitado a los denominados “análisis conjugados”. El análisis bayesiano conjugado se define como aquel donde la distribución a posteriori de los parámetros de interés tiene la misma forma que la distribución a priori. Veamos un ejemplo sencillo de análisis conjugado:

Supongamos que disponemos de información sobre el número de éxitos en una determinada intervención quirúrgica. En este caso los datos seguirían una distribución binomial $X \sim Binom(n, \theta)$, y la verosimilitud sería:

$$p(x|\theta) = \binom{n}{x} \theta^x (1 - \theta)^{n-x} \propto \theta^x (1 - \theta)^{n-x} \quad (2)$$

dado que $\binom{n}{x}$ no depende del parámetro θ , podemos considerarlo una constante de normalización y eliminarlo.

El siguiente paso del análisis bayesiano sería definir la distribución a priori del parámetro θ . Al tratarse de un parámetro que toma valores entre 0 y 1 se propone la distribución Beta, es decir $\theta \sim \text{Beta}(a, b)$. A través de los valores de los hiperparámetros a y b se podrá definir la información a priori disponible.

$$p(\theta) = \frac{\Gamma(a + b)}{\Gamma(a) \Gamma(b)} \theta^{a-1} (1 - \theta)^{b-1} \propto \theta^{a-1} (1 - \theta)^{b-1} \quad (3)$$

Finalmente, la distribución a posteriori del parámetro de interés θ se obtiene a partir de la aplicación del teorema de Bayes:

$$p(\theta|x) \propto p(\theta) \cdot p(x|\theta) \propto \theta^{x+a-1} (1 - \theta)^{n-x+b-1} \quad (4)$$

Se observa, en esta ocasión, como la distribución a posteriori del parámetro θ sigue también una distribución Beta con parámetros $\theta|x \sim \text{Beta}(x + a, n - x + b)$, por lo que estaríamos ante un caso de análisis conjugado.

Ilustremos este caso con un ejemplo numérico con datos de $n = 30$ intervenciones, de las cuales $x = 10$ fueron un éxito. Se decide no incorporar información a priori empleando como distribución a priori la distribución $\text{Beta}(1,1)$, que equivale a una distribución uniforme, $p(\theta) \propto 1$. En este caso la distribución a posteriori del parámetro θ es analíticamente tratable y se corresponde con una distribución $\text{Beta}(11,21)$. A partir de dicha distribución podríamos concluir que su media a posteriori es 0,3435, la moda 0,3333, o, por ejemplo, que el parámetro θ está entre 0,2134 y 0,5137 con una probabilidad del 95 %.

Sin embargo, a medida que aumenta la complejidad del problema es más probable que la distribución a posteriori obtenida a través del

teorema de Bayes no tenga una expresión tratable analíticamente. Para estos casos se han desarrollado en la última década del siglo xx métodos de simulación, tales como los métodos Markov Chain Monte Carlo (MCMC) (1). El desarrollo de estos métodos de simulación, junto con la evolución de los equipos informáticos que permiten hacer estas simulaciones en poco tiempo, y el desarrollo de software específico para realizar estas simulaciones (WinBUGs o su versión abierta OpenBUGS) (17,18) han sido las principales claves de la expansión de la metodología bayesiana.

Para una primera aproximación de los métodos bayesianos en la evaluación de tecnologías sanitarias se recomienda la lectura de O'Hagan y Luce (19), que resume así las cinco principales ventajas de la metodología bayesiana:

1. Los métodos bayesianos resuelven la inferencia de forma más natural y útil, interpretando la incertidumbre sobre los parámetros estimados a través de distribuciones de probabilidad.
2. Los métodos bayesianos hacen uso de toda la información disponible, no únicamente de la información proporcionada por los datos.
3. Los métodos bayesianos pueden acometer problemas más complejos con los métodos frecuentistas.
4. Los métodos bayesianos son ideales para la toma de decisiones, ya que pueden concluir con qué probabilidad una determinada hipótesis es cierta.
5. Los métodos bayesianos son, en general, más transparentes que los métodos frecuentistas sobre las hipótesis que se asumen en el proceso de inferencia, recogidas en la verosimilitud asumida para los datos, y la distribución a priori para los parámetros del modelo.



Modelos analíticos de decisión bayesianos

Los modelos analíticos de decisión se pueden beneficiar de la metodología bayesiana desde diferentes perspectivas. En esta sección repasaremos algunas de ellas:

1. Meta-análisis
2. Estimación de las probabilidades de transición
3. Análisis de sensibilidad

1. Meta-análisis

En muchos casos los modelos analíticos de decisión aplicados a la evaluación económica de tecnologías sanitarias están basados en datos de coste o efectividad de trabajos previos publicados u otras fuentes de información secundarias. Las revisiones sistemáticas, y en concreto, el meta-análisis tienen como objetivo extraer la información de estas distintas fuentes de información y cuantificar el conocimiento disponible hasta ese momento de un determinado parámetro de interés.

Uno de los aspectos principales del meta-análisis es el tratamiento de la heterogeneidad entre estudios. Desde una perspectiva frecuentista, los denominados “modelos de efectos aleatorios” reconocen que existe dicha heterogeneidad, por lo que existirá un efecto para cada estudio y la estimación del efecto global deberá ser algún tipo de media ponderada entre estos efectos individuales. Es en este apartado donde se aprecia la principal diferencia entre la perspectiva frecuentista y bayesiana. En el análisis bayesiano, los parámetros propios de cada estudio son considerados variables aleatorias que siguen una determinada distribución de probabilidad cuya media, o meta-parámetro, mediría el efecto global. A continuación, se muestra a modo ilustrativo

el meta-análisis bayesiano para el caso normal. Se asume, por tanto, que los datos de cada estudio siguen una distribución normal,

$$x_{ij}|\theta_j, \sigma_j \sim N(\theta_j, \sigma_j), \forall i = 1, \dots, n_j, \quad (5)$$

y que los parámetros θ_j se distribuyen también normalmente entre los estudios, siendo su media el parámetro θ que mide el efecto global, y la varianza σ que mide la heterogeneidad entre-estudios.

$$\theta_j \sim N(\theta, \sigma) \quad \forall j = 1, \dots, k \quad (6)$$

Para finalizar la formulación del modelo bayesiano se deben definir las distribuciones a priori de los parámetros del modelo, a través de los hiperparámetros (a, b, c, d) . Un ejemplo de distribución a priori no informativa sería $a = 0, b = 10^5, c = c_j = d = d_j = 0.001, \forall j = 1, \dots, k$.

$$\theta \sim N(a, b)$$

$$\sigma^2 \sim \text{Inversa-Gamma}(c, d)$$

$$\sigma_j^2 \sim \text{Inversa-Gamma}(c_j, d_j) \quad \forall j = 1, \dots, k.$$

Este ejemplo muestra también como la perspectiva bayesiana ofrece un tratamiento natural de la incertidumbre asociado a la estimación del meta-parámetro, que incluye la variabilidad intra-estudio y la entre-estudios (20,21). Además, la flexibilidad del análisis bayesiano ha favorecido su aplicación para resolver problemas en el meta-análisis como el tratamiento de datos discretos (22), meta-análisis multivariantes (23), o el meta-análisis en red (24).

2. Estimación de las probabilidades de transición

Las probabilidades de transición entre eventos son uno de los parámetros fundamentales para la realización de los modelos analíticos de decisión. Lee y Zellner (25) presentaron hace más de 50 años la aproximación bayesiana a la estimación de las probabilidades de transición en modelos de Markov estacionarios. En este trabajo se compararon los estimadores bayesianos, de máxima verosimilitud y de mínimos cuadrados, concluyendo para todos los tamaños muestrales que el estimador bayesiano era superior.

La flexibilidad de la metodología bayesiana ha permitido que muchos autores hayan adoptado esta metodología para resolver problemas como la estimación de probabilidades de transición con información incompleta (26). Pelzer y Eisinga (27) discutieron las ventajas de la estimación bayesiana frente al estimador de máxima verosimilitud para estimar probabilidades de transición con datos de encuestas transversales repetidas. La comparación entre metodologías se ilustró evaluando un programa de prevención del VIH entre consumidores de droga.

Uno de los argumentos adicionales para emplear metodología bayesiana es que, en ocasiones, los resultados provenientes de la literatura, o de un meta-análisis, no pueden ser utilizados directamente en el modelo de decisión. Imaginemos por ejemplo que la información disponible está medida en Riesgos Relativos u Odds ratios, y que para nuestro modelo necesitamos conocer las probabilidades de transición. Será necesario realizar la transformación a probabilidades (28). El principal problema surge en el tratamiento de la incertidumbre. Es de esperar que los valores obtenidos de la literatura, o de un meta-análisis, hayan sido estimados con incertidumbre, y esta incertidumbre debe trasladarse a la estimación de las probabilidades de transición. El método delta (29) o bootstrap (30) han sido las aproximaciones más utilizadas para la estimación de esta incertidumbre. Sin embargo,

esta estimación puede realizarse de forma sencilla desde la perspectiva bayesiana. A partir de la distribución de un determinado parámetro puede estimarse la distribución de cualquier transformación de dicho parámetro a través de los métodos MCMC.

3. Análisis de sensibilidad

Los análisis de sensibilidad permiten estudiar la robustez de los resultados de un modelo de decisión ante variaciones sobre los distintos supuestos realizados. Los análisis de sensibilidad también han sido propuestos como un medio de cuantificar la incertidumbre asociada a la toma de decisiones, principalmente los análisis de sensibilidad probabilísticos, que asignan distribuciones de probabilidad a los distintos parámetros del modelo. Los modelos de decisión bayesianos ya incorporan toda la incertidumbre asociado a la decisión a través de las distribuciones de probabilidad a posteriori. Las conclusiones finales, ya sean en términos de efectividad, coste o coste–efectividad, incorporan en su distribución a posteriori la incertidumbre en la estimación de todos los parámetros intermedios (probabilidades de transición, utilidades, costes, etc.) (31). Por este motivo, los modelos de decisión bayesianos se han asimilado a los análisis de sensibilidad probabilísticos, aunque difieren en la forma en la que se definen las distribuciones de probabilidad (32).

Sin embargo, los modelos bayesianos incorporan una fuente de incertidumbre que sí debe ser explorada a través del análisis de sensibilidad: las distribuciones a priori. El análisis bayesiano permite incorporar información a priori en el análisis y puede ser relevante analizar cómo varían los resultados finales ante la variación de la distribución a priori. Este análisis se le denomina “análisis de robustez bayesiano” (33). Incluso cuando se emplean distribuciones a priori no informativas puede ser relevante realizar análisis de robustez

bayesiano (34). No siempre es sencillo definir distribuciones completamente no informativas, principalmente para los parámetros de dispersión (35).

El análisis bayesiano permite además añadir una fuente de incertidumbre extra en la toma de decisiones, la incertidumbre asociada al modelo. En muchos modelos de decisión se pueden plantear dudas sobre qué modelo siguen los datos. Por ejemplo, en modelos de supervivencia son varias las distribuciones de probabilidad que podemos asumir que siguen los datos (36). Sin embargo, desde la perspectiva bayesiana no es necesario seleccionar un único modelo entre todos los posibles. El análisis bayesiano permite estimar la probabilidad a posteriori para cada uno de los modelos en discusión, y que la decisión final se obtenga a través de una media ponderada de los resultados observados en cada modelo. Este análisis se conoce como BMA (*Bayesian Model Averaging*) (37). Un ejemplo ilustrativo de la aplicación del BMA en la extrapolación de datos de supervivencia se puede encontrar en Negrín y col. (38).

4. Ejemplo práctico de elaboración de un modelo bayesiano

Para finalizar este capítulo mostraremos un ejemplo práctico de un modelo bayesiano de árbol de decisión. En concreto se realizará un estudio coste–efectividad del uso de inhibidores de la neuraminidasa (NIs) para reducir la incidencia de la gripe A y B en adultos sanos. Este ejemplo ilustrativo ha sido extraído de Cooper y col. (13), uno de los artículos pioneros en mostrar la aproximación bayesiana a los modelos analíticos de decisión. En este capítulo se mostrarán los pasos a seguir para realizar el análisis y los resultados que obtuvieron los autores. Los códigos necesarios para la ejecución del modelo en WinBUGS/OpenBUGS están disponibles como Anexo en el artículo Cooper y col. (13).

El análisis aquí mostrado se trata de un árbol de decisión simplificado basado en el modelo de decisión desarrollado por la Agencia de Evaluación de Tecnologías Sanitarias del NHS (National Health Service) del Reino Unido (39). El objetivo del estudio es analizar el coste–efectividad de prescribir NIs a la población adulta (entre 12 y 65 años) durante una epidemia de gripe. El árbol de decisión correspondiente a este análisis se muestra en la Figura 1.



Figura 1. Árbol de decisión



Meta–análisis

La efectividad de la prescripción de NIs es estimada a través de un meta–análisis que incluye 6 estudios resumidos en la Tabla 1.

Tabla 1. Estudios previos empleados en el meta-análisis

Estudio	N placebo	Gripe placebo (%)	N NIs	Gripe NIs (%)
Estudio 1	554	34 (6,1 %)	553	11(2,0 %)
Estudio 2	423	40 (9,5 %)	414	7 (1,7 %)
Estudio 3	144	9 (6,3 %)	144	3 (2,1 %)
Estudio 4	268	19 (7,1 %)	268	3 (1,1 %)
Estudio 5	251	6 (2,4 %)	252	3 (1,2 %)
Estudio 6	462	34 (7,4 %)	493	4 (0,8 %)

A partir de esta información se desarrolló un meta-análisis bayesiano cuyo resultado vendrá medido a través de odds ratio (OR) (20). La especificación del meta-análisis es:

$$r_j^c \sim \text{Binomial}(n_j^c, p_j^c) \quad \forall j = 1, \dots, k$$

$$r_j^t \sim \text{Binomial}(n_j^t, p_j^t) \quad \forall j = 1, \dots, k$$

$$\mu_j = \text{logit}(p_j^c)$$

$$\text{logit}(p_j^t) = \mu_j + \delta_j$$

$$\delta_j \sim N(\Delta, \tau)$$

$$OR = \exp(\Delta)$$

donde $k = 6$ se refiere al número de estudios; r_j^c se refiere al número de casos de gripe en el grupo placebo, de los n_j^c , casos estudiados; similarmente r_j^t y n_j^t corresponden al grupo intervención NIs. p_j^c y p_j^t son las tasas de infección del placebo y del grupo NIs en el estudio j , respectivamente; μ_j es la transformación logarítmica del odds del grupo control para el estudio j , mientras que δ_j se interpreta como la transformación logarítmica del Odds Ratio (OR) entre ambos tratamientos para el

estudio j . El parámetro Δ recoge la transformación logarítmica del OR global, no condicionado al estudio, y τ es la desviación típica que mide la variabilidad entre estudios.

El siguiente paso para la definición completa del modelo bayesiano es asignar distribuciones a priori sobre los parámetros del modelo, en este caso:

$$\mu_j \sim N(0,10)$$

$$\Delta \sim N(0,10)$$

$$\tau^2 \sim \text{Inversa-Gamma}(0.001, 0.001)$$

En este ejemplo se ha optado por especificar distribuciones a priori vagas o no informativas. En concreto, distribuciones normales con varianzas grandes para los log(odds) de cada estudio y el log(OR), y la distribución inversa-gamma para la varianza con ambos parámetros cercanos a 0 (40).

Estimación de las probabilidades de transición

En este ejemplo, los resultados del meta-análisis anterior no se expresan directamente como probabilidades. Sin embargo, las probabilidades de contraer gripe en el grupo control (p^c) y en el grupo NIs (p^i) pueden obtenerse a partir de la estimación del OR del meta-análisis. En concreto, podemos partir de la proporción de población del Reino Unido que contrae gripe durante un periodo epidémico sin hacer uso de inhibidores NIs. Esta proporción es de $p^c = 0.0542$, lo que equivale a un odds de 0.0573. La probabilidad de contraer gripe en el grupo de tratamiento NIs se obtendría a partir de las siguientes fórmulas:

$$odds^t = OR \cdot odds^c$$

$$p^t = \frac{odds^t}{1 + odds^t}$$

Notar que no se estimará un único valor de p^t sino que se aproximará la distribución a posteriori de dicho parámetro a partir de la distribución a posteriori del OR. Esta forma de proceder permite que la incertidumbre en la estimación del OR se traslade a la estimación de p^t .

Otros parámetros del modelo

Para realizar el análisis coste–efectividad debemos definir los costes. En este ejemplo se asume que el coste medio de un episodio de gripe para un adulto es de £16.78 (SD=2.34), que incluye las visitas al médico general, visitas ambulatorias e ingresos hospitalarios. Se asume una distribución normal a estos costes (39). Se asume también que el tiempo medio para eliminar los síntomas de la gripe es de 8.2 días (SD = $\sqrt{2}$) (39), y que sigue una distribución log–normal. Para los pacientes en tratamiento NIs se añadirá el coste del tratamiento. Se asume un seguimiento a los pacientes de 6 semanas, y el coste del tratamiento se considera fijo de £2.40 diarios, es decir, £100.80 en las seis semanas, que tras añadir impuestos y otros gastos generales se fija en £118.44. Además, a la visita al médico general donde se prescriben los NIs se le asigna un coste de £19 (41). Todos los precios están expresados en precios corrientes de 2000/2001. En este ejemplo, y por simplicidad, se han asumido que muchos de los costes estudiados son conocidos (fijos). Por supuesto, se podría proponer un modelo más complejo donde los costes fueran modelizados a través de distribuciones de probabilidad que recogieran la incertidumbre sobre los mismos.

Análisis de sensibilidad

En este apartado se podría analizar la robustez de los resultados del modelo ante variaciones en cualquiera de los parámetros del mismo, como por ejemplo el apartado costes. Sin embargo, centraremos este análisis de sensibilidad a estudiar el impacto de los resultados a distintas especificaciones no informativas de las distribuciones a priori, en concreto, la distribución a priori asumida al parámetro de varianza τ^2 , que mide la variabilidad entre estudios en el meta-análisis. En la literatura se han propuesto distintas distribuciones a priori no informativas para los parámetros de varianza y, en ciertos modelos jerárquicos, los resultados pueden variar (34). En este ejemplo se muestran tres distribuciones a priori no informativas alternativas:

1. $\tau^2 \sim \text{Inversa-Gamma}(0.001, 0.001)$
2. $\tau \sim \text{Uniforme}(0, 10)$
3. $\tau^2 \sim \text{Uniforme shrinkage}$ o encogimiento

La primera distribución a priori es la distribución no informativa más utilizada. Esta distribución asigna probabilidades a priori similares para un amplio rango de valores positivos, pero muestra un peso algo superior a los valores cercanos a 0. El segundo ejemplo de distribución a priori no informativa es la uniforme, que asigna el mismo peso a valores de τ entre 0 y 10. Por último, la distribución Uniforme *shrinkage* utiliza la variación muestral observada entre los estudios (σ^2) como referencia para estimar los valores a priori de la varianza $\tau^2 = \sigma^2 \cdot (1 - B)/B$, donde $B \sim \text{Uniforme}(0, 1)$ (42).

Otras distribuciones a priori no informativas podrían también incluirse en el análisis de sensibilidad, o incluso distribuciones a priori informa-

tivas construidas a partir de la opinión de expertos o estudios previos. De esta forma se puede evaluar la robustez de los resultados a la distribución a priori.

Estimación del modelo

El modelo aquí planteado se estima a través de métodos de simulación MCMC. En concreto se propone una cadena con 25000 simulaciones, de las cuales 5000 son descartados como muestra *burn-in*. La simulación se ha realizado utilizando el programa OpenBUGS (18).

Resultados

La Tabla 2 muestra los resultados del modelo. Del meta-análisis podemos concluir que el OR del tratamiento NIs en relación al control es estimado en 0.2041 de media, con un intervalo de credibilidad al 95 % de (0.1239, 0.3138), por lo que la probabilidad de contraer gripe es inferior en el grupo NIs. A partir de este OR se estima una probabilidad de contraer gripe en el grupo NIs de 0.0116, con un intervalo de credibilidad al 95 % de (0.0071, 0.0177). El coste de los pacientes en el grupo control es estimado en media en £19.91, frente al coste de £137.6 del grupo NIs. Por otro lado, el número de días de gripe evitados se estima en 0.0890 de media. La ratio coste–efectividad incremental (RCEI) muestra un coste de £1340.0 por día evitado de gripe, IC 95 % (1102.0, 1725.0).

Tabla 2. Estudios previos empleados en el meta-análisis

Variable	Media	Desviación Típica	Error MC	Intervalo credibilidad 95 %
OR	0.2041	0.0533	0.0012	(0.1239, 0.3138)
p^c	0.0542	.	.	.
p^t	0.0116	0.0030	0.00007	(0.0071, 0.0177)
Coste control	19.91	0.1268	0.0009	(19.66, 20.16)
Coste NIs	137.6	0.0573	0.0012	(137.5, 137.8)
Diferencia costes	117.7	0.1119	0.0014	(117.5, 117.9)
Días evitados	0.0890	0.0099	0.0002	(0.0683, 0.1068)

Del análisis de sensibilidad podemos concluir que la estimación de la RCEI es bastante robusta. Para la distribución a priori 2 se obtienen una RCEI algo menos precisa, con una media de £1350.0 por día evitado de gripe, IC 95 % (1086.0, 1879.0); mientras que para la distribución a priori 3 las diferencias son menores, con una media de £1354.0 por día evitado de gripe, IC 95 % (1099.0, 1785.0).



Bibliografía

1. Gilks W, Richardson S, Spiegelhalter D. Markov Chain Monte Carlo in Practice. Chapman and Hall, London. 1996.
2. Eddy DM, Hasselblad V, Shachter R. An introduction to a Bayesian method for meta-analysis: The confidence profile method. *Med Decis Making*. 1990;10:15-23.
3. Eddy D, Hasselblad V, Shachter R. Meta-analysis by the Confidence Profile Method: The Stastical Synthesis of Evidence. Academic Press, Boston. 1992.
4. Spiegelhalter D, Feedman L, Parmar M. Bayesian approaches to randomized trial (with discussion). *J R Stat Soc Series A*. 1994;157:357-416.
5. Jones D. Bayesian approach to the economic evaluation of health care technologies. In B. Spiker, ed., *Quality of Life and Phamaeconomics in Clinical Trials*. Philadelphia Lippincott-Raven, 1996, p. 1189-1196..
6. O'Hagan A, Stevens JW, Montmartin J. Inference for the cost-effectiveness acceptability curve and cost-effectiveness ratio. *PharmacoEconomics*. 2000;17:339-49.
7. O'Hagan A, Stevens JW. Bayesian assessment of sample size for clinical trials of cost-effectiveness. *Med Decis Making*. 2001;21:219-30.
8. O'Hagan A, Stevens JW. A framework for cost-effectiveness analysis from clinical trial data. *Health Econ*. 2001;10:303-15.
9. Palmer S, Samsa G, Ancukiewicz M, et al. Assessing uncertainty in cost-effectiveness analyses. *Med Decis Making*. 1997;17:390-401.
10. Parmigiani G. *Modelling in Medical Decision Making: A Bayesian Approach (Statistics in Practice)*. Chichester Wiley. 2002.
11. Fryback DG, Chinnis JO Jr, Ulvila JW. Bayesian cost-effectiveness analysis. An example using the GUSTO trial. *Int J Technol Assess Health Care*. 2001;17:83-97.
12. Cooper NJ, Sutton AJ, Abrams KR. Decision analytical economic modelling within a Bayesian framework: application to prophylactic antibiotics use for caesarean section. *Stat Methods Med Res*. 2002;11:491-512.

13. Cooper NJ, Sutton AJ, Abrams KR, et al. Comprehensive decision analytical modelling in economic evaluation: a Bayesian approach. *Health Econ.* 2004;13:203-26.
14. Bayes T. An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London.* 1763;53:370-418. <https://royalsocietypublishing.org> (acceso verificado 5 de Nov de 2023)
15. Berger J. The case for objective bayesian analysis. *Bayesian Anal.* 2006;1:385-402.
16. Consonni G, Fouskakis D, Liseo G, et al. Prior distributions for objective bayesian analysis. *Bayesian Anal.* 2018;13:627-79.
17. Spiegelhalter D, Thomas A, Best N. Winbugs version 1.3. User manual, Biostatistics Unit, Cambridge Medical Research Council. 2000.
18. Lunn D, Spiegelhalter D, Thomas A, et al. The bugs project: Evolution, critique and future directions. *Stat Med.* 2009;28:3049-67.
19. O'Hagan A, Luce B. A Primer on Bayesian Statistics in Health Economics and Outcomes Research. Bayesian Initiative in Health Economics & Outcomes Research, Centre of Bayesian Statistics in Health Economics, Sheffield. 2003.
20. Sutton AJ, Abrams KR. Bayesian methods in meta-analysis and evidence synthesis. *Stat Methods Med Res.* 2001;10(4):277-303.
21. Schmid C, Carlin B, Welton N. Bayesian methods for meta-analysis. In C. Schmid, T. Stijnen y I. White, eds. 1st Ed. *Handbook of Meta-Analysis.* New York Chapman and Hall. 2020; p. 38
22. Moreno E, Vázquez-Polo FJ, Negrín MA. Objective Bayesian meta-analysis for sparse discrete data. *Stat Med.* 2014;33:3676-92.
23. Nam I, Mengersen K, Garthwaite P. Multivariate meta-analysis. *Stat Med.* 2003;22:2309-33.
24. Greco T, Landoni G, Biondi-Zoccai G, et al. A Bayesian network meta-analysis for binary outcome: how to do it. *Stat Methods Med Res.* 2016;25:1757-73.
25. Lee T, Zellner A. Maximum likelihood and bayesian estimation of transitions probabilities. *J Am Stat Assoc.* 1968;63:1162-79.

26. Welton NJ, Ades AE. Estimation of markov chain transition probabilities and rates from fully and partially observed data: uncertainty propagation, evidence synthesis, and model calibration. *Med Decis Making.* 2005;25:633-45.
27. Pelzer B, Eisinga R. Bayesian estimation of transition probabilities from repeated cross sections. *Stat Neerl.* 2001;56:23-33.
28. Miller DK, Homan SM. Determining transition probabilities: confusion and suggestions. *Med Decis Making.* 1994;14:52-8.
29. Oehlert G. A note on the delta method. *Am Stat.* 1992;46:27-9.
30. M, Betensky RA. Simultaneous Confidence Intervals Based on the Percentile Bootstrap Approach. *Comput Stat Data Anal.* 2008;52:2158-65.
31. Briggs A. Handling uncertainty in cost-effectiveness models. *Pharmacoeconomics.* 2000;17:479-500.
32. Boshuizen H, van Baal P. Probabilistic sensitivity analysis: Be a bayesian. *Value Health.* 2009;12:1210-4.
33. Berger J. Robust bayesian analysis: sensitivity to the prior. *J Stat Plan Inference.* 1990;25:303-28.
34. Lambert P, Sutton A, Burton P, et al. How vague is vague? a simulation study of the impact of the use of vague prior distributions in mcmc using winbugs. *Stat Med.* 2005;24:2041-28.
35. Lemoine N. Moving beyond noninformative priors: why and how to choose weekly informative priors in Bayesian analyses. *Oikos.* 2019;128:912-28.
36. Ibrahim J, Chen M, Sinha D. *Bayesian Survival Analysis.* Springer series in statistics. Springer-Verlag, Berlin. 2001.
37. Hoeting J, Madigan D, Raftery A et al. Bayesian model averaging: A tutorial. *Stat Sci.* 1999;14:382-427.
38. Negrín MA, Nam J, Briggs AH. Bayesian Solutions for Handling Uncertainty in Survival Extrapolation. *Med Decis Making.* 2017;37:367-76.
39. Turner D, Wailoo A, Nicholson K, et al. Systematic review and economic decision modelling for the prevention and treatment of influenza A and B. *Health Technol Assess.* 2003;7(35): 1-170.

40. Gelman A. Prior distributions for variance parameters in hierarchical models. *Bayesian Anal.* 2006;1:515-33
41. Netten J, Rees T, Harrison G. Unit cost of health and social care 2001. Tech rep. 2001. Personal Social Services Research Unit (PSSRU). <https://www.ukc.ac.uk/PSSRU> (acceso verificado 5 Nov 2023)
42. DuMouchel W, Normand, S. Computer-modeling and graphical strategies for meta-analysis. In D. Stangl y D. Berry, eds., *Meta-analysis in Medicine and Health Policy*. New York. Marcel Dekker 2000; p. 127–178.



15

SOFTWARE PARA ANÁLISIS DE DECISIONES EN MEDICINA

Francisco Javier Díez^a, Carmen María Yago^a, Jorge Pérez-Martín^a,
Manuel Luque^b, Manuel Arias^a

^a *Dpto. Inteligencia Artificial*

^b *Dpto. Estadística, Investigación Operativa y Cálculo Numérico
Universidad Nacional de Educación a Distancia (UNED), Madrid*



Introducción

En el primer capítulo de este libro hemos visto por qué los modelos analíticos son útiles en la toma de decisiones, especialmente cuando se trata de determinar la relación de coste-efectividad de una o varias intervenciones. En la práctica los modelos se construyen y se evalúan con computadores (generalmente, ordenadores personales), pues por su complejidad resultaría imposible analizarlos a mano o con una calculadora. En este capítulo vamos a estudiar los programas de ordenador más utilizados en la evaluación económica de tecnologías sanitarias.



Tipos de modelos

Los capítulos precedentes han presentado los modelos de decisión utilizados habitualmente en evaluación económica. Entre los modelos atemporales, es decir, aquéllos que no incluyen una representación explícita del tiempo ni de la evolución del paciente, los árboles de decisión, explicados en el capítulo 9, son los más comunes, pero también pertenecen a este grupo varios tipos de modelos gráficos probabilistas (1), como las redes bayesianas (2), los diagramas de influencia (3-5) y las redes de análisis de decisiones (6,7).

Entre los modelos temporales, los más comunes son los de transición entre estados, estudiados en los capítulos 10 y 11, que en la práctica coinciden con los modelos de Markov. Un ejemplo especialmente interesante, al que haremos referencia varias veces, es el de Paulden y col. (8) para el tratamiento del cáncer de mama. Inicialmente se implementó en Excel. Posteriormente Hollman y col. (9) lo implementaron

también en R, MATLAB y TreeAge Pro; las cuatro versiones del modelo están disponibles como material complementario en la página web del artículo.^a Hay una quinta versión, implementada en OpenMarkov (10). En las próximas secciones citaremos este modelo como “el de Paulden”.

También son modelos temporales los de simulación de eventos discretos (SED; cf. capítulo 12). El modelo para el tratamiento de la osteoporosis desarrollado por la Decision Support Unit (DSU) del British National Institute for Health and Care Excellence (NICE) (11), al que nos referiremos en las próximas secciones como “el de la DSU”, tiene varias implementaciones: tanto las versiones en R, Excel/VBA, Excel/DICE, Simul8 y TreeAge Pro, implementadas por la propia DSU,^b como las de Arena y OpenMarkov, implementadas por Carmen M. Yago, están disponibles en Internet.^c

Los modelos de transmisión dinámicos, presentados en el capítulo 13, constituyen el tercer tipo de modelo temporal.

Por otro lado, el capítulo 5 ha analizado los distintos tipos de incertidumbre que surgen al construir modelos de evaluación económica. Entre ellos se encuentra la incertidumbre sobre los valores numéricos de los parámetros, la cual puede representarse asignando distribuciones de probabilidad —denominadas “de segundo orden”— a los parámetros. Todos los modelos (ya sean árboles de decisión, diagramas de influencia, modelos de Markov, SED, etc.) se consideran *bayesianos* cuando tienen distribuciones de segundo orden. El capítulo 14 estudia algunos de ellos.

