

## Problemas metodológicos más relevantes en los estudios de mortalidad y progresión de la enfermedad en pacientes con EPOC

Josep Andrés Martínez, Rosa Palomino Meneses y Belén Santiago Josefat

GOC Networking. Barcelona. España.

La enfermedad pulmonar obstructiva crónica (EPOC) es una entidad clínica caracterizada por una limitación progresiva y no completamente reversible al flujo en las vías aéreas, que se asocia a una respuesta inflamatoria anómala de los pulmones, principalmente frente al humo del tabaco. El mejor indicador de la progresión y la gravedad de la enfermedad es la medida de la obstrucción al flujo aéreo (volumen espiratorio forzado en el primer segundo), expresada en porcentaje respecto al valor de referencia.

La mayoría de los tratamientos existentes para la EPOC no han mostrado tener un efecto claro en la progresión y mortalidad de la enfermedad, debido probablemente al hecho de que la EPOC es un proceso heterogéneo de muy larga duración y con una amplia variedad de fenotipos de pacientes, situaciones clínicas y evolución de la enfermedad.

En los últimos años el número de estudios sobre la progresión y mortalidad en la EPOC ha crecido de manera muy significativa, convirtiéndose el análisis de supervivencia en una de las herramientas más importantes de explotación de los datos resultantes. Sin embargo, ciertos factores metodológicos asociados a este tipo de investigaciones, tales como el diseño del estudio, las variables de medida del efecto y la determinación de la muestra, han influido en las conclusiones finales de dichos estudios. Además, el análisis de la progresión y de la tasa de mortalidad suele llevar asociada la necesidad de una larga duración de seguimiento y un alto número de participantes, lo que limita especialmente la financiación y el seguimiento de este tipo de estudios. En el caso de los estudios de mortalidad asociada a la EPOC, se suman además las dificultades a la hora de reconocer la causa específica de la muerte y de identificar los factores pronóstico de mortalidad de la enfermedad, que son los principales factores metodológicos que limitan la posibilidad de conseguir datos concluyentes.

The most important methodological problems in studies of mortality and disease progression in patients with chronic obstructive pulmonary disease

Chronic obstructive pulmonary disease (COPD) is characterized by progressive and incompletely reversible airflow obstruction associated with an anomalous inflammatory response of the lungs, mainly to tobacco smoke. The best indicator of disease progression and severity is measurement of airflow obstruction (forced expiratory volume in 1 second), expressed as a percentage of the predicted value derived from a healthy reference population. Most of the treatments available for COPD have not been shown to clearly affect disease progression or mortality, probably because COPD is a heterogeneous, longstanding process and because there is wide variety in patients' phenotypes, clinical situations, and clinical course. In the last few years, the number of studies on disease progression and mortality in COPD has markedly increased, making survival analysis one of the most important tools for exploiting the resulting data. However, certain methodological factors associated with this type of research, such as the study design, the variables used to measure effect, and determination of the sample, have influenced the conclusions of these studies. Moreover, analysis of disease progression and mortality rates usually entails prolonged follow-up periods and a large number of participants, limiting the funding and continuity of this type of study. In the case of studies of mortality associated with COPD, there are the additional difficulties of identifying the specific cause of death and of identifying the prognostic factors of mortality from the disease, which are the main methodological factors that hamper the possibility of obtaining conclusive data.

**Palabras clave:** EPOC. Mortalidad. Progresión de la enfermedad. Métodos de investigación. Análisis de supervivencia.

**Key words:** COPD. Mortality. Disease progression. Research methods. Survival analysis.

## Introducción

La enfermedad pulmonar obstructiva crónica (EPOC) es un trastorno permanente y progresivo que se caracteriza por la disminución del flujo en las vías aéreas y que está causado, principalmente, por una reacción inflamatoria frente al humo del tabaco<sup>1</sup>. Según la Organización Mundial de la Salud, todos los años la EPOC provoca la muerte de unos 3 millones de personas en el mundo. En España se estima que cada año mueren más de 18.000 personas por EPOC, que es la quinta causa de fallecimiento más común entre los varones, con una tasa anual de 55 muertes por 100.000 habitantes, y la séptima para las mujeres, con una tasa anual de 18 muertes por 100.000 habitantes<sup>2</sup>. Se estima que para el año 2020 la EPOC será la tercera causa mundial de defunción.

