

Evaluación de la relación
beneficio-riesgo de los medicamentos:
evidencia, cuantificación e incertidumbre

Hacia la objetividad y la transparencia

ANDREW MAGUIRE
FRANCISCO J. DE ABAJO

SUMARIO

1. Introducción	350
2. Relación beneficio-riesgo: concepto y componentes	350
3. Métodos cuantitativos	352
3.1. Relación entre medidas de impacto de beneficio y de daño	353
3.2. Relación beneficio-riesgo incremental	354
3.3. Análisis de decisión de criterios múltiples	356
4. Una nueva aproximación: la simulación de cohortes de pacientes	358
4.1. Una aproximación sencilla: estimación del número esperado de casos	359
4.2. Una aproximación más compleja: simulación del tiempo hasta el acontecimiento	363
5. Conclusiones	366
Referencias	368

LO QUE VD. APRENDERÁ EN ESTE CAPÍTULO

1. La evaluación de la relación beneficio-riesgo de los medicamentos es una labor crucial de las agencias de regulación de medicamentos.
2. Las opiniones de expertos han sido hasta ahora la manera más factible de evaluar todas las pruebas y considerar si los efectos beneficiosos de los fármacos justificarían los posibles efectos adversos de los mismos. Este proceso, aunque se presume riguroso e imparcial, tiene un alto componente subjetivo.
3. La cuantificación del balance beneficio-riesgo tiene un solo objetivo: maximizar los beneficios de las opciones terapéuticas y proteger a las poblaciones de riesgos innecesarios.
4. Existen hoy en día múltiples iniciativas para dotar al proceso de evaluación de la relación beneficio-riesgo de una mayor objetividad y transparencia. Y ello pasa por incorporar la cuantificación de los beneficios y de los riesgos, y utilizar medidas que los integren (efectos netos). La tarea es, no obstante, ardua porque en muchas ocasiones se trata de comparar acontecimientos de índole diferente.
5. Un punto clave en cualquier método es cómo se escoge, y cómo se justifica, el conjunto de acontecimientos que se incluirán en la evaluación beneficio-riesgo.
6. Se revisan diversos métodos que integran beneficios y riesgos en medidas cuantitativas. Entre ellos, el análisis de decisión de criterios múltiples que es el propuesto por la Agencia Europea de Medicamentos.
7. Se propone una nueva forma de realizar esta cuantificación consistente en simular, con pacientes reales procedentes de bases de datos clínicos automatizadas, los diferentes acontecimientos, beneficiosos y perjudiciales, y evaluar en el conjunto de dicha población el efecto neto. Se explicará a través de varios ejemplos.

1. INTRODUCCIÓN

Afirmaba James Black que los medicamentos son «venenos útiles»,¹ modernizando de alguna manera la famosa sentencia de Paracelso («solo la dosis hace el veneno»). Con ello trataba de hacer entender al gran público, de un modo sencillo, que no existe el medicamento inocuo, que todos, si son activos, tienen la potencialidad de producir daño. En consecuencia, se impone ponderar el conjunto de los beneficios potenciales de los medicamentos con el de sus posibles riesgos, actividad que conocemos como evaluación de la relación beneficio-riesgo.²

Desde un punto de vista estrictamente ético, solo los medicamentos con una relación beneficio-riesgo favorable deberían estar disponibles para su uso. La evaluación de dicha relación es la labor fundamental de las agencias de regulación de medicamentos antes de decidir la autorización o no de la comercialización de un medicamento, pero continúa a lo largo de toda la vida del mismo, a medida que se genera nuevo conocimiento que pueda modificar alguno de los componentes del beneficio o del riesgo.²

Esta evaluación ha consistido básicamente en un juicio de valor que los expertos hacen tomando como base toda la información disponible. Pero, aunque se presupone rigor, imparcialidad y prudencia en los expertos, es bien sabida la falibilidad de las decisiones humanas y los sesgos de sus valoraciones.³ Por eso, en los últimos años, se están produciendo iniciativas para tratar de cuantificar en la medida de lo posible la relación

beneficio-riesgo, incorporando la incertidumbre,⁴ y haciendo, de este modo, que el proceso de toma de decisiones sea más científico, más objetivo y más transparente. Estas iniciativas deben mucho al desarrollo y convergencia de la farmacoepidemiología⁵ y la farmacoeconomía.⁶

En este capítulo repasaremos en primer lugar el concepto de la evaluación del balance beneficio-riesgo y sus distintos componentes, para después adentrarnos en algunas de las aproximaciones metodológicas más importantes que se han aplicado, y finalizar con una propuesta novedosa. Estos modelos pueden llegar a tener cierta complejidad. No obstante, se ha procurado no usar formulas para centrarse más en la descripción del proceso y de los conceptos que forman los modelos.

2. RELACIÓN BENEFICIO-RIESGO: CONCEPTO Y COMPONENTES

En términos generales, la relación beneficio-riesgo es el balance entre los riesgos y beneficios de una terapia, relativos a otras opciones terapéuticas, incluyendo la opción de no utilizar ningún tratamiento. Este concepto se ha usado en varios campos en los cuales se evalúa el riesgo.⁷ El riesgo es la probabilidad de que ocurra un efecto nocivo de un fármaco, en tanto que el daño es la manifestación de dicho riesgo. Asimismo el beneficio se plasma operativamente en la efectividad («eficacia del fármaco en las condiciones de uso reales»).

Los riesgos y los beneficios tienen unas cualidades que dan lugar a varios tipos de balance. Una cualidad es la gra-

vedad del acontecimiento, y la otra la forma cómo se mide. Por ejemplo, los bisfosfonatos tienen como beneficio la reducción de fracturas, pero pueden dañar el esófago y causar úlceras.⁸ En este ejemplo, tanto el beneficio como los riesgos son variables discretas dicotómicas (0: no acontecimiento; 1: acontecimiento), y pueden ser contados como número de acontecimientos. Sin embargo, en el caso de los antiinflamatorios no esteroideos (AINE), se puede contar el riesgo en número de acontecimientos (hemorragia gastrointestinal y efectos cardiovasculares) pero el beneficio es una variable continua (mejora del dolor y aumento de la movilidad) que no admite el mismo tipo de recuento. Los modelos de simulación que se describirán más adelante se aplican mejor cuando los efectos (positivos o negativos) se pueden contar como número de acontecimientos. Teniendo en cuenta estas dos cualidades, pueden describirse cuatro escenarios para la evaluación de riesgos y beneficios:

1. *Riesgo*: acontecimientos adversos graves. *Beneficio*: reducción de un acontecimiento grave

Como ejemplo se puede poner a los fármacos para reducir la glucemia en personas con diabetes. En estos casos tanto el beneficio como los riesgos asociados con las terapias pueden considerarse simultáneamente; algunos fármacos (v.gr., rosiglitazona) han sido asociados con un aumento de riesgo cardiovascular, mientras que uno de los objetivos de estos fármacos, a través del control de la glucemia, es la reducción del riesgo cardiovascular.⁹

2. *Riesgo*: acontecimientos adversos graves. *Beneficio*: mejora de la calidad de vida

Un ejemplo puede ser el tratamiento del dolor y la inflamación con AINE. Estos fármacos tienen dos riesgos relevantes: el cardiovascular y el de hemorragia digestiva alta. Estos acontecimientos pueden poner en peligro la vida del paciente y, por lo tanto, la letalidad del acontecimiento debido al fármaco tiene que ser incorporada en el modelo.¹⁰ La parte positiva del balance, el beneficio, no afecta a la mortalidad o supervivencia de los pacientes. Para poder integrar los beneficios con los riesgos se tendría que crear un acontecimiento basado en el cambio de la calidad de vida, como una dicotomía basada en el porcentaje de mejora del dolor. En este escenario, no obstante, podría estar justificado comparar las opciones terapéuticas solo en términos del riesgo que conlleva cada opción si las pruebas indican que no hay diferencias en la efectividad de los fármacos, o bien si la preocupación principal es la seguridad. Este tipo de aplicación de los métodos se puede denominar modelo de «daños comparativos».