^a <https://link.springer.com/article/10.1007%2Fs40273-017-0510-8>

^b <https://www.sheffield.ac.uk/nice-dsu/tsds/patient-level-simulation>

^c <https://www.mdpi.com/2227-7390/11/7/1602>



Criterios para evaluar las ventajas e inconvenientes

La elección de una herramienta de software para toma de decisiones en medicina ha de hacerse teniendo en cuenta las perspectivas de cada parte implicada. Por un lado tenemos a aquellos que crean el modelo, generalmente investigadores académicos, agencias de evaluación de tecnologías sanitarias o empresas especializadas. Por otro lado, están quienes reciben el modelo, normalmente las mismas agencias de evaluación de tecnologías sanitarias, que examinan detalladamente la evidencia clínica y el modelo construido, aconsejando a los gobiernos qué intervenciones debe cubrir el sistema público de salud.

Los criterios para elegir una herramienta u otra afectan a todas las partes implicadas, aunque no en igual medida. Aquellos que implican principalmente a los creadores del modelo son:

- Capacidad expresiva: El software elegido tiene que permitir implementar el tipo de modelo más adecuado a la naturaleza del problema y a los datos disponibles; dicho de otro modo, el modelo debe condicionar la elección de la herramienta, y no al revés.
- Esfuerzo requerido para construir y modificar el modelo: La herramienta debe permitir construir y depurar el modelo en un tiempo aceptable. Además, muchas veces los modelos evolucionan por motivos de cualquier índole y la herramienta debe permitir modificarlo con un esfuerzo razonable, sin tener que rehacerlo desde cero.
- Aceptación por el receptor: La herramienta debe estar entre aquellas que el destinatario acepta.
- Precio: la herramienta debe ser viable económicamente dado el presupuesto de que se disponga.

Para el receptor del modelo son de especial importancia estos criterios:

- **Transparencia:** La herramienta debe ofrecer la posibilidad de examinar, de forma relativamente sencilla y sin demasiado esfuerzo, si las hipótesis del modelo son adecuadas y si la implementación es correcta.
- **Capacidad de verificación y validación,** que está directamente relacionada con la transparencia: La herramienta debe permitir determinar que el modelo implementado se corresponde con el modelo conceptual, y que éste se corresponde con la realidad.

Por último, algunos criterios afectan directamente a ambas partes son:

- **Curva de aprendizaje:** La toma de contacto inicial no debe ser tan costosa que haga preferir el uso de herramientas más sencillas pero quizá menos adecuadas.
- **Eficiencia:** la herramienta no debe precisar un consumo excesivo de recursos —como memoria o tiempo de computación— para obtener los resultados una vez que el modelo está implementado.

Con estos criterios vamos a juzgar cada una de las herramientas que mostraremos a continuación.



Software utilizado para análisis de decisiones

Software de propósito general

En esta sección presentamos varios lenguajes de programación y programas de ordenador de propósito general, es decir, aplicables —en principio— a cualquier problema informático, que se utilizan habitualmente para construir modelos de evaluación económica.

1.1. Lenguajes y entornos de programación

Dado que el análisis de decisiones sanitarias implica la realización de numerosos cálculos de probabilidad y estadística, muchos de los modelos se construyen utilizando lenguajes y herramientas orientados al cálculo matemático.

Uno de ellos es el lenguaje R, cada vez más popular en numerosos campos de aplicación. En el análisis de decisiones sanitarias se ha utilizado para implementar todo tipo de modelos.

Un ejemplo clásico es el análisis de coste-efectividad de diferentes fármacos para la epilepsia, de Hawkins y col. (12) El motivo para implementarlo en R fue la dificultad de incluir en Excel probabilidades de transición que variaban a lo largo del tiempo; es lo que a veces se denomina *modelo semi-markoviano* (*semi-Markov model*, en inglés). Otro ejemplo de modelo de Markov para análisis de coste-efectividad es el de Paulden, ya citado.

Además, el grupo *Decision Analysis in R for Technologies in Health (DARTH)* ha desarrollado y recopilado numerosos recursos muy interesantes, accesibles desde su página web.^d En particular, han publicado dos tutoriales: uno explica cómo construir modelos semi-markovianos de cohortes (13), mientras que el otro explica cómo implementarlos mediante microsimulación (14).

Análogamente, la *Innovation and Value Initiative*^e agrupa a varios investigadores que tienen como objetivo crear modelos abiertos (*open-source*) para el sistema sanitario de Estados Unidos, que puedan ser criticados y enriquecidos con múltiples aportaciones (como en los programas de software libre). Este grupo ha publicado dos modelos de Markov: uno para artritis reumatoide y otro para cáncer de pulmón no microcítico; cada modelo fue implementado en R y en C++, y ambas versiones están disponibles en abierto (15). Dos miembros de este grupo, Incerti y Jansen, han

^d <https://darthworkgroup.com>

^e <https://www.thevalueinitiative.org>

implementado además una librería en R para análisis de coste-efectividad llamada *hesim*.^f

Como ejemplo de simulación de eventos discretos implementado en R podemos citar el de Gray y col. (16) para el cribado de cáncer de mama. El motivo para utilizar una SED fue la heterogeneidad de las pacientes y la variedad de eventos que pueden ocurrirles en diferentes momentos de su vida, como la detección en una ronda de cribado, la detección clínica (por síntomas), la muerte por otras causas, etc. Del mismo modo, Glover y col. (17) implementaron en R una SED para el análisis de coste-efectividad del cribado de aneurisma aórtico abdominal, tras constatar que un modelo de Markov construido previamente era incapaz de modelizar la complejidad de este problema.

En algunas ocasiones R se ha utilizado en combinación con otras herramientas. Por ejemplo, un modelo de Markov para el tratamiento de niños infectados con el VIH (18) fue implementado en R por la complejidad de sus ecuaciones matemáticas en que se basaba; en concreto, las probabilidades de transición dependían de la edad en que cada niño entraba en el modelo y del tiempo en que había permanecido en cada estado. Sin embargo, se construyó una interfaz de entrada-salida en Excel, para que los expertos en esta enfermedad, que no sabían programar en R, pudieran modificar los parámetros del modelo y analizar los resultados.

Otro lenguaje de programación de propósito general utilizado a veces para análisis de decisiones en medicina es C++. Por ejemplo, Chhatwal y col. (19) implementaron en este lenguaje un modelo markoviano de microsimulación para el análisis de coste-efectividad y de impacto presupuestario de los nuevos antivirales para la hepatitis C. Posteriormente, Hodgson y col. (20) publicaron un análisis de coste-efectividad de los tratamientos del virus sincitial respiratorio (VSR), cuyo código en C++ está disponible en internet.

Hay además varios programas comerciales para cálculo matemático en los que es posible desarrollar programas complejos utilizando las numerosas

^f <https://cran.r-project.org/web/packages/hesim>

funciones y librerías que proporcionan. Uno de ellos es MATLAB, desarrollado por MathWorks.^g Como hemos indicado, existe una versión en MATLAB del modelo de Paulden, disponible en Internet.

Stata es otro programa comercial para cálculo matemático y el análisis estadístico, especialmente orientado a la estadística, creado por StataCorp LLC.^h En medicina se ha utilizado sobre todo para analizar datos de pacientes individuales, obtenidos en ensayos clínicos. Por ejemplo, Bachman y col. (21) realizaron un análisis de coste-efectividad con datos de estudios aleatorizados por grupos (*cluster randomized trials*). Utilizando las funciones matemáticas de Stata, construyeron dos modelos de regresión, uno de mínimos cuadrados y otro bayesiano, para comparar los resultados. (El mismo artículo explica cómo hacer lo mismo en WinBUGS, tal como veremos en la sección de modelos bayesianos.) Por su parte, Leurent y col. (22) explican en un tutorial cómo realizar un análisis de sensibilidad sobre un estudio de coste-efectividad cuando no se puede suponer que los datos ausentes se deben al azar (*not-at-random missing data*). El código de su método, en Stata, está disponible como material suplementario de su artículo. Más recientemente, Mutubuki y col. (23) muestran la importancia de realizar un análisis estadístico correcto en los análisis de coste-efectividad, evitando suposiciones simplistas. También ellos ofrecen el código en Stata como material suplementario.

Mathematica, de Wolfram Research Inc.,ⁱ es un programa capaz de realizar cálculo simbólico (es decir, puede operar con símbolos, y no sólo con números; ello le permite, por ejemplo, calcular la integral de una función). En economía de la salud se ha utilizado sobre todo para el análisis de coste-efectividad mediante modelos dinámicos de transmisión; por ejemplo, para la vacunación contra el virus del papiloma humano (VPH) (24) y contra la varicela (25).

^g <https://www.mathworks.com/products/matlab.html>

^h <https://www.stata.com>

ⁱ <https://www.wolfram.com/mathematica>

Ventajas e inconvenientes

La principal ventaja de utilizar un lenguaje de programación es la absoluta flexibilidad en cuanto a los modelos que pueden implementarse. El principal inconveniente es la necesidad de escribir el código del programa. Esta tarea puede verse ayudada cuando se utiliza un lenguaje de programación, como R, MATLAB, Stata o Mathematica, que incorpora numerosas funciones matemáticas, algunas elementales (por ejemplo, la multiplicación de matrices) y otras orientadas a la estadística (por ejemplo, cálculos con distribuciones de probabilidad), con las que poder construir programas sofisticados.

C++ tiene la ventaja de la rapidez, por ser un lenguaje más cercano al *código máquina*, es decir, a las instrucciones que va a ejecutar el procesador, pero esto acarrea dos inconvenientes. El primero es que la programación en C++ es mucho más compleja que en otros lenguajes de alto nivel (es decir, menos parecidos a las instrucciones del procesador), como los mencionados en el párrafo anterior. El segundo inconveniente es que no incluye las funciones matemáticas y estadísticas necesarias para el análisis de decisiones. Aunque este problema se puede paliar utilizando librerías especializadas, crear modelos en C++ sigue siendo complicado, y por ello, en la práctica sólo utilizan este lenguaje equipos que cuentan con ingenieros informáticos u otros expertos en programación.

En cuanto a la legibilidad del modelo, este aspecto depende, como en cualquier programa informático, del estilo del programador. Si el código está bien estructurado en funciones (llamadas métodos o procedimientos en algunos lenguajes) de tamaño razonable, si los nombres de las variables y las funciones están bien escogidos y se han insertado en el código numerosos comentarios explicativos, es mucho más fácil entender el modelo. Lamentablemente, aunque estas recomendaciones están bien detalladas en todos los libros de ingeniería del software, muchos programadores no las tienen en cuenta, sobre todo por la prisa de hacer que el código funcione. Ello hace que, como ocurre con otros

muchos programas, a menudo resulta difícil comprender los modelos implementados en estos lenguajes de programación; incluso el propio equipo que lo ha desarrollado puede tener dificultades para depurarlo o actualizarlo unos meses más tarde.

En cuanto al coste económico, existen compiladores y entornos de programación gratuitos para C++ y R, lo cual supone una ventaja importante. En cambio, las licencias de programas comerciales como MATLAB, Stata o Mathematica son bastante caras; los precios, que cambian con frecuencia y dependen del usuario que los adquiere (por ejemplo, suelen ser más baratos para el entorno académico que para las empresas) se pueden consultar en sus respectivos sitios web.

1.2. Hojas de cálculo

Las hojas de cálculo (en inglés, *spreadsheets*), creadas inicialmente para tareas de contabilidad, se utilizan hoy en día en infinidad de campos. Lotus 1-2-3 fue el primer programa de este tipo en alcanzar popularidad, pero unos años más tarde fue desplazado por Microsoft Excel, que es actualmente el más utilizado, hasta convertirse en un estándar *de facto*.

La cantidad de modelos de análisis de decisiones implementados en Excel es incalculable. Se utiliza sobre todo para construir árboles de decisión y modelos de Markov. Un ejemplo clásico es el de Chancellor y col. (26) para el tratamiento del SIDA, que aparece como caso ilustrativo en varios libros y artículos; en particular, la implementación en Excel se ofrece como material de apoyo del libro de Briggs y col. (27) [ejercicio 4.7].¹ Otro ejemplo más sofisticado es el de Briggs y col. (28) para la artroplastia de cadera; aparece igualmente en dicho libro, como ejercicio 5.7, y también está disponible en Excel en la misma página web. También la versión en Excel del modelo de Paulden está disponible en Internet, como ya hemos indicado.

¹ <https://www.herc.ox.ac.uk/downloads/decision-modelling-for-health-economic-evaluation>

Aunque Excel incorpora numerosas funciones matemáticas —por ejemplo, varias distribuciones de probabilidad—, puede ser útil utilizar complementos (*add-ins*) para tareas específicas. En concreto Crystal Ball, de Oracle,^k se utiliza a veces para el análisis de sensibilidad probabilista con árboles de decisión y modelos de Markov (29-31). Otros autores usan @RISK, de Palisade Decision Tools,^l una herramienta que la Organización Mundial de la Salud recomendaba en 2003 (32-34).

Un complemento más complejo para Excel es DICE, del que hablaremos en la sección de programas específicos para medicina.

Ventajas e inconvenientes

La principal ventaja de una hoja de cálculo es que integra en un mismo archivo las entradas, el código y la salida, de modo que quien recibe un modelo implementado en Excel puede ver fácilmente no sólo sus parámetros y el resultado final, sino también los resultados intermedios y las fórmulas utilizadas para calcularlos.

El principal problema de Excel es el tamaño de las tablas que se requieren para implementar modelos complejos. Así, el modelo de Paulden ocupa 27 MB en Excel; está organizado en cuatro hojas, la mayor de las cuales tiene 615 filas y 1.518 columnas, es decir, casi un millón de celdas (aunque algunas de ellas están en blanco; se han introducido para hacer el modelo más legible).

Otro problema es la complejidad de las fórmulas. Por ejemplo, el contenido de la celda BDW13 de ese modelo es:

^k <https://www.oracle.com/applications/crystalball>

^l <https://www.palisade.com/risk>

$$=(BDB13*((BDF7)+(1@Pdeathother)*(BDH7)*(Pdeathchemo)*((BDG7*6)+C_sepsis)+(1@Pdeathother)*(BDH7)*(1Pdeathchemo)*(Pchemotox)*(P_ER_chemo_reason_1* C_neutropenia+P_ER_chemo_reason_2*C_injuries+P_ER_chemo_reason_3*C_unspec+P_ER_chemo_reason_4*C_pain+P_ER_chemo_reason_5*C_nausea+P_ER_chemo_reason_6*C_gastro+ P_ER_chemo_reason_7*C_chest))))@discount$$

Un modelo de Markov construido por una agencia de evaluación española para un análisis de coste-efectividad, tenía 8 hojas; cada una de las 6 hojas correspondientes a las intervenciones analizadas tenía 85 filas y 92 columnas. En total, el archivo contenía más de 47.000 celdas, algunas tan complejas como ésta:

$$=SUMA(AP7:BC7)*cCitHC2*BUSCARV(\$C7;Variables!\$A\$4:\$H\$21;7;VERDADERO)+ BUSCARV(\$C7;Variables!\$A\$4:\$H\$21;7;VERDADERO)*SUMA(AP7:AR7)*(1E_CitHC2)* (cColposcopia+cBiopsia)+BUSCARV(\$C7;Variables!\$A\$4:\$H\$21;7;VERDADERO)* SUMA(AS7:BC7)*S_CitHC2*(cColposcopia+cBiopsia)+SUMA(AS7:AW7)*S_CitHC2* BUSCARV(\$C7;Variables!\$A\$4:\$H\$21;7;VERDADERO)*cConizacion+(SUMA(AX7:AY7)* BUSCARV(\$C7;Variables!\$A\$4:\$H\$21;7;VERDADERO)*S_CitHC2+(SUMA(AX6:AY6)*sLCC))* cHisterectomía+(SUMA(AZ7:BA7)*BUSCARV(\$C7;Variables!\$A\$4:\$H\$21;7;VERDADERO)+ (SUMA(AZ6:BA6)*sRCC))*(cHisterectomía+cRadio)+(SUMA(BB7:BC7)* BUSCARV(\$C7;Variables!\$A\$4:\$H\$21;7;VERDADERO)+(SUMA(BB6:BC6)*sDCC))*(cHisterectomía+cRadio)$$

Obviamente, para el receptor del modelo es imposible entender el significado de estas fórmulas y comprobar que la implementación es correcta. De hecho, incluso para el propio programador puede resultar muy difícil localizar y depurar los errores.^m Un motivo es la imposibilidad de examinar una por una tal cantidad de celdas. Otro motivo es que las fórmulas suelen contener referencias a otras celdas y a conjuntos de celdas (en los ejemplos anteriores aparecen las referencias BDB13, BDF7, AP7:BC7, $\$A\$4:\$H\21 , etc.); aunque las versiones modernas de Excel incluyen facilidades para trazar las referencias cruzadas, depurar las fórmulas suele ser mucho más complicado que analizar un fragmento de un lenguaje de programación, en que se pueden insertar comentarios, cada variable tiene un nombre, cada ecuación puede dividirse en varios pasos (realizando cálculos intermedios) y escribirse en varias líneas... Por todo ello, también en el caso de las hojas de cálculo el estilo del programador influye decisivamente en la legibilidad del modelo, al menos tanto como al usar algunos de los lenguajes que hemos mencionado.

Otro inconveniente de Excel es que el análisis de sensibilidad requiere escribir macros en Visual Basic Analytics (VBA), un lenguaje de programación integrado en las aplicaciones de Microsoft Office. Las macros suelen ser muy sencillas, pero es necesario conocer la sintaxis de este lenguaje, que no es trivial.

^m Es famoso el caso de dos artículos publicados por Carmen Reinhart y Kenneth Rogoff, de la Universidad de Harvard, en los que demostraban que el aumento de la deuda pública de un país reduce su crecimiento económico. Su estudio sirvió para justificar drásticos recortes de gasto público en muchos países. Más tarde, un estudiante de doctorado, Thomas Herndon, descubrió que el archivo Excel creado por Reinhart y Rogoff contenía un error en una fórmula, al dejar fuera 5 filas que debería haber incluido. Al corregir ese error se llegaba a la conclusión contraria. En Internet hay gran cantidad de páginas sobre el tema.

El documento "The risks of using spreadsheets for statistical analysis", disponible en <https://www.ibm.com/downloads/cas/7YEX9BKK>, cita éste y otros ejemplos de errores de Excel que han tenido graves consecuencias económicas, y describe algunos errores frecuentes en el uso de hojas de cálculo. Recomendamos su lectura. (En cambio, la segunda parte de ese documento presenta el programa estadístico SPSS, de IBM, como la solución para tales problemas, lo cual es discutible.)

En cuanto al coste económico, Microsoft Excel es un programa comercial, lo cual en principio tendría un impacto a la hora de realizar un análisis de coste-efectividad. Este problema podría paliarse utilizando hojas de cálculo de software libre, como OpenOffice o LibreOffice. Sin embargo, la mayor parte de los ordenadores personales tienen instalado Excel —al menos para los usuarios de Windows— por ser uno de los programas más utilizados para muchas tareas, de modo que a la hora de escoger una herramienta para el análisis de decisiones el coste económico de este programa nunca es relevante en la práctica.

2. Programas para modelado

En esta sección presentamos varios programas para modelado, que no son específicos de la medicina ni del análisis de coste-efectividad. Los hemos agrupado en tres tipos:

1. programas para inferencia bayesiana,
2. programas para simulación de eventos discretos y
3. programas para modelado de sistemas dinámicos.

2.1. Programas para inferencia bayesiana

Existen varios programas de ordenador para construir modelos probabilistas a partir de datos. El usuario debe especificar la estructura del modelo, que consta de un conjunto de variables (en medicina, por ejemplo, podrían ser las enfermedades, los síntomas, las intervenciones terapéuticas, etc.) y un grafo dirigido acíclico que establece relaciones entre ellas; es habitual representar también algunos de los parámetros del modelo (por ejemplo, la prevalencia de una enfermedad) como variables. Combinando el modelo y los datos, estos programas calculan, mediante

diferentes algoritmos de simulación estocástica, la distribución de probabilidad a posteriori —aproximada— de cada variable.

Entre estos programas destacan los creados dentro del proyecto BUGS (*Bayesian inference Using Gibbs Sampling*), de la Universidad de Cambridge, Reino Unido.ⁿ Como indica su nombre, estos programas utilizan el muestreo de Gibbs, un algoritmo de tipo MCMC (Markov chain Monte Carlo) para la inferencia bayesiana.^o El más conocido y utilizado de esta familia es WinBUGS, para el sistema operativo Windows. OpenBUGS es una versión posterior, de código abierto, que funciona en Windows y en Linux. Ambas versiones están “congeladas”, es decir, los autores ya no se ocupan de introducir mejoras en ellas. La única versión que sigue activa se llama MultiBUGS (35).^p También es software libre y funciona en Windows y en Linux. La principal ventaja frente a sus predecesores es que automáticamente realiza las simulaciones en paralelo, lo cual reduce drásticamente el tiempo de ejecución (pues hoy en día todos los procesadores tienen varios núcleos, incluso en los ordenadores personales).

JAGS (*Just Another Gibbs Sampler*)^q es un programa similar, desarrollado independientemente, pero compatible con la familia BUGS, ya que usa (casi) el mismo lenguaje formal para especificar los modelos. Existen versiones para Windows, Linux y Mac OS X.

Stan^r es otro programa de estadística bayesiana, similar en muchos aspectos a los anteriores, pero utiliza algoritmos de inferencia que en general convergen con mayor rapidez que el muestreo de Gibbs, es decir, permiten calcular las probabilidades a posteriori (aproximadas) en menos tiempo. También tiene versiones para Windows, Linux y Mac.

Para realizar análisis de coste-efectividad resulta útil la librería

ⁿ <https://www.mrc-bsu.cam.ac.uk>

^o https://en.wikipedia.org/wiki/Markov_chain_Monte_Carlo

^p <https://www.multibugs.org>

^q <https://mcmc-jags.sourceforge.io>

^r <https://mc-stan.org>

BCEA (*Bayesian Cost-Effectiveness Analysis*),^s de Gianluca Baio, que permite utilizar WinBUGS para la inferencia bayesiana y R para analizar los resultados.

Numerosos estudios de medicina y epidemiológica utilizan estas herramientas. En particular, en Internet se puede encontrar el código en BUGS de muchos modelos para análisis de coste-efectividad: el modelo de Markov de Spiegelhalter y Best (36) para prótesis de cadera,^t los ejercicios resueltos del libro de Spiegelhalter y col. (37),^u los ejemplos y ejercicios del libro de Lunn y col. (38), varios de ellos sobre coste-efectividad,^v así como un buen número de estudios del *National Institute for Health and Care Excellence* (NICE, Reino Unido) británico.^w

Por otro lado, hemos hablado ya de la *Innovation and Value Initiative*, cuyo objetivo es crear modelos abiertos (*open-source*) que puedan ser criticados y enriquecidos con múltiples aportaciones. Un análisis de coste-efectividad de varios medicamentos fue implementado JAGS (39); el PDF que describe el modelo se enriquece en cada nueva versión con numerosos comentarios de diversas personas.

Por último, el libro de Baio y col. (40) contiene numerosos ejemplos de aplicación de la librería BCEA.

^s <https://gianluca.statistica.it/software/bcea/>

^t <https://www.mrc-bsu.cam.ac.uk/software/bugs/the-bugs-project-new-winbugs-examples>

^u <https://www.mrc-bsu.cam.ac.uk/software/bayeseval/bayesian-approaches-examples>

^v <https://www.mrc-bsu.cam.ac.uk/software/bugs/the-bugs-project-the-bugs-book>

^w <https://www.google.com/search?q=%C2%A8cost-effectiveness%22+%22winbugs%22+site%3Anice.org.uk>

Ventajas e inconvenientes

La principal ventaja de estos programas es la posibilidad de construir modelos que combinen conocimiento experto (tanto sobre las relaciones causales entre las variables como estimaciones a priori de los valores de los parámetros) y datos (generalmente, datos sobre pacientes individuales, procedentes de estudios epidemiológicos).

WinBUGS cuenta con una interfaz gráfica para construir y evaluar los modelos, que puede ser útil para algunos análisis preliminares, pero para un análisis de decisiones —en particular, para un análisis de coste-efectividad— se hace necesario utilizar un lenguaje de programación. De hecho, estos programas de inferencia bayesiana pueden ejecutarse en línea de comandos o desde otros lenguajes y entornos, como R, MATLAB, Stata, Mathematica, Python, Julia... Para los analistas que dominan alguno de estos lenguajes, eso es una gran ventaja, porque les permite una gran flexibilidad en la construcción de los modelos y en el análisis de los resultados.

El principal inconveniente de estas herramientas de inferencia probabilista estocástica es que requieren conocimientos de estadística bayesiana y de algún lenguaje de programación, lo cual puede ser una barrera para muchos expertos en epidemiología y en economía de la salud.

En cuanto al coste económico, una ventaja de estas herramientas es que son todas gratuitas, al igual que los entornos para programar en R, Python o Julia.

2.2. Programas para simulación de eventos discretos (SED)

La SED puede realizarse con herramientas de simulación de propósito general, más fáciles de usar de lo que sería implementar un simulador

en código. Entre ellas destacamos Arena y Simul8 por su capacidad, flexibilidad y popularidad. De hecho, Gul y col. (41) citan ambas como las herramientas preferidas para tomar decisiones con respecto a las políticas de gestión de unidades de urgencias. Ambas tienen interfaces gráficas de usuario intuitivas, con menús y diálogos, a través de los cuales los modelos se construyen utilizando bloques predefinidos y conectándolos.

Arena Simulation es una plataforma comercial desarrollada por Rockwell Automation.^x Este software permite diseñar modelos en una interfaz gráfica del estilo “arrastrar y soltar”, que se traducen a un lenguaje interno llamado SIMAN. Arena tiene un conjunto de plantillas (*templates*) con bloques para construir el modelo de simulación, como puede verse en la **Figura 1**, y permite al usuario definir nuevas plantillas. Para ofrecer mayor flexibilidad y permitir modelos más complejos, Arena permite programar parte del modelo utilizando lenguajes de propósito general como Microsoft Visual Basic o C.

Arena es una aplicación nativa Microsoft Windows, por lo que su interfaz resulta reconocible y fácil de usar para los usuarios de dicho sistema operativo. Esto también facilita el intercambio de datos entre Arena y otras aplicaciones de esta plataforma, como procesadores de texto, hojas de cálculo y otros.

Además, Arena proporciona animaciones gráficas en 2D y 3D para ver cómo evoluciona el modelo, así como diversos medios de presentación de resultados, incluyendo gráficas e informes.

Simul8 es otra herramienta comercial, propiedad de SIMUL8 Corporation.^y Al igual que Arena, permite desarrollar y ejecutar modelos SED mediante una interfaz de “arrastrar y soltar”, conectando bloques cuya funcionalidad también se puede personalizar con el lenguaje de *script* Visual Logic. En la **Figura 2** se puede ver el aspecto de un modelo sencillo en Simul8.

^x <https://www.rockwellautomation.com/es-es/products/software/arena-simulation.html>

^y <https://www.simul8.com>

Figura 1. Implementación en Arena del modelo de la DSU para osteoporosis (11)

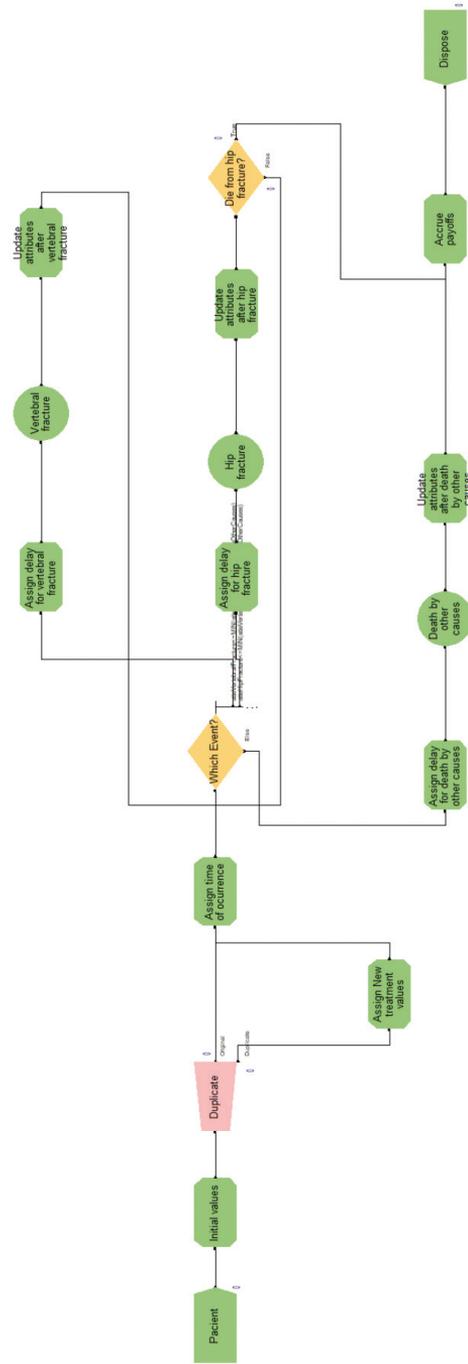
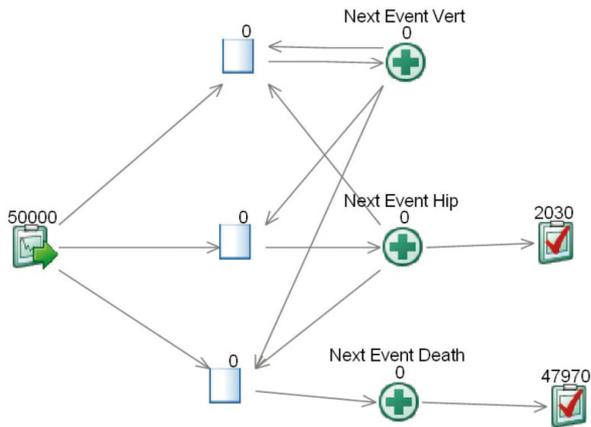


Figura 2. Implementación en Simul8 del modelo de la DSU para osteoporosis (11)



La progresión de la simulación también se puede ver mediante animación, en este caso 2D, y hay además herramientas de presentación de resultados. Se puede integrar con aplicaciones Python o C#.

Simul8 se puede utilizar como una aplicación de escritorio o mediante una interfaz web a través del navegador, la cual permite compartir simulaciones a través de Internet sin que los participantes tengan que instalar ni descargar ningún programa.

Como hemos dicho, existen versiones del modelo de la DSU (11) en Arena y en Simul8. También está implementado en Arena el modelo de Igarashi y col. (42), mientras que el de Bessen y col. (43) se implementó en Simul8. Todos ellos están disponibles en Internet.

Ventajas e inconvenientes

Arena y Simul8 son herramientas flexibles de propósito general que proporcionan los medios adecuados para simular la planificación de servicios (por ejemplo, un servicio de urgencias, al que los pacientes llegan

con más frecuencia a ciertas horas o ciertos días de la semana). Ambas tienen interfaces gráficas que permiten construir modelos ensamblando bloques predefinidos, aunque para modelos complejos, dichos bloques suelen necesitar especificar parte de la operativa mediante código (fragmentos de programa de ordenador) o a través de fórmulas. Los diagramas de bloques y las animaciones permiten capturar y hacerse a la idea de cómo funcionaría en el mundo real el sistema simulado. Además, ambas herramientas son rápidas al ejecutar los modelos. De ahí que sean tan populares.