Para diagnosticar la EPOC es preciso realizar una espirometría, en condiciones basales y tras la inhalación de un agente broncodilatador, que muestre un valor del cociente FEV<sub>1</sub>/FVC menor de 0,70. La medida de la obstrucción al flujo aéreo (FEV<sub>1</sub>), expresada en porcentaje respecto al valor de referencia, es el mejor indicador de la progresión y la gravedad de la enfermedad. La Global Initiative for Chronic Obstructive Lung Disease (GOLD), en su edición de 2007, ha propuesto una clasificación de la gravedad de la EPOC en 4 estadios en función de los valores de la espirometría tras broncodilatación<sup>3</sup> (fig. 1). A medida que la enfermedad progresa, la mayor gravedad de la obstrucción al flujo aéreo se asocia con más síntomas, peor calidad de vida y una mayor incidencia de exacerbaciones. La mayoría de los tratamientos existentes para la EPOC son sintomáticos y no han mostrado tener un efecto claro en la progresión y mortalidad de la enfermedad, probablemente debido al hecho de que la EPOC es un proceso heterogéneo de muy larga duración, con una gran variedad de fenotipos de pacientes, de situaciones clínicas y de evolución de la enfermedad.

En los últimos años el número de estudios sobre la progresión y mortalidad en la EPOC ha crecido de manera muy significativa. Sin embargo, ciertos factores metodológicos asociados a este tipo de estudios han influido en sus conclusiones. Además, el análisis de la progresión y de la tasa de mortalidad suele llevar aso-

ciada la necesidad de un largo seguimiento y un alto número de participantes, lo cual limita especialmente la financiación y el seguimiento de este tipo de estudios.

Uno de los principales retos a la hora de estudiar la progresión de la enfermedad es definir los parámetros indicadores de dicha progresión. Históricamente, el descenso anual del valor del FEV<sub>1</sub> ha sido el parámetro más utilizado a la hora de establecer la progresión de la EPOC. Mientras que en una persona sana este descenso anual suele ser de 20-30 ml, en pacientes con EPOC<sup>4</sup> la media se encuentra en 60 ml. No obstante, se trata de un parámetro vinculado al estado de la función pulmonar y recientemente se ha cuestionado su correspondencia con la evolución fisiopatológica de la enfermedad<sup>5-7</sup>.

En el caso de los estudios de mortalidad asociada a la EPOC, las dificultades que entraña reconocer la causa específica de la muerte e identificar factores pronósticos de mortalidad de la enfermedad son los principales factores metodológicos que limitan la posibilidad de conseguir datos concluyentes. En cuanto a la dificultad a la hora de asignar la causa específica de muerte, diversos estudios han mostrado que la EPOC está infradiagnosticada y que en más de la mitad de los pacientes con EPOC muy grave ésta no aparecía en el certificado de defunción<sup>8</sup>.

El estudio TORCH<sup>9</sup> fue el primer estudio internacional cuyo objetivo principal fue analizar todas las causas de mortalidad, así como las causas específicas relacionadas con la EPOC, en pacientes con EPOC moderada y grave. Las causas de muerte fueron analizadas por un comité independiente formado por 3 expertos, que obtuvieron una coincidencia del 83% en la adjudicación de la causa de muerte de los casos asignados y un 52% de coincidencia con la causa descrita en las historias médicas de los pacientes.

Con el propósito de encontrar factores pronósticos de mortalidad en pacientes con EPOC, se ha estudiado la asociación entre varios parámetros de esta enfermedad y la tasa de mortalidad. En un estudio epidemiológico, de 3 años de seguimiento, se describió una mortalidad de un 23% y la edad y el valor inicial del FEV<sub>1</sub> antes de la administración del broncodilatador fueron factores pronósticos de la enfermedad. Además, tras ajustar el riesgo a la edad y al valor de FEV<sub>1</sub>, la mortalidad se relacionó positivamente con la capacidad pulmonar total

Estadio	Recomendaciones
EPOC leve (1) FEV <sub>1</sub> /FVC < 0,70 FEV <sub>1</sub> ≥ 80% referencia	Evitar factores de riesgo; vacuna antigripal; broncodilatadores de corta acción cuando se precisen
EPOC moderada (2) FEV <sub>1</sub> /FVC < 0,70 50% ≤ FEV <sub>1</sub> < 80% referencia	Tratamiento de fondo con uno o más broncodilatadores de acción prolongada y rehabilitación pulmonar
EPOC grave (3) FEV <sub>1</sub> /FVC < 0,70 30% ≤ FEV <sub>1</sub> < 50% referencia	Esteroides inhalados en caso de exacerbaciones frecuentes
EPOC muy grave (4) FEV <sub>1</sub> /FVC < 0,70 FEV <sub>1</sub> < 30% referencia o FEV <sub>1</sub> < 50% referencia con insuficiencia respiratoria crónica o insuficiencia cardíaca derecha	¿O <sub>2</sub> y cirugía?