3. *Riesgo*: acontecimientos adversos leves. *Beneficio*: reducción de un acontecimiento grave

Es poco probable que sea necesario aplicar métodos como los modelos de cuantificación de beneficios y riesgos si no existen efectos adversos graves asociados a las opciones terapéuticas. La prueba principal se basaría en la eficacia determinada por los ensayos clínicos. No obstante, sigue habiendo debate so-

bre la traducción de la eficacia a la efectividad.

4. *Riesgo*: acontecimientos adversos leves. *Beneficio*: mejora de calidad de vida

Por los motivos explicados en el escenario anterior, es poco probable que estos métodos sean necesarios para fármacos cuyos efectos, positivos y negativos, no son graves.

Hasta hace poco, las agencias reguladoras empleaban evaluaciones cualitativas de los beneficios y riesgos. Esto se debía a la falta de una herramienta y un marco que permitieran tener en cuenta conjuntos de acontecimientos adversos y beneficiosos.¹¹ Las opiniones de expertos en el área terapéutica era la manera más factible de evaluar todas las pruebas y considerar si los efectos beneficiosos de los fármacos justificarían los posibles efectos adversos de los mismos.

El Consejo de Organizaciones Internacionales de Ciencias Médicas, (CIOMS) afirmaba en 1998 que «un aspecto frustrante de la evaluación beneficio-riesgo es que no hay ningún algoritmo o medida que pueda proporcionar comparaciones cuantitativas directas de las opciones terapéuticas, lo cual podría ayudar a la toma de decisiones».¹² Recientemente, la Agencia Europea de Medicamentos (EMA) se ha involucrado en una revisión de métodos y ha llegado a recomendar la aplicación de una de estas opciones.¹³ Asimismo, ha habido un cambio en la legislación y la obligación de las compañías farmacéuticas de presentar evaluaciones de beneficio-riesgo se ha hecho más exigente.¹⁴

3. MÉTODOS CUANTITATIVOS

Existe una gama de métodos con grados diferentes de complejidad, pero todas las aproximaciones, complejas y sencillas, siguen el mismo principio: el de estimar efectos netos.

Si bien existen claras ventajas en usar el método más sencillo disponible, como la facilidad en la comunicación de resultados, la sencillez conlleva suposiciones que pueden hipotecar la credibilidad del mismo. Por otra parte, las medidas epidemiológicas tradicionales pueden resultar insuficientes, dado que no tienen en cuenta algunos conceptos necesarios para la evaluación de múltiples riesgos y beneficios.

El concepto epidemiológico de la «fracción atribuible» o «riesgo atribuible» es un ejemplo sencillo de una medida de efecto neto.¹⁵ Este concepto resta del riesgo, o del número de casos de acontecimientos adversos entre las personas expuestas al medicamento de interés, la parte esperable entre las personas no expuestas. Como ejemplo tenemos el conocido caso del antiinflamatorio rofecoxib, retirado del mercado mundial en septiembre de 2004 por la compañía farmacéutica.¹⁶ Para calcular el impacto neto sobre el riesgo cardiovascular, la agencia de regulación de medicamentos norteamericana (FDA) aplicó la prevalencia del uso de rofecoxib y el riesgo relativo de enfermedad coronaria grave asociado al uso del mismo, concluyendo: «probablemente ha ocurrido un exceso de casos de enfermedad coronaria grave de entre 88.000 y 140.000 en EEUU durante el tiempo que rofecoxib ha permanecido en el

mercado y [si se asume una tasa de letalidad de un 44%] muchos de los casos atribuibles al rofecoxib produjeron la muerte del paciente».¹⁶

Conviene no perder de vista el alcance del problema: según el citado estudio, en un solo país se estima que entre 40.000 y 70.000 personas perdieron la vida debido a un fármaco, y, en parte, podría haberse evitado si se hubieran utilizado otras alternativas terapéuticas más seguras. Obviamente, queda la duda de cuándo se tuvo la evidencia clara de que los riesgos, o daños, eran mayores que los beneficios. Finalmente, otro elemento interesante de este caso es que el riesgo podría haber estado ligado a la duración de la exposición, por tanto, si la duración de la exposición a rofecoxib hubiese sido más corta, el daño causado habría sido menor. Se volverá a examinar la importancia de esta última faceta del riesgo en la próxima sección.

3.1. Relación entre medidas de impacto de beneficio y de daño

Una forma sencilla de calcular los efectos netos es a través de la estimación de las medidas de impacto NNT y NNH.¹⁷ El NNT, del inglés *Number Needed to Treat*, es una medida de beneficio e indica el número de pacientes que tendrían que ser tratados para conseguir la respuesta positiva al tratamiento en un paciente. Por tanto, cuanto más bajo el NNT, menor sería el número de pacientes que se necesitaría para que uno de ellos presentara el efecto deseado y, en consecuencia, más eficaz resultaría el fármaco. El NNH o *Number Needed to Harm*, representaría el número de per-

sonas que tendrían que ser expuestas a un fármaco para que aparezca un acontecimiento adverso en un paciente debido a dicho fármaco. Por tanto, cuanto más alto es el NNH más seguro es el fármaco. En consecuencia, si ponemos ambos conceptos en relación obtendríamos una cuantificación razonable del balance beneficio-riesgo: para que una opción terapéutica tenga una relación beneficio-riesgo favorable es preciso que el número de personas que haya que tratar para conseguir un acontecimiento beneficioso adicional sea menor al número de personas expuestas para que ocurra un acontecimiento dañino adicional. Esto se puede describir como una resta (NNH-NNT), o bien, como un cociente (NNH/NNT). En ambos casos, cuanto más alto sea el valor, más favorable será la relación beneficio-riesgo.

Aunque este método puede cuantificar la relación de beneficios y riesgos y es fácilmente comprensible, queda limitado por tres factores: 1) compara solo un acontecimiento de ambos tipos; 2) no representa la incertidumbre que hay en las pruebas; y 3) no pondera la importancia relativa del beneficio respecto al riesgo, es decir, ¿cómo se pueden poner en la misma escala acontecimientos de índole tan diferente como un infarto (expresado por el NNH) y un aumento en una escala de calidad de vida (expresado por el NNT)? Para resolver esta limitación se ha propuesto aplicar «utilidades», tanto al riesgo como al beneficio. Las «utilidades» son valores subjetivos y representarían la calidad de vida asociada a un estado de enfermedad, oscilando entre 0 (peor estado posible) y 1 (salud perfecta).¹⁸ No obstante, esta aproximación ha

sido calificada por la EMA como un método demasiado simple para ser útil en la evaluación del balance beneficio-riesgo de los medicamentos.¹³

3.2. *Relación beneficio-riesgo incremental*

Otro método que se ha usado proviene del trabajo realizado por economistas y, en particular, de aquellos que se han dedicado a las evaluaciones de coste-efectividad.^{19,20} El método se conoce como Relación Beneficio-Riesgo Incremental (IRBR) y describe el balance beneficio-riesgo utilizando la probabilidad de cada tipo de acontecimiento. Por lo tanto se adapta mejor cuando la variable de resultado es un acontecimiento, en lugar de un cambio cuantificado en una escala progresiva, como es el caso de la escalas de calidad de vida. Este método es también ideal cuando hay que comparar fármacos en términos de un solo acontecimiento dañino y un solo acontecimiento beneficioso. No puede aplicarse con facilidad cuando hay acontecimientos múltiples. No obstante, el punto fuerte de estos métodos es la clara comunicación de la incertidumbre a través del uso de la probabilidad de concluir que el balance es positivo. Este grado de información, que cuantifica dudas usando probabilidades, queda dentro del marco bayesiano.¹⁹ La aproximación toma como punto de partida la diferencia de probabilidades del acontecimiento beneficioso entre el tratamiento de interés y el comparador; al seguir las reglas de la probabilidad la diferencia entre probabilidades variará entre -1,0 y 1,0. Asimismo, se aplica a la diferencia de probabilidades de un