Igual que ocurre con los lenguajes de programación, los inconvenientes se derivan en parte de sus ventajas: al ser de propósito general, necesitan un mayor trabajo de personalización para aplicarlas al ámbito sanitario. Esto se evidencia más en modelos para economía de la salud, donde se hace necesaria mucha personalización y escritura de código. Además, el ser plataformas tan potentes, con tantas opciones, hace que requieran un mayor esfuerzo de aprendizaje, que quizás no compense a la hora de construir modelos sencillos.

El hecho de que sean herramientas comerciales tiene pros y contras: la ventaja está en un servicio de atención al cliente rápido y preparado, mientras que el inconveniente está en su elevado coste. Rockwell Automation ofrece una versión de evaluación de Arena por tiempo indefinido, aunque con capacidad limitada. También se dispone de una licencia de estudiante —pensada para usar en conjunción con el libro de texto “Simulation with Arena” (44)—, que no permite su uso comercial. Por su parte, SIMUL8 Corporation permite evaluar su herramienta durante cinco días, aunque este periodo puede prorrogarse si se argumenta una causa justificada. Para estudiantes y proyectos de doctorado se puede solicitar una licencia gratuita durante un año, petición que valorará la empresa para decidir sobre su concesión.

2.3. Programas para modelos dinámicos de transmisión

Existen numerosos programas para construir modelos de transmisión, como los estudiados en el capítulo 13.^z En esta sección vamos a presentar los dos más utilizados para evaluación económica: Stella y Berkeley Madonna.

STELLA (inicialmente escrito con mayúsculas, por ser las siglas de “Systems Thinking, Experimental Learning Laboratory with Animation”) fue introducido por Barry Richmond en 1985 para modelizar sistemas dinámicos mediante ecuaciones diferenciales y ecuaciones en diferencias finitas. Actualmente es distribuido por ISEE Systems.^{aa} Es un lenguaje de programación visual que, al igual que su precursor, DYNAMO, permite construir modelos utilizando cuatro tipos de componentes gráficos: existencias (stock), flujo, convertidor y conector. A partir del modelo, el software produce ecuaciones en diferencias finitas y permite al usuario seleccionar un método de análisis numérico para resolverlas. Los usuarios también pueden especificar un paso de tiempo y un tiempo de ejecución para la simulación. Stella puede generar datos en forma gráfica o tabular. Entre los formatos de archivo utilizados para guardar el modelo se encuentra el estándar basado en XML para modelos de sistemas dinámicos, XMILE. En 2012, Naimi y Voinov (45) lanzaron StellaR, un software que puede traducir modelos de Stella al lenguaje de programación R.

Berkeley Madonna es un paquete de software de modelado matemático que resuelve numéricamente ecuaciones diferenciales ordinarias y ecuaciones en diferencias finitas; fue desarrollado originalmente en la Universidad de Berkeley para ejecutar modelos creados con Stella, aunque actualmente funciona de forma independiente. Tiene una sintaxis relativamente sencilla para definir ecuaciones diferenciales, junto con una interfaz de usuario fácil de usar. Las posibilidades gráficas le permiten al usuario evaluar rápidamente si un tipo de modelo sirve para describir los

^z https://en.wikipedia.org/wiki/Comparison_of_system_dynamics_software

^{aa} <https://www.iseesystems.com>

datos a analizar y así facilita comunicar los resultados al experto del dominio que ha de tomar las decisiones. Se ha convertido en un estándar en el desarrollo y comunicación de modelos de farmacología cuantitativa, si bien su uso para modelar la dinámica de sistemas se ha expandido a otras áreas como la física de sistemas, epidemiología, salud ambiental y ecología de poblaciones.

Stella ha sido utilizado en aplicaciones para distintos campos como la biología, la ecología y el medio ambiente. En el campo de la salud, Stella ha sido utilizado, por ejemplo, para construir modelos para 1) el tratamiento con naloxona de la sobredosis de opioides, 2) simular distintas políticas epidémicas en relación al COVID-19 y 3) para la prevención del suicidio.^{ab} Los avances en software están permitiendo la aplicación de la dinámica de sistemas en enfoques orientados a apoyo en la toma de decisiones en economía de la salud.

El grupo de Investigación Operativa del Departamento de Salud en Inglaterra ha utilizado el modelado de sistemas dinámicos con Stella para el desarrollo e implementación de programas en varias áreas, como la detección del cáncer de cuello de útero y de la clamidia (46).

Un ejemplo de aplicación de Berkeley Madonna a la economía de la salud es el modelo para coste-efectividad del cribado del VIH en la atención primaria, de Baggaley y col. (47).

Ventajas e inconvenientes

Tanto Stella como Berkeley Madonna utilizan una sintaxis bastante simple para especificar las ecuaciones, además de interfaces gráficas para construir y evaluar modelos. La interfaz de Berkeley Madonna, por ejemplo, permite poner cada parámetro en una barra de desplazamiento (en inglés, *slider*) y ver cómo varían los resultados al modificarlo. Otra ventaja es que los métodos numéricos que utilizan para resolver las ecuaciones son sumamente rápidos. El principal inconveniente es que, al ser herramientas

^{ab} <https://www.iseesystems.com/solutions/showcase>

de propósito general, hay que implementar todos los cálculos necesarios para estimar los valores de interés (por ejemplo, el coste y la efectividad) a partir de los resultados obtenidos al resolver las ecuaciones.

En cuanto al precio, ambas herramientas ofrecen versiones gratuitas con capacidades limitadas y versiones profesionales de pago, con precios reducidos para estudiantes y para instituciones académicas.

3. Programas para análisis de decisiones en medicina

Existen además programas de ordenador diseñados específicamente para análisis de decisiones en medicina, que incluyen la posibilidad de realizar análisis de coste-efectividad.

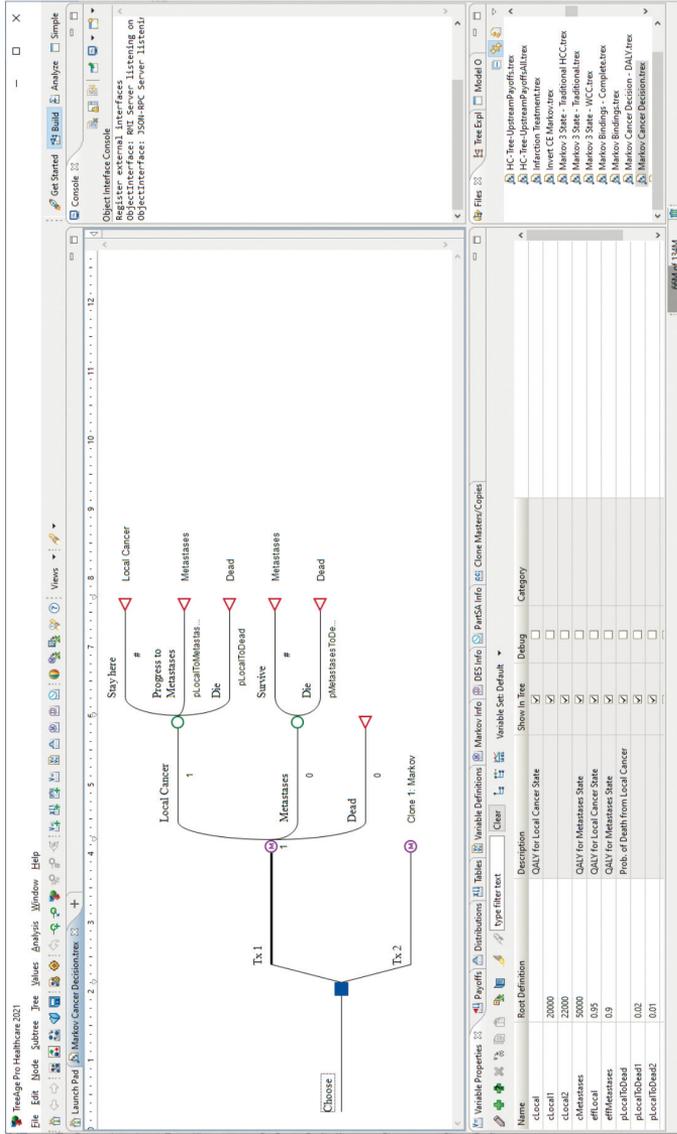
3.1. TreeAge

La empresa TreeAge Software LLC proporciona el conjunto de herramientas comerciales TreeAge Pro, orientado a la toma de decisiones. Entre ellas se encuentra TreeAge Pro Healthcare (en adelante, TreeAge), que permite comparar estrategias basadas en coste, efectividad y otros resultados.^{ac}

Con su interfaz gráfica se pueden construir y analizar modelos basados en árboles de decisión, como los que se explican en el capítulo 9 de este libro. Además, TreeAge permite implementar modelos de Markov y de simulación de eventos discretos manteniendo la estructura de árbol pero añadiendo nodos específicos para estos tipos de modelos. A cada nodo se le pueden añadir los cálculos que precise, pudiendo también definir variables cuyo valor se puede ir recalculando en los diferentes nodos. La **Figura 3** muestra un ejemplo de modelo de Markov implementado como árbol de decisión, con algunos parámetros definidos como variables.

^{ac} <https://www.treeage.com>

Figura 3. Modelo de Markov implementado en TreeAge Pro HealthCare



Una vez construido el modelo, TreeAge puede realizar automáticamente análisis de coste-efectividad, cálculo del valor esperado de la información perfecta y análisis de sensibilidad. También puede mostrar mediante gráficas e informes los resultados obtenidos.

Una funcionalidad destacable de TreeAge es su interoperabilidad con Microsoft Excel, con el que se pueden importar y exportar resultados, e incluso una versión evaluable del propio modelo.

TreeAge se usa ampliamente para tomar decisiones relativas a economía de la salud. Como ejemplos tenemos el modelo de Markov de Paulden y la simulación de eventos discretos de la DSU, ya citados, así como otros disponibles en el sitio web de la empresa.

Ventajas e inconvenientes

TreeAge es una herramienta ampliamente reconocida para toma de decisiones en medicina por varias razones: los árboles de decisión resultan familiares para muchos profesionales de la salud, la interfaz gráfica es fácil de utilizar y de entender (al menos para modelos relativamente sencillos), interactúa con Microsoft Excel y es aceptada por todas las agencias evaluadoras de tecnologías sanitarias.

El principal inconveniente es que todos sus modelos tienen estructura de árbol, lo cual lleva a replicar la información y puede dificultar la forma en que se tratan las variables. Aunque la inclusión de “clones” subsana en parte esta situación, sigue sin conseguirse un modelo compacto. Esto se aprecia, por ejemplo, en el modelo de Paulden donde, a pesar del uso abundante de clones, es casi imposible obtener una visión global del árbol. Además, la simulación de modelos es más lenta que en las demás herramientas para DES, como puede verse en (48).

Aunque TreeAge ofrece un conjunto muy amplio de funciones, podría darse el caso de que el diseñador del modelo necesitara alguna que no esté implementada; por ejemplo, una nueva distribución de probabilidad.

En este caso el usuario podría programarla como un *script* de Python e integrarla en TreeAge, pero para ello es necesario tener ciertos conocimientos de programación.

Por otro lado, al ser un software comercial, TreeAge presenta las mismas ventajas e inconvenientes que otras herramientas ya estudiadas: buen servicio de atención al cliente y precio elevado. Existe una versión de evaluación gratuita, pero sólo permite construir y analizar árboles de hasta 25 nodos y caduca a los 14 días. Además, a diferencia de otros programas comerciales, aunque el usuario haya adquirido una licencia, si no paga el coste anual de mantenimiento el programa puede dejar de funcionar en cualquier momento y ya no se puede instalar ni en el mismo ordenador ni en otro.

3.2. DICE

El marco *Discretely Integrated Condition Event (DICE)* fue propuesto por Jaime Caro [2016] como un método especialmente diseñado para la evaluación económica de tecnologías sanitarias, que, en principio, se puede implementar en cualquier lenguaje de programación.

DICE se viene desarrollado dentro de la compañía Evidera, la cual ofrece una versión gratuita y de código abierto programada en Microsoft Excel con VBA.^{ad} Existe además una implementación en C# (49), que permite evaluar modelos más complejos; se están desarrollando además una versión en Python de código abierto y un formato para especificar los modelos en JSON.

Los elementos principales de DICE son las condiciones (*conditions*), que representan los aspectos del modelo que persisten a lo largo del tiempo, los eventos (*events*), que reflejan las ocurrencias instantáneas que pueden modificar algunas condiciones o el instante en el que suceden otros eventos, y los perfiles (*profiles*), que permiten introducir los datos de

^{ad} <https://www.evidera.com/dice>

los pacientes. Al evaluar el modelo se van computando los resultados en términos de costes, efectividad y cualquier elemento de interés definido por el usuario, como número de fracturas o de altas hospitalarias.

Con estos conceptos, DICE permite construir modelos de eventos discretos, así como modelos de Markov implementados mediante cohortes o mediante microsimulación (*patient-level simulation*). Para todos estos casos se puede realizar un análisis de sensibilidad.

A pesar de ser una herramienta muy reciente, DICE ha suscitado mucho interés y está siendo cada vez más utilizada. Ejemplos de su uso son el modelo de Pennington y col. (50) para el tabaquismo o el de Deniz y col. (51), que evalúa diversos tratamientos para el carcinoma avanzado de células renales.



Figura 4. Tabla de condiciones en un modelo DICE

tblConditions			
Name	Initial Level	Notes	CurCond Level
ID	0	Read from profiles	1
Age	0	Read from profiles	62.2323503
Time	0	The "clock", current model time	6294.11594
TimePrevEvent	Never	Makes calculations easier	4060.7373
IntervNum	1	System, used when multiple interventions	2
Tmt	INDEX(tblTx[Name],IntervNum)		NoInterv
rndHip	rand()		0.81551364
rndVert	rand()		0.14495987
rndVert2	rand()		0.52282853
rndDeath	rand()		0.95943063
rndDeathHip	rand()		0.90508987
HadVert	0		1
HadHip	0		1
UtilBase	INDEX(tblUtilities[Value],1)		0.7
UtilHip	1		0.75
UtilVert	1		0.9
TimePast	0		6.11465747
DiscFactC	0	To simplify calcs in the events	3.76177451
CostTreatment	0		0
NextEventTime	Never	System, keeps track of time for next event	6294.11594
NextEvent	1	System, keeps track of which event is next	6



Figura 5. Tabla de condiciones en un modelo DICE

Event: HipFract (name: tblHipFract)

Type	Name	Expression
Condition	Time	HipFract
Event	HipFract	Never
Condition	TimePast	(Time-TimePrevEvent)/YearLength
Condition	Age	Age+TimePast
Condition	DiscFactC	(EXP(-Time/YearLength*LN(1+DR))-EXP(-TimePrevEvent/YearLength*LN(1+DR)))/(-1 ^{DR})
Condition	HadHip	1
Output	QALYs	QALYs+(UtilHip*UtilVert*UtilBase)*TimePast
Output	QALYsDisc	QALYsDisc+(UtilHip*UtilVert*UtilBase)*DiscFactC
Condition	UtilHip	INDEX(tblUtilities[Value],3)
Output	TxCost	TxCost+CostTreatment*TimePast
Output	HipCost	HipCost+cFracHip
Output	LYs	LYs+(Time-TimePrevEvent)/YearLength
Output	TxCostDisc	TxCostDisc+CostTreatment*DiscFactC
Output	HipCostDisc	HipCostDisc+cFracHip/(1+DR)^(Time/YearLength)
Output	LYsDisc	LYsDisc+DiscFactC
Event	DeathHipFract	IF(rndDeathHip<DeathFromHipFrac,Time,Never)
Condition	TimePrevEvent	Time
Condition	NextEventTime	Min(CurEventTime)
Condition	NextEvent	Match(NextEventTime,CurEventTime,0)

Ventajas e inconvenientes

El éxito de DICE radica en que presenta importantes ventajas a la hora de formular modelos en economía de la salud, ya que ofrece un marco estandarizado adaptado a las necesidades de la toma de decisiones en medicina mediante conceptos sencillos, integrando las características de los pacientes. Además, la implementación en Microsoft Excel, que es gratuita, facilita la comunicación con las agencias evaluadoras de tecnologías sanitarias. Otra ventaja es que permite implementar tanto modelos de eventos discretos como de transición de estados.

Entre los inconvenientes cabe citar que no existe, en el momento de escribir este capítulo, una interfaz gráfica para construir los modelos. La implementación en Microsoft Excel requiere conocer en detalle la sintaxis de esta herramienta y aun así es fácil cometer errores en alguna celda. Además, por las características de esta hoja de cálculo, es muy lenta, pudiéndose necesitar varias horas para completar una simulación.

3.3. OpenMarkov

OpenMarkov^{ae} es un programa de código abierto para modelos gráficos probabilistas (MGPs) (1) desarrollado por el Centro de Investigación sobre Sistemas Inteligentes de Ayuda a la Decisión (CISIAD) de la UNED, en Madrid. Los MGPs, desarrollados desde los años 1980 en el campo de la inteligencia artificial, incluyen las redes bayesianas (2) y los diagramas de influencia (3-5), ampliamente utilizados en aplicaciones médicas, así como nuevos tipos de modelos creados por el grupo de la UNED para el análisis de coste-efectividad en medicina, entre los que se encuentran los diagramas de influencia markovianos (6,7), las redes de análisis de decisiones (52) y las DESnets (48).

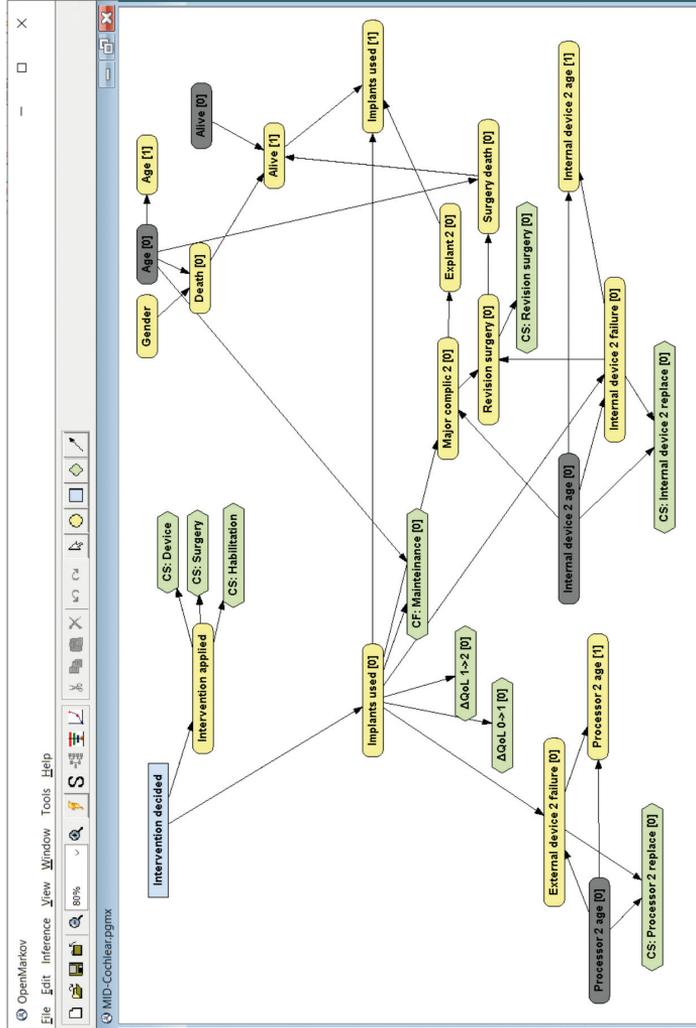
Ventajas e inconvenientes

La principal ventaja de OpenMarkov frente a los lenguajes y entornos de programación (R, C++, MATLAB, etc.) y frente a las hojas de cálculo (Excel) es que su interfaz gráfica permite construir modelos complejos y evaluarlos sin tener que escribir ninguna línea de código ni fórmulas complejas. La especificación de los modelos también es mucho más sencilla que en TreeAge porque no es necesario definir cada parámetro como una variable. Además, al no tener que escribir código ni fórmulas, se evitan los errores de programación, ahorrando así tiempo de depuración y asegurando la corrección del modelo; recordemos que la curva de aprendizaje y el esfuerzo requerido para construir y modificar los modelos son dos de las características deseables a la hora de escoger una herramienta de modelado.

Otra característica deseable es la capacidad expresiva, es decir, la posibilidad de construir modelos complejos. OpenMarkov destaca en este aspecto debido a la utilización de MGPs. Por ejemplo, los

^{ae} <https://www.openmarkov.org>

Figura 6. Modelo de Markov para el análisis de coste-efectividad del implante coclear bilateral (53)



diagramas de influencia markovianos han permitido realizar análisis de coste efectividad para problemas como la estadificación mediastínica del cáncer de pulmón (54) y la artroplastia total de rodilla (55); los árboles de decisión equivalentes a estos modelos tendrían decenas de miles de ramas cada uno, por lo que habría sido imposible construirlos. Del mismo modo, las redes de análisis de decisiones permiten realizar análisis de coste-efectividad cuando no se conoce previamente el orden óptimo de las decisiones, lo cual es imposible con otros tipos de modelos (6,7).

Para el receptor del modelo es más sencillo examinarlo cuando está implementado en OpenMarkov porque el grafo causal pone de manifiesto las hipótesis estructurales del modelo (en concreto, las relaciones de dependencia e independencia probabilista), mientras que con otras herramientas estas hipótesis están implícitas en el código. Además, el modelo está separado del programa que lo evalúa —que es código abierto y común a todos los modelos— y ha sido comprobado con técnicas de ingeniería del software; por tanto, el receptor no necesita comprobar que este código es correcto. En cambio, cuando el modelo está implícito en el código o en las fórmulas que lo evalúan, resulta más difícil comprobar que la implementación es correcta.

Una ventaja adicional de OpenMarkov es que está escrito en Java y, por tanto, funciona en todos los sistemas operativos: Windows, Linux, Mac OS, etc.

En cuanto al precio, OpenMarkov tiene la ventaja de ser software libre. Los módulos de análisis de sensibilidad y de simulación de eventos discretos son gratuitos por el momento, aunque no de código abierto; están distribuidos por DeciSupport AI,^{af} una *spin-off* de la UNED que ofrece además cursos y consultoría sobre análisis de decisiones para medicina, especialmente para OpenMarkov. Sin embargo, esto conlleva el inconveniente de que, por ser un prototipo de investigación de una universidad, ofrecido como software libre, no está tan pulido y depurado como los programas comerciales.

^{af} <https://www.decisupport.com>

Una desventaja de OpenMarkov frente a los lenguajes de programación es que, como en el caso de TreeAge, su capacidad para construir modelos y evaluarlos está limitada por las funcionalidades que ofrece la interfaz gráfica. Dado que OpenMarkov es software libre, un programador experto en Java podría añadir las funcionalidades que se necesiten en cada caso, pero la tarea es bastante compleja.



Conclusión

En este capítulo hemos presentado varias herramientas que permiten construir modelos para evaluación económica en medicina. ¿Cuál es la más adecuada? La respuesta depende del tipo de problema, de los datos disponibles y de los conocimientos del usuario. Recomendamos al lector que consulte los artículos publicados por la “task force” conjunta de las sociedades ISPOR y SMDM en la revista *Medical Decision Making*.^{ag} Aquí ofrecemos unas breves indicaciones.

En primer lugar hay que examinar el tipo de problema. Por ejemplo, si queremos evaluar una intervención para una enfermedad que se contagia de persona a persona (por ejemplo, una vacuna para una enfermedad de transmisión sexual), lo más adecuado sería utilizar un modelo dinámico de transmisión (cf. cap. 13), y la herramienta elegida podría ser Stella o Berkeley Madonna, siempre que sea posible pagar el precio de la licencia.

Luego hay que examinar los datos disponibles. Si hay datos de pacientes individuales, lo más adecuado es construir un modelo bayesiano (cf. cap. 14), que podría implementarse usando MultiBUGS (aunque sólo funciona en Windows y en Linux) o bien JAGS o Stan, en combinación con la librería BCEA de Gianluca Baio para R. Todas estas herramientas son gratuitas.

^{ag} <https://journals.sagepub.com/toc/mdm/32/5>

Si se trata de construir un árbol de decisión, un modelo de Markov o una simulación de eventos discretos (cf. caps. 9 a 12), OpenMarkov es, con diferencia, la herramienta más sencilla de utilizar, y en general permite construir con un esfuerzo relativamente pequeño modelos mucho más complejos que en Excel o en TreeAge; de hecho, en vez de un árbol de decisión resulta mucho más fácil construir un diagrama de influencia o una red de análisis de decisiones (pues, como ya hemos dicho, es posible implementar en OpenMarkov modelos cuyos árboles de decisión contendrían decenas de miles de ramas). Pero si el modelo tiene características “atípicas” que impiden usar OpenMarkov, lo más adecuado sería implementarlo en R, siempre que el usuario tenga los conocimientos de programación necesarios. Tanto OpenMarkov como R son gratuitos.

Por último, hay que tener en cuenta qué herramientas acepta el destino del modelo, que, como hemos dicho, generalmente es una agencia gubernamental de evaluación de tecnologías sanitarias.



Bibliografía recomendada

A lo largo de este capítulo hemos ofrecido abundante bibliografía y recursos para conocer y comparar las diferentes para análisis de decisiones. Aquí ofrecemos algunas referencias adicionales que pueden ser de interés general, especialmente algunos libros.

Por ser R el lenguaje que tiene la mayor cantidad de código aplicable a modelado de economía de la salud (56), existen multitud de recursos sobre el mismo. Son de destacar el libro de Peng (57) y el tutorial de Green y col. (58).

El libro de Lunn y col. (38) “The BUGS Book” es un excelente material de referencia para aprender el lenguaje, a la vez que explica las diferencias de uso entre WinBUGS, OpenBUGS y JAGS, así como su forma de interactuar con R.

El libro de Caro y col. (59) explica la simulación de eventos discretos y cómo implementar los modelos utilizando Arena y Excel. En cuanto a DICE, se puede encontrar abundante material en la web de Evidera y en la página web del proyecto iMPACT-HTA.^{ah}

A su vez, el libro de Shiftlet y Shiftlet (60) proporciona un conjunto de tutoriales y modelos de sistemas dinámicos implementados con diferentes herramientas, como Stella, Berkeley Madonna, R, MATLAB y otros.

Por último, se pueden encontrar manuales de usuario, tutoriales, ejemplos y mucha más información en las webs de las herramientas estudiadas en este capítulo, cuyos enlaces hemos dado en las secciones correspondientes.

^{ah} <https://www.evidera.com/dice>, <https://www.impact-hta.eu/work-package-2>



Bibliografía

1. Koller D, Friedman N. Probabilistic Graphical Models: Principles and Techniques. The MIT Press, 2009.
2. Pearl J. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann, 1988.
3. Howard RA, Matheson JE. Influence diagrams. En: Howard RA, Matheson JE (Eds.), Readings on the Principles and Applications of Decision Analysis. Menlo Park, CA: Strategic Decisions Group. 1984; p. 719-62.
4. Pauker S, Wong J. The influence of influence diagrams in medicine. *Decis Anal.* 2005;2: 238-44.
5. Arias M, Díez FJ. Cost-effectiveness analysis with influence diagrams. *Methods Inf Med.* 2015;54:353-358.
6. Díez FJ, Luque M, Bermejo I. Decision analysis networks. *Int J Approx Reason.* 2018;96:1-17.
7. Díez FJ, Luque M, Arias M, et al. Cost-effectiveness analysis with unordered decisions. *Artif Intell Med.* 2021;117:102064.
8. Paulden M, Franek J, Pham B, et al. Cost-effectiveness of the 21-gene assay for guiding adjuvant chemotherapy decisions in early breast cancer. *Value Health.* 2013;16:729-39.
9. Hollman C, Paulden M, Pechlivanoglou P, et al. A comparison of four software programs for implementing decision analytic cost-effectiveness models. *PharmacoEconomics.* 2017;35:817-30.
10. Mora A. Comparación de Openmarkov con cuatro herramientas de análisis de coste-efectividad. Tesis de máster. Dpto. Inteligencia Artificial, UNED, Madrid. 2019.
11. Davis S, Stevenson M, Tappenden P, et al. Nice DSU Technical Support Document 15: Cost-effectiveness modelling using patient-level simulation. London: National Institute for Health and Care Excellence (NICE); 2014.

12. Hawkins N, Schulpher M, Epstein D. Cost-effectiveness analysis of treatments for chronic disease: Using R to incorporate time dependency of treatment response. *Med Decis Making*. 2005;25:511-9.
13. Alarid-Escudero F, Krijkamp E, Enns EA. A tutorial on time-dependent cohort state-transition models in R using a cost-effectiveness analysis example. *Med Decis Making*. 2023;43:21-41.
14. Krijkamp EM, Alarid-Escudero F, Enns EA, et al. Microsimulation modeling for health decision sciences using R: A tutorial. *Med Decis Making*. 2018;38:400-22.
15. Jansen JP, Incerti D, Linthicum MT. Developing open-source models for the US health system: Practical experiences and challenges to date with the Open-Source Value Project. *PharmacoEconomics*. 2019;37:1313-20.
16. Gray E, Donten A, Karssemeijer N, et al. Evaluation of a stratified national breast screening program in the United Kingdom: An early model-based cost-effectiveness analysis. *Value Health*. 2017;20:1100-09.
17. Glover MJ, Jones E, Masconi KL, et al. Discrete event simulation for decision modeling in health care: Lessons from abdominal aortic aneurysm screening. *Med Decis Making*. 2018;38:439-51.
18. Ryan M, Griffin S, Chitah B, et al. The cost-effectiveness of cotrimoxazole prophylaxis in HIV-infected children in Zambia. *AIDS*. 2008;22:749-757.
19. Chhatwal J, Kanwal F, Roberts MS, et al. Cost-effectiveness and budget impact of hepatitis C virus treatment with sofosbuvir and ledipasvir in the United States. *Ann Intern Med*. 2015;162:397-406.
20. Hodgson D, Pebody R, Panovska-Griffiths J, et al. Evaluating the next generation of RSV intervention strategies: a mathematical modelling study and cost-effectiveness analysis. *BMC Med*. 2020;18:348.
21. Bachmann MO, Fairall L, Clark A, et al. Methods for analyzing cost effectiveness data from cluster randomized trials. *Cost Eff Resour Alloc*. 2007;5:12.
22. Leurent B, Gomes M, Faria R, et al. Sensitivity analysis for not-at-random missing data in trial-based cost-effectiveness analysis: A tutorial. *PharmacoEconomics*. 2018; 36:889-901.