Fig. 1. Criterios de Global Initiative for Chronic Obstructive Lung Disease (GOLD) para clasificar los estadios de la enfermedad pulmonar obstructiva crónica (EPOC). FEV<sub>1</sub>: volumen espiratorio máximo en el primer segundo; FVC: capacidad vital forzada; insuficiencia respiratoria crónica: presión parcial de oxígeno arterial (PaO<sub>2</sub>) < 60 mmHg (8,0 kPa) con o sin presión parcial de CO<sub>2</sub> arterial (PaCO<sub>2</sub>) > 50 mmHg (6,7 kPa), respirando aire ambiente y al nivel del mar.

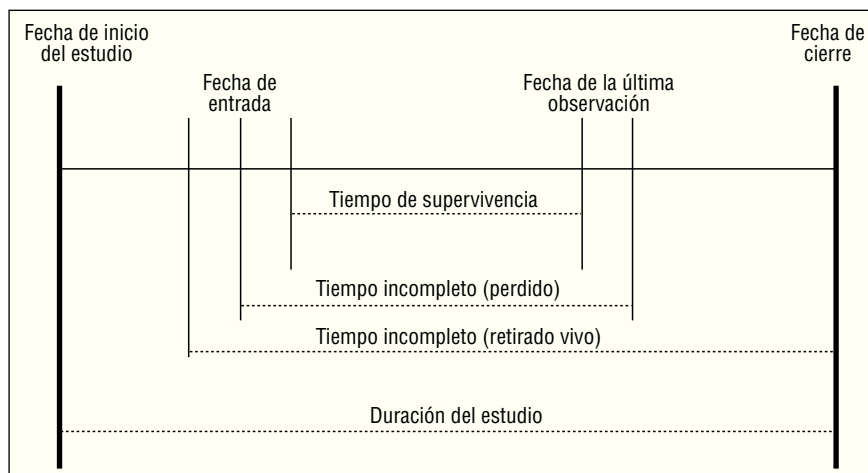


Fig. 2. Esquema general de los estudios.

y la frecuencia cardíaca, y negativamente con la tolerancia al ejercicio<sup>10</sup>. Otro estudio, que incluía a una cohorte de 5.542 personas seleccionadas de la población general, tras 22 años de seguimiento demostró que los pacientes con EPOC moderada o grave según la clasificación GOLD presentaban una supervivencia inversamente relacionada con el valor del FEV<sub>1</sub> que tenían al entrar en el estudio (razón de riesgo relativo de 1,4 a 2,0 para la EPOC moderada y de 2,1 a 3,5 para la grave)<sup>11</sup>. Sin embargo, en un análisis prospectivo de 3 años de seguimiento en pacientes con EPOC hospitalizados con exacerbaciones agudas, se determinó una asociación positiva entre mortalidad y duración de la enfermedad (riesgo relativo [RR] = 1,158), valores bajos de albúmina (RR = 0,411) y un índice de masa corporal bajo (RR = 0,830), mientras que los valores del FEV<sub>1</sub> no mostraban una asociación significativa<sup>12</sup>. También se ha descrito que hasta que el FEV<sub>1</sub> es inferior al 50% del valor teórico, este parámetro no es un buen predictor de la mortalidad. De igual forma, en pacientes con FEV<sub>1</sub> muy bajo esta medida tenía poco valor predictivo<sup>10</sup>.

Otro parámetro fisiológico que se ha considerado factor pronóstico de la EPOC es la capacidad inspiratoria (IC). Se ha descrito que la relación entre la IC y la capacidad pulmonar total (IC/TLC) podría ser una variable pronóstica de la evolución de la enfermedad, debido a su relación inversa con la mortalidad. En una cohorte de 689 pacientes con EPOC, con un 27% de fallecimientos durante los 5 años de seguimiento, se determinó que había una relación directa entre mortalidad y mayor edad, menor índice de masa corporal, menor tolerancia al ejercicio, más disnea y mayor relación IC/TLC. Tras realizar un análisis multivariante, el índice IC/TLC resultó un buen factor de mortalidad en pacientes con EPOC, independiente de otras variables<sup>13</sup>.

Otra razón que podría explicar la falta de datos concluyentes sobre la asociación entre los parámetros de función pulmonar y la mortalidad de la enfermedad en pacientes con EPOC es que ésta se asocia a importantes consecuencias sistémicas, tales como inflamación, desnutrición y debilidad muscular, lo que podría contribuir a la aparición de síntomas y a la morbimortalidad de la

EPOC, sobre todo en sus fases avanzadas. Así, la categorización de los pacientes con EPOC en función del grado de disnea es más predictiva de supervivencia a los 5 años que la categorización según el FEV<sub>1</sub><sup>14</sup>. El peso corporal<sup>15,16</sup>, la capacidad para realizar ejercicio y la percepción del estado de salud, medida por cuestionarios de calidad de vida, son igualmente parámetros que han mostrado una asociación con la mortalidad en la EPOC<sup>17-20</sup>.