acontecimiento adverso. Por lo tanto, la evidencia queda reducida a unas coordenadas que pueden dibujarse sobre un plano, con el eje horizontal representando la diferencia de las probabilidades de beneficio, y el eje vertical la diferencia de las probabilidades de riesgo (figura 1). El plano, por tanto, queda dividido en cuatro cuadrantes que se pueden denominar, en el sentido de las agujas del reloj, como nordeste, sureste, sudoeste y noroeste. Si el estudio, o ensayo clínico, apunta a que el fármaco de interés es más seguro y más eficaz que el comparador entonces las coordenadas estarán en el cuadrante sureste. Si hiciéramos múltiples estudios tendríamos coordenadas que variarían en su posición, lo cual crearía una nube de puntos en el plano. Si la mayoría caen en el cuadrante sureste entonces el conjunto de pruebas sigue apuntando a que el balance beneficio-riesgo global es favorable para el fármaco de interés. La nube de puntos sería la visualización de la incertidumbre de la evidencia disponible. La consideración de la incertidumbre es imprescindible para los responsables de la toma de decisiones. También se puede visualizar la incertidumbre de un solo estudio al simular la distribución de la diferencia de probabilidades aplicando los errores estándares de cada estimación; para ello se usan métodos conocidos como Monte Carlo. Si la mayor parte de la nube de puntos, que representa la distribución de evidencia del balance beneficio-riesgo, cae en el cuadrante sureste, o bien en el noroeste, entonces es fácil llegar a una conclusión sobre el balance beneficio-riesgo: sureste claramente bueno; noroeste claramente malo. Pero, ¿qué pasa si la nube cae en el cuadrante nor-

deste? Este fue justamente el caso presentado por Lynd²⁰ (figura 1).

En este ejemplo se comparó el beneficio (evitar una trombosis venosa profunda), el riesgo (padecer un sangrado grave entre el fármaco de interés (enoxaparina) y el comparador (heparina no fraccionada a dosis bajas). La estimación de riesgo fue mayor para la enoxaparina, pero la probabilidad de evitar una trombosis venosa profunda también fue mayor. Por lo tanto la diferencia de probabilidades cae en el cuadrante nordeste, cuadrante que representa mayor beneficio pero a la vez peor perfil de riesgo. La dispersión de puntos que forma la nube de la incertidumbre se estima al aplicar la variabilidad cuantificada por los errores estándar. La decisión sobre el balance

beneficio-riesgo dependerá de dónde cae la mayor parte de la nube y nos obligará a definir un umbral de aceptabilidad. Por ejemplo, la línea diagonal ($\mu=1$) implica que un acontecimiento adverso tiene la misma importancia que un acontecimiento beneficioso y si la mayor parte de la nube —la evidencia— se encuentra por debajo de esta línea, entonces concluimos que el balance es positivo bajo la suposición de igual importancia de los acontecimientos. Pero es poco probable que los expertos den la misma importancia a acontecimientos diferentes. Debido a esto se puede cambiar el umbral; la línea con más pendiente ($\mu=2$) implica que el acontecimiento beneficioso tiene el doble de peso —importancia— que el riesgo y viceversa para la línea más pla-

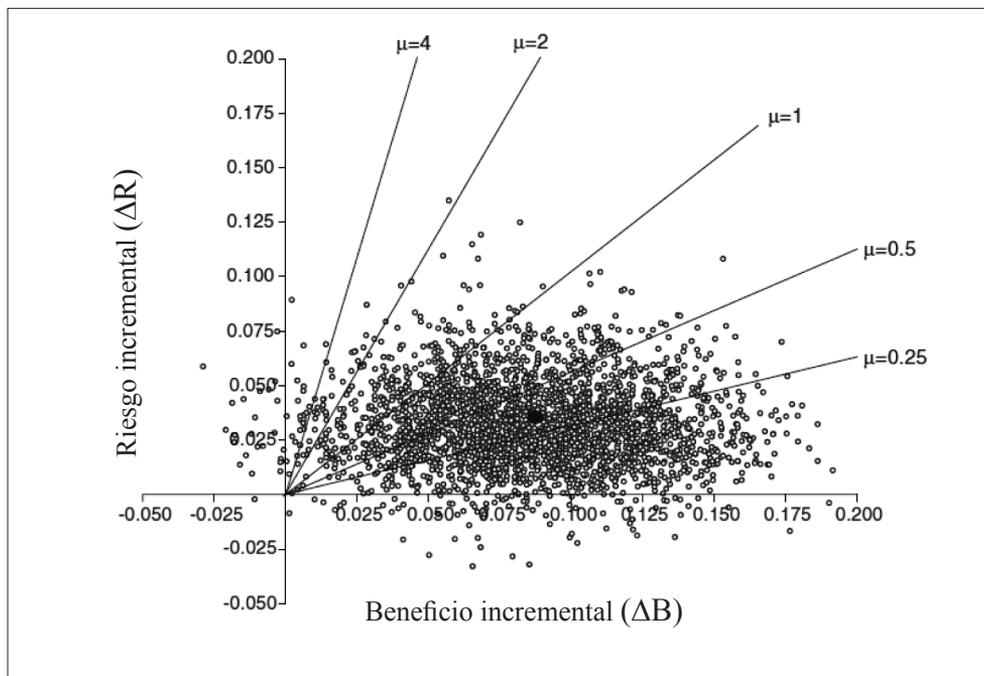


Figura 1: Distribución de las diferencias de probabilidades de beneficio (ΔB) y las de riesgo (ΔR) con los umbrales de aceptabilidad definidos según su pendiente ($\mu=0,5$, $\mu=1$, $\mu=2$, etc.). (Tomado de Lynd con permiso)²⁰

na ($\mu=0,5$). Precisamente, si los expertos opinasen que el riesgo tiene el doble de peso que el beneficio, representado por la línea $\mu=0,5$, entonces la decisión sobre el balance beneficio-riesgo sería más complicada pues casi la mitad de la evidencia sobre el balance caería encima de este umbral de aceptabilidad. Por lo tanto, se puede expresar la probabilidad de aceptar que hay un balance positivo en función de la proporción de los puntos que están por debajo del umbral de aceptabilidad.

La ventaja de esta aproximación es que los expertos y los responsables de la toma de decisiones tienen la libertad de cambiar el umbral de aceptabilidad, y observar cómo la probabilidad de aceptar un balance positivo varía según se cambie la opinión sobre la importancia relativa de los acontecimientos. Este método también proporciona una forma clara de comunicar la incertidumbre inherente a las estimaciones de riesgo y de beneficio. Sin embargo, su mayor limitación es que, en realidad, la comparación solo se puede basar en un par de acontecimientos. Como se ha resaltado anteriormente, tampoco tiene en cuenta la duración de la exposición y su efecto sobre la probabilidad del acontecimiento.

3.3. Análisis de decisión de criterios múltiples

Los dos métodos que se han descrito anteriormente se incluyen entre los 18 métodos revisados por el grupo de trabajo de la EMA.¹³ Este grupo concluyó que el método más recomendable es el de los modelos de «análisis de decisión de criterios múltiples» (MCDA), dado que

estos modelos sí que pueden incluir varios sucesos de índole totalmente diferente.²² Tras decidir los acontecimientos que definirían tanto el beneficio como los posibles riesgos, el investigador crea un árbol de decisión: las terminaciones de cada rama son los acontecimientos mientras que la división principal de las ramas es la agrupación de los acontecimientos beneficiosos y de los riesgos (figura 2).