23. Mutubuki EN, El Alili M, Bosmans JE, et al. The statistical approach in tri-
albased economic evaluations matters: Get your statistics together! *BMC
Health Serv Res.* 2021;21:475.
24. Yamabe K, Singhal PK, Abe M, et al. The cost-effectiveness analysis of
a quadrivalent human papillomavirus vaccine (6/11/16/18) for females in
Japan. *Value Health. Reg Issues* 2013;2:92-7.
25. Wolfson LJ, Daniels VJ, Pillsbury M, et al. Cost-effectiveness analysis of
universal varicella vaccination in Turkey using a dynamic transmission
model. *PLoS One.* 2019;14:e0220921.
26. Chancellor JV, Hill AM, Sabin CA, et al. Modelling the cost effectiveness
of lamivudine/zidovudine combination therapy in HIV infection. *Pharma-
coEconomics.* 1997;12:54-66.
27. Briggs A, Claxton K, Sculpher M. *Decision Modelling for Health Economic
Evaluation.* Oxford University Press; 2006.
28. Briggs A, Sculpher M, Dawson J, et al. The use of probabilistic decision
models in technology assessment: the case of total hip replacement. *Appl
Health Econ Health Policy.* 2004;3:79-89.
29. Miller T, Hallfors D, Cho D, et al. Cost-effectiveness of school sup-
port for orphan girls to prevent HIV infection in Zimbabwe. *Prev Sci.*
2013;14:503-12.
30. Jiang Y, Cai D, Chen D, et al. The cost-effectiveness of conducting three
versus two reverse transcription-polymerase chain reaction tests for diag-
nosing and discharging people with COVID-19: Evidence from the epidem-
ic in Wuhan, China. *BMJ Glob Health.* 2020;5:e002690.
31. Nguyen JT, Cole AL, Leech AA. Cost-effectiveness of first-line tyrosine
kinase inhibitor therapy initiation strategies for chronic myeloid leukemia.
Value Health. 2020;23:1292-9.
32. Edejer TTT, Baltussen R, Adam T, et al. *WHO Guide to Cost-Effectiveness
Analysis.* Geneva, Switzerland: World Health Organization; 2003.
33. Sicuri E, Bardají A, Nhampossa T, et al. Cost-effectiveness of intermittent
preventive treatment of malaria in pregnancy in southern Mozambique.
PLoS One. 2010;15:e13407.

34. Conover CJ, Bailey J. Certificate of need laws: A systematic review and cost-effectiveness analysis. *BMC Health. Serv Res.* 2020;20:748.
35. Goudie RJB, Turner RM, De Angelis D, et al. MultiBUGS: A parallel implementation of the BUGS modelling framework for faster Bayesian inference. *J Stat Softw.* 2020;95:1-20.
36. Spiegelhalter DJ, Best NG. Bayesian approaches to multiple sources of evidence and uncertainty in complex cost-effectiveness modelling. *Stat Med.* 2003;22:3687-709.
37. Spiegelhalter DJ, Abrams KR, Myles JP. *Bayesian Approaches to Clinical Trials and Health-Care Evaluation.* Wiley, 2004.
38. Lunn D, Jackson C, Best N, et al. *The BUGS Book – A Practical Introduction to Bayesian Analysis.* CRC/Chapman and Hall, 2012.
39. Incerti D, Jansen JP. A description of the IVI-RA model v2.0. <https://innovationvalueinitiative.github.io/IVI-RA/model-description/model-description.pdf>. (Acceso verificado el 14 de Jul de 2023).
40. Baio G, Berardi A, Heath A. *Bayesian Cost-Effectiveness Analysis with the R package BCEA.* Springer, 2017.
41. Gul M, Guneri AF. A comprehensive review of emergency department simulation applications for normal and disaster conditions. *Comput Ind Eng.* 2015;83:327-44.
42. Igarashi A, Goto R, Suwa K, et al. Cost-effectiveness analysis of smoking cessation interventions in Japan using a discrete-event simulation. *Appl Health Econ Health Policy.* 2016;14:77-87.
43. Bessen T, Karnon JA. Patient-level calibration framework for evaluating surveillance strategies: A case study of mammographic follow-up after early breast cancer. *Value Health.* 2014;17:669-78.
44. Kelton WD, Sadowski R, Zupick N. *Simulation with Arena.* 6^a ed., McGraw Hill; 2014.
45. Naimi B, Voinov A. StellaR: A software to translate Stella models into R open-source environment. *Environ Modell Softw* 2012;38:117-8.

46. Royston G, Dost A, Townshend J, et al. Using system dynamics to help develop and implement policies and programmes in health care in England. *Syst Dyn Rev.* 1999;15:293-315.
47. Baggaley RF, Irvine MA, Leber W, et al. Cost-effectiveness of screening for HIV in primary care: A health economics modelling analysis. *Lancet HIV.* 2017;4:e465-e474.
48. Yago CM, Díez FJ. DESnets: A graphical representation for discrete event simulation and cost-effectiveness analysis. *Mathematics.* 2023;11:1602.
49. Caro JJ, Maconachie R, Woods M, et al. Leveraging DICE (Discretely Integrated Condition Event) simulation to simplify the design and implementation of hybrid models. *Value Health.* 2020;23:1049-55.
50. Pennington B, Filby A, Owen L, et al. Smoking cessation: A comparison of two model structures. *PharmacoEconomics.* 2018;36:1101-12.
51. Deniz B, Ambavane A, Yang S, et al. Treatment sequences for advanced renal cell carcinoma: A health economic assessment. *PLoS One.* 2019;14:e0215761.
52. Díez FJ, Yebra M, Bermejo I, et al. Markov influence diagrams: A graphical tool for cost-effectiveness analysis. *Med Decis Making.* 2017;37:183-95.
53. Pérez-Martín J, Artaso MA, Díez FJ. Cost-effectiveness of pediatric bilateral cochlear implantation in Spain. *Laryngoscope.* 2017;127:2866-72.
54. Luque M, Díez FJ, Disdier C. Optimal sequence of tests for the mediastinal staging of non-small cell lung cancer. *BMC Med Inform Decis Mak.* 2016;16:1-14.
55. León D. Modelo gráfico probabilista para artroplastia total de rodilla. Tesis de máster, Dpto. Inteligencia Artificial, UNED, Madrid, 2011.
56. Incerti D, Thom H, Baio G, Jansen JP. R you still using Excel? The advantages of modern software tools for health technology assessment. *Value Health.* 2019;22:575-9.
57. Peng RD. *R Programming for Data Science.* 5ª edición, Leanpub, 2016. <https://bookdown.org/rdpeng/rprogdatascience> (Acceso verificado el 14 de Jul de 2023)

58. Green N, Lamrock F, Naylor N, et al. Health economic evaluation using Markov models in R for Microsoft Excel users: A tutorial. *Pharmacoeconomics*. 2023;41:5-19.
59. Caro JJ, Möller J, Karnon J, et al. *Discrete Event Simulation for Health Technology Assessment*. Chapman and Hall/CRC, 2015.
60. Shiflet AB, Shiflet GW. *Introduction to Computational Science: Modeling and Simulation for the Sciences*. Princeton University Press; 2014.



16

RETOS FUTUROS DE LOS MODELOS ANALÍTICOS DE DECISIÓN

Marta Trapero-Bertran^a, Mireia Díaz Sanchis^b,
Ana Magdalena Vargas-Martínez^c

^a *Departamento de Economía y Empresa. Facultad de Derecho, Economía y Turismo.
Universitat de Lleida*

^b *Programa de Investigación en Epidemiología del Cáncer (PREC)
Instituto Catalán de Oncología, IDIBELL, CIBERESP*

^c *Departamento de Enfermería. Facultad de Enfermería, Fisioterapia y Podología.
Universidad de Sevilla*

El futuro del uso de los modelos analíticos de decisión en la evaluación de tecnologías sanitarias está regido por los retos generados del correcto uso de la metodología y, por las limitaciones que estos modelos tienen. Es decir, el futuro de los modelos analíticos de decisión no solo debe afrontar las limitaciones usuales de los métodos de modelización, sino también su complejidad la cual provoca una baja implicación de ciertos actores, la falta de transparencia y verificación de los modelos, la interpretación y comunicación de los resultados, la incorporación de nuevas técnicas y metodologías o la dificultad para simular ciertos eventos o enfermedades. Algunos de estos aspectos han sido discutidos ampliamente y existe literatura científica al respecto, aunque siguen representando un reto dado que **no se ha llegado a un consenso**. Posiblemente, algunos de ellos, se irán **definiendo** claramente en un **futuro próximo** mientras otros se **solucionarán por sí solos** con el paso del tiempo a medida que las evaluaciones económicas vayan tomando protagonismo en la toma de decisiones. En cambio, otros aspectos son de difícil solución y todavía tienen un largo recorrido entre las discusiones de los modeladores.

A continuación, se identifican una serie de problemas y retos en las diversas etapas de desarrollo de un modelo analítico de decisión encontrados en la literatura científica que se plantean con el uso de los modelos analíticos de decisión para evaluar la eficiencia de cualquier tecnología sanitaria, entendida ésta como cualquier prueba o método diagnóstico, tratamiento, intervención quirúrgica o de carácter preventivo e incluso cambios organizativos en la prestación de servicios o políticas sanitarias. No pretende ser una lista exhaustiva, y algunos de los temas se han abordado en capítulos previos, pero aquí intentamos dar una visión global.

El reto principal será identificar los problemas e incertidumbres en nuestro modelo y buscar el recurso adecuado para cada uno de ellos.

El tema sobre la identificación de la incertidumbre y su manejo, aunque es uno de los grandes retos de la modelización, dado que se evalúa ampliamente en el capítulo 1 y 5, no se abordará específicamente en este capítulo, aunque se mencionan en algún momento algunas de las fuentes clásicas de incertidumbre.



Retos de los modelos analíticos de decisión

A continuación, se identifican una serie de retos en los modelos analíticos de decisión utilizados en el ámbito de las políticas y programas sanitarios. Estos retos se clasifican según nos encontremos en la etapa **pre-modelización (PREM)**, donde se diseña y contextualiza el modelo, en la etapa de **desarrollo del modelo (DESM)**, donde se lleva a cabo la consecución del modelo, y la etapa **post-modelización (POSTM)**, donde se evalúa y monitoriza el impacto del modelo. Estas tres etapas cubren: **PREM** - (a) análisis del contexto para **diseñar la pregunta a responder por el modelo**; (b) aspectos sobre las fuentes de datos y el acceso a la información; (c) las **particularidades** de los modelos en las **decisiones sanitarias**; **DESM** - (d) los **aspectos metodológicos**, tanto de tipo de modelos como de elementos de una evaluación económica; (e) **POSTM** (f) el proceso de **validación**; y, (g) los **resultados, extrapolación y su interrelación con la toma de decisiones**.

1 Etapa Pre-modelización

1.1 Establecimiento del contexto de la modelización

El contexto se refiere a la identificación de los límites del sistema a modelizar, las condiciones económicas, ambientales, políticas, sociales y

tecnológicas que subyacen en la creación del sistema que dirigirán las cuestiones a abordar dentro de los límites a las que se debe dar respuesta, así como los valores e intereses de las diferentes partes involucradas (1). El contexto del modelo se determina normalmente en la etapa inicial de encuadre del problema y es crucial para ayudar a la toma de decisiones, ya que determina las cuestiones que se abordarán, la selección de los resultados de interés que se estimarán con el modelo, y la potencial incertidumbre que pueda generar el contexto en los resultados de una evaluación económica.



Figura 1

Figura 1A. Definiendo los límites del sistema

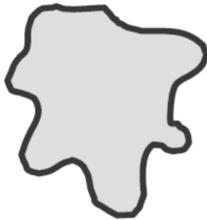
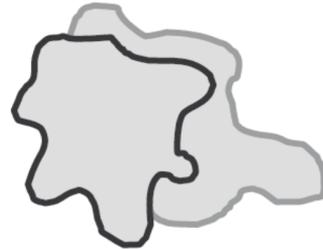


Figura 1B. En un contexto de incertidumbre: ambigüedad en la definición de los límites del sistema



En esta etapa, se deberían involucrar todas las partes interesadas del proceso de definición del problema dado que a menudo tienen diferentes percepciones de la realidad, habitualmente debido a que están relacionadas con sus diferentes perspectivas. En ocasiones se observa que el problema se enmarca en un contexto donde solo se ajusta a los valores tácitos de los expertos y/o se ajusta a las herramientas que los expertos pueden utilizar para proporcionar una “solución” al problema, conduciendo a un “apoyo a la toma de decisiones” sesgado y manipulado.

Decidir sobre el encuadre adecuado es una parte importante del problema y debe prestarse la máxima atención hasta tal punto que se incorporen en el análisis contextos alternativos razonables. El concepto y la metodología de la validación del contexto propuestos por Dunn pueden ayudar a evitar los problemas que surgen de una formulación incorrecta del problema (2).

Todos los actores que han intervenido en la definición del problema deben asegurarse de que el modelo conceptual está bien definido y que el analista tiene una buena comprensión del sistema, de las relaciones entre datos de entrada y resultados, asunciones y formulación del modelo para que lleve a cabo una buena codificación (3).

Los modelos analíticos de decisión habitualmente se utilizan para explorar y comparar el impacto de diversas alternativas políticas de salud en los resultados de interés teniendo en cuenta una gran variedad de escenarios. El modelo deberá representar las relaciones causa-efecto de las diferentes características del sistema que se expresaran como funciones matemáticas. Como ya hemos comentado anteriormente, estos modelos suelen contener componentes inherentes de subjetividad que se manifiestan ya en la fase conceptual cuando se toman decisiones sobre qué elementos se incluirán en el análisis y cuáles se dejarán de lado. Esta subjetividad afecta la manera en que los modeladores traducirán el modelo conceptual en ecuaciones matemáticas.



Figura 2

Figura 2A. Estructura del modelo: las relaciones dominantes en el sistema

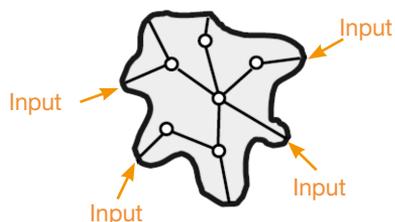


Figura 2B. Incertidumbre en la estructura del modelo: diferentes interpretaciones de qué son las relaciones dominantes en el sistema (en relación a Fig.2A)



También nos podemos encontrar con los eventos inesperados y de alto impacto que se dan frecuentemente en enfermedades infecciosas. Los modelos no pueden anticipar eventos raros, como la aparición del SARS-coV2. Sin embargo, los modelos pueden usarse para prepararse frente a eventos de baja probabilidad y alto impacto, aunque se debe ir con cuidado a identificar cuando sucede el problema para no crear escenarios catastrofistas. A este reto de por sí, habrá que añadir el hecho del poco conocimiento y la falta de este tipo de eventos y sus causas.

Ejemplo (4). Un ejemplo clásico de un evento de alto impacto de baja probabilidad es una pandemia rara y letal. En la primera mitad de la última década hubo un gran interés en la preparación y planificación para una pandemia de gripe, estimulada por la propagación esporádica de algunas infecciones y muertes por la gripe aviar H5N1. Muchos de los modelos mostraron una utilidad potencial de los antivirales en el control o respuesta a esta pandemia, e incluso la posible contención si la cepa emergente se fuera detectada con suficiente anticipación. Estos modelos se centraron en la aparición de una cepa de gripe extremadamente virulenta, y la detección temprana de la transmisión, pero cuando finalmente ocurrió la esperada pandemia en 2009, resultó ser leve y cuando se detectó ya se había extendido ampliamente en México. El proceso de planificación dejó a mucha gente con la impresión de que todas las pandemias de gripe pueden ser eventos catastróficos y con una alta letalidad. Preparar adecuadamente la planificación sin alimentar las percepciones erróneas del riesgo es un reto difícil, particularmente porque, por definición, habrá muy poca información real sobre cómo será un evento de este tipo

Sin embargo, en contraposición, tendríamos la situación que se ha vivido recientemente con el SARS-coV2. Quizás el hecho que ocurrió con la gripe aviar, pudo haber relajado el grado de atención y planificación de una posible pandemia que desafortunadamente ha ocurrido.

Quizás, los primeros indicios de la enfermedad no crearon la alerta suficiente y llevaron a cierta despreocupación inicial para evitar el pánico causado con la gripe aviar. Igual una alerta superior y mayor planificación nos hubiese llevado a la misma situación, pero quizás hubiera dibujado un escenario diferente. Insistimos en puntualizar que estas situaciones presentan retos muy difíciles de afrontar.



Recomendaciones

- En el proceso de contextualización es indispensable saber desde el inicio a qué pregunta debe responder el modelo y quién es el que formula esta pregunta, estableciendo una vinculación desde el inicio con el proceso de toma de decisiones.
- Deberían involucrarse todas las partes interesadas del proceso de definición del problema dado que a menudo tienen diferentes percepciones de la realidad.

1.2. Participación ciudadana en investigación sanitaria

Pearson y col. resaltan la situación creciente en cuanto al interés internacional de los potenciales beneficios que puede aportar la participación ciudadana en la investigación (PPI), destacando el requerimiento de muchos financiadores de su inclusión como requisito previo para la financiación (5). La involucración de pacientes y usuarios tanto en la planificación como en el desarrollo y evaluación de los servicios de salud y asistencia social es un hecho en algunos países, y en parte, en algunas de nuestras Comunidades Autónomas. Se plantean diversas cuestiones en cuanto a la no inclusión de la PPI afectando al diseño y validación de modelos por razones éticas y económicas. Pearson y col. muestra un par de casos

relacionados con la atención neonatal y la fractura del cuello del fémur en los que destaca la contribución de la PPI en el diseño y validación del modelo, y aporta una serie de obstáculos o retos de su inclusión (5).

Relativo a los obstáculos enumerados por Pearson y col. destacarían los siguientes: a) la **ausencia de un lenguaje común entre los modeladores y los pacientes** de forma que pueden entenderse mutuamente; b) dilemas sobre si se requiere una **aprobación de los comités de ética antes de la inclusión del PPI**, aunque la percepción general es que no sería preciso dado que éstos no actúan de la misma forma que los sujetos a investigación; c) el **procedimiento para la selección de los pacientes y usuarios** para participar en este proceso; y por último, d) el **nivel de interés o comprensión de los investigadores sobre lo que aportan los pacientes en las partes del desarrollo más metodológico** del modelo (5).



Recomendaciones

- La participación ciudadana en la investigación sanitaria y en el proceso de decisiones sanitarias puede aportar muchos beneficios en el diseño y validación de un modelo, pero los pacientes y ciudadanos deben recibir formación para compartir el lenguaje con los modeladores y entender bien los modelos.

1.3. Aspectos sobre los datos de entrada

Son muchos los retos que se presentan en la **recopilación y gestión de datos** requerida para parametrizar o calibrar un modelo, ya sea **por falta o por exceso de información**. En la mayoría de ocasiones, un mismo modelo sintetiza datos biológicos, epidemiológicos y económicos de diversas fuentes y de distintos períodos. Así pues, normalmente, el modelo

integra datos que son heterogéneos entre sí y que pueden provenir de fuentes que contienen datos incompletos, incorrectos o inconsistentes, o simplemente, los investigadores no registran e incorporan todos los datos relevantes. Todo ello, finalmente, puede crear una imagen incompleta o sesgada del sistema a modelizar (6). Sin embargo, la recopilación y gestión de los datos es una **fase fundamental de la modelización dado que esta información será la base de los resultados que ayudarán a la toma de decisiones**. Algunos aspectos sobre los datos de entrada pueden encontrarse en los capítulos 1 y 7 del libro.

1.3.1. Identificación de la evidencia científica

Los métodos de la mayoría de guías sobre evaluaciones económicas o evaluaciones de tecnologías sanitarias (ETS) apuntan que primeramente debe identificarse toda la evidencia científica relevante para la modelización del problema en cuestión, incluyendo desde la información proveniente de los ensayos clínicos hasta la opinión de expertos (7-14). Sin embargo, **no existen pautas ni una metodología clara apropiada para identificar la evidencia científica utilizada para informar los parámetros de un modelo de decisión**, excepto en el caso de las revisiones sistemáticas de los ensayos clínicos aleatorizados de efectividad clínica donde existen métodos de búsqueda establecidos y ciertas recomendaciones para cada uno de los parámetros clave del modelo (15). Una vez identificada esta evidencia, debería **seleccionarse o rechazarse según ciertos criterios**, preferiblemente predefinidos, y es entonces cuando de nuevo surgen multitud de dudas de como filtrar la diferente información y cual sería suficiente para obtener resultados fiables. Cuando hablamos de evidencia “suficiente” debemos preguntarnos por si la incorporación de información adicional tendría algún efecto en la reducción de la incertidumbre del parámetro y si se obtendrían resultados sustancialmente diferentes (16). La cuestión de la evidencia “suficiente” es importante para determinar en qué partes

del sistema existen lagunas y poder sugerir investigaciones futuras de datos relevantes para la toma de decisiones.

Sin embargo, se cree que muchos investigadores utilizan lo que perciben como la evidencia más relevante o conveniente para cada parámetro de entrada del modelo en lugar de primero identificar toda la evidencia relevante disponible y luego seleccionar la más apropiada. El enfoque exhaustivo de identificación de la evidencia científica tiene importantes implicaciones de tiempo y presupuesto, por esta razón, rara vez se lleva a cabo en la práctica. Aunque algún estudio alerta sobre la práctica de la búsqueda exhaustiva (17). Las **dificultades para especificar adecuadamente los criterios de inclusión y exclusión**, especialmente cuando la evidencia requerida no es el enfoque principal de un estudio, pueden conllevar a una identificación insuficiente de la evidencia de interés.

Ejemplo (17). Este estudio investigó las características de los ensayos clínicos encontrados en revisiones sistemáticas y el impacto en los metaanálisis de la exclusión de los ensayos difíciles de localizar o de baja calidad. Los autores recomiendan que los investigadores consideren previamente qué tipo de búsqueda bibliográfica y el grado de exhaustividad apropiados para la revisión en cuestión, teniendo en cuenta las limitaciones presupuestarias y de tiempo. Consideran que en situaciones en las que los recursos son limitados, las evaluaciones de calidad meticulosas deben tener prioridad sobre las búsquedas exhaustivas de literatura y las traducciones de artículos. Aunque la búsqueda de la literatura en lengua inglesa maximizó los resultados, obviamente, la traducción de artículos puede ser esencial en la toma de decisiones en países de habla no inglesa.

Ejemplo (18). Este estudio investigó la viabilidad y efectividad de la búsqueda en distintas bases de datos de literatura científica sobre la información requerida para informar un modelo analítico de decisión.

La estrategia más apropiada dependía de la pregunta y la búsqueda específica en múltiples bases de datos resultó más efectiva que la búsqueda exhaustiva en una sola base de datos.



Recomendaciones

- La cuantificación de cómo la información adicional reduce la incertidumbre en la decisión se puede estimar mediante el valor esperado de la información perfecta (EVPI) que corresponde al coste medio que se estaría dispuesto a pagar para tener información completa y perfecta, es decir, sería el coste de eliminar toda la incertidumbre del análisis (19).
- Para evitar una insuficiente identificación de la evidencia científica o que sea de conveniencia o sesgada se recomiendan varios puntos. 1) todo el proceso debe estar completamente documentado para ayudar a la transparencia, reproducibilidad y actualización. 2) Establecer a priori las estrategias de búsqueda apropiadas para cada parámetro y definir los criterios de inclusión y exclusión que pueden no ser evidentes al inicio. 3) El punto 2 necesita de la comunicación continuada y fluida entre los diferentes actores que intervienen en la construcción del modelo. 4) Publicar esta información juntamente con los resultados. Debido a las restricciones de espacio de los artículos científicos, puede incluirse como material complementario de forma online.
- Un enfoque más eficiente para las búsquedas de información es realizarlas de forma estructurada y centrada en torno a los parámetros que se espera que tengan mayor influencia en los resultados. La identificación de dichos parámetros se realizará a través de la opinión de expertos, estudios previos y/o modelos iniciales. También pueden realizarse búsquedas iterativas (16)

1.3.2. Fuentes de información

La utilización de bases de datos individuales de pacientes es siempre preferible a datos de estudios ecológicos pues permiten análisis más detallados. Sin embargo, obtener información de datos individuales no siempre es posible y, además, estos datos requieren de un manejo estadístico anterior a su utilización (*por ejemplo, un metaanálisis*). Así que, en muchas ocasiones deben utilizarse datos de estudios epidemiológicos agregados, o también denominados, estudios ecológicos (16). Un uso habitual de datos de estudios ecológicos es en la evaluación del impacto posterior a la implementación de una intervención mediante datos de vigilancia ecológica. Una limitación importante de estos datos es la incapacidad de atribuir causalidad a la disminución en las tasas de la enfermedad (20). La utilización de estos datos en la calibración de un modelo, haciendo ciertas asunciones sobre el impacto en la población, puede conllevar a una simulación errónea del estado de la enfermedad. Por lo tanto, uno de los retos sería decidir hacer **uso de datos individuales o agregados** en un modelo.

Ejemplo. Supongamos que tenemos un modelo que simula la historia natural del cáncer de cuello de útero y queremos calibrar el modelo a las tasas de incidencia de esta enfermedad en ausencia de intervención. Los datos históricos de estas tasas ya contemplan una actividad de cribado oportunista, aunque fuera mínima, y se desconoce la cobertura e impacto. Si se utilizan estos datos para la calibración suponiendo que son tasas en ausencia de intervención los resultados del modelo estarán sesgados y no sabremos en qué grado.

También obtendremos una imagen errónea del problema si se usan datos agregados poblacionales para un grupo de individuos que no tienen exactamente las mismas características. *Por ejemplo, utilizar la incidencia de cáncer de endometrio de la población general en mujeres con*

síndrome de Lynch que tienen un riesgo mayor. Aun así, en ocasiones, los datos individuales también carecen de suficiente calidad. Los ensayos clínicos pueden tener una asignación aleatoria insuficiente, falta de ciego, o incumplimiento y/o abandono, mientras que los estudios observacionales pueden sufrir de sesgo de selección o sesgo de no respuesta y/o confusión (16). Por lo tanto, aunque se priorice un tipo de datos, es importante poder evaluar la calidad de la evidencia que deberemos combinar en los modelos de decisión, ya que como hemos comentado anteriormente, los resultados obtenidos dependerán de la calidad de la entrada de datos. Aun resultará más importante si primero se desea identificar toda la evidencia relevante para luego incorporar esta información en el análisis, entonces será esencial poder evaluar la calidad del diseño de los estudios.



Recomendaciones

- Aunque se requiere metodología y mayor orientación de cómo abordar la calidad de los datos, sí que es recomendable que los investigadores califiquen cualitativa o cuantitativamente la calidad de la evidencia incorporada en el modelo. Se recomienda que se excluya la información de poca calidad y se evalúe a posteriori la información más incierta mediante análisis de sensibilidad.

1.3.3. Inteligencia artificial

Hoy en día, gracias a los avances en las tecnologías de la información y digitalización, muchos modelos pueden utilizar datos extraídos de los Sistemas de Información Sanitaria o de bases de datos en línea sobre

estadísticas de enfermedades. Esto, por un lado, ayuda a obtener y comparar información de diferentes procedencias y con características diferentes, pero surge el **reto de que los datos almacenados en registros electrónicos están dispersos, son heterogéneos y no se ajustan directamente al objetivo de la simulación** (21). Aquí es donde juega un papel importante **la inteligencia artificial (IA)** o computacional que **permite analizar y extraer información para darles valor** (22). Las aplicaciones de los procesos de inteligencia artificial como el **'big data'**, **'data mining'**, o **'machine learning'** en el campo sanitario son múltiples y entre las más comunes están la extracción y gestión de bases de datos, el rastreo web o el análisis estadístico con finalidades predictivas o de clasificación (23-24). Aun así, esta disciplina está actualmente en auge y en los próximos años aparecerán nuevas utilidades que podrán aplicarse a los modelos analíticos de decisión.

El término *'big data'* se ha popularizado y extendido como el nombre genérico de técnicas específicas como la minería de datos (*data mining*) o el aprendizaje automático (*machine learning*) (25). Sin embargo, el *'big data'* se refiere habitualmente al modo de almacenar, gestionar y procesar conjuntos de datos estructurados y no estructurados que son extremadamente grandes o complejos en intervalos de tiempo razonables mediante técnicas y sistemas especiales (26). **La combinación de técnicas de 'big data' con el análisis de datos** mediante la minería de datos o el aprendizaje automático es el que da un potencial enorme a esta tecnología. Estos métodos **permiten la construcción** automatizada **de modelos analíticos** que habitualmente **se conoce como** modelización de datos **'data modeling'** y se centra en la representación y uso de correlaciones entre los datos, principalmente, de dos formas distintas: mediante *data mining* o mediante *machine learning* (27). Básicamente, la minería de datos busca patrones repetitivos o asociaciones desconocidas que expliquen el comportamiento de grandes conjuntos de datos, mientras que el aprendizaje automático encuentra asociaciones, aprende de ellas y hace predicciones.

MODELOS	Modelos analíticos basados en datos	Modelos analíticos basados en simulación
PRINCIPAL FORTALEZA	Capaces de hacer predicciones robustas e informativas a partir de correlaciones entre datos.	Pueden representar relaciones causales entre el conjunto de datos de entrada y los resultados.
LIMITACIONES	<ul style="list-style-type: none"> • No pueden detectar causalidad entre los datos de entrada y los resultados dada la falta de conocimiento teórico del modelo sobre el sistema que analiza. • La robustez de las predicciones viene dada por el volumen de datos disponible. • No pueden hacer predicciones bajo condiciones distintas a las observadas ni bajo eventos inesperados que no estén representados en los mismos datos. • Útil para realizar análisis descriptivos y predictivos. 	<ul style="list-style-type: none"> • Es necesario conseguir el conocimiento suficiente para describir la mecánica del sistema. • La robustez de las predicciones viene dada, en gran medida, por la calidad del conocimiento y los datos. • En muchos casos, el conocimiento se implementa basándose en asunciones. • Suelen ser más costosos computacionalmente que los basados en datos. Esto limita su utilización para la toma de decisiones en tiempo real. • Útil para realizar análisis prescriptivos que informen de cómo lograr los resultados deseados.

Fuente: Adaptado de (27)

Conociendo las fortalezas y limitaciones de las dos metodologías, **un gran reto consiste en el uso de la modelización combinada** que busca integrar las fortalezas de ambos tipos de modelización y compensar las limitaciones para mejorar los resultados de cada uno de ellos individualmente. Kim y col. presentan la combinación de ambas metodologías mediante un ejemplo de control de un invernadero (27).

Los retos que afrontan los responsables de las políticas sanitarias en el contexto de la IA son importantes. **Estos retos aparecen a tres niveles:** un nivel macro, como las consideraciones políticas relacionadas con las **posibles amenazas a la intimidad y la privacidad que surgen de las empresas de inteligencia artificial** que administran datos confidenciales; a un nivel meso, como la **falta o singularidad de las reglamentaciones de políticas** para todo el mercado; y a un nivel micro, como la **falta de regulación legal de responsabilidad** para los equipos que utilizan la IA para la toma de decisiones (22).

Los responsables de la formulación de políticas gubernamentales destacan la importancia de que los datos confidenciales no estén en manos de empresas de inteligencia artificial y que, más allá de experimentos limitados, el gobierno no apoyará a estas empresas (22). Es decir, que haya un control de uso de la IA en la atención sanitaria. Un segundo reto está relacionado con la falta de regulación de la IA, en concreto, no existe una definición oficial compartida en el mercado de lo que es la tecnología de IA, y no existen estándares oficiales compartidos en la industria sobre cómo las organizaciones pueden utilizar la IA ni cómo se debe evaluar su desempeño. Esto genera incertidumbre del uso de la IA por parte de los hospitales. El tercer reto está relacionado con la falta de reglas de responsabilidad en el uso de la IA para la toma de decisiones. Dado que la tecnología de IA puede reemplazar partes del proceso de toma de decisiones tradicionalmente realizado por humanos, no existe una regulación sobre cómo incluir actores no humanos en el sistema de responsabilidad legal (22). El reto es establecer las responsabilidades y los estándares o regulaciones.