### Diseño del estudio

Dada la naturaleza de las hipótesis en cuestión y de las dudas planteadas por los estudios ya realizados, es muy poco probable que nuevos estudios transversales y/o de casos y controles proporcionen más luces que sombras. Así pues, en la actualidad sólo se contemplan 2 opciones: los diseños de cohortes y los ensayos clínicos aleatorizados (ECA).

En esencia, el esquema general para ambos tipos de investigación es el mismo (fig. 2); esto es, entre las fechas de inicio y cierre del estudio se observan ciertas características de una muestra de efectivos que se incluyen, en general, de forma no simultánea, durante períodos diversos, dado que habrá un contingente de pacientes supervivientes al cierre del estudio, un segundo grupo que habrá fallecido durante el período de observación y un tercero que habrá abandonado antes de finalizar dicho período. La principal diferencia, que no única, entre ambos diseños reside en que el estudio de cohortes es observacional y el investigador no interviene en la asignación de individuos expuestos al factor en estudio (tratamiento farmacológico o no), mientras que en un ECA el investigador asigna la exposición al factor en estudio de forma aleatorizada y, cuando es posible, a doble ciego. De ahí se deriva la principal limitación de los diseños observacionales puros: ¿hasta qué punto los pacientes tratados son comparables a los que no han recibido tratamiento? ¿Cuál es el impacto del efecto placebo? Además, si la cohorte es retrospectiva, también será necesario resolver otras cuestiones metodológicas relevantes que guardan relación con la calidad y disponibilidad de los datos previamente registrados, y con los

eventuales cambios en las definiciones de las variables (criterios diagnósticos, técnicas quirúrgicas o de laboratorio, tratamientos farmacológicos, etc.).

Sin duda el ECA es el diseño que mejor resuelve los problemas de validez interna, puesto que la intervención del investigador optimiza la comparabilidad de los pacientes incluidos en cada grupo. No obstante, aun realizados mediante protocolos de investigación impecables, tampoco están exentos de problemas de validez. En efecto, el hecho de que la tasa de abandonos sea elevada y significativamente distinta en cada grupo<sup>9,21,22</sup> condiciona la validez interna, y tanto la autoselección de los pacientes participantes, necesariamente voluntaria, como los criterios de inclusión y exclusión apropiados para evitar factores de confusión en la medida del efecto alejan las características de la muestra analizada de la población de referencia; esto es, en el caso de la EPOC, los pacientes atendidos en la práctica clínica habitual.

### Variables de medida del efecto

La primera cualidad que debe poseer cualquier variable es la de medir con la máxima exactitud y precisión el fenómeno que representa y, como es lógico, sea cual fuere el tipo de estudio, las variables incluidas en los cuadernos de recogida de datos deben obedecer a definiciones precisas. Las investigaciones más recientes sobre la EPOC miden la evolución de la enfermedad mediante pruebas espirométricas, la frecuencia y la gravedad de las exacerbaciones, y el deterioro del estado de salud y la calidad de vida de los pacientes.

Los criterios GOLD<sup>3</sup>, cuyo esquema se reproduce en la figura 1, se han utilizado para clasificar el grado de progresión de la EPOC en investigaciones recientes. Sin embargo, este parámetro muestra una desviación estándar muy grande<sup>6,7,23</sup>. Otra limitación que presenta el uso del descenso anual del FEV<sub>1</sub> como indicador de progresión de la EPOC es que la espirometría plantea problemas en cuanto a la calibración de los aparatos y la competencia y homogeneidad de los investigadores respecto a la técnica, el procedimiento y las condiciones de realización de la prueba, problemas que suelen paliarse mediante la formación específica y el uso del mismo modelo de instrumento, algo de especial dificultad en estudios multicéntricos y/o internacionales.

Otros parámetros propuestos como indicadores de progresión de la enfermedad son la disminución de la capacidad de realizar ejercicio, que es el resultado de la combinación de la disminución de la función pulmonar y los efectos extrapulmonares de la EPOC (problemas cardiovasculares y debilidad muscular)<sup>24</sup>; la disminución de la calidad de vida<sup>25</sup> y la utilización de los recursos sanitarios<sup>26,27</sup>. Estos parámetros también llevan asociadas importantes dificultades por lo que respecta a su empleo en estudios de larga duración y/o con un alto número de individuos. La frecuencia de las exacerbaciones también se ha utilizado para medir la progresión de la EPOC. Sin embargo, existen diversos problemas metodológicos asociados, como la variabilidad en la definición y la adjudicación de exacerbación<sup>28</sup>, la depen-

dencia de la situación basal del paciente<sup>29</sup> y la variabilidad en la medicación previa administrada a cada individuo<sup>28,30</sup>, por lo que es necesario conocer en detalle las definiciones empleadas para interpretar correctamente los resultados de una investigación. Asimismo, la intervención de comités de expertos que supervisen de forma centralizada los criterios efectivamente utilizados por los investigadores clínicos es una fórmula acertada para mejorar la calidad de esta variable.