Se le asigna un valor a cada acontecimiento —terminación de rama— que representa su peso en la puntuación beneficio-riesgo, y después se suman estos valores de forma ponderada para crear una única puntuación para cada opción terapéutica. Entonces se comparan las puntuaciones para inferir sobre el balance de beneficios y riesgos. Sin embargo, esta explicación obvia lo difícil que es asignar un valor a cada acontecimiento y luego aplicar una ponderación para poder sumar y crear la puntuación. El método propone crear «valores de preferencia»; para estos se crea una «curva de preferencia» que asocia un dato observado en un estudio con un valor de preferencia de entre 0 (menos preferible) y 100 (más preferible). Por ejemplo, lo más preferible (100) para el aumento del intervalo QTc causado por un fármaco sería 0 msec, pero como aumentos menores a 8 msec no tendrían un impacto clínico, la curva de preferencia seguiría marcando valores altos de preferencia para este rango de aumentos.¹³ A partir de aquí, con el aumento de QTc la curva de preferencia descendería llegando a 0 (menos preferible) a unos 20 msec (figura 2). Si un estudio concluye que el fármaco de interés se asocia con un aumen-

to de 9,2 msec, entonces se estima que el correspondiente valor de preferencia a través de la curva podría ser aproximadamente de 85. Cuando hay acontecimientos como, por ejemplo, el insomnio o la depresión, lo mejor es que no ocurran. Por tanto, se traza una línea recta de 100 a 0 y se asocia la incidencia de pacientes libres del acontecimiento directamente con un valor de 0 a 100; una incidencia del acontecimiento del 5% se asociaría con un valor de preferencia del 95. De

esta manera el valor 100 siempre se considera bueno, tanto si es de riesgo como si es beneficioso. Por lo tanto, los valores siempre tienen el mismo significado facilitando una suma teóricamente fácil. No obstante, la suma tiene que tener en cuenta la importancia relativa de cada acontecimiento, ya que, por ejemplo, el insomnio aislado es probablemente menos grave que la depresión. Así pues, se asignan pesos a cada acontecimiento, y para ello, se precisará la opinión de ex-

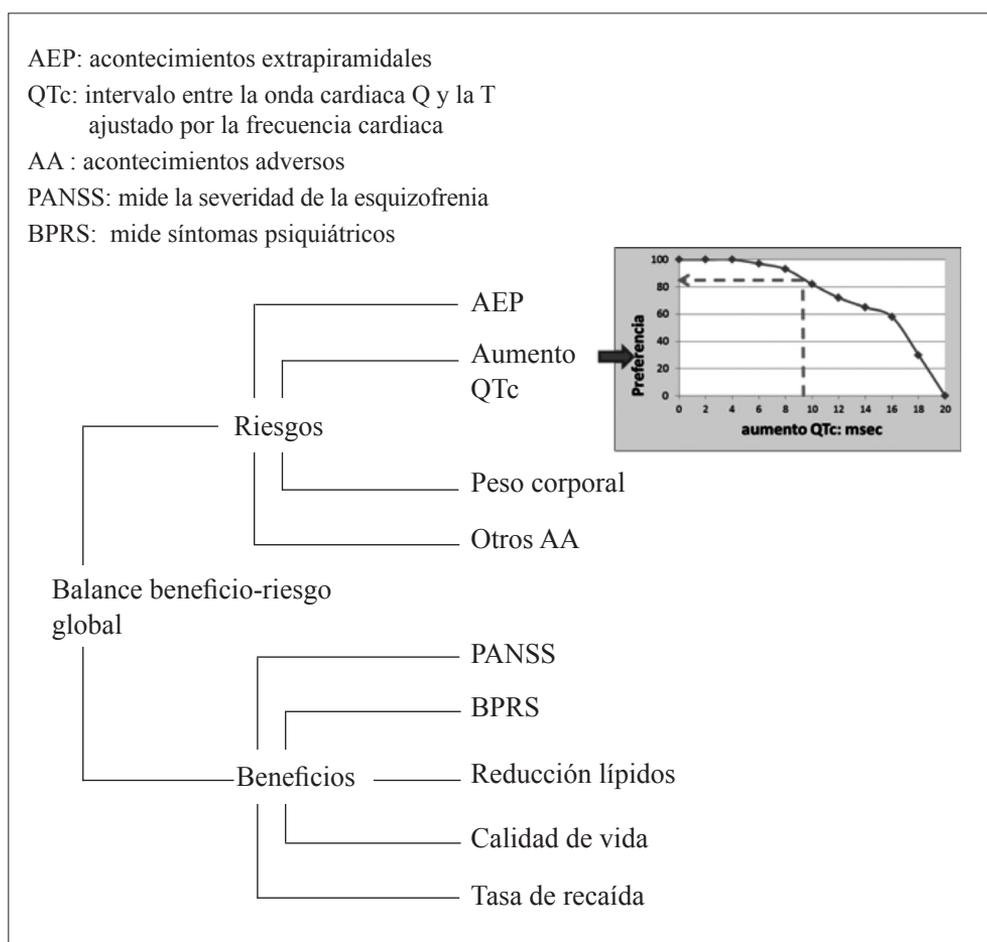


Figura 2. El árbol de decisiones usado para el análisis de decisión de criterios múltiples (MCDA) de un antipsicótico y la curva de preferencia para un acontecimiento

pertos en el área. Esto y la asignación de las curvas de preferencia son los puntos clave y problemáticos de esta aproximación. Una vez acordadas las ponderaciones y asignados los valores de preferencia, se aplica una estandarización de las ponderaciones y después se realiza una suma ponderada de todos los valores de cada acontecimiento. La estandarización se aplica de manera que la suma final para el fármaco queda entre 100 y 0; cuanto más cerca está del 100 mejor es el balance beneficio-riesgo. El próximo paso es la comparación. El método permite cierto grado de análisis de sensibilidad sobre las ponderaciones relativas. Sin embargo, no se contempla cómo puede afectar la comparación a la forma de las curvas de preferencia, las cuales son subjetivas. Tampoco se contempla la variabilidad de los datos procedentes de estudios ni el impacto de la variabilidad de las ponderaciones; el análisis de sensibilidad solo contempla variar una sola ponderación. La duración de la exposición al fármaco tampoco puede considerarse en este marco. No obstante, tanto el hecho de poder incorporar varios acontecimientos simultáneamente, como la importancia relativa de cada uno de ellos y, además, el ser un método relativamente fácil, son los puntos que llevaron a la EMA a recomendar esta aproximación.

4. UNA NUEVA APROXIMACIÓN: LA SIMULACIÓN DE COHORTES DE PACIENTES

Tal vez no haya una aproximación que pueda abarcar todos los tipos de beneficios y riesgos. El MCDA, o una versión modificada, podría ser lo más adecuado

cuando tienen que aplicarse preferencias asociadas a estados de salud. Cuando las variables son dicotómicas (v.gr., evitar un trombo o padecer un infarto agudo de miocardio), o bien se puedan convertir en variables dicotómicas (v.gr., una disminución de dolor de un 25%), métodos como el IRBR pueden resultar muy apropiados, si bien se topa con la limitación de poder incorporar solo dos acontecimientos en el cálculo, uno de beneficio y uno de riesgo. Una faceta interesante del IRBR es la de poner en manos de los expertos la libertad de variar la importancia relativa de los acontecimientos de riesgos y los de beneficios, mientras el analista solo les facilita la información necesaria. Ese no es el caso con los MCDA y otros métodos parecidos, puesto que la importancia de los acontecimientos (preferencias, utilidades) es parte de la estructura del modelo y, por tanto, la validez de las conclusiones depende de la validez de las preferencias o utilidades.

Un método que puede abordar escenarios con varios acontecimientos simultáneamente, incluyendo la muerte, y tiene la capacidad de informar a los expertos sin tener que incorporar preferencias o utilidades a priori es la microsimulación. La microsimulación también tiene en cuenta la duración de la exposición, y asimismo, los periodos en los cuales aparecen los acontecimientos asociados a la exposición al fármaco. El método consiste en crear cohortes de pacientes, los cuales pueden ser reales, pero se simula la aparición de los acontecimientos durante un seguimiento bajo unas suposiciones de exposición a las opciones terapéuticas. Es decir, lo que se simula es el seguimiento para cada

paciente, y es por el hecho de llegar hasta el nivel del paciente por lo que se denomina microsimulación. En adelante trataremos de dar una explicación breve de este método sin usar formulas, pero al lector que esté interesado en profundizar en este tema le recomendamos consultar textos sobre la teoría de colas que es la que proporciona el fundamento de esta aproximación, o bien sobre la Simulación de Acontecimientos Discretos (*Discrete Event Simulation*) que representa el tipo de modelo que se sugiere.^{23,24}