En el contexto de la IA, también existen retos organizativos y de gestión que son de gran relevancia para los responsables de la formulación de políticas gubernamentales, pero igual no tanto para el resto de decisores sanitarios. *Por ejemplo, la falta de planes estratégicos para el desarrollo de la IA, la falta de personal cualificado y las continuas amenazas percibidas por el reemplazo de personal* (22). **A nivel de gestión, se percibe que el**

principal reto es el de la resistencia de los centros sanitarios al intercambio de datos que requiere la tecnología de IA (22). Actualmente, los datos de un paciente son propiedad del centro en el que se ha tratado al paciente. Para un centro sanitario, los datos significan valor, incluso si los gerentes aún no saben claramente cómo usar los datos que posee el centro para obtener ganancias. **Una de las cuestiones clave es establecer quién debería ser el propietario de estos datos**, que probablemente, en España, se podría canalizar a través de un centro de datos impulsado por los mismos gobiernos regionales o nacionales. También existe una tensión entre la necesidad de integración de datos y el interés de algunas organizaciones individuales de restringir el acceso a los datos. Como concluye Lenz y col. (28) aún no existen soluciones tecnológicas que integren completamente tanto aspectos de la gestión como del abordaje médico para una toma de decisiones bien informada. Además, resume algunas de las problemáticas existentes que acusan su ausencia en las siguientes: integración de sistemas heterogéneos, problemas sociales, procesos en evolución y, seguimiento de procesos. **A nivel de recursos humanos, uno de los principales retos para la adopción de IA es la falta de las habilidades requeridas para implementarla (29).** La adopción de la IA en la atención médica necesita personal con conocimientos interdisciplinarios tanto de las disciplinas tecnológicas como médicas.



Recomendaciones

- La IA puede jugar un papel importante en los modelos analíticos de decisión mediante la modelización basada en el 'big data', aunque para obtener el mayor potencial y poder hacer análisis prescriptivos requiere de la combinación con la modelización basada en la simulación la cual incorpora las relaciones causales de la dinámica del sistema.



Recomendaciones

- Dados los retos que presenta la IA para los responsables de la formulación de políticas sanitarias, es recomendable un control de su uso en la atención sanitaria, así como una regulación para la toma de decisiones basada en la IA, el establecimiento de la propiedad de los datos a través, por ejemplo, de un centro de datos impulsado por el gobierno regional o nacional, y la dotación de recursos humanos con conocimientos interdisciplinarios.

1.3.4. Simulación de la no intervención

Otra situación en la que a menudo **faltan datos** es en la simulación de la “no intervención”, es decir, la simulación de la situación previa a la implementación de una intervención (20). Todos los modelos de evaluación económica (tanto antes de la implementación, pre-implementación, como después de la implementación, *post-implementación*) deben estimar un comparador para permitir una evaluación de los costes y beneficios en salud incrementales de la intervención. La obtención de datos en estos casos no siempre es evidente.

Ejemplo. Supongamos que queremos evaluar el impacto del cribado oportunisto de cáncer de cuello de útero en España, para ello, necesitamos calcular el coste y el beneficio en salud de esta situación comparada con lo que ocurría previamente sin el cribado. Sin embargo, el cribado oportunisto de cuello de útero hace muchos años que se aplica en España y no hay datos claros previos en ausencia de intervención.



Recomendaciones

Los métodos más apropiados para estimar la “no intervención” dependerán de cada situación y de los datos disponibles previos a la implementación:

- En los modelos de pre-implementación, habitualmente se estima proyectando las tasas recientes de enfermedad asumiendo que no hay cambios en el periodo de proyección o solo cambios demográficos relevantes esperados.
- En los modelos post-implementación, se puede modelar la carga de enfermedad actual bajo la intervención y estimar las tasas de “no intervención” aplicando algún factor de inflación a las tasas actuales. También se puede utilizar información de la de poblaciones similares a las que se quiere simular. En algunas circunstancias y según los datos disponibles, pueden utilizarse proyecciones con modelos estadísticos que incorporen predictores del comportamiento pasado a lo largo del tiempo. En el caso de infecciones, también se puede simular la transmisión subyacente de la infección usando un modelo dinámico.

1.3.5. Obtención de los costes

La inclusión de qué costes deben incluirse en los modelos analíticos de decisión puede considerarse una decisión metodológica y se abordará más adelante, pero uno de los retos importantes residirá en la identificación y obtención de estos costes, teniendo en cuenta la perspectiva del análisis. Está descrito que en los modelos analíticos de decisión en salud se presta más atención a la recopilación de datos biológicos que a la

recopilación de datos económicos; por esta razón, es posible que **la calidad de la recopilación de datos económicos no siempre cumpla las expectativas o satisfaga las necesidades de la modelización económica** (30). La mayoría de las veces, los datos económicos provienen de fuentes alternativas que no están vinculadas en paralelo a los datos clínicos. Canalizar todos estos datos dispares directamente en modelos de decisión para examinar cuestiones económicas con horizontes temporales prolongados y en diferentes subgrupos de población es complejo y propicia la aparición de sesgos en los resultados. En ocasiones, para los datos económicos se utilizan fuentes de “Real-World data” que capturan algunos aspectos de la atención médica y los costes asociados, pero generalmente, no incluyen datos sobre detalles clínicos importantes (*por ejemplo, el estadio del cáncer que en la mayoría de ocasiones determina la magnitud del coste*). Por el contrario, las bases de datos de registros médicos que capturan elementos específicos de la enfermedad, es posible que no incluyan otros aspectos más generales de la utilización de recursos (*por ejemplo, las hospitalizaciones relacionadas con la enfermedad*). En cualquier caso, incluso si los datos se acercaran al punto de ser adecuados, sería difícil calibrarlos para el mundo real sin poblaciones de referencia del mundo real (30).



Recomendaciones

- Dado que en la mayoría de ocasiones, es difícil recabar todos los datos de la misma fuente, se recomienda usar y contrastar diferentes fuentes y entender que conceptos de costes se incluyen en cada una.
- Revisar las guías generales y específicas de algunas enfermedades sobre evaluaciones económicas o ETS, además de estudios previos para comprobar y ampliar las fuentes de datos económicos disponibles.

1.4. Retos que plantean las particularidades de la modelización de algunos eventos de salud

Existen algunas características propias de la disciplina sanitaria que generan retos en cuanto al diseño y desarrollo de un modelo analítico de decisión. Por ejemplo, las enfermedades transmisibles y sus consecuencias, la medicina de precisión, la medicina personalizada, la perinatología, los programas de cribado o enfermedades multisistémicas. También la temporalidad de los eventos clínicos, la necesidad de participación ciudadana en las decisiones sanitarias o los sistemas de información sanitaria.

1.4.1. Enfermedades transmisibles y sus consecuencias

En cuanto a las enfermedades transmisibles, y la pandemia de la COVID-19, la investigación se enfrenta a numerosos retos basados en múltiples aspectos que afectan a la incertidumbre, como son la limitación de datos disponibles e incompletos, incertidumbre en la gravedad y duración de las medidas de distanciamiento social dificultando la modelización en las primeras etapas (31). En esta línea se encuentra la **particularidad de las evaluaciones económicas de las vacunas o enfermedades infecciosas**, en la que se debe tener presente que **éstas afectan a más partes interesadas** que a la que se le administra y brindan beneficios adicionales que no se pueden capturar fácilmente, es decir, **generan externalidades** y éstas han de ser incorporadas en el modelo (32). Particularmente, **el conocimiento en detalle del curso temporal de la infección en los modelos**, resulta fundamental para la predicción de la trayectoria de la infección en los individuos y, por lo tanto, inferir a la población (33). Si partimos de modelos subyacentes mal definidos, la modelización posterior será necesariamente especulativa. Este problema es un reto particular en las infecciones emergentes, donde hay escasez de evidencia sobre el curso temporal de la enfermedad (34).

En cuanto a los límites de los modelos en salud pública, es conocido el **buen funcionamiento de los mismos en enfermedades infecciosas cuando se contrastan escenarios hipotéticos muy generales** (la estrategia A es más exitosa que la estrategia B); pero, sin embargo, el modelo **no se caracteriza por su predicción cuantitativa de los resultados en contextos más particulares o específicos** (la estrategia A reducirá los casos en un 24,5%). Por tanto, parece importante, como resalta Metcalf y col. (4), distinguir entre “escenarios” y “predicciones”, entendiendo por “escenarios” aquellos que pueden suceder si “x”, “y” y “z” se mantienen, y por “predicciones” la mejor suposición de lo que sucederá en realidad dado lo que sabemos sobre el estado del sistema (es decir, la epidemia) en este momento. A menudo, **los decisores de políticas sanitarias requieren números concretos** para usar en acciones concretas (*por ejemplo, un umbral de inmunidad de grupo o cobertura de vacunación*). Por tanto, deben encontrarse formas en las que el modelo de predicción se pueda incorporar en el campo de la salud pública y las enfermedades infecciosas.

Otra de las limitaciones, además de ser **un reto de estos modelos, es la predicción de horizontes temporales largos**, desarrollando predicciones durante décadas, necesarias para poder obtener efectos o cambios de las intervenciones en el sistema de forma significativa. *Algunos ejemplos de la literatura se refieren al cambio en la curva de incidencia por edad después de la vacunación contra la rubeola (35); o la existencia de períodos de ‘luna de miel’^a en la infección del sarampión como consecuencia de un control exitoso que puede ir seguido de grandes brotes (36)*. Esto trae consigo la **dificultad de validar las predicciones**, ya que resulta poco probable que muchos supuestos del modelo se mantengan a lo largo del tiempo necesario para que se cumplan dichas predicciones.

^a Períodos de control aparente.



Recomendaciones

- Debido a las externalidades que generan las enfermedades transmisibles, se deben tener presente a las partes interesadas y/o afectadas.
- Deben encontrarse formas en las que los modelos predictivos puedan ser incorporados en el campo de la salud pública, dada la demanda por parte de los decisores políticos en acciones concretas.
- Dada la dificultad de que muchos supuestos del modelo se mantengan con el tiempo, es necesario desarrollar métodos para actualizar y validar los supuestos del modelo con horizontes temporales largos a medida que aumenta la disponibilidad de datos o estos supuestos pierden validez (4).

1.4.2. Medicina de precisión

El concepto de **“medicina de precisión”** hace alusión al **abordaje individualizado y adecuado para cada paciente**, teniendo en cuenta el tipo de tratamiento según las características específicas de la persona y del momento (37). Estos enfoques persiguen lograr dichas terapias individualizadas **a partir de métodos cuantitativos y algoritmos** que permitan estratificar, monitorizar y diseñar el tratamiento para el paciente (38-41).

Aunque guarda relación con aspectos de la genómica, la medicina de precisión se expande más allá de ésta, abarcando una cantidad de fuentes de datos cada vez mayor y con mayor disponibilidad para la medicina. Esta integración de aspectos diagnósticos y de tratamiento con datos genómicos, fisiológicos y otros datos médicos ha sido llamado de muy diversas formas: medicina de sistemas, medicina

personalizada, biomedicina de sistemas computacionales, P4 (predictiva, preventiva, personalizada, participativa) y medicina de precisión, entre otras (42).

En la literatura se establecen **una serie de retos importantes a los que se enfrenta este abordaje**, algunos de éstos de carácter técnico **relacionado con la recopilación, procesamiento, almacenamiento e intercambio de datos** (37,43-46), y otros relacionados con **aspectos clínicos**, sociales y éticos (47-51).

Aunque se han producido avances en este campo, como hemos comentado anteriormente, sigue resultando una tarea complicada (37). Uno de los dilemas surge acerca de **quiénes deben gestionar los datos** y si esto debe tener un coste para los investigadores u otros usuarios que necesiten su acceso. Y por supuesto, el **adecuado aprovechamiento de estos extensos volúmenes de datos requiere de herramientas sofisticadas** para asegurar su correcta visualización, gestionar la calidad de los datos y su almacenamiento (42). *Un ejemplo de este tipo de programas es el UK Biobank (2016) que permite mediante acceso abierto el acceso a datos anónimos sobre salud de 500.000 voluntarios.*

Otro aspecto importante a tener en cuenta guarda relación con aspectos éticos como es el requerimiento de consentimiento informado sobre esta recopilación y uso de datos, pudiendo ser utilizados de forma dinámica y en expansión (52). En esta línea, **una adaptación del uso ético y transparente de estos datos se deberá desarrollar** evitando que las decisiones que se tomen puedan generar una discriminación negativa, siendo importante como señala Colijn y col. la incorporación del juicio humano junto a los datos (37). Aun no existe una herramienta que identifique características relevantes de los pacientes, las integre dando lugar a predicciones interpretables, abordajes terapéuticos óptimos y nuevas políticas, si bien existen datos multivariados de la salud de las personas a largo plazo (37). Resulta necesario el **desarrollo de modelos matemáticos** capaces de **describir de forma concisa los procesos** que relacionen **ocurrencia simultánea de eventos, comorbilidades** y

patrones de tiempo tanto en enfermedades como en cualquier proceso relacionado con la salud separando la salud individual de su contexto social y conductual.

Un ejemplo de este tipo de medicina de precisión que genera retos para los modelos es la conducta antivacunación la cual guarda relación con la epidemiología pudiendo interferir en la inmunidad de grupo (53). Esto requiere la necesidad de incorporar dinámicas de creencias y conductas junto a datos de la propagación de la infección.



Recomendaciones

- Es necesario el adecuado aprovechamiento de los extensos volúmenes de datos que utiliza la medicina de precisión mediante herramientas sofisticadas, como las ya existentes en diferentes países.
- Con el objeto de evitar que las decisiones tomadas a partir de los métodos cuantitativos y algoritmos usados en la medicina de precisión den lugar a una discriminación negativa, es necesaria una adaptación del uso ético y transparente de estos datos.
- La investigación paralela de modelos matemáticos que den respuesta a los requerimientos de la medicina de precisión junto a la investigación sobre esta misma, requiere de una perspectiva no solo individual sino comunitaria.

1.4.3. Medicina personalizada (datos genómicos)

Otro reto creciente guarda relación con la disponibilidad de datos genómicos o más generalmente el big-data genético utilizado para el diagnóstico, tratamiento, vigilancia de brotes, evaluación de riesgos e investigación

epidemiológica (53). Aquí, autores como Grenfell y col. (55) hablan de un nuevo tipo de modelo al que denominan “**filodinámica**”. Este término es acuñado a la **combinación de la inmunodinámica, epidemiología y biología evolutiva** requerida para la síntesis de una amplia variedad de patógenos. *Un ejemplo de este tipo de modelo es la composición de la vacuna de la gripe estacional ante la predicción de las futuras cepas con la ayuda de modelos evolutivos e información sobre secuenciación del genoma* (56). Estos desarrollos en la metodología de modelización plantean no solo **problemas técnicos en cuanto a disponibilidad del hardware y software**, sino también la **necesidad de desarrollar conocimiento** en todas las partes involucradas dentro del equipo para utilizar este tipo de datos. Para el campo de las matemáticas, esto se traduce en una **mayor importancia de las herramientas de computación científica** requiriendo, por tanto, una mayor profesionalización en esta área (54).

1.4.4. Perinatología: dualidad materno infantil

Otro campo que plantea retos en los modelos de decisión es la perinatología, entendida como el período comprendido entre las 20 o 28 semanas de gestación, consideradas como período de viabilidad, hasta las 4 semanas después del nacimiento (57). A diferencia de otras especialidades médicas, la perinatología **involucra a dos personas intrincadamente relacionadas** que en ocasiones presentan un **conflicto de interés materno-fetal o intereses en competencia** en cuanto a la salud de ambos (58). Esto quiere decir que las decisiones tomadas por la madre, además de afectar a ella misma afectan a la salud de su descendencia, así como su calidad de vida, incurriendo en costes tanto para el sistema sanitario como para la sociedad en general. Las decisiones clínicas de los profesionales que los atienden a veces se encuentran en el dilema de que dicha decisión favorezca a uno a expensas del otro (57).

Ejemplos de estas decisiones podrían ser la no prescripción de medicamentos que salvan la vida de la madre debido al riesgo de teratogenicidad fetal o cuando se prolongan embarazos de riesgo maternal severo para intentar reducir las complicaciones derivadas de la prematuridad en el feto. Además, se ha de tener en cuenta que el feto no puede tomar decisiones, por lo que éstas recaen en la madre o el proveedor de atención sanitaria. Un ejemplo más concreto de estos dilemas, sería la elección del anticoagulante óptimo en mujeres embarazadas con válvulas cardíacas mecánicas. En la práctica actual existen cuatro estrategias de tratamiento (1. antagonistas de la vitamina K, 2. heparina de bajo peso molecular, 3. heparina no fraccionada o tratamiento secuencial prescribiendo heparina de bajo peso molecular o no fraccionada en el primer trimestre y periparto, y 4. antagonistas de la vitamina K en el segundo y tercer trimestre), cada una de éstas con efectos diversos a corto y largo plazo para ambos (57). En este ejemplo se refleja la importancia del uso adecuado del horizonte temporal, teniendo en cuenta que tanto un sangrado abundante en la mujer como un coágulo sanguíneo podrían provocar un accidente cerebrovascular y afectar a su vida en general y a la de su familia. Así mismo, las consecuencias en el feto como la parálisis cerebral o la discapacidad física. Esta díada afectaría al uso de las utilidades de salud conjuntas, los umbrales de coste-efectividad referidos a la díada materno-infantil, las medidas de resultado clínico y el horizonte temporal, incluyendo costes y resultados en salud durante un periodo de la vida de las dos personas. No hay consenso en cómo proceder o presentar los resultados en estos casos. En relación con el uso de las utilidades, algunos estudios reportan esta información de forma separada distinguiendo entre utilidades para estados de salud de la madre y del feto, y otros que las usan de forma combinada.



Recomendaciones

- Los modelos de decisión que evalúan intervenciones o programas de perinatología deben contemplar no sólo los costes y resultados en salud de la madre sino también las consecuencias para el feto.

1.4.5. Programas de cribado

Uno de los retos del uso del análisis de decisión y evaluación económica de programas de cribado recae en la **definición del alcance del análisis**. Esto guarda relación con la dificultad para calcular el coste-efectividad del cribado para una única enfermedad o afección, ya que podría detectarse a través de otras condiciones, afecciones o hallazgos incidentales relacionados con el objetivo principal u original del cribado en evaluación. Por tanto, si estos hallazgos son informados y se les hace un seguimiento, deberían ser incluidos en los análisis, y definir el alcance de éstos.

Otro de los retos, y en particular del cribado neonatal, está relacionado con la metodología, dada la **falta de estandarización** para la **asignación de utilidades** a los estados de salud infantil (59-60). Añadir que, aunque no es un reto particular y exclusivo de los programas de cribado, en el caso del programa de cribado neonatal, existe una escasez de datos en cuanto a resultados a largo plazo. Autores como Prosser y col. (61) proponen como referentes, análisis previos de otros tipos de programas de cribado como son la mamografía en la detección del cáncer de mama o la colonoscopia para el cáncer de colon para una posible solución.

Como ya hemos comentado anteriormente, la **identificación de datos para la intervención de comparación** se suma a los retos de los análisis de decisión de los programas de cribado. La estrategia de

comparación frente un cribado universal podría ser un cribado oportunista o la ausencia de cribado. En cualquiera de los dos casos, puede haber un vacío de información tanto de efectividad como de costes.



Recomendaciones

- Dada la necesidad de definir explícitamente el alcance del análisis, se propone evaluar, por un lado, la eficiencia del cribado teniendo en cuenta únicamente el evento para el cual se lleva a cabo como objetivo primario y, por otro lado, incluir los otros hallazgos y la información sobre los mismos, pudiendo comparar a posteriori su eficiencia en ambas situaciones (58).

1.4.6. Enfermedades multisistémicas y afecciones complejas

El abordaje de enfermedades multisistémicas representa un reto dado que tanto el diagnóstico como el manejo y progresión de las mismas es complejo. En este campo, se presentan dificultades estructurales en relación a la inclusión o exclusión de estados de enfermedad y eventos, o la caracterización de las vías de atención sanitaria, entre otros. *El lupus eritematoso sistémico es un ejemplo de este tipo de enfermedad multisistémica, en la que los pacientes pueden experimentar una gran variedad de manifestaciones clínicas en diferentes órganos a lo largo del tiempo planteando un reto dentro de los modelos analíticos de decisión que simulan una cohorte homogénea (62).*

En línea con la evaluación económica de afecciones complejas, se encuentran las intervenciones complejas consideradas aquellas cuyos beneficios van más allá de la salud, como es el caso de la “prescripción

social^b” (63). En este caso, el reto es la inclusión o exclusión de impactos fuera de la disciplina de la salud.



Recomendaciones

- Es importante incluir los impactos más allá de la disciplina de la salud en los análisis de decisión, modelos de enfermedades multisistémicas y afecciones complejas.

2. Etapa de Desarrollo del modelo

2.1. Aspectos metodológicos

Como ya hemos comentado anteriormente, ya en la fase conceptual de la modelización se hacen asunciones sobre el sistema que afectarán la manera en que los modeladores traducirán el modelo conceptual en ecuaciones matemáticas. Además, para modelizar el problema, deben determinarse una serie de **características técnicas que harán que el modelo reproduzca el problema con mayor o menor veracidad**. En muchos casos, estas decisiones no suelen tomarse en base a la mejor reproducibilidad del sistema sino por la consecución de su simplicidad, por tanto, la metodología en sí presenta diferentes retos

^b Este tipo de intervención es definida por Byford and Sefton (54) como aquella que: “se dirige a pacientes con problemas sociales y de salud complejos que son de naturaleza crónica, tiene un impacto en múltiples áreas de la vida de una persona (y, potencialmente, de la vida de su familia) e impone costes sociales significativos; se caracteriza por un alto grado de participación ciudadana; tiene múltiples componentes personalizados individualmente; sus destinatarios son heterogéneos; tiene múltiples objetivos complejos y múltiples resultados complejos (enfocados en la salud física, pero también problemas sociales como la deuda, la vivienda y los derechos de asistencia social y resultados de “bienestar” más abstractos); tiene participación de múltiples agencias; y los mecanismos a través de los cuales se logran los resultados son inciertos.

que en ocasiones se identifican con posibles soluciones y en otras, son cuestiones de amplio debate. Una de las características más relevantes que debe tener un modelo es la transparencia, para asegurar la reproducibilidad de un modelo. En diferentes ocasiones la transparencia puede ser la solución o parte de la solución al problema que plantea el análisis de decisión.

2.1.1. Temporalidad del modelo

El tiempo es un aspecto muy importante en los modelos analíticos de decisión dado que **el momento y la duración de los eventos clínicos, las intervenciones y sus consecuencias afectan la estimación de los costes y los efectos en salud futuros** (64). Algunas cuestiones relacionadas con el tiempo de modelización que plantean retos son: el número de cohortes, el tiempo de simulación, la estabilidad en el tiempo, contemplar tiempo discreto o continuo, duración de los ciclos, conversión adecuada de las tasas y probabilidades, y el descuento de los costes y efectos en salud futuros a valores presentes (64).

2.1.2. Número de cohortes

Uno de los retos que afrontan los modelos, en relación a la temporalidad, es la elección correcta del número de cohortes a considerar en el modelo. Una cohorte de simulación representa una serie de individuos agrupados con fines analíticos que se simula durante un período de tiempo. Muchos modelos analíticos de decisión utilizan **una única cohorte** suponiendo que la cohorte modelizada representa no solo la primera cohorte, sino cualquiera de las sucesivas cohortes que incorporan la intervención (64). Tal suposición es razonable si no hay diferencias relevantes entre cohortes, no hay efectos notables compartidos entre cohortes o si no se prevén grandes cambios en los costes unitarios o innovaciones tecnoló-

gicas en un futuro próximo. La utilización de modelos de una sola cohorte puede ser adecuado en determinadas circunstancias de asignación de recursos, pero la ventaja principal reside en el beneficio de la simplicidad conceptual. De hecho, el uso generalizado de modelos de una sola cohorte puede reflejar una suposición implícita, en muchas ocasiones errónea, de que el impacto económico y en salud en una cohorte determinada es representativa de otras cohortes, tanto en el presente como en el futuro. Por el contrario, los **modelos de múltiples cohortes** son analíticamente más complejos y simulan varias cohortes de individuos que se suelen diferenciar por características como el año de nacimiento, el sexo, la raza o el año de diagnóstico, y, por lo tanto, el perfil de riesgo de la enfermedad asociado a esta característica. Los modelos multicohortes pueden ser cerrados si simulan las cohortes introducidas en el período inicial, o abiertos si permiten la entrada de nuevas cohortes a lo largo del tiempo (65).

Ejemplo. Los modelos utilizados en evaluaciones económicas de programas de cribado pueden incluir una o varias cohortes de nacimiento. Dado que estos programas involucran múltiples cribados durante varios años para cada cohorte de nacimiento, algunos investigadores abogan por modelar todas las cohortes en lugar de una sola cohorte de nacimiento, argumentando que representa con mayor precisión la implementación real.

Sin embargo, algunos autores indican que las estimaciones de los resultados en modelos multicohorte deben darse de forma separada por cohorte y no de forma agregada dado que, si no, no se tiene en cuenta la heterogeneidad entre cohortes (66). Tal heterogeneidad puede ser relevante para la toma de decisiones cuando hay una variación considerable en los resultados entre cohortes.



Recomendaciones

- Determinar a priori el número de cohortes a modelizar adecuados al problema concreto y no por simplicidad.
- El uso de una sola cohorte puede ser adecuada si no hay importantes dependencias entre cohortes o no se prevén grandes cambios en un futuro próximo.
- En el caso de utilizar varias cohortes, explorar la relevancia de dar estimaciones individuales por cohorte o agregadas.

2.1.3. Horizonte temporal

Otro reto importante, en cuanto a temporalidad, es la elección correcta del horizonte temporal. El número de cohortes modelizadas y la simulación de sus efectos tienen implicaciones sobre **cuánto tiempo debe simularse una intervención y cuánto tiempo deben evaluarse sus efectos** (64). El **tiempo durante el cual se aplica la intervención** por persona o cohorte es inherente a la intervención en sí, y **puede ser breve** (por ejemplo, el *tratamiento con una antitérmico para la fiebre cuando se tiene gripe*) o **de por vida** (por ejemplo, la *administración de insulina para pacientes con diabetes tipo 1*). El tiempo de simulación se define con el horizonte analítico que es el tiempo en el que se evaluarán los costes y los efectos en salud en el futuro y, por lo general, debería incluir la duración de la intervención y ser lo suficientemente largo como para captar todas las diferencias significativas en costes y efectos en salud entre las alternativas consideradas. **Por lo general, se modela hasta la finalización de la intervención**, aunque en el caso de múltiples cohortes, algunas veces se limita el tiempo de simulación afectando al tiempo de implementación de la intervención para algunas cohortes. Debe tenerse en cuenta que la reducción de la duración

podría conllevar a estimaciones de los costes y efectos en salud no representativos de la realidad. Por ejemplo, en el caso de evaluar el coste-efectividad de intervenciones en enfermedades crónicas, se recomienda un horizonte temporal durante toda vida de las cohortes modelizadas, sin embargo, en otros casos dependerá del curso de la enfermedad y del objetivo del análisis.



Recomendaciones

- El horizonte temporal debe ser adecuado al objetivo del estudio y a los eventos simulados. Se recomienda que el tiempo de simulación incluya la duración de la intervención y pueda captar las diferencias en costes y efectos en salud entre las diferentes alternativas.
- En el caso de utilizar un horizonte temporal corto, hay que ser consciente de y coherente con las limitaciones en los resultados.

2.1.4. Simulación del tiempo

Un aspecto crucial en la modelización es **cómo simular el tiempo de los eventos clínicos**. El nivel de detalle utilizado para simular el tiempo de los eventos también determinará la precisión con la que se estima la cantidad de tiempo o duración (*por ejemplo, la duración de una enfermedad, los años de vida ajustados por calidad o la duración de una intervención*), y, como consecuencia, los costes (67).

Hay diferentes técnicas de modelización que se han explicado en los capítulos 1 y 9 a 14 del libro. Entre los métodos más comúnmente utilizados se incluyen los árboles de decisión, los modelos de transición de estados individuales y de cohortes (Markov), la simulación

de eventos discretos, los modelos de procesos estocásticos basados en ecuaciones diferenciales o modelos dinámicos de transmisión y los modelos híbridos que combinan diferentes enfoques (68). Hay que tener en cuenta que la distinción entre técnicas de modelización no siempre es clara y en muchas ocasiones se superponen. La elección de la técnica de modelización es importante dado que afectará la forma en la que se simula el tiempo. Por ejemplo, los modelos basados en árboles de decisión simulan los eventos y el orden en que ocurrirán, pero no simulan explícitamente el tiempo (69). Los modelos de transición de estados simulan una o varias cohortes a lo largo del tiempo mientras se desplazan entre un número finito de posibles estados de salud mutuamente excluyentes. Por lo general, estos modelos se evalúan en tiempo discreto, es decir, el factor tiempo se asocia con la probabilidad de que los pacientes se muevan por ciertos estados de salud en una serie de períodos de tiempo discretos (70). En los modelos de Markov, estos períodos se denominan ciclos y los movimientos entre estados de salud solo ocurren al final de cada ciclo. La duración de estos ciclos depende de la enfermedad y de las intervenciones que se estén evaluando, pudiendo ser ciclos mensuales o anuales, por ejemplo. Sin embargo, es importante señalar que el tiempo tiene una naturaleza continua y es como suelen ocurrir los eventos clínicos.

Por lo tanto, los modelos de Markov en tiempo discreto solo son una aproximación al proceso de progresión de una enfermedad y cuanto mayor es el ciclo entre estados, menor es su semejanza con la realidad. Esta **aproximación del tiempo continuo por un tiempo discreto puede producir errores en las estimaciones** (infraestimación de la duración del evento), **sobre todo con ciclos largos**, dado que se asume que la pertenencia a un estado es constante en cada ciclo (71).



Recomendaciones

- Se recomienda la elección de la duración del ciclo no únicamente en función de la disponibilidad de los datos, sino teniendo en cuenta una duración de ciclo lo suficientemente corta como para que no se requiera aplicar la corrección de medio ciclo.

2.1.5. Efectos en salud

La **medición de los beneficios en salud** es otro tema clave en las evaluaciones económicas, sin embargo, **no hay un consenso en la elección de la métrica** (72). Existe muy escasa literatura empírica que explore las diferencias del uso de las diferentes medidas y de su potencial impacto en los resultados (72). **Para que la medición de los beneficios en salud alcance su mayor potencial en la toma de decisiones, debería basarse en un sistema genérico en el que las ganancias o pérdidas se pudieran comparar entre el mayor número de intervenciones posibles.** Los años de vida ajustados por calidad (AVAC), los años de vida ajustados por discapacidad (AVAD) y los años de vida ganados (AVG) son medidas habituales en las evaluaciones económicas de intervenciones sanitarias (73). Si bien, esta última es una medida que refleja la mortalidad, los AVAC y los AVAD son medidas que combinan la mortalidad con la morbilidad, un ejercicio que implica un equilibrio entre la cantidad y la calidad de salud. Algunos autores argumentan que la mortalidad y la morbilidad son dimensiones totalmente diferentes, y combinarlas en una sola unidad numérica no tiene sentido (73). Otros, sin embargo, argumentan que el ejercicio es necesario para convertir los principios de asignación de recursos en criterios que se puedan utilizar de manera coherente.

Los AVAC que representan años de vida saludable ponderados y penalizados por un factor de utilidad, se han convertido en una de las medidas más usadas y recomendadas por la comunidad académica y muchos sistemas sanitarios como The National Institute for Health and Care Excellence (NICE) del Reino Unido (74). En contraposición, algunas instituciones internacionales, como la Organización Mundial de la Salud (OMS) o el Banco Mundial, a menudo basan sus evaluaciones económicas usando los AVAD. Tanto el cálculo de los AVAC como de los AVAD ha sido cuestionado en diversas ocasiones (73). Además, las utilidades para todos los estados de salud de cada una de las enfermedades para los diferentes contextos no siempre están disponibles. Los AVG, sin embargo, evitan ponderar los años de vida ganados por la intervención.