Recientemente se ha publicado la validación del modelo BODE como predictor de exacerbaciones, hospitalizaciones y mortalidad en la EPOC<sup>31,32</sup>. Este índice incorpora 4 parámetros: función pulmonar (FEV<sub>1</sub>), síntomas (disnea), capacidad de ejercicio (prueba de la marcha) y estado nutricional (índice de masa corporal), lo que apoya la necesidad de una evaluación multidimensional de la EPOC. Este concepto permitirá un mejor conocimiento y manejo de la EPOC, así como el desarrollo de nuevos tratamientos combinados que actúen sobre todos los elementos implicados tanto en las manifestaciones pulmonares como en las sistémicas de esta enfermedad.

Puhan et al<sup>33</sup> mostraron que los cuestionarios específicos, como el St. George's Respiratory Questionnaire, proporcionan medidas más sensibles y específicas del estado de salud y la calidad de vida de los pacientes con EPOC que los cuestionarios generales como es el SF-36, que tiene en el SF-12 una versión reducida. Se trata de una variable de gran interés, dado que la EPOC afecta progresivamente y de forma importante el bienestar del paciente.

### Representatividad de la muestra

Conviene recordar que la representatividad o exactitud de una muestra nada tiene que ver con su tamaño ni con su precisión; es decir, tal como ilustra la figura 3, una muestra puede ser grande y sesgada, pequeña y representativa, o cualquier otra combinación de ambos atributos. De acuerdo con la teoría de muestreo, la representatividad de una muestra depende de que todos los componentes de la población diana hayan gozado de igual probabilidad de formar parte de la muestra efectivamente seleccionada. Así, el único método que garantiza la igualdad de probabilidad exigida es el aleatorio simple, que, sin embargo, suele presentar problemas de aplicación insalvables en la investigación clínica y epidemiológica. Por ello comúnmente se recurre a procedimientos de selección pseudoaleatorios –que se aproximen, cuanto más mejor, a los aleatorios–, definidos a partir de ciertos criterios de inclusión y de exclusión, concebidos para sustentar razonablemente la comparabilidad entre grupos (validez interna) y la generalización de los resultados (validez externa): 2 centros de interés que, por otra parte, colisionan con frecuencia. Basta revisar los criterios de exclusión de ECA recientes<sup>9,22</sup> para comprobar que, en legítimo interés de la validez interna, estos estudios excluyen a sujetos que forman parte de los pacientes tratados en la práctica clínica habitual, lo que, sumado al inevitable proceso de auto-selección de voluntarios, implica que los resultados de

los ECA pueden no reproducirse exactamente en la práctica clínica; tanto más cuanto los criterios de inclusión no son los mismos en todos los estudios y, por consiguiente, sus conclusiones no son generalizables a poblaciones de referencia idénticas.

También debe tenerse en cuenta la posibilidad de incurrir en el sesgo de Berkson, descrito en 1946 a partir de la observación de una asociación negativa entre el cáncer y la tuberculosis pulmonar. Consiste en un sesgo de selección que puede darse cuando –como suele ser el caso– los pacientes investigados son seleccionados en centros hospitalarios bien sea para un ECA o un estudio de cohortes. Se deriva del hecho circunstancial de que la exposición a alguno de los factores de riesgo del estudio –léase tabaquismo o tratamiento esteroideo– aumenta la probabilidad de enfermar en general y, en particular, de ingresar en un hospital, en cuyo caso la tasa de exposición en el grupo control podría estar sobrestimada. Obsérvese en este sentido que los estimadores obtenidos en muestras sesgadas están sistemáticamente desviados en una u otra dirección y que, por lo tanto, son inexactos y pueden –aunque no necesariamente– inducir a conclusiones falaces, salvo que el sesgo se haya identificado, medido y corregido.