Para realizar la simulación del seguimiento de un conjunto de pacientes, lo primero que se requiere son las características de los pacientes que serían la diana para la evaluación del balance beneficio-riesgo. No nos interesan todas las características de los pacientes, sino solo las que sean determinantes de los acontecimientos de interés. Los pacientes diana pueden encontrarse en bases de datos clínicas automatizadas como las de atención primaria (BIFAP en España, o bien, THIN o CPRD en Reino Unido).²⁵⁻²⁷ También pueden usarse encuestas de salud como el NHANES de los EEUU.²⁸

La segunda pieza de información que hace falta para realizar la simulación es la probabilidad (riesgo) de cada acontecimiento. Los datos de las probabilidades deberían estar suficientemente detallados como para dar riesgos específicos asociados con cada nivel de los determinantes de los acontecimientos. Por ejemplo, si modelizamos la aparición de infarto agudo de miocardio, entonces deberíamos disponer de los riesgos de infarto según los grupos de edad y según la presencia de enfermedad cardiovascular y diabetes. Estas probabilidades pueden

proceder de la población general siempre que la indicación del fármaco objeto de estudio no tenga relación directa con los acontecimientos de interés. Por ejemplo, podemos usar datos de la población general para obtener riesgos de infarto agudo de miocardio —según edad y sexo— si modelizamos el efecto de un fármaco para tratar el glaucoma. Por el contrario, estos riesgos no sirven si estamos modelizando las opciones terapéuticas para la diabetes tipo 2, debido al aumento de riesgo cardiovascular asociado con esta enfermedad.²⁹

Una vez que tengamos un grupo de pacientes que puedan representar los usuarios del fármaco bajo estudio y hayamos ligado la probabilidad de cada acontecimiento, beneficioso o perjudicial, a cada paciente, entonces podremos tratar de estimar como irían apareciendo los acontecimientos a lo largo del seguimiento. Una complicación que merece la pena evitar es el envejecimiento de la cohorte, por lo que es aconsejable no simular un seguimiento más allá de cinco años.

Antes de centrarnos en la microsimulación vamos a comentar un método muy sencillo pero que también proporcionaría el número de acontecimientos que aparecerían en el seguimiento. Tras describirlo presentaremos un ejemplo desarrollado a partir de datos provenientes de la base de datos THIN.

4.1. Una aproximación sencilla: estimación del número esperado de casos

Un teorema básico de las leyes de la probabilidad dice que la suma de las proba-

bilidades da el número de sucesos que pueden esperarse;³⁰ un jugador de dados realista esperaría, al tirar su dado un total de seis veces, que el número tres apareciera solo una vez. De la misma manera, si hay 1.000 pacientes cuyo riesgo de un acontecimiento (probabilidad) es de 0,1 por año, entonces esperaríamos un total de 100 casos a lo largo de un año. Quizás más, quizás menos. Esa variabilidad reflejaría la incertidumbre estadística.

El ejemplo que vamos a usar para describir los métodos de generación de acontecimientos se basa en el tratamiento del glaucoma primario de ángulo abierto. Las opciones terapéuticas de interés son el maleato de timolol y el latanoprost, cuyo objetivo es disminuir la presión intraocular.³¹ El timolol es un betabloqueante de uso tópico, pero, a pesar de esto, ha habido evidencia de absorción sistémica y, por tanto, existe la posibilidad de efectos adversos sistémicos de tipo beta-bloqueante.³² La otra terapia principal, latanoprost, es un análogo de las prostaglandinas. También es de uso tópico pero los estudios de farmacocinética han demostrado que no tiene absorción sistémica.^{33,34} Por lo tanto, de entrada, se puede considerar que el uso de latanoprost no puede asociarse con efectos adversos a órganos alejados del ojo. Este caso es un ejemplo del balance beneficio-riesgo limitado a considerar solo los efectos adversos; en esencia se trata de un ejemplo de los «daños comparativos» que anteriormente hemos mencionado. Así mismo, el ejemplo es más sencillo de lo habitual ya que podemos suponer que el latanoprost no se asocia con efectos adversos graves ajenos al ojo y, por ello, no tiene contra-

indicaciones relevantes.³⁴ El timolol, sin embargo, debido a su absorción sistémica ha sido asociado a un empeoramiento de la insuficiencia cardiaca y a la exacerbación del asma y de la enfermedad pulmonar obstructiva crónica.³¹⁻³⁶ Por lo tanto, está contraindicado en pacientes que ya padecen estas enfermedades.³⁴ A pesar de la diferencia en el perfil de seguridad, los dos fármacos siguen siendo opciones de primera línea en Europa. Resulta obvio que el timolol tendrá un mayor número de acontecimientos adversos, pero la pregunta es: ¿cuál es la diferencia neta entre las opciones, y cuál es una diferencia aceptable?

Tal y como se decía antes, en primer lugar necesitamos definir los pacientes. En este caso son los pacientes tratados para el glaucoma. Con este fin, se realizó en la base de datos THIN de Reino Unido un estudio transversal en el cual se identificaron todos los pacientes con un diagnóstico de glaucoma y que recibían tratamiento para disminuir la presión intraocular. La fecha del estudio transversal fue el 1 de julio de 2007, y se estudió a 17.392 pacientes que habían recibido tratamiento en los últimos 60 días. La edad mediana de 77 años, con una alta prevalencia de hipertensión (51%), ya nos indicaba que podríamos esperar un número elevado de acontecimientos como la insuficiencia cardiaca. Las mujeres sumaban un poco más de la mitad (52%). El número de pacientes que ya padecían insuficiencia cardiaca fue del 4%, mientras que había un 13% con asma y un 7% con enfermedad pulmonar obstructiva crónica. Estas características, además del tabaquismo, son los determinantes de los acontecimientos.

Para el siguiente paso, el de estimar las probabilidades anuales de los acontecimientos en la población general, también se usaron los datos de THIN. Para ello se calcularon las tasas de incidencia a partir de todos los adultos de la base de datos seguidos durante un año desde el 1 de enero de 2006. Estos pacientes proporcionaban un denominador de un total de 1,53 millones de años-persona. Las tasas de incidencia para cada acontecimiento fueron calculadas para cada nivel de los factores determinantes, incluyendo si eran pacientes con antecedentes de los acontecimientos, es decir, si eran casos prevalentes. El hecho de usar una base de datos grande ofrece las siguientes ventajas:

1. Puede proporcionar la granularidad necesaria para asignar una probabilidad a cada nivel de los determinantes; en este ejemplo se necesitó una estimación para cada quinquenio de edad (12 grupos) por sexo, y según la historia de hipertensión y la historia de insuficiencia cardiaca, lo cual proporcionó un total de 96 estratos (12x2x2x2) con su correspondiente probabilidad. Es poco probable que la información disponible en la bibliografía aporte esta granularidad. Pero a pesar de tratarse de una base de datos grande y disponer de un denominador de más de 1,5 millones de años-persona, fue preciso aplicar métodos para suavizar las estimaciones puesto que para algunos grupos se disponía de números reducidos.
2. Se puede tener en cuenta la correlación entre los acontecimientos, lo cual es poco probable encontrar en

la bibliografía. Si tuviéramos como acontecimientos de interés la insuficiencia cardiaca y el infarto agudo de miocardio entonces podríamos esperar que hubiera correlación entre estos dos acontecimientos.

De todas formas, como medida de calidad, es preciso contrastar las estimaciones con las que se aportan en la bibliografía. En este caso las estimaciones sí que se parecían a lo publicado.