NICE, una de las organizaciones más influyentes en el uso de la toma de decisiones en salud, sigue discutiendo, ajustando y considerando apreciaciones en el uso de los AVAC. Por ejemplo, reconoce que otros factores, además de los beneficios en salud expresados en términos de AVAC, podrían ser relevantes para las decisiones, o considera que los AVAC al final de la vida pueden tener un valor 2,5 veces superior que los AVAC estándar, o plantea ajustes cuando se evalúan tratamientos para enfermedades raras, o valora incorporar las desigualdades en salud en las evaluaciones económicas, o establece que metodología para calcular las utilidades es la más apropiada (75).

Así pues, aunque el uso generalizado en las actuales evaluaciones económicas es el AVAC y parece haber un consenso bastante amplio al respecto, es un tema abierto que plantea retos importantes en un futuro.



Recomendaciones

- La mayoría de las guías de evaluaciones económicas y ETS, recomiendan el uso del AVAC.

Para favorecer la comparabilidad de diferentes estudios se recomiendan los siguientes puntos en referencia a la métrica de los beneficios en salud:

- La definición, descripción y transparencia en el uso de las utilidades en los AVAC.
- Revisar la bibliografía sobre el evento en concreto para realizar un análisis con la medida de beneficios en salud que ayude a la comparación entre estudios. Por ejemplo, en los países en vías en desarrollo con el uso del AVAD.
- Es deseable realizar varios análisis con diferentes medidas de beneficios en salud (por ejemplo, ajustando o no por calidad, AVAC y AVG) si hay dudas sobre la calidad de las utilidades o si ello favorece la comparabilidad con otros estudios.

2.1.6. Uso de recursos y costes

Qué costes deben considerarse en los modelos analíticos de decisión sigue siendo un tema controvertido entre los modeladores y expertos en evaluaciones económicas. Así pues, un concepto clave en las evaluaciones económicas es el punto de vista o la perspectiva desde la cual se puede realizar el análisis y condiciona el tipo de costes y resultados incluidos (76). Este concepto se ha explicado en los capítulos 1 y 3 de este libro. La selección de por sí de la perspectiva, en principio, no es ningún reto dado que los modeladores escogerán la perspectiva que más se adecue al objetivo del estudio.

Existe un **amplio consenso que deben incorporarse todos los costes relacionados con las intervenciones evaluadas dentro del horizonte temporal del análisis**. Sin embargo, todavía hay un **amplio debate sobre si deben incorporarse los costes futuros**, es decir, aquellos que se producen fuera del horizonte temporal del análisis y que son consecuencia de los años de vida ganados por la implementación de una intervención (77-78). Habitualmente, este tipo de costes se ignora, independientemente de la perspectiva elegida para el análisis. Los costes futuros se dividen en costes médicos (relevantes tanto para la perspectiva social como sanitaria) y costes no médicos (solo relevantes para la perspectiva social)^c. Los modeladores coinciden que los costes médicos relacionados deberían incluirse en los modelos, pero hay mucha discrepancia en cuanto los costes médicos no relacionados. Entre los que abogan por su inclusión, hay partidarios de su incorporación con el fin de obtener resultados consistentes con la maximización de la utilidad de los medicamentos administrados a lo largo de la vida de los pacientes. Sin embargo, los partidarios de no incluirlos basan su decisión en que, en su opinión, estos costes futuros no son consecuencia directa de los medicamentos administrados y se van a producir de manera similar en todos los grupos de estudio, y, por lo tanto, no es necesario tenerlos en cuenta. Además, **la obtención de estos costes puede ser una tarea difícil** y, si al final se consigue, no es fácil verificar su exactitud, lo que además podría sesgar el resultado (78). Por otra parte, se ha comprobado que **la inclusión de estos costes médicos futuros no relacionados suele tener un impacto pequeño en el resultado final** ya que el ajuste

^c Los costes futuros no médicos se refieren a los costes no sanitarios relacionados con el consumo durante los años de vida ganada (por ejemplo, costes de vivienda, alimentación y ocio) o también las ganancias de producción futuros que serían costes futuros negativos (beneficios del trabajo). Los costes futuros médicos, a su vez, se dividen en costes médicos relacionados y no relacionados. Los costes futuros médicos relacionados serían aquellos afines al tratamiento o seguimiento de la patología basal (por ejemplo, costes de controles realizados por un oncólogo después de un cáncer de mama). Los costes futuros médicos no relacionados serían aquellos que se producen como consecuencia de otra patología diferente a la basal ni consecuencia de ella durante los años de vida ganado (por ejemplo, los costes de tratar un ataque cardíaco).

temporal va a hacer que la magnitud final sea de pequeña cuantía y, por lo tanto, de poca relevancia para contestar a la pregunta de si el nuevo medicamento en estudio va a ser más eficiente.

Los análisis, generalmente, asumen que la intervención y las tecnologías relacionadas permanecen estables durante el horizonte temporal analítico, al igual que los precios o costes unitarios de los diferentes procesos del sistema. Es aconsejable utilizar precios unitarios lo más recientes posible y, aunque es poco probable que se mantengan constantes, no se espera que cambios relativamente pequeños alteren las conclusiones (79). Sin embargo, sí **se debe prestar atención a cambios abruptos de precios que puedan impactar sustancialmente las decisiones**. Aun así, reiteramos que predecir estos cambios es difícil y todavía más las consecuencias de tales cambios y no está claro cómo deberían manejarse mejor en la modelización.

Los costes necesarios también se ven afectados por la incertidumbre en la identificación y obtención de datos, tal como se ha mencionado anteriormente en los aspectos relacionados con los datos de entrada.



Recomendaciones

- La mayoría de las guías de evaluaciones económicas y ETS, recomiendan aplicar la perspectiva social.
- La perspectiva social debe abarcar la mayoría de costes, pero, en cualquier caso, debe quedar claro qué categorías de costes se incluyen y cuales no para facilitar la transparencia y comparabilidad.
- Algunos proveedores y ciertas circunstancias requieren un análisis con costes más restrictivos. En estos casos, se recomienda que, si es posible, se realice un doble análisis que incluya la perspectiva social.

2.1.7. Tasa de descuento

La forma apropiada de aplicar el descuento en los estudios de evaluaciones económicas ha recibido considerable atención tanto en la literatura científica como en las guías nacionales de evaluaciones económicas dado que puede tener una gran influencia en los resultados.

A menudo, los costes y beneficios en salud considerados en una evaluación económica no solo se incurren en el año en curso, sino que se materializan más allá. El momento cuando se producen las consecuencias de las intervenciones en una evaluación económica es relevante porque las personas generalmente valoran menos los costes y beneficios en salud futuros que los costes y beneficios en salud actuales y, además, su valor disminuye cuanto más distantes están de la implementación de la intervención (80). Los mayores puntos de debate se centran en la base misma para la aplicación de una tasa de descuento en los beneficios en salud, la magnitud de la tasa, si los costes y beneficios en salud deben descontarse por igual, si deben disminuir o permanecer constantes en el tiempo, si deben ser retardado respecto al tiempo de la intervención o si deben aplicarse descuentos diferentes para el efecto intra e intergeneracional (81).

Aunque actualmente para los economistas es un tema baladí, durante muchos años la tasa de descuento ha recibido muchas críticas, y, de hecho, todavía resulta difícil para muchas personas entender porque los costes futuros, pero sobre todo los beneficios en salud futuros deben descontarse y tener menor valor que los costes y beneficios en salud actuales.

Actualmente, todas las guías de evaluaciones económicas recomiendan realizar descuento a los beneficios en salud, además de los costes, aunque no hay consenso en cómo aplicarlo y qué tasa utilizar (82). **La mayoría de los países recomiendan descontar los costes y beneficios en salud a una tasa común y constante en el tiempo para costes y beneficios en salud entre el 3 y el 5% según las propias recomendaciones.**

La elección de las tasas de descuento también tiene **implicaciones de equidad, así como de eficiencia**. *Por ejemplo, si el horizonte temporal de una evaluación económica es largo, entonces el descuento constante tiene el efecto de dejar un peso cercano a cero para las generaciones futuras, lo que se argumenta que es una característica indeseable e injusta del descuento constante dado que perjudica los intereses de las generaciones futuras (80).*

Para algunos eventos en salud, pasa **mucho tiempo antes de ver el impacto** entre la implementación de la intervención y los beneficios en salud frente a una enfermedad. Esto hace que los beneficios en salud se vean muy afectados por el descuento en contraposición a intervenciones inmediatas (*por ejemplo, el alivio sintomático del dolor cabeza*) en que los beneficios en salud no se ven prácticamente afectados por el descuento.

Para las enfermedades en que las intervenciones tienen implicaciones en cohortes futuras (*por ejemplo, las enfermedades infecciosas con las vacunas*) algunos autores diferencian las **implicaciones intrageneracionales e intergeneracionales**, y se cuestionan si el descuento también debería ser diferente (81). Es decir, no es igual la comparación de los efectos en salud de un individuo de edad a en el momento t con el mismo individuo de edad $a + 1$ en el momento $t + 1$ (*descuento intrageneracional*), que de un individuo de edad a en el momento t con otro individuo de edad a en el momento $t + 1$ (*descuento intergeneracional*). El descuento reducirá el valor no solo de la salud futura y ganancias de consumo de la generación actual, sino también el valor total de toda la salud y el consumo de una cohorte futura en comparación con la actual. La mayoría de las evaluaciones económicas de intervenciones solo necesitan tener en cuenta la cohorte que recibe la intervención y los beneficios de salud en las cohortes futuras son independientes de las decisiones tomadas en las cohortes anteriores.

Ejemplo (81). En la modelización de la vacunación frente enfermedades infecciosas, a menudo, se incluyen cohortes futuras con el fin de captar mejor los beneficios indirectos (y los daños) como la

inmunidad de grupo. Los beneficios para la salud de las cohortes futuras a través de la inmunidad de grupo dependen de las decisiones tomadas en cohortes anteriores.

El **descuento intrageneracional** podría basarse en las preferencias temporales individuales, mientras que el **descuento intergeneracional** implica cuestiones más amplias de equidad (81). Las generaciones futuras no pueden participar en las decisiones presentes que de algún modo les acabarán afectando.



Recomendaciones

- Seguir las recomendaciones nacionales al respecto sobre la utilización de la tasa de descuento.
- En ausencia de tasa de descuento nacional, seguir las recomendaciones de las guías de evaluaciones económicas y ETS próximas al contexto donde se realiza el análisis.
- En caso de discordancia entre guías sobre la tasa de descuento, contemplar diferentes posibilidades en los análisis de sensibilidad.

2.1.8. Tipos de modelos

La selección del tipo de modelo para la modelización del problema en cuestión también ofrece diversos retos, sobre todo el de no adecuación al modelo conceptual. La complejidad de los modelos analíticos de decisión representa una barrera importante en el área de la salud. La difícil comprensión de la metodología unida a la complejidad de la programación del modelo hace que, sobre todo, los clínicos se desentiendan de las evaluaciones económicas con el uso de modelos analíticos. Si bien

existen algunas interfaces y paquetes específicos de los métodos más comúnmente utilizados, faltan herramientas robustas y fáciles de usar por médicos, administradores de centros sanitarios y decisores en salud (3).

Hay muchos tipos de posibles modelizaciones y su clasificación es confusa porque en la mayoría de casos son combinables. Sí se pueden establecer claramente características de los modelos que deben decidirse previamente a la modelización. Por ejemplo, los modelos pueden ser **estáticos o dinámicos, abiertos o cerrados, agregados o individuales** y **deterministas o estocásticos**. Cada característica definirá un tipo de modelo muy concreto que deberá ajustarse al problema a modelizar. Sin embargo, en muchas ocasiones, las características que se aplican a un problema concreto no se escogen por la mayor adecuación al modelo conceptual sino por practicidad y simplicidad.

Además, aun escogiendo las características adecuadas que se ajusten al problema en cuestión, existen algunas limitaciones matemáticas para poder representar la realidad de los problemas sanitarios (54; 81-84). En los capítulos 9 a 14 se definen las características básicas de los modelos más habituales y la elaboración de estos modelos.

Aunque existen diferentes guías para escoger el tipo de modelo a desarrollar, diseñar el modelo e incluso validarlo, no ha habido ningún avance en los últimos años sobre **herramientas que fomenten y ayuden a la comparabilidad de los modelos existentes**. La idea sería poder comparar si modelos distintos que parten de los mismos supuestos llegan a obtener resultados similares. Este ejercicio, se lleva a cabo muy pocas veces dado que la inversión de tiempo y de recursos económico es grande y resulta más rentable diseñar de novo otro tipo de modelo. Habitualmente, estos ejercicios de comparabilidad se llevan a término en un contexto más académico, para desarrollar las habilidades en el desarrollo y reproducibilidad de modelos que para informar las decisiones sanitarias. En algunas ocasiones, se han comparado modelos simples con modelos más complejos que simulan el mismo evento en contextos con suficientes datos de entrada para modelar diferentes asunciones con el

objetivo posterior de utilizar los modelos simples en contextos con datos de entrada limitados (85-87).



Recomendaciones

- Escoger el tipo de modelo que mejor se ajuste al problema a modelizar y no por practicidad y/o simplicidad.

3 Etapa Post-modelización

3.1. Validación

Dado que la modelización simplifica la realidad con el objeto de estimar y predecir resultados que posteriormente se inferirán de nuevo a la realidad, es necesario llevar a cabo un proceso de validación de los modelos para **evaluar la precisión** y hacer comprender el valor de los mismos para determinar, predecir o representar el mundo real (88). Existen diferentes guías que pueden facilitar la validación de un modelo analítico de decisión (89-96). Podemos distinguir dos tipos de validación de un modelo: la **interna o dependiente**, en el que se comprueba si el modelo no tiene errores o incongruencias en el diseño y su operatividad y si los resultados obtenidos parecen coherentes, y la **externa**, donde haciendo uso de datos reales se comprueba la capacidad real de predicción del modelo y se calibra o ajusta el modelo para que produzca estimaciones lo más cercanas posible a la realidad.

Cualquier desarrollo de un modelo debe ir acompañado de un proceso de validación. Este proceso es imprescindible para aumentar la confianza en la relevancia de un modelo analítico de decisión, medir con la mayor

precisión posible los costes y resultados de los eventos, evaluar la robustez de la distribución muestral de los parámetros de entrada del modelo, así como confirmar la ausencia de errores de implementación. Se han desarrollado algunas herramientas que ayudan al proceso del análisis de decisión de intervenciones en salud pública. Un ejemplo, sería la Assessment of the Validation Status of Health-Economic decision models (AdViSHE) (96). Esta herramienta nace con el objetivo de ayudar a priorizar los esfuerzos en la validación de los modelos y se centra en cinco aspectos: validación del modelo conceptual, validación de los datos de entrada, validación del modelo computarizado, validación operativa y otras técnicas de validación. En cuanto a la validación del modelo conceptual, AdViSHE sugiere una prueba de validez aparente que sea compartida con expertos y una prueba de validez cruzada comparando con otros modelos conceptuales existentes en la literatura. Respecto a la validación de los datos de entrada del modelo, esta herramienta propone también una prueba de validez aparente compartiendo los datos con expertos como parte de la revisión por pares, así como una prueba de ajuste del modelo, como sería la comprobación mediante pruebas estadísticas adecuadas de aquellos parámetros de entrada basados en modelos de regresión. Referente a la validación del modelo computarizado, Vemer y col. (96) señalan la necesidad de una revisión externa, realizar pruebas con valores extremos para detectar si los resultados son plausibles o cualquier error de codificación, pruebas de identificación de que todos los posibles patrones de comportamiento de los individuos estén simuladas garantizando la diversidad de la inclusión de toda la cohorte modelizada, y pruebas unitarias basadas en análisis separados, como por ejemplo, en el caso del estudio de Briggs y col. (86) donde evalúa el efecto de una intervención en un factor de riesgo conductual, y cómo ese factor afecta a los resultados de enfermedad. Existen diferentes enfoques para llevar a cabo este proceso de validación (88,91,97).

El proceso de validación resulta especialmente un **reto en modelos que abordan enfermedades conocidas como no transmisibles o crónicas**, como el cáncer, la enfermedad obstructiva crónica o la diabetes

entre otras, debido a la naturaleza de sus resultados. Estas enfermedades requieren de un **horizonte largo**, y además, dependen de numerosos factores lo que dificulta la identificación de efectos independientes a la intervención (98).



Recomendaciones

- Se debe realizar un proceso de validación interna que evalúe el modelo, su operatividad y los resultados obtenidos.
- Siempre que sea posible también se llevará a cabo un proceso de validación externa si existen datos reales que se puedan utilizar.

3.2. Retos en la monitorización y verificación de un modelo

La monitorización y validación de un modelo son elementos importantes del proceso post-desarrollo de un modelo, sin embargo, se les presta muy poca atención y se les asignan pocos recursos.

La monitorización del modelo analítico de decisión consiste en ajustar la toma de decisiones a lo largo del tiempo según la evolución, disponibilidad y acceso a los datos. Por tanto, permite hacer una estimación y predicción dinámica de la eficiencia de una intervención.

La verificación del modelo permite confirmar que está correctamente implementado con respecto al modelo conceptual. Durante la verificación, el modelo se prueba para encontrar y corregir errores en la implementación del modelo. Se pueden utilizar varios procesos y técnicas para asegurar que el modelo coincide con las especificaciones y supuestos del modelo conceptual. Estos incluyen, entre otros, verificar el modelo por medio de un experto, realizar diagramas de flujo lógicos que incluyan cada acción posible, examinar minuciosamente los resultados

del modelo para determinar su razonabilidad en una variedad de parámetros de entrada (99). En muchas ocasiones, el proceso de verificación de un modelo no es transparente ni explícito y se obvia por la falta de recursos, tanto monetarios como de tiempo.



Recomendaciones

- Se recomienda realizar la verificación del modelo con algunas de las técnicas existentes y ser transparente en el proceso.
- Para asegurar la eficiencia de una intervención a lo largo del tiempo, se recomienda realizar una monitorización del modelo.

3.3. Generalización de los modelos

Otro tema clave se relaciona con la generalización de los modelos. Este concepto se refiere a la **capacidad de un modelo para adaptarse a la aparición de nuevos datos**, es decir, la capacidad de absorber o hacer uso de nuevos datos sin perder precisión en las predicciones.

Una primera pregunta que deberemos hacernos cuando obtenemos datos para la modelización es si la población a estudio, **los resultados y las intervenciones encontrados en la bibliografía son diferentes a los que se consideran en la evaluación**. Una situación muy común sería la utilización de la eficacia cuando realmente necesitamos la efectividad. Frente a esta situación, también debemos preguntarnos **cómo extrapolamos apropiadamente entre los datos requeridos y los datos reales proporcionados en la literatura**. Hay que tener en cuenta que, si adecuamos los datos al caso de interés haciendo ciertas asunciones, los resultados del modelo serán específicos para la situación y no generalizables. Si utilizamos los datos reales en un escenario diferente, los

resultados del modelo pueden estar sesgados. Un modelo no generalizable hará predicciones inexactas cuando se le proporcionen un nuevo conjunto de datos que no sea similar al conjunto inicial, lo que hará que el modelo sea inhábil a pesar de que sea capaz de hacer predicciones precisas para unos datos determinados. Así pues, en algunas ocasiones, sería deseable poder generalizar los modelos a otros entornos o perspectivas cambiando simplemente los datos de entrada del modelo (16).



Recomendaciones

- Se debe sopesar a priori la aplicabilidad, transferibilidad y generalización de los modelos según los objetivos y los datos disponibles.

3.4. Repositorios

Uno de los retos pendientes en la última década de la modelización en el campo sanitario es el desarrollo de una **base de datos de modelos analíticos de decisión, para optimizar el tiempo entre el periodo de desarrollo de un modelo y la inmediatez en la toma de decisiones sanitarias**. La implementación de un repositorio de modelos está lejos aún, aunque su desarrollo haría acortar mucho los tiempos de desarrollo, agilizaría la toma de decisiones y fomentaría el trabajo en red entre equipos de investigación diferentes. Sin embargo, este es un tema que genera muchas resistencias entre el colectivo académico e investigador. Los derechos de autor en los proyectos de investigación, las barreras administrativas que puedan existir en el acceso, la ubicación de un repositorio nacional e internacional, o los recursos económicos para monitorizar y mantener este repositorio son algunas de las cuestiones que hacen que este sea un reto muy importante para afrontar a largo plazo.



Recomendaciones

- Sería necesario y conveniente superar las barreras actualmente existentes para conseguir un repositorio de modelos analíticos de decisión con el objetivo de agilizar la toma de decisiones.

3.5. Modelos con conflictos de interés

En el desarrollo de un modelo analítico de decisión es imprescindible que prevalezca el interés académico y de investigación si realmente quiere utilizarse esta herramienta para informar la toma de decisiones. En cuanto aparecen intereses comerciales o privados, o incluso políticos en el desarrollo de un modelo, la aparición de conflictos de interés está asegurada. Por ejemplo, la industria farmacéutica es una gran productora de modelos analíticos de decisión para evaluar la eficiencia de tecnologías sanitarias, principalmente fármacos. Frecuentemente, los modelos analíticos de decisión que se publican no muestran transparencia suficiente en los procesos de validación interna y externa de sus modelos, y, además, los resultados obtenidos generalmente favorecen su interés comercial. Esto puede causar reticencias a los investigadores a reproducir y utilizar modelos diseñados y desarrollados por los modeladores provenientes del sector privado. Además de una transparente **declaración de conflictos de interés**, es de vital importancia hacer un **proceso de validación interna y externa** para transmitir confianza a los potenciales usuarios del modelo y los encargados de desarrollar las políticas sanitarias.



Recomendaciones

- Los modelos de decisión realizados con intereses comerciales deben ser manifiestos con la declaración de conflictos de interés, además de escrupulosos y transparentes en los procesos de validación interna y externa de los modelos.

3.6. Aspectos sobre los resultados del modelo, extrapolación y su interrelación con la toma de decisiones

Aunque en los últimos años se ha publicado gran cantidad de trabajos en el ámbito de los modelos analíticos de decisión, es sorprendente y decepcionante que no se haya transformado en una mayor aplicación, principalmente por una **barrera de comunicación y falta de confianza entre modeladores y decisores de políticas sanitarias** (3). Es probable que pocos modeladores comprendan realmente la atención médica y pocos médicos o decisores sanitarios comprendan realmente un modelo (3).

El procedimiento de toma de decisiones a menudo se realiza a través de muchos debates y negociaciones entre un grupo de diferentes actores interesados en el tema. Los tomadores de decisiones pueden tener argumentos a favor o en contra de cursos de acción alternativos. Asimismo, pueden adoptar y, en consecuencia, proponer sus propias estrategias que cumplan unas metas a un nivel específico. Los conflictos de intereses son inevitables y se requiere apoyo para lograr consenso y compromiso. Además, el conocimiento de los hechos no siempre es suficiente para tomar una decisión. Los juicios de valor, según los antecedentes, el tipo de personalidad, la cognición, el estilo de decisión, la posición en la empresa y los objetivos de cada responsable de la toma de decisiones, suelen ser los temas más críticos (100).

Por último, pero no menos importante, los responsables de la toma de decisiones no suelen ser competentes en matemáticas o informática. El sistema debe proporcionarles las herramientas adecuadas para participar en la discusión de una manera “natural”, al mismo tiempo que se debe facilitar una confianza mutua. Esto está de acuerdo con la visión de los pioneros en el campo de los sistemas de soporte de decisiones, es decir, al respaldar y no reemplazar el juicio humano, el sistema ocupa el segundo lugar y los usuarios primero.

3.6.1. Comunicación compartida: integración de los modeladores en el proceso de toma de decisiones

La comunicación de los modeladores de cómo se pueden y deben usar los modelos para dar respuesta a las cuestiones de política sanitaria es imprescindible. A veces, las necesidades de los decisores sanitarios son relativamente simples, pero lo que realmente se está modelando es innecesariamente complicado y no tiene un gran interés político. Por este motivo, es muy importante en el proceso de investigación **reducir la brecha entre la “necesidad de saber” y “lo que se está haciendo”**. La **interacción y comunicación continuada entre los modeladores y los responsables de la formulación de políticas** sanitarias contribuirá en gran medida a este objetivo (4). Por ejemplo, el sistema sanitario inglés y holandés tienen integrados dentro de los sistemas nacionales de vigilancia los equipos de modelización (4). Las interacciones frecuentes crean la posibilidad de un uso más eficaz de los modelos predictivos y prescriptivos basados en diferentes escenarios. La **posibilidad de tener disponibles modelos que puedan hacer prescripciones**, tanto a corto como a largo plazo, y proporcionar pronósticos continuamente actualizados que capturen la incertidumbre y ganen precisión a medida que se disponga de nueva información es muy interesante para diseñar políticas públicas realistas y sostenibles. Estos **modelos** deben ser lo más **simples** posibles, pero no tan simples como para perder el realismo de los

supuestos y alterar las recomendaciones de políticas. El análisis de sensibilidad puede ayudar a evaluar diferentes escenarios que incorporen más realismo al análisis.

Los retos en la comunicación pueden mitigarse estableciendo un **diálogo inicial sobre lo que la modelización tiene para ofrecer y qué datos pueden ser valiosos para la toma de decisiones** sanitarias, particularmente a la luz de tecnologías cambiantes y metodologías novedosas. Sin embargo, para resolver el problema hay que crear **relaciones a largo plazo que fomenten la comprensión y la confianza** para evitar un enfoque “único” de la modelización. Esto desarrollaría un ciclo en el que los modelos informan las políticas, lo que conduce a nuevos datos y preguntas sobre políticas, y retroalimenta el diseño y desarrollo del modelo. Para ello, se necesita una **continua formación técnica de los modeladores** para mantener los resultados actualizados y los **modelos** suficientemente **flexibles** para poder actualizarlos constantemente. Esto nos llevaría a un ciclo continuo de decisión y monitorización de las decisiones sanitarias.



Recomendaciones

- Es imprescindible la inclusión en el equipo de trabajo de los modeladores, algún responsable de política sanitaria para poder tener en cuenta su opinión en el desarrollo de los modelos y evitar las barreras de comunicación entre modeladores y decisores. La comunicación entre los dos grupos de profesionales debe ser fluida desde el inicio de la evaluación o proyecto.

3.6.2. Umbrales de coste-efectividad

Los responsables de la toma de decisiones médicas se enfrentan al reto de identificar la asignación óptima de recursos para optimizar los beneficios en salud. Sin embargo, debido a las constantes limitaciones presupuestarias, la mayoría de los sistemas de salud no pueden poner a disposición todas las nuevas intervenciones desarrolladas para la detección, prevención y tratamiento de enfermedades (101). Cada vez más, se exige a los gobiernos que administren estratégicamente estos recursos limitados, invirtiendo solo en aquellos servicios que ofrecen los mejores resultados de salud y al mismo tiempo que una atención asequible, efectiva, segura y centrada en el paciente (102). En respuesta a estas presiones, los responsables de la toma de decisiones en salud utilizan las evaluaciones económicas para identificar aquellas intervenciones que brindan el mejor valor y ayudan a la asignación de recursos. Es muy común en este contexto el uso de los análisis de coste-efectividad donde la comparación de los costes y los beneficios en salud entre intervenciones se expresa usando la razón de coste-efectividad o de coste-utilidad incremental, según los beneficios en salud que se utilicen. En este contexto, se tiende a utilizar un umbral de coste-efectividad formal o informal por debajo del cual una intervención se considera coste-efectiva. Este umbral ayuda en el proceso de la toma de decisiones, sin embargo, hay una gran discusión sobre el verdadero valor de este umbral (75,102-104).

Los umbrales de coste-efectividad son específicos por país y existe controversia si incluso deberían haber varios umbrales por país para diferentes situaciones según si emplean para la valoración monetaria de la sociedad de los beneficios en salud o para el coste de oportunidad resultante de la desinversión necesaria cuando se adopta una nueva tecnología (105-106). Tampoco parece haber consenso sobre el método más apropiado para calcular los umbrales de coste-efectividad (104). En los últimos años, se han desarrollado o adaptado otros métodos. Por ejemplo, el método de la frontera de eficiencia

usado por países como Alemania, o la heurística por parte de la Organización Mundial de la Salud (OMS) que usa entre una y tres veces el producto interior bruto (PIB) per cápita del país (9). Establecer este umbral con cierto sentido es importante y puede influir en la toma de decisiones. Muchos países de rendas bajas utilizan los umbrales de coste-efectividad recomendados por el programa CHOosing Interventions that are Cost-Effective (WHO-CHOICE) de la OMS para evaluar las intervenciones de salud. Algunos con rentas más altas también los utilizan cuando el país carece de un umbral específico (107).

Así pues, la determinación de los umbrales de coste-efectividad representa también un reto para los responsables de la toma de decisiones en salud.



Recomendaciones

- Seguir las recomendaciones de las guías de evaluaciones económicas y ETS nacionales.
- En caso de no disponer de un umbral en el contexto específico, utilizar diversos umbrales que se utilicen en entornos similares.

3.6.3. Falta de correlación/interacción entre las necesidades de los decisores en salud y los investigadores que desarrollan los análisis de decisión

La ETS se define como un campo multidisciplinario de análisis de políticas, donde se estudia las implicaciones médicas, sociales, éticas y económicas del desarrollo, la difusión y el uso de la tecnología sanitaria. La ETS tiene diferentes partes (por ejemplo, evaluaciones económicas y de efectividad clínica) y diferentes enfoques (cualitativo: revisión narrativa y

tablas de evidencia; cuantitativo: metaanálisis y análisis de decisiones) (108). En los últimos años, en un número creciente de informes de ETS, los cuáles informan las decisiones de política sanitarias, las agencias de ETS van más allá de revisar y resumir cualitativamente la evidencia de estudios internacionales publicados, como se ha hecho tradicionalmente, y ahora elaboran modelos para garantizar que los resultados reflejen el contexto del sistema de atención de salud del país investigado.

Se necesita un periodo considerable de tiempo para poder desarrollar un modelo de decisión desde cero y aplicarlo o adaptarlo a los datos nacionales. Es por este motivo, que la Agencia Alemana para la Evaluación de Tecnologías Sanitarias (DAHTA) encargó un proyecto metodológico que incluía el desarrollo de un marco para (a) identificar los modelos de decisión existentes de alta calidad, transparencia y flexibilidad, (b) transferirlos, extenderlos y adaptarlos en cooperación con los autores originales al ámbito nacional y (c) establecer una red de colaboración entre organizaciones de ETS de diferentes países que promueva y auspicie el intercambio de estos modelos entre agencias de ETS. Esta idea e iniciativa ha sido respaldada por países como el Reino Unido y Canadá, ayudando a evitar la duplicación de trabajos en el desarrollo de modelos y lograr la sinergia entre equipos de investigación y desarrollo de modelos.