### Tamaño muestral

El número de efectivos estudiados determina la precisión de los estimadores y de los tests de hipótesis estadísticas. Respecto a estos últimos, la precisión se mide mediante la probabilidad de error que se asume en caso de rechazarse la hipótesis nula o de no asociación (error alfa o de tipo I), y la probabilidad de error que implica aceptar la hipótesis nula (error beta o de tipo II), con frecuencia indicado en forma de potencia estadística ( $1 - \beta$ ) expresada en porcentaje. Así, se trata de errores aleatorios que disminuyen conforme se aumenta el tamaño muestral. Nótese, no obstante, que su reducción no tiene ningún efecto en eventuales sesgos, y que alfa y beta no son probabilidades complementa-

rias –esto es, no tienen por qué sumar la unidad–; por consiguiente, es posible que, en contrastes dados, la inferencia estadística no permita ni afirmar ni negar la hipótesis nula asumiendo probabilidades de error suficientemente pequeñas para ambos enunciados.

Limitaciones presupuestarias aparte, el tamaño muestral idóneo para una precisión dada depende de la incidencia de las variables dependientes (a menor tasa, mayor número de efectivos), de la diferencia que se quiera observar entre grupos (a menor diferencia, mayor muestra) y de la proporción esperada de pérdidas y abandonos. Sobre este punto se destaca que la precisión de un estudio se calcula a partir de los datos efectivamente observados, con independencia de las estimaciones realizadas en el protocolo de la investigación –con frecuencia, optimistas.

### Duración del estudio

Para estudiar adecuadamente la progresión de la enfermedad y la supervivencia en pacientes con EPOC, el sentido común apunta que no es posible analizar un fenómeno si no se observa en su conjunto. Por consiguiente, los períodos de observación deberían extenderse tanto como dura la historia natural de la enfermedad desde su inicio hasta la defunción del paciente. Sin embargo, llevar a cabo una investigación que cumpla con este requisito implica costes muy elevados, dificultades logísticas de primera magnitud y, en el caso de los ECA, acaso limitaciones éticas insoslayables. No sorprende, pues, que los ECA recientes y en curso limiten su duración a 2<sup>22</sup>, 3<sup>9</sup> y 4 años<sup>21</sup>. Con todo, estos períodos de investigación dejan una cuestión en el aire: ¿el comportamiento de los supervivientes después de cerrar el estudio será semejante al observado?

Por otra parte, aumentar el tamaño de la muestra no resuelve este problema, aunque pueda disimularlo. Lo cierto es que acumular, por ejemplo, 6.000 años-persona mediante 1.000 pacientes seguidos durante 6 años o con 6.000 pacientes seguidos durante un año propor-

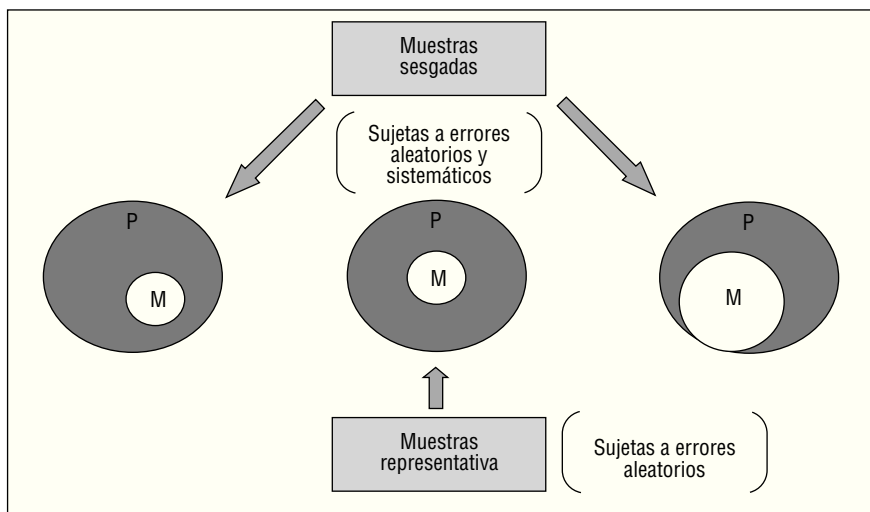


Fig. 3. Representatividad o exactitud de una muestra.