Tras obtener la información requerida para generar los acontecimientos en la población de pacientes a los que se les podría administrar tratamiento para el glaucoma, el siguiente paso es asignar la probabilidad de cada suceso a cada miembro de la cohorte de pacientes. Es decir, tenemos 17.392 pacientes reales y cada uno tiene una probabilidad de padecer un episodio agudo de insuficiencia cardiaca y una probabilidad de padecer una exacerbación de asma/enfermedad pulmonar obstructiva crónica. También se les asigna a los pacientes la probabilidad de morir durante un año, dato que se obtuvo de las estadísticas nacionales. A partir de esta información es posible estimar los sucesos por microsimulación o bien por el método sencillo de números esperados. Al tratarse de probabilidades anuales, solo se tienen que sumar las probabilidades de los pacientes para estimar el número de casos que se esperarían de cada acontecimiento a lo largo de un año. Estas frecuencias serían las que se presentan en situación basal y representan, por tanto, lo que se esperaría en ausencia de un efecto nocivo del fármaco, ya que son producto de los riesgos en la población general. Para estimar lo que

podría pasar bajo la exposición a timolol se aplican unos riesgos relativos hallados en estudios previos. Se supone que la exposición a timolol podría aumentar el riesgo de insuficiencia cardiaca por un factor de 1,20,³⁶ mientras que el riesgo de exacerbación de asma/enfermedad pulmonar obstructiva crónica aumentaría por un factor de 1,47.^{37,38} También se asume que el riesgo de debutar con asma/enfermedad pulmonar obstructiva crónica aumentaría por un factor de 1,20.³⁷ Estos riesgos relativos definen los escenarios de exposición. Puesto que no hay absorción sistémica de latanoprost, suponemos unos riesgos relativos de 1,0 (es decir, no hay aumento de riesgo) asociados a la exposición a esta opción terapéutica. Para estimar las frecuencias de acontecimientos en la población de pacientes solo tenemos que multiplicar los riesgos basales de cada paciente por los riesgos relativos que definen exposición y luego sumar. En la tabla 1 se muestra el número de pacientes esperado para una cohorte de 10.000 pacientes con glaucoma sometidos a tratamiento

En términos netos podríamos esperar unos 67 casos de más de asma/enfermedad pulmonar obstructiva crónica al tratar 10.000 pacientes con timolol durante

un año. Asimismo esperaríamos unos 19 casos de más de insuficiencia cardiaca. En estos momentos, el lector puede estar preguntándose si son números de episodios o números de personas afectas y qué pasa con la mortalidad. La respuesta a la primera duda es que es difícil separar estas dos definiciones de acontecimiento. La respuesta a la segunda es que no se ha considerado el posible impacto sobre la mortalidad. No se ha considerado el efecto sobre la mortalidad porque no se dispone de riesgos relativos correspondientes, aunque es totalmente concebible que un aumento de los acontecimientos debiera llevar a un aumento de defunciones debidas directamente al hecho de padecer un acontecimiento grave. Estas se pueden estimar al aplicar la probabilidad de muerte tras sufrir un acontecimiento, y esta probabilidad se conoce como letalidad. Sin embargo, este método supone que un paciente queda expuesto a lo largo de un año, pero en realidad puede morir durante ese año, debido al fármaco o no, y por tanto no puede haber acontecimientos tras la muerte. Todos estos problemas, el que un paciente presente distintos episodios, la consideración de la mortalidad en el modelo y la discontinuación del segui-

Tabla 1: Número de acontecimientos esperados a lo largo de un año para 10.000 pacientes tratados para el glaucoma

	Latanoprost	Timolol	Diferencia
Asma/enfermedad pulmonar obstructiva crónica	145	212	+67
Insuficiencia cardiaca	94	113	+19

miento, pueden resolverse en la microsimulación, que es lo que vamos a tratar a continuación.

4.2. Una aproximación más compleja: simulación del tiempo hasta el acontecimiento

La clave de la microsimulación para resolver los problemas anteriormente citados es generar el tiempo hasta que ocurra el acontecimiento de interés. El tiempo se simula a partir del riesgo del acontecimiento para cada paciente. Para entender el método empezaremos por explicar cómo se simula el tiempo hasta un solo acontecimiento, después pueden añadirse múltiples acontecimientos del mismo, o de diferentes tipos para un

solo paciente, y entonces se agregan los pacientes para así simular la experiencia de una cohorte de pacientes. El acontecimiento de más importancia para el modelo es la muerte ya que determina el fin del seguimiento.

El método de simular un solo acontecimiento se parece al análisis de supervivencia. La curva va del 1 al 0 en el eje vertical con una pendiente inicial más aguda que se va aplanando en el tiempo (eje horizontal), toma su forma según el riesgo del acontecimiento; cuanto más probable más pendiente tiene la curva (ver inserto en figura 3). El momento en el que el paciente va a tener el acontecimiento se puede asumir que se comportaría como un proceso estocástico. Operativamente, por tanto, se

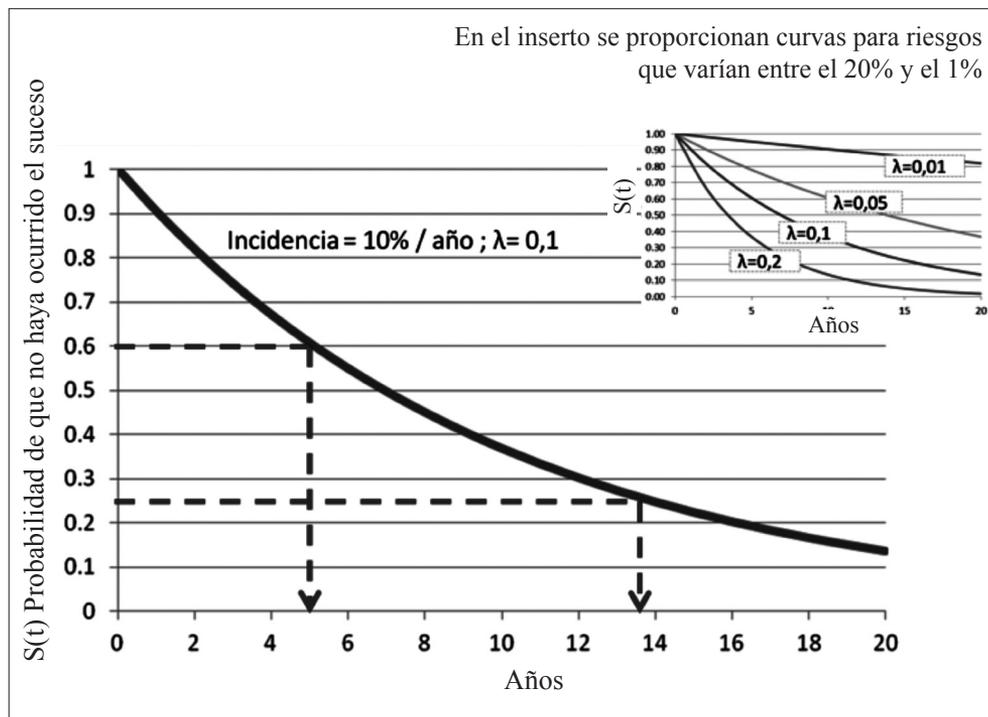


Figura 3: Simulación del tiempo hasta un único acontecimiento cuando el riesgo de sufrir dicho acontecimiento es del 10% por año ($\lambda = 0,1$)

puede generar para cada paciente un número aleatorio entre 0 y 1 que se correspondería con la probabilidad de que no ocurra el acontecimiento (eje vertical). Cada valor de probabilidad en dicho eje vertical se correspondería con un tiempo de espera (eje horizontal). La curva que se obtiene, denominada «curva de decadencia exponencial» se usa en la simulación de colas.²³ Gráficamente, esta manera de generar tiempos hasta que ocurra el acontecimiento, dado un nivel de riesgo, queda reflejada en la figura 3.

Los valores de los números aleatorios para $S(t)$ cerca del 1,0 indican que el acontecimiento ocurre pronto, mientras que valores cerca del 0 corresponden a esperas prolongadas. De hecho, números aleatorios muy pequeños pueden corresponder a acontecimientos que ocurrirían de aquí a 100 años, lo cual es poco realista y por ello se aconseja limitar el plazo de seguimiento, de tal modo que concluiremos que el acontecimiento no ocurre si el tiempo hasta el acontecimiento es superior al seguimiento. Cuando ocurre durante el seguimiento, entonces necesitamos saber cuándo acontece. En general, cuanto mayor es el riesgo, más rápidamente aparecerán los acontecimientos y, en consecuencia, más elevada será la frecuencia observada de acontecimientos en la población durante el seguimiento.