Muchos modelos o análisis de decisión de políticas de salud utilizan una cohorte homogénea, hipotética o fija (sin tener en cuenta la incidencia) de pacientes y no muestran los valores esperados de los resultados y costes de salud totales absolutos para una población heterogénea en un país en particular. Es más probable que dicha información surja de los llamados modelos de políticas de salud (108). Aunque este término no se define de manera uniforme, la mayoría de estos modelos consideran la heterogeneidad de la población real de un país (u otra entidad) con respecto a la edad, el género, la etnia, la gravedad de la enfermedad o la distribución de los factores de riesgo y otras características relevantes para el problema de decisión. Estos

modelos deben utilizar un enfoque de cohorte dinámico, es decir, que permita que los pacientes entren y salgan del grupo de decisiones (es decir, la población dinámica en un punto de tiempo dado, incluidos todos los pacientes para quienes se tomó una decisión) durante el horizonte de tiempo analítico (108). Este enfoque puede llevar a estimaciones de coste-efectividad conservadoras, porque el numerador de la relación coste-efectividad incremental puede incluir algunos gastos relacionados con la intervención en los últimos años del horizonte de tiempo analítico sin incluir el efecto beneficioso de la intervención en la salud futura, o el ahorro de costes debido a la reducción de la morbilidad y la mortalidad (108). Un ejemplo de un modelo de política de salud dinámica sería el modelo de política de tabaquismo que se utilizó para describir el impacto en salud pública y el coste-efectividad de un programa mejorado de educación antitabaco en las escuelas de los Estados Unidos a nivel nacional (109-110). Este modelo simula el nacimiento, la muerte, el envejecimiento y los cambios en el estado del tabaquismo en la población general. Este modelo calculó los costes del programa, los costes médicos, los años de vida totales, los AVAC y la ratio de coste-efectividad incremental para el programa antitabaco frente al statu quo después de 25 y 50 años.



Recomendaciones

- Es imprescindible que los decisores sanitarios consulten y colaboren estrechamente con expertos externos en modelización, a quién deberían comunicar las necesidades de evaluación que existen actualmente en el país para poder informar mejor la política sanitaria.

3.6.4. Importancia de incluir un análisis económico en un modelo analítico de decisión

Para la toma de decisiones, lo que generalmente se requiere no es simplemente una estimación media de lo que es probable que suceda, o un análisis basado en escenarios, sino más bien un **análisis de los costes y beneficios de un conjunto de posibles políticas de salud pública** para orientar la decisión a favor de una frente las demás (4). Esta estimación de costes y beneficios recae en el análisis económico y comúnmente **es una dimensión no incorporada en la mayoría de los modelos de enfermedades infecciosas**. De la misma manera, los modelos analíticos de decisión realizados por los economistas rara vez están diseñados para tener en cuenta los efectos no lineales que surgen de la dinámica de las enfermedades infecciosas (4). En los próximos años, este tema jugará un **papel esencial en el desarrollo de modelos más útiles para la toma de decisiones y el desarrollo de políticas sanitarias**. De hecho, la toma de decisiones en salud pública y el coste y beneficio relativo a diferentes alternativas no debería ser competencia de los modeladores, ni siquiera de los economistas; sino que debería ser **responsabilidad de los grupos de salud pública gubernamentales e internacionales** (111). Esta dualidad y colaboración entre los modeladores y economistas puede proporcionar un poderoso apoyo a la toma de decisiones que no sería posible si se combina el trabajo individual de ambos.



Recomendaciones

- Es imprescindible incorporar la dimensión económica en los modelos analíticos de decisión para una toma de decisiones informada en política sanitaria.

3.6.5. Integración de diferentes puntos de vista en la modelización matemática dentro de las instituciones de salud pública

Los investigadores que desarrollan un modelo deben plantearse seriamente cómo interactúan e integran la participación ciudadana. Esta interacción debería tener en cuenta los siguientes aspectos (5): a) debería **familiarizarse más a los pacientes y ciudadanía con el uso de modelos**, incluso haciendo uso de modelos fuera del sector médico para que su uso en la atención sanitaria fuera de más fácil comprensión y se pudiera transferir el aprendizaje de otras áreas o disciplinas al sistema de salud; b) **facilitar la participación ciudadana en el diseño de modelos** a lo largo de un periodo continuo en el tiempo en el que puedan aportar información, por ejemplo, sobre los límites del modelo conceptual y recibir una **retroalimentación de lo que es y no factible** representar sobre su problema de salud; c) **intercambiar conocimiento entre el modelizador y el paciente** o ciudadano que facilite ampliar las perspectivas con las que se analiza el problema; d) **flexibilizar y hallar puntos de encuentro de opiniones diferentes**; e) fomentar una **perspectiva evaluadora sobre los resultados en salud** que perdure en el tiempo.

En las etapas en la que se diseña y desarrolla el modelo es importante tener en cuenta las siguientes premisas (112). El diseño del modelo debe ser **flexible**, donde los valores de las variables puedan cambiarse fácilmente, y el diseño permita un amplio uso del modelo en diferentes instituciones. Es importante **involucrar a los pacientes y ciudadanía en todo momento** y presentar e introducir los modelos a las personas específicas de las instituciones involucradas para **generar credibilidad y ayudar a que los modelos reflejen los procesos de la vida real**. También es importante tener en cuenta que, aunque los modelos deben mantenerse lejos de los posibles intereses políticos de los decisores sanitarios, **deberían comprender y aceptar las relaciones políticas** en el contexto donde se vaya a tomar la decisión informada por el modelo. Una disponibilidad y **asignación de recursos**

suficiente para llevar a cabo el modelo es clave para invertir no solo en recursos humanos, sino también en hardware, software y otros recursos de computación necesarios.

Una vez el modelo ya se ha realizado, es importante realizar una **revisión del proyecto para evaluar** si el proyecto funcionó como se esperaba, cumplió con los objetivos y qué podría haberse hecho de manera diferente, y **promocionar los resultados** dentro del contexto sanitario.

Existen dos posibilidades para poder incorporar personas expertas en diseñar y desarrollar modelos analíticos de decisión en el área sanitaria (54): **centralizar**, acumulando la masa crítica de modeladores con la ventaja que pueden intercambiar ideas e información sobre los aspectos técnicos de modelización, y **descentralizar**, teniendo estos profesionales integrados en grupos de trabajo con investigadores de otros campos (biología, epidemiología, inmunología, ciencias ambientales), asegurando la comunicación entre ellos desde el inicio.

La tendencia actual es **ir hacia una organización más descentralizada de estos profesionales** (54). Esto puede mejorar la sinergia de trabajo de estos profesionales en la recopilación de datos y el desarrollo del modelo, y generar ideas de investigación más innovadoras, aunque podrían aislarse de otros modeladores y obtener una retroalimentación insuficiente sobre los aspectos técnicos de su trabajo (54).

Los modeladores trabajan en una amplia gama de temas en proyectos interdisciplinarios y, por lo tanto, a menudo se encuentran dispersos en diferentes organizaciones. Kallioniemi estableció dos niveles organizativos para el desarrollo de una institución de bioinformática: una unidad central que opere en la institución y otras unidades periféricas más pequeñas integradas en diferentes grupos de investigación (113). De esta manera, el proceso de desarrollo del modelo de decisión analítico combina la estrecha colaboración de los modeladores con científicos de otras disciplinas en unidades periféricas y puede disponer de suficiente masa crítica en una organización para la retroalimentación metodológica de los profesionales.

Para poder llevar a cabo este proyecto hay que establecer algunas premisas y/o requisitos. La unidad central debe garantizar la **transparencia** y mantener ciertas tareas organizativas de **coordinación y responsabilidad** en los proyectos. Se necesita disponer de **recursos monetarios** para poder desarrollar las actividades. Es imprescindible la **buena y constante comunicación** entre todos los implicados en el desarrollo del modelo. Es imprescindible poder **disponer de fondos para mantener y actualizar el modelo** una vez se ha acabado el proyecto de investigación. Para que un modelo se utilice y genere confianza en los decisores de política sanitaria es imprescindible que se haga una **correcta difusión y comunicación de la importancia de las aportaciones conseguidas**, aunque el modelo debe mantenerse independiente de las cuestiones de política e investigación específicas. A menudo esta contribución en la formulación de políticas de salud pública no se reconoce explícitamente, aunque ha sido fundamental para generar proyecciones y estimaciones del impacto de diversas intervenciones en salud. La aparición de la COVID-19 ha cambiado ligeramente este aspecto, aunque puede ser efímero, ya que incluso una parte de la ciudadanía está pendiente de las predicciones de los modelos epidemiológicos.

La **comunicación entre los decisores y los profesionales que diseñan y desarrollan los modelos debe estar organizada** bajo el objetivo común de resolver un problema de salud común específico. Por otro lado, las instituciones deben **proteger las condiciones laborales de los profesionales** que realizan estos modelos analíticos de decisión con estrategias de contratación a largo plazo en las instituciones para atraer investigadores con talento y mantener la masa crítica. Además, es necesaria la **educación continua de estos profesionales** mediante cursos o su participación en jornadas o conferencias, ya que las herramientas de modelización matemática deben estar actualizadas en todo momento en términos de formulación matemática, implementación en código informático, algoritmos numéricos y enfoques estadísticos (114). Por último, pero no menos importante, las **preguntas de investigación**

que se persiguen deben estar **inspiradas en preguntas de política de salud pública reales** junto con preguntas de interés para la ciudadanía expresadas por los medios de comunicación. La aplicabilidad de los resultados en la formulación de políticas debe ser directa (115).



Recomendaciones

- Un equipo de modelizadores matemáticos encargados de desarrollar un modelo para evaluar políticas o programas/intervenciones sanitarias debe incluir a un paciente y/o ciudadano, a profesionales sanitarios y decisores de política sanitaria, desde el inicio del proyecto, para asegurar que el modelo de análisis de decisión que se construya dé respuesta a las necesidades y/o cuestiones planteadas por los diferentes agentes involucrados en el problema de salud.



Recomendaciones o propuesta de soluciones

Etapa PRE-modelización

Establecimiento del contexto de la modelización

- En el proceso de contextualización es indispensable saber desde el inicio a qué pregunta debe responder el modelo y quién es el que formula esta pregunta, estableciendo una vinculación desde el inicio con el proceso de toma de decisiones.
- Deberían involucrarse todas las partes interesadas del proceso de definición del problema dado que a menudo tienen diferentes percepciones de la realidad.

Etapa PRE-modelización

Participación ciudadana y usuarios en investigación sanitaria

- La participación ciudadana en la investigación sanitaria y en el proceso de decisiones sanitarias puede aportar muchos beneficios en el diseño y validación de un modelo, pero los pacientes y ciudadanos deben recibir formación para compartir el lenguaje con los modeladores y entender bien los modelos.

Aspectos sobre los datos de entrada

- La cuantificación de cómo la información adicional reduce la incertidumbre en la decisión se puede estimar mediante el valor esperado de la información perfecta (EVPI) que corresponde al coste medio que se estaría dispuesto a pagar para tener información completa y perfecta, es decir, sería el coste de eliminar toda la incertidumbre del análisis (19).
- Para evitar una insuficiente identificación de la evidencia científica o que sea de conveniencia o sesgada se recomiendan varios puntos. 1) todo el proceso debe estar completamente documentado para ayudar a la transparencia, reproducibilidad y actualización. 2) Establecer a priori las estrategias de búsqueda apropiadas para cada parámetro y definir los criterios de inclusión y exclusión que pueden no ser evidentes al inicio. 3) El punto 2 necesita de la comunicación continuada y fluida entre los diferentes actores que intervienen en la construcción del modelo. 4) Publicar esta información juntamente con los resultados. Debido a las restricciones de espacio de los artículos científicos, puede incluirse como material complementario de forma online.
- Un enfoque más eficiente para realizar la búsqueda de información es realizarla de forma estructurada y centrada en torno a los parámetros que se espera que tengan mayor influencia en los resultados. La identificación de dichos parámetros se realizará a través de la opinión de expertos, estudios previos y/o modelos iniciales. También pueden realizarse búsquedas iterativas (16).
- Aunque se requiere metodología y mayor orientación de cómo abordar la calidad de los datos, es recomendable que los investigadores califiquen cualitativa o cuantitativamente la calidad de la evidencia incorporada en el modelo. Se recomienda que se excluya la información de poca calidad y se evalúe a posteriori la información más incierta mediante análisis de sensibilidad.

Etapa PRE-modelización

Aspectos sobre los datos de entrada *cont.*

- La inteligencia artificial puede jugar un papel importante en los modelos analíticos de decisión mediante la modelización basada en el ‘big data’, aunque para obtener el mayor potencial y poder hacer análisis prescriptivos requiere de la combinación con la modelización basada en la simulación la cual incorpora las relaciones causales de la dinámica del sistema.
- Dados los retos que presenta la inteligencia artificial para los responsables de la formulación de políticas sanitarias, es recomendable un control de su uso en la atención sanitaria, así como una regulación para la toma de decisiones basada en la inteligencia artificial, el establecimiento de la propiedad de los datos a través, por ejemplo, de un centro de datos impulsado por el gobierno regional o nacional, y la dotación de recursos humanos con conocimientos interdisciplinarios.
- Los métodos más apropiados para estimar la “no intervención” dependerán de cada situación y de los datos disponibles previos a la implementación:
 - En los modelos de pre-implementación, habitualmente se estima proyectando las tasas recientes de enfermedad asumiendo que no hay cambios en el periodo de proyección o solo cambios demográficos relevantes esperados.
 - En los modelos post-implementación, se puede modelar la carga de enfermedad actual bajo la intervención y estimar las tasas de “no intervención” aplicando algún factor de inflación a las tasas actuales. También se puede utilizar información de la de poblaciones similares a las que se quiere simular. En algunas circunstancias y según los datos disponibles, pueden utilizarse proyecciones con modelos estadísticos que incorporen predictores del comportamiento pasado a lo largo del tiempo. En el caso de infecciones, también se puede simular la transmisión subyacente de la infección usando un modelo dinámico.
- Dado que en la mayoría de ocasiones, es difícil recabar todos los datos de la misma fuente, se recomienda usar y contrastar diferentes fuentes y entender que conceptos de costes se incluyen en cada una.
- Revisar las guías generales y específicas de algunas enfermedades sobre evaluaciones económicas o ETS, además de estudios previos para comprobar y ampliar las fuentes de datos económicos disponibles.

Etapa PRE-modelización

Retos que plantean las particularidades de la modelización de algunos eventos de salud

Enfermedades transmisibles y sus consecuencias

- Debido a las externalidades que generan las enfermedades transmisibles, se deben tener presente a las partes interesadas y/o afectadas.
- Deben encontrarse formas en las que los modelos predictivos puedan ser incorporados en el campo de la salud pública, dada la demanda por parte de los decisores políticos en acciones concretas.
- Dada la dificultad de que muchos supuestos del modelo se mantengan con el tiempo, es necesario desarrollar métodos para actualizar y validar los supuestos del modelo con horizontes temporales largos a medida que aumenta la disponibilidad de datos o estos supuestos pierden validez (4).

Medicina de precisión

- Es necesario un adecuado aprovechamiento de los extensos volúmenes de datos que utiliza la medicina de precisión mediante herramientas sofisticadas como las ya existentes en otros países.
- Con el objeto de evitar que las decisiones tomadas a partir de los métodos cuantitativos y algoritmos usados en la medicina de precisión den lugar a una discriminación negativa, es necesaria una adaptación del uso ético y transparente de estos datos.
- La investigación paralela de modelos matemáticos que den respuesta a los requerimientos de la medicina de precisión junto a la investigación sobre esta misma, requiere de una perspectiva no solo individual sino comunitaria.

Medicina personalizada (datos genómicos)

- Dada la necesidad de definir explícitamente el alcance del análisis, se propone evaluar, por un lado, la eficiencia del cribado teniendo en cuenta únicamente el evento para el cual se lleva a cabo como objetivo primario y, por otro lado, incluir los otros hallazgos y la información sobre los mismos, pudiendo comparar a posteriori su eficiencia en ambas situaciones (58).
-

Perinatología: dualidad materno infantil

- Los modelos de decisión que evalúan intervenciones o programas de perinatología deben contemplar no sólo los costes y resultados en salud de la madre sino también las consecuencias para el feto.

Etapa PRE-modelización

Retos que plantean las particularidades de la modelización de algunos eventos de salud *cont.*

Programas de cribado

- Dada la necesidad de definir explícitamente el alcance del análisis, se propone evaluar, por un lado, la eficiencia del cribado teniendo en cuenta únicamente el evento para el cual se lleva a cabo como objetivo primario y, por otro lado, incluir los otros hallazgos y la información sobre los mismos, pudiendo comparar a posteriori su eficiencia en ambas situaciones (58).

Enfermedades multisistémicas y afecciones complejas

- Es importante incluir los impactos, más allá de la disciplina de la salud, en los análisis de decisión, modelos de enfermedades multisistémicas y afecciones complejas.

Etapa de Desarrollo del modelo

Aspectos metodológicos

- Determinar a priori el número de cohortes a modelizar adecuados al problema concreto y no por simplicidad.
- El uso de una sola cohorte puede ser adecuada si no hay importantes dependencias entre cohortes o no se prevén grandes cambios en un futuro próximo.
- En el caso de utilizar varias cohortes, explorar la relevancia de dar estimaciones individuales por cohorte o agregadas.
- El horizonte temporal debe ser adecuado al objetivo del estudio y a los eventos simulados. Se recomienda que el tiempo de simulación incluya la duración de la intervención y pueda captar las diferencias en costes y efectos en salud entre las diferentes alternativas.
- En el caso de utilizar un horizonte temporal corto, hay que ser consciente de y coherente con las limitaciones en los resultados.

Etapa de Desarrollo del modelo

Aspectos metodológicos *cont.*

- Se recomienda la elección de la duración del ciclo no únicamente en función de la disponibilidad de los datos, sino teniendo en cuenta una duración de ciclo lo suficientemente corta como para que no se requiera aplicar la corrección de medio ciclo (66).
- La mayoría de las guías de evaluaciones económicas y ETS, recomiendan el uso del AVAC.
- Para favorecer la comparabilidad de diferentes estudios se recomiendan los siguientes puntos en referencia a la métrica de los beneficios en salud:
 - La definición, descripción y transparencia en el uso de las utilidades en los AVAC.
 - Revisar la bibliografía sobre el evento en concreto para realizar un análisis con la medida de beneficios en salud que ayude a la comparación entre estudios. Por ejemplo, en los países en vías en desarrollo con el uso del AVAD.
 - Es deseable realizar varios análisis con diferentes medidas de beneficios en salud (por ejemplo, ajustando o no por calidad, AVAC y AVG) si hay dudas sobre la calidad de las utilidades o si ello favorece la comparabilidad con otros estudios.
- Seguir las recomendaciones nacionales sobre la utilización de la tasa de descuento.
- En ausencia de tasa de descuento nacional, seguir las recomendaciones de las guías de evaluaciones económicas y ETS próximas al contexto donde se realiza el análisis.
- En caso de discordancia entre guías sobre la tasa de descuento, contemplar diferentes posibilidades en los análisis de sensibilidad
- Escoger el tipo de modelo que mejor se ajuste al problema a modelizar y no por practicidad y/o simplicidad.

Etapa Post-modelización

Validación

- Se debe realizar un proceso de validación interna que evalúe el modelo, su operatividad y los resultados obtenidos.
- Siempre que sea posible también se llevará a cabo un proceso de validación externa si existen datos reales que se puedan utilizar.

Etapa Post-modelización

Retos en la monitorización y verificación de un modelo

- Se recomienda realizar la verificación del modelo con algunas de las técnicas existentes y ser transparente en el proceso.
- Para asegurar la eficiencia de una intervención a lo largo del tiempo, se recomienda realizar una monitorización del modelo

Generalización de los modelos

- Se debe sopesar a priori la aplicabilidad, transferibilidad y generalización de los modelos según los objetivos y los datos disponibles.

Repositorios

- Sería necesario y conveniente superar las barreras actualmente existentes para conseguir un repositorio de modelos analíticos de decisión con el objetivo de agilizar la toma de decisiones.

Modelos con conflictos de interés

- Los modelos de decisión realizados con intereses comerciales deben ser manifiestos con la declaración de conflictos de interés.
- Además, deben ser escrupulosos y transparentes en los procesos de validación interna y externa de los modelos.

Aspectos sobre los resultados del modelo, extrapolación y su interrelación con la toma de decisiones y comunicación compartida: integración de los modeladores en el proceso de toma de decisiones

- Es imprescindible la inclusión en el equipo de trabajo de los modeladores algún responsable de política sanitaria para poder tener en cuenta su opinión en el desarrollo de los modelos y evitar las barreras de comunicación entre modeladores y decisores. La comunicación entre los dos grupos de profesionales debe ser fluida desde el inicio de la evaluación o proyecto.

Umbral de coste-efectividad

- Seguir las recomendaciones de las guías de evaluaciones económicas y ETS nacionales.
- En caso de no disponer de un umbral en el contexto específico, utilizar diversos umbrales de entornos similares.

Etapa Post-modelización

Falta de correlación/interacción entre las necesidades de los decisores en salud y los investigadores que desarrollan los análisis de decisión

- Es imprescindible que los decisores sanitarios consulten y colaboren estrechamente con expertos externos en modelización, a quién deberían comunicar las necesidades de evaluación que existen actualmente en el país para poder informar mejor la política sanitaria.

Importancia de incluir un análisis económico en un modelo analítico de decisión

- Es imprescindible incorporar la dimensión económica en los modelos analíticos de decisión para una toma de decisiones informada en política sanitaria.

Integración de diferentes puntos de vista en la modelización matemática dentro de las instituciones de salud pública

- Un equipo de modelizadores matemáticos encargados de desarrollar un modelo para evaluar políticas o programas/intervenciones sanitarias debe incluir a un paciente y/o ciudadano, a profesionales sanitarios y decisores de política sanitaria desde el inicio del proyecto, para asegurar que el modelo de análisis de decisión que se construya dé respuesta a las necesidades y/o cuestiones planteadas por los diferentes agentes involucrados en el problema de salud.



Bibliografía

1. Walker W, Harremoës P, Rotmans J, et al. Defining Uncertainty: A Conceptual Basis for Uncertainty Management in Model-Based Decision Support. *Integrated Assessment*. 2003;4-Issue 1.
2. Dunn WN. Using the method of context validation to mitigate type III errors in environmental policy analysis. In Hisschemöller M, Hoppe RA, Dunn WN, et al. *Knowledge, power and participation in environmental policy analysis*. 1st Ed. New Brunswick, NJ: Transaction Publishers 2001; p. 417-436.
3. Eldabi T, Paul RJ, Young T. Simulation modelling in healthcare: reviewing legacies and investigating futures. *J Oper Res Soc*. 2007;58:262–70.
4. Metcalf CJE, Edmunds WJ, Lessler J. Six challenges in modelling for public health policy. *Epidemics*. 2015;10:93–6.
5. Pearson M, Monks T, Gibson A, et al. Involving patients and the public in healthcare operational research—The challenges and opportunities. *Oper Res Health Care*. 2013;2:86-9.
6. Maier HR, Ascough II JC, Wattenbach M, et al. Uncertainty in Environmental Decision Making: Issues, Challenges and Future Directions. In: Jakeman AJ, Voinov AA, Rizzoli AE, et al, editors. *Developments in Integrated Environmental Assessment*. Elsevier. 2008; p. 69–85.
7. NICE. Guide to the Methods of Technology Appraisal 2013 [Internet] [Internet]. PubMed. 2013 [cited 2021 Jun 18]. <https://pubmed.ncbi.nlm.nih.gov/sire.ub.edu/27905712/> (Acceso verificado el 27 Sept 2023)
8. Institute for Quality and Efficiency in Health Care. General Methods. Version 4.2 2015. Germany. https://www.iqwig.de/methoden/iqwig_general_methods_version_204-2.pdf (Acceso verificado el 27 Sept 2023)
9. Zozaya N, Martínez L, Alcalá B, Hidalgo-Vega A. Evaluación, financiación y regulación de los medicamentos innovadores en los países desarrollados. Ed. Fundación Weber. 2018. https://weber.org.es/wp-content/uploads/2017/11/interactivo_informe_evaluacion_financiacion_y_regula-

- cion_de_los_medicamentos_innovadores_en_los_paises_desarrollado.pdf ((Acceso verificado el 27 Sept 2023)
10. Haute Autorité de Santé (HAS). Doctrine de la commission de la transparence (CT) Principes d'évaluation de la CT relatifs aux médicaments en vue de leur accès au remboursement. Adopté par la CT le 15 février 2023. https://www.has-sante.fr/upload/docs/application/pdf/2021-03/doctrine_ct.pdf (Acceso verificado el 27 Sept 2023)
 11. Healthcare Improvement Scotland. Guidance to Submitting Companies for Completion of New Product Assessment form (NPAF). Supplement for medicines eligible for the interim acceptance decision option. 2021. Scottish Medicines Organization (SMO). <https://www.scottishmedicines.org.uk/media/6229/20210901-guidance-supplement-interim-acceptance.pdf> (Acceso verificado el 27 Sept 2023)
 12. Italian Medicines Agency. Criteria for ranking therapeutic innovation of new drugs and elements for supplementing the dossier for admission to the reimbursement system. Working group on innovative drugs. 2007. http://www.agenziafarmaco.gov.it/allegati/integral_document.pdf (Acceso verificado el 27 Sept 2023)
 13. The dental and pharmaceutical benefits agency. General guidelines for economic evaluations from the Pharmaceutical Benefits Board (LFNAR). Pharmaceutical Benefits Board. Suecia. 2003. <https://www.tlv.se/download/18.2e53241415e842ce95514e9/1510316396792/Guidelines-for-economic-evaluations-LFNAR-2003-2.pdf> (Acceso verificado el 27 Sept 2023)
 14. Patented Medicine Prices Review Board Canada. Compendium of policies, guidelines and procedures. Conseil d'examen du prix des médicaments brevetés Canada. Updated February 2017. https://www.pmprb-cepmb.gc.ca/CMFiles/Compendium_Feb_2017_EN.pdf (Acceso verificado el 27 Sept 2023)
 15. Paisley S. Identification of Evidence for Key Parameters in Decision-Analytic Models of Cost Effectiveness: A Description of Sources and a Recommended Minimum Search Requirement. *Pharmacoeconomics*. 2016;34: 597-608.

16. Cooper NJ, Sutton AJ, Ades AE, et al. Working Group on the Use of Evidence in Economic Decision Models. Use of evidence in economic decision models: practical issues and methodological challenges. *Health Econ.* 2007;16:1277–86.
17. Egger M, Juni P, CJ B, et al. How Important are Comprehensive Literature Searches and the Assessment of Trial Quality in Systematic Reviews? Empirical Study. *Health Technol Assess.* 2003;7:1-6.
18. Golder S, Glanville J, Ginnelly L. Populating decision-analytic models: the feasibility and efficiency of database searching for individual parameters. *Int J Technol Assess Health Care.* 2005;21:305-11.
19. York Health Economics Consortium (YHEC). Expected Value of Perfect Information (EVPI) (2016). York; York Health Economics Consortium; 2016. <https://yhec.co.uk/glossary/expected-value-of-perfect-information-evpi/> (Acceso verificado el 27 Sept 2023)
20. Newall AT, Reyes JF, Wood JG, et al. Economic evaluations of implemented vaccination programmes: key methodological challenges in retrospective analyses. *Vaccine.* 2014;32:759–65.
21. Almagoshi S. Simulation Modelling in Healthcare: Challenges and Trends. *Procedia Manuf.* 2015;3:301–7.
22. Sun TQ, Medaglia R. Mapping the challenges of Artificial Intelligence in the public sector: Evidence from public healthcare. *Gov Inf Q.* 2019;36:368-83.
23. Dinov ID. Methodological challenges and analytic opportunities for modeling and interpreting Big Healthcare Data. *Gigascience.* 2016;5:12.
24. Chan C-L, Chang C-C. Big Data, Decision Models, and Public Health. *Int J Environ Health Res.* 2020;17:6723.
25. Hong L, Luo M, Wang R, et al. Big Data in Health Care: Applications and Challenges. *Data and Information Management.* 2018;2.
26. Bakker L, Aarts J, Uyl-de Groot C, et al. Economic evaluations of big data analytics for clinical decision-making: a scoping review. *J Am Med Inform Assoc.* 2020;27:1466-75.

27. Kim BS, Kang BG, Choi SH, et al. Data modeling versus simulation modeling in the big data era: case study of a greenhouse control system. *Simulation*. 2017;93:579–94.
28. Lenz R, Peleg M, Reichert M. Healthcare process support: achievements, challenges, current research. DBIS Epub (CORE), 2012. <https://core.ac.uk/reader/299376498>. (Acceso verificado el 27 Sept 2023)
29. Liu Q, Diao L, Tu G. The Application of Artificial Intelligence in Mobile Learning. In: *International Conference on System Science, Engineering Design and Manufacturing Informatization*. Yichang, China: IEEE; 2010 ; p. 80–3. <http://ieeexplore.ieee.org/document/5640149/> (Acceso verificado el 27 Sept 2023)
30. Miller JD, Foley KA, Russell MW. Current challenges in health economic modeling of cancer therapies: a research inquiry. *Am Health Drug Benefits*. 2014;7:153-62.
31. Bertozzi AL, Franco E, Mohler G, et al. The challenges of modeling and forecasting the spread of COVID-19. *Proc Natl Acad Sci USA*. 2020;117:16732–8.
32. Standaert B, Rappuoli R. 3. How comprehensive can we be in the economic assessment of vaccines? *J Mark Access Health Policy*. 2017;5:1336044.
33. Wallinga J, Lipsitch M. How generation intervals shape the relationship between growth rates and reproductive numbers. *Proc R Soc B*. 2007;274: 599–604.
34. Lloyd-Smith JO, Funk S, McLean AR, et al. Nine challenges in modelling the emergence of novel pathogens. *Epidemics*. 2015;10:35–9.
35. Knox EG. Strategy for Rubella Vaccination. *Int J Epidemiol*. 1980;9:13–23.
36. Mclean AR, Anderson RM. Measles in developing countries. Part II. The predicted impact of mass vaccination. *Epidemiol Infect*. 1988;100: 419-42.
37. Colijn C, Jones N, Johnston IG, et al. Toward Precision Healthcare: Context and Mathematical Challenges. *Front Physiol*. 2017;8:136.
38. Chan P, Zuckenberg M. Advancing human potential and promoting equal opportunity. Chan Zuckenberg Initiative. 2017. <https://chanzuckenberg.com/newsroom/lessons-in-philanthropy-2017/> (Acceso verificado el 27 Sept 2023)

39. Collins FS, Varmus H. A New Initiative on Precision Medicine. *N Engl J Med*. 2015;372:793-5.
40. White House and Gates Foundation to Convene Precision Public Health Summit at UCSF | UC San Francisco. <https://www.ucsf.edu/news/2016/05/403091/white-house-andgates-foundation-convene-precision-public-health-summit-ucsf> (Acceso verificado el 27 Sept 2023)
41. White House Precision Medicine Initiative. Precision Medicine Initiative | The White House [Internet]. 2016. <https://obamawhitehouse.archives.gov/precision-medicine> (Acceso verificado el 27 Sept 2023)
42. Duffy DJ. Problems, challenges and promises: perspectives on precision medicine. *Brief Bioinform*. 2016;17:494-504.
43. Garber AM, Tunis SR. Does Comparative-Effectiveness Research Threaten Personalized Medicine? *N Engl J Med*. 2009;360:1925-7.
44. Palmisano A, Zhao Y, Li M-C, et al. OpenGeneMed: a portable, flexible and customizable informatics hub for the coordination of next-generation sequencing studies in support of precision medicine trials. *Brief Bioinform*. 2017; 18: 723-734.
45. Sboner A, Elemento O. A primer on precision medicine informatics. *Brief Bioinform*. 2016;17:145-53.
46. Servant N, Roméjon J, Gestraud P, et al. Bioinformatics for precision medicine in oncology: principles and application to the SHIVA clinical trial. *Front Genet*. 2014;5:152.
47. Castaneda C, Nalley K, Mannion C, et al. Clinical decision support systems for improving diagnostic accuracy and achieving precision medicine. *J Clin Bioinform*. 2015;5:4.
48. Cohn EG, Henderson GE, Appelbaum PS. Distributive justice, diversity, and inclusion in precision medicine: what will success look like? *Genet Med*. 2017;19:157-9.
49. Juengst ET, Flatt MA, Settersten RA. Personalized Genomic Medicine and the Rhetoric of Empowerment. *Hastings Center Report*. 2012;42:34-40.
50. Khoury MJ, Gwinn ML, Glasgow RE, et al. A Population Approach to Precision Medicine. *Am J Prev Med*. 2012;42:639-45.