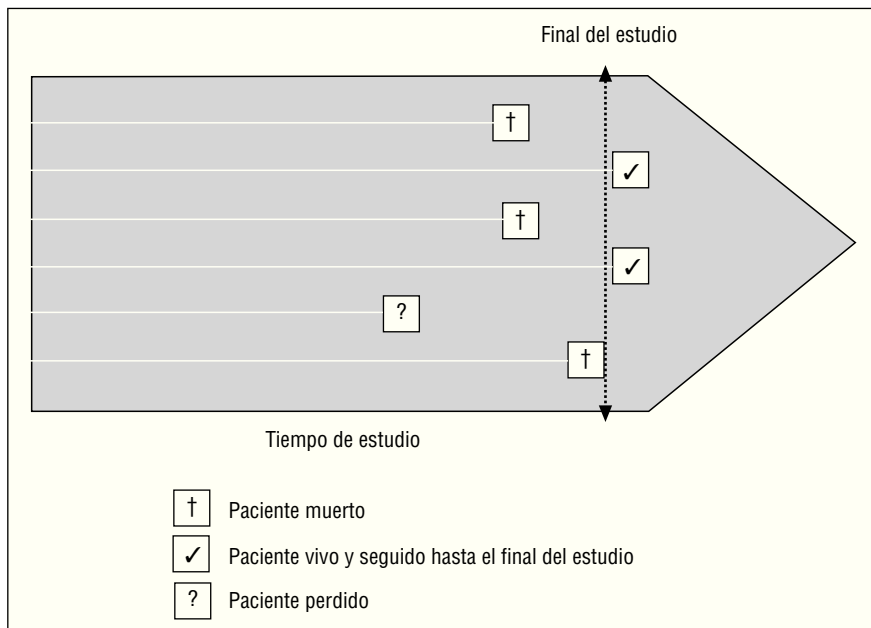


Fig. 4. Tipología de los datos obtenidos.

na la misma precisión estadística, a pesar de no significar científicamente ni mucho menos lo mismo.

### Pérdida de efectivos

Es el problema logístico por excelencia de los estudios prospectivos en general, que arrecea cuanto mayor es el período de seguimiento y que conlleva 2 consecuencias secundarias de primer orden: disminución de la precisión preestablecida cuando las pérdidas son mayores que las pronosticadas, e introducción de sesgos cuando las causas de las retiradas no son independientes del efecto que se desea observar.

Los datos correspondientes a los efectivos que han abandonado un estudio pueden tratarse de diversas maneras; por cierto, ninguna satisfactoria. Hacer caso omiso de ellos es una solución hoy en desuso, dado que procediendo así no se tiene en cuenta que los sujetos retirados han sobrevivido hasta la fecha del abandono; en consecuencia, el riesgo se sobrestimaré. Por el contrario, dar por supuesto que han sobrevivido –otra solución felizmente en desuso– implica subestimar el riesgo. Asimismo, considerar que su comportamiento habrá sido semejante a los observados también es una opción escasamente sustentable, puesto que conlleva aceptar a ciegas hipótesis indisociables de las que se pretende contrastar.

En la actualidad, el tratamiento común que reciben los datos de los abandonos es considerarlos censurados, lo que se traduce en dar por supuesto que las causas de las retiradas no están relacionados con la enfermedad; es decir, que la probabilidad de abandono es independiente del riesgo investigado, lo que tampoco está exento de controversia, ya que de hecho es posible que la decisión de abandonar sea consecuencia del empeoramiento del estado físico, de un cambio de residencia en

busca de un clima mejor o de un cambio en la actividad laboral al objeto de evitar –si es el caso– el factor de riesgo en estudio. En este sentido, es claro que el valor veritativo de los resultados de los estudios INSPIRE (un 38,6% de abandonos) y TORCH (un 39,1%) está fuertemente condicionado por la conformidad de la hipótesis subyacente a la censura de los datos, tanto más cuanto en ambos ensayos se han observado mayores pérdidas en el grupo placebo<sup>9</sup>.

No hay una solución ideal para el tratamiento de la pérdida de efectivos, salvo que se produzcan en una proporción lo suficientemente pequeña para no sesgar los resultados de la investigación sea cual fuera su destino real. Por ello, incorporar a los proyectos de investigación de larga duración programas de motivación e incentivación dirigidos a pacientes, y/o acciones complementarias de rastreo y localización de submuestras de casos perdidos se vislumbran como fórmulas adecuadas para paliar el impacto de este problema<sup>21</sup>.

### Análisis estadístico

El conjunto de técnicas estadísticas que se emplean para analizar la progresión de la enfermedad en pacientes con EPOC son las que se conocen generalmente como “análisis de supervivencia”. El término “supervivencia” se debe a que en un principio estos métodos se concibieron para estudiar el tiempo hasta la muerte de un paciente (fundamentalmente en el campo de la oncología).