Cuando hay múltiples acontecimientos para un solo paciente se podrían generar los tiempos para cada uno de los acontecimientos. Pero esto, con frecuencia, no es tan sencillo, ya que estos acontecimientos podrían estar relacionados y cambiar el perfil de riesgo del paciente a partir del momento en que acontece uno

de ellos. Por ejemplo, si simulamos un infarto agudo de miocardio en el seguimiento, entonces, a partir del momento en que haya ocurrido, debemos aumentar tanto el riesgo de morir como el riesgo de padecer otro infarto. La forma de solucionarlo en la microsimulación es detener el seguimiento en el primer acontecimiento y actualizar el perfil de riesgo. Una vez actualizado el conjunto de riesgo del paciente, se puede volver a iniciar la simulación de tiempos a partir de este momento, teniendo en cuenta que queda menos tiempo para que termine el seguimiento. Esto se repite hasta que el seguimiento se agota si el paciente no muere antes. Como se ha mencionado previamente, la mortalidad es un acontecimiento primordial para esta aproximación ya que, como es obvio, no puede haber acontecimientos tras la muerte. Para simular la exposición a un fármaco, se aplican los riesgos relativos de cada acontecimiento que se encuentran en la bibliografía al conjunto de riesgos que tiene el paciente. Igual que se hizo al usar el método sencillo descrito en el epígrafe anterior.

Merece la pena mencionar ahora cómo incorporar la incertidumbre que existe tanto para los riesgos en situación basal como para los riesgos relativos que definen los escenarios de exposición a fármacos. Esto es relativamente fácil e involucra un segundo nivel de simulación. En lugar de asignar al paciente el riesgo, se usa una distribución de probabilidades cuyo punto central es el riesgo estimado, y la distribución alrededor de este punto viene definida por el error estándar de la estimación del riesgo. Se escoge un valor aleatorio

sobre esta distribución y es este valor el que queda asignado al paciente. Por lo tanto, al repetir la simulación el riesgo asignado también variará. Los riesgos relativos también son estimaciones y vienen con sus errores estándares y, por tanto, también se puede generar un valor aleatorio alrededor del valor central; como ejemplo el riesgo relativo de exacerbación de la insuficiencia cardíaca es 1,16 y el intervalo de confianza del 95% es de 0,74 a 1,82.³⁶ Entonces multiplicamos el riesgo en situación basal generado según la distribución de probabilidad por el riesgo relativo generado a partir de otra distribución de probabilidades. Si las estimaciones están basadas en números pequeños y, por tanto, existe un grado grande de incertidumbre; esto quedaría reflejado en errores estándares grandes y habrá, por consiguiente, una gran imprecisión o variabilidad en las estimaciones de riesgo.

De este modo se pueden simular varios acontecimientos para un paciente y se puede introducir la incertidumbre inherente en los parámetros. Este proceso se repite para cada uno de los pacientes que integran la cohorte. Después de ejecutar la simulación para la cohorte entera se pueden sumar los acontecimientos que se observan en el seguimiento. Dada la naturaleza de la simulación creada a base de números aleatorios —que se usan tanto para simular tiempos hasta el acontecimiento como para generar los riesgos absolutos y los riesgos relativos—, cada vez que se ejecuta una simulación el número de acontecimientos variará. Por lo tanto, no sería adecuado usar una sola ejecución de la simulación, sino que hay que repetirla

hasta que la distribución de los números de acontecimientos sea estable. El número de repeticiones de la simulación dependerá de la variabilidad de la simulación, pero suele ser de varios miles hasta que la estimación de la varianza no cambie. Es evidente que todos estos cálculos son posibles hoy día gracias a la enorme potencia de los ordenadores (en unas pocas horas se pueden realizar un millar de simulaciones para cada escenario de exposición).

Este método fue aplicado para la cuantificación de los riesgos y beneficios de celecoxib comparado con los AINE tradicionales (naproxeno, ibuprofeno y diclofenaco).¹⁰ En este caso, la incertidumbre fue incorporada aplicando un análisis de sensibilidad consistente en variar los riesgos relativos en función del valor del error estándar.¹⁰ Los acontecimientos de sangrado gastrointestinal, enfermedad coronaria y mortalidad fueron modelizados para escenarios de exposición en los que se comparaba celecoxib con los tres AINE tradicionales. Ya que celecoxib tiene un menor riesgo de sangrado gastrointestinal comparado con los AINE tradicionales, esto se interpretó como un beneficio. El riesgo fue la probabilidad de enfermedad cardiovascular. La conclusión fue que celecoxib tenía un mejor balance beneficio-riesgo que los AINE tradicionales.

A través del reajuste de los conjuntos de riesgos para cada paciente, es posible modelizar el impacto de la discontinuación del fármaco tras sufrir un acontecimiento adverso. Asimismo, se pueden modelizar periodos de exposición al fármaco y escenarios en los cuales el impacto de la exposición no es uniforme a

lo largo del tiempo. Un ejemplo puede ser modelizar el impacto de los bisfosfonatos sobre la disminución del riesgo de fractura; el efecto no es inmediato, sino que el bisfosfonato tiene un efecto gradual sobre la densidad ósea que a su vez produce una disminución lenta del riesgo de fractura. Así, esta aproximación puede resolver la limitación que presentan otros métodos al no tener en cuenta el impacto que sobre el balance beneficio-riesgo puede tener la duración de la exposición que se mencionó al inicio de este capítulo.

La microsimulación no depende de utilidades o preferencias, pero sería fácil aplicar tales ponderaciones a cambios en el estado de salud tras sufrir un acontecimiento. De hecho, al saber cuándo aparecen los acontecimientos, se pueden asignar las ponderaciones a los periodos que corresponden a los distintos estados de salud. Por último, este método puede aplicarse también para predecir el impacto de las medidas de minimización de riesgos: por ejemplo, conocer el impacto que tendría la inclusión de una contraindicación en el perfil de reacciones adversas en la población de pacientes candidata a recibir el tratamiento.

5. CONCLUSIONES

La aplicación de métodos cuantitativos en las decisiones sobre el balance beneficio-riesgo es un proceso en evolución. De aquí a unos pocos años se pueden esperar importantes cambios en este campo a medida que nuevos modelos sean evaluados, rechazados o aceptados. No obstante, es poco probable que pueda haber una sola

aproximación metodológica que abarque todos los escenarios de riesgo y beneficio. Los promotores de modelos tendrán que afrontar varios retos: 1) la comunicación clara de la incertidumbre inherente a los parámetros estimados; 2) cómo aplicar ponderaciones a los acontecimientos y cómo evaluar la incertidumbre en estas ponderaciones; 3) la validación de los métodos, y 4) cómo conseguir la aceptación entre profesionales de los distintos campos.

En este capítulo hemos abordado el concepto y los componentes de la relación o balance beneficio-riesgo, además de describir algunos métodos para su posible cuantificación, pero un punto clave que hay que destacar es cómo se escoge, y cómo se justifica, el conjunto de acontecimientos que se incluirán en dicha evaluación beneficio-riesgo. La justificación es clave puesto que es posible influir en el balance según el conjunto que uno decida apropiado. No importa la elegancia de los modelos aplicados si el conjunto de acontecimientos está potencialmente equivocado. Por este motivo las aportaciones por parte de expertos en el área terapéutica seguirán siendo imprescindibles. Esta justificación es parte del proceso denominado PROACT-URL por la EMA,⁶ que pretende proporcionar una estructura al proceso al describir una serie de procedimientos a seguir. El modelizar el balance es solo una parte de este proceso.

La cuantificación del balance beneficio-riesgo tiene un solo objetivo: maximizar los beneficios de las opciones terapéuticas al tiempo que se protege a las poblaciones de riesgos innecesarios. Por lo tanto, además de informar

a los responsables de las decisiones, la cuantificación debería tener la misión de optimizar el balance; por ejemplo, se tendría que aplicar en distintas subpoblaciones para identificar en cuál de ellas el balance es más o menos favorable y actuar en consecuencia. Por

ejemplo, tal vez rofecoxib estuviese aún disponible si, una vez identificados sus posibles riesgos, se hubiera limitado su uso a poblaciones de bajo riesgo cardiovascular. Pero esto es ya un contrafactual sobre el que solo cabe especular.