51. Schork NJ. Personalized medicine: Time for one-person trials. *Nature*. 2015;520:609-11.
52. Khoury MJ, Lademarco MF, Riley WT. Precision Public Health for the Era of Precision Medicine. *Am J Prev Med*. 2016;50:398-401.
53. Salathé M, Bonhoeffer S. The effect of opinion clustering on disease outbreaks. *J R Soc Interface*. 2008;5:1505-8.
54. Kretzschmar M. Disease modeling for public health: added value, challenges, and institutional constraints. *J Public Health Policy*. 2020;41:39-51.
55. Grenfell BT. Unifying the Epidemiological and Evolutionary Dynamics of Pathogens. *Science*. 2004;303:327-32.
56. Neher RA, Bedford T. Real-Time Analysis and Visualization of Pathogen Sequence Data. *J Clin Microbiol*. 2018;56:e00480-18.
57. D'Souza R, Shah PS, Sander B. Clinical decision analysis in perinatology. *Acta Obstet Gynecol Scand*. 2018;97:491-9.
58. D'Souza R, Bonasia K, Shah PS, et al. Clinical decision analysis and model-based economic evaluation studies in perinatology: A systematic review. *Acta Obstet Gynecol Scand*. 2019;98:967-75.
59. Prosser LA, Grosse SD, Kemper AR, et al. Decision analysis, economic evaluation, and newborn screening: challenges and opportunities. *Genet Med*. 2012;14:703-12.
60. Grosse SD, Prosser LA, Asakawa K, et al. QALY weights for neurosensory impairments in pediatric economic evaluations: case studies and a critique. *Expert Rev Pharmacoecon Outcomes Res*. 2010;10:293-308.
61. Prosser LA, Hammitt JK, Keren R. Measuring Health Preferences for Use in Cost-Utility and Cost-Benefit Analyses of Interventions in Children: Theoretical and Methodological Considerations. *Pharmacoeconomics*. 2007;25: 713-26.
62. Gavan S, Bruce I, Payne K. Generating evidence to inform health technology assessment of treatments for SLE: a systematic review of decision analytic model-based economic evaluations. *Lupus Sci Med*. 2020;7: e000350.
63. Wildman J, Wildman JM. Combining Health and Outcomes Beyond Health in Complex Evaluations of Complex Interventions: Suggestions for Economic Evaluation. *Value in Health*. 2019;22:511-7

64. O'Mahony JF, Newall AT, van Rosmalen J. Dealing with Time in Health Economic Evaluation: Methodological Issues and Recommendations for Practice. *PharmacoEconomics*. 2015;33:1255-68.
65. Kim JJ, Brisson M, Edmunds WJ, et al. Modeling cervical cancer prevention in developed countries. *Vaccine*. 2008;26 Suppl 10:K76-86.
66. O'Mahony JF, van Rosmalen J, Zauber AG, et al. Multicohort models in cost-effectiveness analysis: why aggregating estimates over multiple cohorts can hide useful information. *Med Decis Making*. 2013;33:407-14.
67. O'Mahony JF, Newall AT, van Rosmalen J. Dealing with Time in Health Economic Evaluation: Methodological Issues and Recommendations for Practice. *PharmacoEconomics*. 2015;33:1255-68.
68. Graves J, Garbett S, Zhou Z, et al. Comparison of Decision Modeling Approaches for Health Technology and Policy Evaluation. *Med Decis Making* 2021; 41:453-64.
69. Barton P, Bryan S, Robinson S. Modelling in the economic evaluation of health care: selecting the appropriate approach. *J Health Serv Res Policy*. 2004;9:110-8.
70. Sato RC, Zouain DM, Sato RC, et al. Markov Models in health care. *Einstein (São Paulo)*. 2010;8:376-9.
71. van Rosmalen J, Toy M, O'Mahony JF. A mathematical approach for evaluating Markov models in continuous time without discrete-event simulation. *Med Decis Making*. 2013;33:767-79.
72. Augustovski F, Colantonio LD, Galante J, et al. Measuring the Benefits of Healthcare: DALYs and QALYs – Does the Choice of Measure Matter? A Case Study of Two Preventive Interventions. *Int J Health Policy Manag*. 2018;7:120.
73. Robberstad B. QALYs vs DALYS vs LYs gained: What are the differences, and what difference do they make for health care priority setting? *Norsk Epidemiologi*. 2005;15:183- 91
74. Eichler H-G, Kong SX, Gerth WC, et al. Use of cost-effectiveness analysis in health-care resource allocation decision-making: how are cost-effectiveness thresholds expected to emerge? *Value Health*. 2004;7:518-28

75. Sculpher M, Palmer S. After 20 Years of Using Economic Evaluation, Should NICE be Considered a Methods Innovator? *PharmacoEconomics*. 2020;38:247-57.
76. Oliva J, Brosa M, Espín J, et al. Cuestiones controvertidas en evaluación económica (I): perspectiva y costes de intervenciones sanitarias. *Rev Esp Salud Pub*. 2015;89:5-14.
77. Kellerborg K, Perry-Duxbury M, de Vries L, et al. Practical Guidance for Including Future Costs in Economic Evaluations in The Netherlands: Introducing and Applying PAID 3.0. *Value Health*. 2020;23:1453-61.
78. Soto J. Incorporación de los costes futuros en los análisis farmacoeconómicos: principios y práctica. *PharmacoEconomics Spanish Research Articles*. 2008;5:72.
79. O'Mahony JF, Kok IMCM de, Rosmalen J van, et al. Practical Implications of Differential Discounting in Cost-Effectiveness Analyses with Varying Numbers of Cohorts. *Value in Health*. 2011;14:438-42.
80. Attema AE, Brouwer WBF, Claxton K. Discounting in Economic Evaluations. *PharmacoEconomics*. 2018;36:745-58.
81. Jit M, Mibei W. Discounting in the evaluation of the cost-effectiveness of a vaccination programme: A critical review. *Vaccine*. 2015;33:3788-94.
82. Reza Davahli MR, Karwowski W, Taiar R. A System Dynamics Simulation Applied to Healthcare: A Systematic Review. *Int J Environ Res Public Health*. 2020;17: 5741.
83. Metcalf CJE, Lessler J. Opportunities and challenges in modeling emerging infectious diseases. *Science*. 2017;357:149-52.
84. Sculpher M, Claxton K, Akehurst R. It's Just Evaluation for Decision Making: Recent Developments in, and Challenges for, Cost-Effectiveness Research. In: Sculpher M. *Health Policy and Economics: Opportunitis and Challenges*: Maidenhead, Open University Press 2005.
85. Goldie SJ, Diaz M, Constenla D, et al. Mathematical models of cervical cancer prevention in Latin America and the Caribbean. *Vaccine*. 2008;26 Suppl 11:L59-72.

86. Goldie SJ, Diaz M, Kim SY, et al. Mathematical models of cervical cancer prevention in the Asia Pacific region. *Vaccine*. 2008;26 Suppl 12:M17-29.
87. Goldie SJ, O'Shea M, Campos NG, et al. Health and economic outcomes of HPV 16,18 vaccination in 72 GAVI-eligible countries. *Vaccine*. 2008;26:4080-93.
88. Briggs ADM, Wolstenholme J, Scarborough P. Estimating the cost-effectiveness of salt reformulation and increasing access to leisure centres in England, with PRIMETIME CE model validation using the AdViSHE tool. *BMC Health Serv Res*. 2019;19:489.
89. Haji Ali Afzali H, Gray J, Karnon J. Model Performance Evaluation (Validation and Calibration) in Model-based Studies of Therapeutic Interventions for Cardiovascular Diseases: A Review and Suggested Reporting Framework. *Appl Health Econ Health Policy*. 2013;11:85-93.
90. de Boer PT, Frederix GWJ, Feenstra TL, et al. Unremarked or Unperformed? Systematic Review on Reporting of Validation Efforts of Health Economic Decision Models in Seasonal Influenza and Early Breast Cancer. *Pharmacoeconomics*. 2016;34:83--45.
91. Eddy DM, Hollingworth W, Caro JJ, et al. ISPOR-SMDM Modeling Good Research Practices Task Force. Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-7. *Med Decis Making*. 2012;32:733-43.
92. Husereau D, Drummond M, Petrou S, et al. ISPOR Health Economic Evaluation Publication Guidelines-CHEERS Good Reporting Practices Task Force. Consolidated Health Economic Evaluation Reporting Standards (CHEERS)- explanation and elaboration: a report of the ISPOR Health Economic Evaluation Publication Guidelines Good Reporting Practices Task Force. *Value Health*. 2013;16:231-50.
93. McCabe C, Dixon S. Testing the validity of cost-effectiveness models. *Pharmacoeconomics*. 2000;17:501-13.
94. National Institute for Health and Care Excellence (NICE). Guide to the process of technology appraisal. Process and methods. 2018. <https://www.nice.org.uk/guidance/TA252/process-and-methods>.

- nice.org.uk/process/pmg19/resources/guide-to-the-processes-of-technology-appraisal-pdf- (Acceso verificado el 27 Sept 2023)
95. Pharmaceutical Benefits Advisory Committee. P Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (PBAC) Version 5.0. September 2016 Australian Government Department of Health. <https://pbac.pbs.gov.au/>. (Acceso verificado el 27 Sept 2023)
 96. Vemer P, Corro Ramos I, van Voorn GAK, et al. AdViSHE: A Validation-Assessment Tool of Health-Economic Models for Decision Makers and Model Users. *PharmacoEconomics*. 2016;34:349-61.
 97. Vemer P, Corro Ramos I, van Voorn GAK, et al. AdViSHE: A Validation-Assessment Tool of Health-Economic Models for Decision Makers and Model Users. *PharmacoEconomics*. 2016;34:349-61. Esta referencia esta repetida con la de arriba, por lo que hay que eliminarla y cambiar el numero de las siguientes referencias
 98. Rutter H, Savona N, Glonti K, et al. The need for a complex systems model of evidence for public health. *The Lancet*. 2017;390:2602-4.
 99. Banks J, Carson II JS, Nelson BL, Nicol DM. *Discrete-Event System Simulation*. Fifth edition. Pearson; 2010.
 100. Adamides ED, Karacapilidis N. Information technology support for the knowledge and social processes of innovation management. *Technovation*. 2006;26:50-9.
 101. Drummond M, Sculpher M. Common methodological flaws in economic evaluations. *Med Care*. 2005;43(Suppl 7):5-14.
 102. Griffiths EA, Hendrich JK, Stoddart SD, et al. Acceptance of health technology assessment submissions with incremental cost-effectiveness ratios above the cost-effectiveness threshold. *Clinicoecon Outcomes Res*. 2015;7:463-76.
 103. Neumann PJ, Cohen JT, Weinstein MC. Updating cost-effectiveness--the curious resilience of the \$50,000-per-QALY threshold. *N Engl J Med*. 2014;371:796-7.
 104. Vallejo-Torres L, García-Lorenzo B, Castilla I, et al. On the Estimation of the Cost-Effectiveness Threshold: Why, What, How? *Value in Health*. 2016;19:558-66.

105. Santos A, Guerra Junior A, Godman B, et al. Cost-effectiveness thresholds: methods for setting and examples from around the world. *Expert Rev Pharmacoecon Outcomes Res.* 2018;18:277-288
106. Claxton K, Martin S, Soares M, et al. Methods for the estimation of the National Institute for Health and Care Excellence cost-effectiveness threshold. *Health Technol Assess.* 2015;19:1-503, v-vi.
107. Bertram MY, Lauer JA, Stenberg K, et al. Methods for the Economic Evaluation of Health Care Interventions for Priority Setting in the Health System: An Update From WHO CHOICE. *Int J Health Policy Manag.* 2021; 10: 673-7.
108. Siebert U. When should decision-analytic modeling be used in the economic evaluation of health care? *Eur J Health Econ.* 2003;4:143-50.
109. Tengs TO, Osgood ND, Chen LL. The Cost-Effectiveness of Intensive National School-Based Anti-Tobacco Education: Results from the Tobacco Policy Model. *Prev Med.* 2001;33:558-70.
110. Tengs TO, Osgood ND, Lin TH. Public Health Impact of Changes in Smoking Behavior: Results From the Tobacco Policy Model. *Med Care.* 2001;39:1131-41.
111. World Health Organization (WHO). Essential public health function Suggested citation. ns, health systems and health security: developing conceptual clarity and a WHO roadmap for action. Geneva: World Health Organization; 2018.
112. Harper PR, Pitt MA. On the challenges of healthcare modelling and a proposed project life cycle for successful implementation. *J Oper Res Soc.* 2004;55:657-61.
113. Kallioniemi O, Wessels L, Valencia A. On the organization of bioinformatics core services in biology-based research institutes. *Bioinformatics.* 2011;27:1345.
114. Caro JJ, Briggs AH, Siebert U, et al. Modeling Good Research Practices—Overview: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force—1. *Med Decis Making.* 2012;32:667-77.
115. Star L, Moghadas SM. The Role of Mathematical Modelling in Public Health Planning and Decision Making. National Collaborating Centre for Infectious Diseases. Purple Paper(22), Issue 22.2010.



17

EL PAPEL DE LOS MODELOS EN LA TOMA DE DECISIONES

Jorge Mestre Ferrándiz

Consultor e Investigador en Economía de la Salud

Profesor Asociado, Universidad Carlos III

En este último capítulo, ofrecemos unas reflexiones finales, a colación con el papel de los modelos analíticos de decisiones en la toma de decisiones sobre evaluación, precio y financiación de medicamentos y tecnologías sanitarias. Para ello, partimos de las siguientes tres cuestiones:

¿Por qué, y para qué, utilizamos modelos analíticos de decisión dentro de la evaluación económica?

¿Qué recomendaciones existen sobre buenas prácticas cuando se vaya a diseñar un modelo analítico de decisión *de novo*, o cuando se quiera valorar uno ya realizado, con el fin de evaluar su calidad, validez y transparencia?

¿Cuál puede ser el futuro de los modelos analíticos de decisión en la toma de decisiones en España?

Pero antes de entrar en detalle, es importante definir dos términos. En primer lugar, ‘modelo’. Así, la RAE incluye 11 definiciones para esta palabra^a, de las cuales la cuarta sería la más apropiada para nuestro caso en cuestión:

“4. m. Esquema teórico, generalmente en forma matemática, de un sistema o de una realidad compleja, como la evolución económica de un país, que se elabora para facilitar su comprensión y el estudio de su comportamiento.”

Por supuesto, existen diferentes tipos de modelos, y de modelos económicos, como bien queda reflejado en este libro. Pero el concepto y lo que persigue, se podrían considerar similar, y en línea con lo que define la RAE —más allá de si estamos evaluando la evolución de una variable macroeconómica, o como en nuestro caso concretamente, evaluar el posible comportamiento de un tratamiento en la vida real.

En segundo lugar, es importante también comentar a lo que nos referimos con modelos analíticos de decisión en nuestro contexto. Para

^a <https://dle.rae.es/modelo> (última visita: 27/01/2023)

ello, partimos de la definición de Soto J (1): “una aproximación sistemática a la evolución más probable de una enfermedad tras la administración de diferentes alternativas terapéuticas para su tratamiento, proyectando los resultados en salud que se van a obtener con cada una de las opciones administradas y los costes asociados en su consecución.”

A partir de aquí, pasamos a comentar las tres preguntas.



¿Por qué, y para qué, utilizamos modelos analíticos de decisión dentro de la evaluación económica?

Como ha quedado claro a lo largo del libro, un elemento clave en la toma de decisiones que nos compete es la incertidumbre; incertidumbre sobre varios factores, entre otras cosas por el vacío entre la evidencia generada sobre eficacia y seguridad (y a veces sobre calidad de vida y coste-efectividad) durante la fase de investigación clínica para la autorización de comercialización por parte del regulador (EMA/FDA), y sus efectos cuando se empiezan a utilizar acorde a las condiciones de uso de la práctica médica rutinaria. Los modelos precisamente nos ayudan a reducir ese vacío de información, permitiendo realizar unas simulaciones predictivas sobre esos resultados posibles en el mundo real.

Queremos remarcar tres aspectos. Primero, es necesario trabajar de forma multidisciplinar, y en colaboración entre clínicos, farmacéuticos, farmacólogos, estadísticos, economistas de la salud, epidemiólogos, entre otros, en el diseño y elaboración del modelo. Segundo, es preferible representar en el modelo la evolución más probable de la enfermedad diana de la manera más sencilla y comprensible que se pueda, pero siempre lo más realista posible (1).

Estos dos aspectos redundarán en un aumento en la calidad y validez de sus resultados, así como su transparencia, lo que permitirá que el modelo sea fiable y preciso (1).

Y, en tercer lugar, este vacío de evidencia sobre el impacto de los nuevos tratamientos, tanto en la práctica clínica como a largo plazo, puede llegar a ser más importante para los nuevos tratamientos/medicamentos/tecnologías sanitarias (en su definición más amplia) que ofrecen posibilidad de curar enfermedades —y donde podemos incluir a las terapias génicas o avanzadas (*advanced therapy medicinal products*, ATMPs por sus siglas en inglés). Esto es debido a que se observan evidencias con un número limitado de pacientes en ensayos clínicos, sin comparador (quizás porque ni siquiera existe o el comparador relevante es ‘cuidado paliativo’), o sin resultados a largo plazo y/o sobre variables finales de efectividad clínica, entre otras. La factibilidad de realizar el ensayo clínico aleatorio contra un comparador activo (que como sabemos sería la principal fuente de la evidencia sobre eficacia) en muchos de estos casos es (casi) imposible, lo que reforzaría la necesidad de un ‘buen’ modelo (y en línea con lo comentado en la siguiente sección).

Por lo tanto, y en este contexto, los modelos analíticos de decisión se utilizan para sintetizar de manera explícita y sistemática la mejor evidencia disponible sobre la historia natural y la evolución de una enfermedad, así como la efectividad, riesgo, costes y coste-efectividad de diferentes opciones terapéuticas existentes para su tratamiento, en una población de interés definida (1).

Para acabar este primer apartado, ya hemos visto la diversidad de modelos existentes (en función de por ejemplo, si interviene el tiempo, si cambia el estado del sistema continuamente o no, si los valores de entrada son aleatorios o no, y como no, si son modelos basado en arboles de decisión o modelos de transición entre estados). Por lo tanto, la elección de un tipo u otro de modelo a la hora de efectuar una evaluación económica depende, básicamente, de la enfermedad que se esté evaluando (aguda, subaguda o crónica), de las alternativas terapéuticas en estudio, del tipo de resultados demandados por los agentes decisores (corto o medio-largo plazo), de la información de calidad disponible, así como de los conocimientos y experiencia del

analista que va a diseñar el modelo (1). Así, ahora pasamos a comentar las recomendaciones sobre buenas prácticas para estos modelos en más detalle.



Recomendaciones sobre buenas prácticas para diseñar nuevos modelos analíticos de decisión o valorar modelos existentes

Hemos identificado dos recomendaciones de buenas prácticas sobre los modelos analíticos de decisión que merecen mención aquí. Ambas están basadas en una revisión sistemática de la literatura y donde la más reciente (2) utiliza como base la más antigua (3), con 26 y 33 artículos incluidos finalmente respectivamente). Hay muchas similitudes entre las dos propuestas en el detalle, si bien Philips y col. (3) estructura el marco acorde a tres aspectos (de estructura, datos y consistencia), y Peñalosa Ramos y col. (2) sobre cinco dimensiones (problema, modelo, síntesis de la evidencia, análisis de la incertidumbre, y transparencia y validación).

Los tres aspectos de Philips y col. (3) son:

- I. De Estructura, en relación con el alcance y estructura matemática del modelo, incluyendo hasta nueve dimensiones: enunciado del problema de decisión, alcance o perspectiva, racional para la estructura, supuestos estructurales, comparadores, tipo de modelo, horizonte temporal, estados de salud, y duración de ciclo.
- II. Datos, que incluyen los métodos para la identificación de datos, y cómo se debe abordar la incertidumbre, con cuatro dimensiones específicas: identificación de datos, análisis de datos pre-modelo, incorporación al modelo, valoración de incertidumbre.
- III. Consistencia: se refiere a la calidad general del modelo, distinguiendo entre consistencia interna y externa.

La estructura de Peñaloza Ramos y col. (2) de buenas prácticas para los modelos analíticos de decisión gira en torno a las siguientes cinco dimensiones (con sus respectivos componentes):

1. Concepción del problema: decisión u objetivo, perspectiva, población diana, resultados en salud, comparadores y horizonte temporal.
2. Concepción del modelo: el tipo de modelo y su estructura.
3. Síntesis de la evidencia: fuentes, las utilidades, duración de ciclo y corrección de medios ciclos, costes/recursos, heterogeneidad en el paciente, y precisión de los parámetros.
4. Análisis de la incertidumbre: alrededor de cómo se incluye y se reporta, estimación e incertidumbre de los parámetros, incertidumbre estructural y otros análisis de incertidumbre.
5. Transparencia y validación del modelo, incluido validez interna y externa.

Se puede observar un gran solapamiento entre los distintos elementos de buenas prácticas, ya sea agrupados en tres o cinco dimensiones.

Por otra parte, Soto Álvarez (1) enumera 13 requisitos mínimos para que los resultados de un modelo sean creíbles, válidos y relevantes para los profesionales sanitarios y otros agentes decisores —y que por supuesto van muy en línea con las dos recomendaciones ya comentadas. Además, Soto Álvarez (1) ofrece 15 recomendaciones a seguir cuando se vaya a diseñar un modelo analítico de decisión *de novo*, o cuando se quiera valorar uno ya realizado, con el fin de evaluar la calidad, validez y transparencia. En concreto, se enumeran la estructura, datos incorporados en el modelo (especialmente los datos de resultados en salud), consistencia del modelo, su validación, su calibración y la reproducibilidad del mismo como aspectos a fijarse especialmente, también en línea con las dos recomendaciones.

Soto Álvarez (1) incluye dos series de recomendaciones adicionales, sobre el formato y los datos que debería aportar el informe donde

se comunican los resultados de un modelo analítico de decisión, y el algoritmo a seguir para elegir el tipo de modelo a emplear. Así, el informe tendría los siguientes epígrafes: Introducción del Informe; Opciones terapéuticas que se comparan en el modelo; Hipótesis y objetivo del modelo; Diseño y métodos empleados en el modelo; Resultados obtenidos; Discusión de los Resultados; Conflicto de interés; Apéndices y referencias. Como argumenta Soto Álvarez (1), una presentación adecuada del modelo es clave de cara a incrementar su transparencia, aspecto crucial para que los resultados obtenidos se utilicen en la toma de decisión en política sanitaria —por lo tanto, el formato a seguir y la información a incluir juegan un rol clave.

La importancia del uso de buenas prácticas en la modelización es más importante si tenemos en cuenta, como argumenta Caro (4), que su uso para pronosticar los resultados de la enfermedad y los efectos de las intervenciones han tenido un importante auge en parte gracias al uso de los modelos durante la COVID-19. Lo que no queda claro es que los modelos hayan sido validados y sometidos a una revisión crítica sistemática, por lo que una de las revistas internacionales más importantes en el ámbito de la economía de la salud ha anunciado en enero 2023 que lanza una serie de artículos sobre revisiones críticas de los enfoques de modelización en áreas de enfermedades específicas. Lo más relevante de esta iniciativa para nuestro capítulo es que en la introducción a la serie, se enumeran los aspectos que las revisiones deberían cubrir: alcance, métodos, conceptualización, estructura, incertidumbre, validación, y transparencia (4). Todos estos aspectos quedan recogidos dentro de las guías de buenas prácticas ya mencionadas.

Y para concluir este apartado sobre buenas prácticas, es importante remarcar la necesidad de un buen gobierno, más allá de los modelos analíticos de decisión, de los sistemas sanitarios, que como bien argumentan Lobo y col. (5), “debe apoyarse en valores fundamentales como son los derechos humanos y el cumplimiento de las normas, así como

en principios rectores de nuestro sistema como son la universalidad, la equidad y la solidaridad” (pp. 13). Más concretamente, estos autores ofrecen cinco principios de buen gobierno, en base a: transparencia, rendición de cuentas, participación, integridad o completitud de procedimientos y reglas, y capacidad operativa.



Reflexiones finales: el futuro de los modelos analíticos de decisión en la toma de decisiones en España

Y para finalizar este último capítulo, para España, ¿qué podemos decir sobre el futuro de los modelos analíticos de decisión? Estos modelos son una parte esencial de la evaluación económica (o de eficiencia); por lo tanto, su futuro va a depender del uso que se haga de la evaluación de la eficiencia durante el proceso de evaluación, financiación y precio. Recientemente, se han publicado y presentado varios informes y análisis muy interesantes y relevantes, sobre el uso de la eficiencia en los procesos de evaluación en España (véase, por ejemplo, Lobo y col. (5), Vallejo-Torres y Oliva (6) y Zozaya y col. (7) Vida y col. (8). Lo importante a remarcar en este capítulo es que, a nuestro parecer, hemos avanzado durante los últimos años, con buenas iniciativas e intenciones; pero aún hay sombras, con demasiada heterogeneidad (en las razones para realizar o no una evaluación económica propia, o en los criterios para concluir si un tratamiento es o no coste efectivo, por ejemplo). Y esto, en parte, implica un uso de la evaluación de la eficiencia bastante limitado. Esto queda plasmado en el análisis de Vallejo-Torres y Oliva (6) sobre el uso de la evaluación económica en los Informes de Posicionamiento (IPTs): a fecha de julio 2022, de 24 IPTs que se les aplica el nuevo procedimiento y que estarían sujetos a evaluación económica, se han publicado 14 en 2021 y 2022. Pero de estos 14, solo cinco incluyen una evaluación económica (análisis de coste utilidad). De los otros nueve, la existencia de

evaluación previas y la ausencia de estudios de efectividad comparada es la razón para no realizar un análisis de coste-efectividad en dos y tres IPTs respectivamente; otros tres realizan un análisis de minimización de costes (se 'asume' o 'concluye' ausencia de valor terapéutico añadido), y el restante concluye que el tratamiento evaluado es dominante. Por poner en contexto, Vallejo-Torres y Oliva (6) identifican más de 70 IPTs con fecha de publicación solo en 2022. Por lo tanto, se puede argumentar que el peso de la evaluación de la eficiencia aun juega un papel muy limitado en los IPTs. Pero preocupa quizá más si cabe la dificultad de evaluar el impacto de la evaluación económica tanto en el posicionamiento como la propuesta de la CIPM al respecto.

Estos trabajos identifican una serie de barreras en el uso de la evaluación económica en España, así como una batería de sugerencias y mejoras que, en nuestra opinión, serían muy adecuadas y bienvenidas. Una sugerencia importante a colación con los modelos analíticos de decisión es la necesidad de unas guías metodológicas sobre la evaluación económica (similares a las que ya existen en muchas agencias y países, incluido algunas en España). Y si bien la evaluación económica va más allá del propio modelo analítico de decisión (ya que este es una herramienta más en la toma de decisiones), sería conveniente generar unas guías de buenas prácticas para estos modelos. Además, como ya se ha comentado, existen varias estructuras en la literatura que pueden servir de base. Por lo tanto, no es de extrañar que una de las recientes recomendaciones del Comité Asesor de Prestación Farmacéutica (CAPF) sobre el procedimiento de evaluación de medicamentos incluya el desarrollo de guías metodológicas al respecto - y mientras tanto, que se utilicen guías ya existentes del Ministerio de Sanidad, del CatSalut y de la Sociedad Española de Farmacia Hospitalaria (9).

Para concluir, los modelos analíticos de decisión son un elemento central dentro del proceso de evaluación de tecnologías sanitarias. En un contexto donde existe una incertidumbre importante, el valor añadido de estos modelos es la posibilidad de ofrecer un enfoque explícito

para sintetizar la evidencia disponible con respecto a la efectividad y los costes de las estrategias alternativas de atención médica. Por lo tanto, la modelización es clave para ayudar o guiar las decisiones durante el proceso de financiación y precio de los medicamentos y tecnologías sanitarias —si bien sabemos que detrás de muchas de estas decisiones hay una evaluación y consideración de aspectos que van más allá de los números que finalmente genere ‘el modelo’.



Bibliografía

1. Soto Álvarez J. Diseño y realización de evaluaciones económicas a través de modelos analíticos de decisión. En: *Evaluación económica de medicamentos y tecnologías sanitarias*. 1ª Ed . Springer Healthcare Iberica, 2012. p. 171-216
2. Peñaloza Ramos, MC, Barton, P, Jowett S, et al. A Systematic Review of Research Guidelines in Decision-Analytic Modeling. *Value Health*. 2015; 18:512-29.
3. Philips Z, Bojke L, Sculpher M, et al. Good practice guidelines for decision analytic modelling in health technology assessment: a review and consolidation of quality assessment. *PharmacoEconomics*. 2006;24:355-71.
4. Caro JJ. Best Practices: A Collection of Systematic Critical Reviews of Modeling Approaches in Specific Disease Areas. *PharmacoEconomics*. 2023;41:119-21.
5. Lobo F, Oliva J, Vida J. La evaluación de la eficiencia de las tecnologías sanitarias en España: origen, evolución y estado actual. 2022. FUNCAS. https://www.funcas.es/documentos_trabajo/la-evaluacion-de-la-eficiencia-de-las-tecnologias-sanitarias-en-espana-origenevolucion-y-estado-actual/ (Acceso verificado el 10 Sept 2023)
6. Vallejo-Torres L, Oliva J. 2022. Economía y políticas de salud, “De la investigación a la acción»: Medicamentos y Economía: La evaluación de la eficiencia en los Informes de Posicionamiento Terapéuticos. Jornada Virtual. 2022. <https://www.funcas.es/eventos/medicamentos-y-economia-la-evaluacion-de-la-eficiencia-en-los-informes-de-posicionamiento-terapeuticos/> (Acceso verificado el 27 Sept 2023)
7. Zozaya N, Villaseca J, Abdalla F, et al. El proceso de evaluación y financiación de los medicamentos en España: ¿dónde estamos y hacia dónde vamos? Fundación Weber. 2022. <https://weber.org.es/publicacion/elproceso-de-evaluacion-y-financiacion-de-los-medicamentos-en-espanadonde-estamos-y-hacia-donde-vamos/> (Acceso verificado el 27 Sept 2023)

8. Vida J, Lobo, F Oliva, J. La (des)organización de la evaluación de la eficiencia de medicamentos y otras tecnologías sanitarias en España: diagnóstico. FUNCAS. https://www.funcas.es/wp-content/uploads/2023/03/La-desorganizacion-de-la-evaluacio%CC%81n_1.pdf (Acceso verificado el 27 Sept 2023)
9. CAPF, 2023. Propuesta de reforma del Procedimiento de evaluación de medicamentos en España. https://www.sanidad.gob.es/profesionales/farmacia/pdf/CAPF_Reforma_Expres_Proced_EE ((Acceso verificado el 27 Sept 2023)



f | *fundación* **porib**