REFERENCIAS

1. Black J. Cómo actúan los medicamentos. En: Laurence DR, Bennett PN (eds). *Farmacología Clínica*. Barcelona: Salvat Editores; 1983.
2. De Abajo FJ, Maciá MA. La contribución de la farmacología clínica a la evaluación de la relación beneficio-riesgo de los medicamentos. En: Sánchez-García P (ed). *Desarrollo de la Farmacología Clínica en España*. Serie monografías. Real Academia Nacional de Medicina. Madrid: Arán Ediciones; 2011, pp. 55-70.
3. Tversky A, Kahneman D. Judgment under uncertainty: heuristics and biases. *Science*. 1974; 185: 1124-31.
4. O'Hagan A, McCabe C, Akehurst R, et al. Incorporation of uncertainty in health economic modelling studies. *Pharmacoeconomics*. 2005; 23: 529-36.
5. Strom B. What is pharmacoepidemiology? En: Strom B (ed). *Pharmacoepidemiology*, 4.^a ed. Chichester: Wiley & Sons; 2005, pp. 3-15.
6. Schulman KA, Glick H, Polsky D, Eisenberg D. *Pharmacoeconomics: economic evaluation of pharmaceuticals*. En: Strom B (ed). *Pharmacoepidemiology*, 4.^a ed. Chichester: Wiley & Sons; 2005, pp. 573-601.
7. Spounge J. A guide for Quantitative risk assessment for offshore installations. Aberdeen: CMPT publication, Publication 99/100a; UK, 1999. Disponible en: <http://www2.energyinst-pubs.org.uk/pdfs/670.pdf> Consultado el 3 de junio de 2013.
8. Medicines Healthcare Products Regulatory Authority. Bisphosphonates. Disponible en: <http://www.mhra.gov.uk/Safetyinformation/Generalsafetyinformationandadvice/Product-specificinformationandadvice/Product-specificinformationandadvice-A-F/Bisphosphonates/index.htm#112> Consultado el 3 de junio de 2013.
9. Lipscombe L, Gomes T, Levesque LE et al. Thiazolidinediones and cardiovascular outcomes in older patients with diabetes. *JAMA*. 2007; 298: 2634-43.
10. Varas-Lorenzo A, Maguire A, Castellsague J, et al. Quantitative assessment of the gastrointestinal and cardiovascular risk-benefit of celecoxib compared to individual NSAIDs at the population level. *Pharmacoepidemiol Drug Safety*. 2007; 16: 366-76.
11. European Medicines Agency. Development and testing of tools and processes for balancing multiple benefits and risks as an aid to informed regulatory decisions about medicinal products. European Medicines Agency, 12 de marzo de 2009. Disponible en: http://www.ema.europa.eu/docs/en_GB/document_library/Report/2011/07/WC500109477.pdf. Consultado el 3 de junio de 2013.
12. CIOMS. Benefit-risk balance for marketed drugs: evaluating safety signals. Report of CIOMS working group IV. Ginebra: CIOMS; 1998.
13. European Medicines Agency. Benefit-risk methodology Project. Work package 2 report: applicability of current tools and processes for regulatory benefit-risk assessment. European Medicines Agency; 2010. Disponible en: http://www.ema.europa.eu/docs/en_GB/document_library/

- Report/2010/10/WC500097750.pdf. Consultado el 3 de junio de 2013.
14. European Medicines Agency. Guideline on good pharmacovigilance practices (GVP). Module VII – Periodic safety update report. European Medicines Agency, 22 de junio de 2012. Disponible en: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2012/06/WC500129136.pdf Consultado el 3 de junio de 2013.
 15. Last JM (ed). A dictionary of epidemiology, 4.^a ed. Oxford: Oxford University Press; 2001.
 16. Graham DJ, Campen D, Hui R, et al. Risk of acute myocardial infarction and sudden cardiac death in patients treated with cyclo-oxygenase 2 selective and non-selective non-steroidal anti-inflammatory drugs: nested case-control study. *Lancet*. 2005; 365: 475-81.
 17. Holden WL, Juhaeri J, Dai W. Benefit-risk analysis: a proposal using quantitative methods. *Pharmacoepidemiol Drug Safety*. 2003; 2: 611-6.
 18. Marschnera, IC, Embersonc J, Irwig L, Walter SD. The number needed to treat (NNT) can be adjusted for bias when the outcome is measured with error. *J Clin Epidemiol*. 2004; 57: 1244-52.
 19. Stevenson MD, Oakley J, Chilcott JB. Gaussian process modeling in conjunction with individual patient simulation modeling: a case study describing the calculation of cost-effectiveness ratios from the treatment of established osteoporosis. *Med Decis Making*. 2004; 24: 89-100
 20. Lynd LD, O'brien BJ. Advance in risk-benefit evaluation using probabilistic simulation methods: an application to the prophylaxis of deep vein thrombosis. *J Clin Epidemiol*. 2004; 57: 795-803.
 21. Spiegelhalter DJ, Best NG. Bayesian approaches to multiple sources of evidence and uncertainty in complex cost-effectiveness modelling. *Statist Med*. 2003; 22: 3687-709.
 22. Department of Communities and Local Government. Multi-criteria analysis: a manual. Londres, enero 2009. Disponible en: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/7612/1132618.pdf. Consultado el 3 de junio de 2013.
 23. Cox DR, Smith WL. Queues: monographs on applied probability and statistics. Londres: Chapman and Hall; 1961.
 24. Banks J, Carson J, Nelson B, Nicol D. Discrete-event system simulation, 4.^a ed. Pearson Prentice Hall; 2005.
 25. Agencia Española de Medicamentos y Productos Sanitarios. BIFAP – Base de datos para la investigación farmacoepidemiológica en atención primaria. Disponible en: <http://www.bifap.org>. Consultado el 3 de junio de 2013.
 26. Clinical Practice Research Datalink. Disponible en: <http://www.cprd.com/home/>. Consultado el 3 de junio de 2013
 27. Cegedim. THIN Data Statistics. Disponible en: <http://csdmruk.cegedim.com/our-data/statistics.shtml>. Consultado el 3 de junio de 2013.
 28. Centers for Disease Control and Prevention. National Health and Nutrition Examination Survey. Disponible en: <http://www.cdc.gov/nchs/nhanes>.

- htm. Consultado el 3 de junio de 2013.
29. Haffner SM, Lehto S, Ronnema T, et al. Mortality from coronary heart disease in subjects with type 2 diabetes and in nondiabetic subjects with and without prior myocardial infarction. *N Engl J Med.* 1998; 339:229-34.
 30. Armitage P, Berry G. *Statistical methods in medical research*, 2.^a ed., Oxford: Blackwell Scientific Publications: 1985.
 31. Alward WLM. Medical management of glaucoma. *N Engl J Med.* 1998; 339: 1298-307.
 32. Stewart WC, Garrison PM. B-blocker-induced complications and the patient with glaucoma, newer treatment to help reduce systemic adverse events. *Arch Intern Med.* 1998; 158: 221-6.
 33. Hugues F, Matte JC, Le Jeune C, Salem A. Effects of beta-adrenoceptors blocking eye drops in patients with chronic bronchitis. *Therapie.* 1992; 47: 211-5.
 34. British National Formulary. Disponible en: <http://www.bnf.org/index.htm>. Edición 54, sección 11.6. Consultado el 3 de junio de 2013.
 35. Confalonieri M, Aiolfi S, Patrini G, et al. Severe bronchial spasm crises induced by topical administration of eyedrops with timolol bas, a non-selective beta blocking agent. *Recenti Prog Med.* 1991; 82: 402-4.
 36. Monane M, Bohn RL, Gurwitz JH et al. Topical glaucoma medications and cardiovascular risk in the elderly. *Clin Pharmacol Ther.* 1994; 55: 76-83.
 37. Huerta C, García-Rodríguez LA, Möller CS, Arellano F. The risk of obstructive airways disease in a glaucoma population. *Pharmacoepidemiol Drug Saf.* 2001; 10: 157-63.
 38. Avorn J, Glynn RJ, Gurwitz JH, et al. Adverse pulmonary effects of topical beta blockers used in the treatment of glaucoma. *J Glaucoma.* 1993; 2: 158-65.