Este método permite estudiar la variable “tiempo transcurrido hasta que se produce un evento” y su dependencia de otras posibles variables explicativas. La variable respuesta no es un valor numérico único, como podría ser la presión arterial, sino que toma la forma de “tiempo transcurrido hasta un suceso” (*time-to-event*) y

contempla la combinación de 2 elementos: si se ha producido o no el suceso (muerte, recidiva, aparición de síntomas, etc.) y cuánto tiempo ha tardado en llegarse al desenlace, por lo que el seguimiento viene definido por una fecha de inicio y una fecha de cierre que lo determinan. La característica más importante de este tipo de datos (tiempo hasta que ocurre un suceso) es que, muy probablemente, al final del período de observación no todos los pacientes habrán presentado el suceso o evento objeto del estudio. Además, puede ocurrir que algunos pacientes se hayan perdido por causas diversas, sin que haya sido posible determinar su estado al final del período de observación. Cuando el suceso es la muerte, el individuo puede haber fallecido por causas diferentes de las que se analizan (p. ej., un accidente). Además, es habitual que los pacientes vayan incorporándose durante todo el período de observación, con fechas de inicio y cierre diferentes para cada uno, por lo que la probabilidad de que ocurra el suceso es menor en las últimas incorporaciones. Cuando los tiempos de supervivencia no se conocen con exactitud, es decir, las observaciones se presentan incompletas, los casos se consideran censurados (fig. 4).

Aunque se llame “análisis de supervivencia”, no siempre tiene que ser la defunción el acontecimiento de desenlace, pero siempre debe tratarse de acontecimientos que, como la muerte, sean irreversibles. El análisis de supervivencia recogerá si se ha producido o no el suceso y el tiempo (en años) desde que cada paciente se incorporó al estudio hasta que murió o se produjo el evento; en caso de que no muera o no se produzca el suceso, se indicarán los años totales de observación. La variable tiempo es una variable continua y no normal, generalmente de distribución asimétrica.

En este tipo de análisis son necesarios ciertos requisitos: definir con exactitud el origen o inicio del seguimiento, definir la escala del tiempo y definir apropiadamente el evento. Además, en este tipo de análisis se asume un supuesto básico –que no siempre se da– y es que los pacientes no censurados representen bien a los censurados.

El análisis puede realizarse mediante técnicas paramétricas y no paramétricas, siendo estas últimas las más utilizadas debido a la naturaleza casi siempre no gaussiana de los datos a analizar.

Para calcular la probabilidad de supervivencia individual acumulada a lo largo del tiempo suele utilizarse el método de Kaplan-Meier<sup>34</sup>, un método no paramétrico que no asume que los datos tengan una distribución particular y asume que los sujetos censurados se habrían comportado del mismo modo que los que han sido seguidos hasta el evento. Para describir la supervivencia por este método hay que disponer del tiempo durante el cual se ha observado al paciente y el estado del paciente al final del seguimiento. El método de cálculo empleado lo convierte en especialmente útil para estudios con un número pequeño de pacientes. En general, este análisis suele acompañarse de la respectiva representación gráfica, llamada curva de Kaplan-Meier (fig. 5), con el objetivo de expresar visualmente cómo va disminuyendo la probabilidad de sobrevivir a medida que pasa el tiempo.

La curva debe interpretarse en su aspecto general, puesto que las conclusiones que se extraigan de los detalles particulares en diferentes puntos de aquélla no son fiables, sobre todo en los extremos, donde la estimación es menos precisa.

En las ocasiones en que no se dispone de observaciones individuales de los eventos, sino de observaciones agrupadas por intervalos de tiempo (p. ej., los datos de mortalidad para un país), se utilizará el cálculo actuarial también conocido como método de Cutler-Ederer, en el que se divide el tiempo en intervalos y se calcula la supervivencia en cada uno de ellos. Este método implica 3 premisas: *a*) que todos los abandonos durante un intervalo dado ocurren aleatoriamente durante dicho intervalo; *b*) que la probabilidad de supervivencia en un tiempo dado es independiente de la probabilidad de supervivencia en los demás períodos, y *c*) que las personas que se retiran del estudio tienen un destino parecido a los que se quedan.

Para comparar 2 o más curvas de supervivencia de forma global se usan diversas pruebas estadísticas de contraste de hipótesis. La hipótesis nula que se contrasta es que los grupos comparados presentan igual supervivencia globalmente, esto es, que no hay diferencias persistentes. La prueba más empleada para comparar curvas de supervivencia es el test de rangos logarítmicos, también llamado test generalizado de Savage<sup>35,36</sup>, que tiene en cuenta las diferencias de supervivencia entre grupos en todos los puntos del tiempo que dura el seguimiento. Además de éste, hay otros tests de comparación de curvas de supervivencia: test generalizado de Wilcoxon (o test de Breslow), test de Tarone-Ware, test de Peto-Peto-Prentice y test de Fleming-Harrington. En general, se recomienda utilizar el test de rangos logarítmicos cuando el evento es poco frecuente o las curvas no se cruzan; en este caso, además, los resultados obtenidos del resto de tests serán similares<sup>37</sup>. Cuando las curvas se cruzan, se recomienda realizar al menos otro test de comparación, preferentemente el de Wilcoxon.

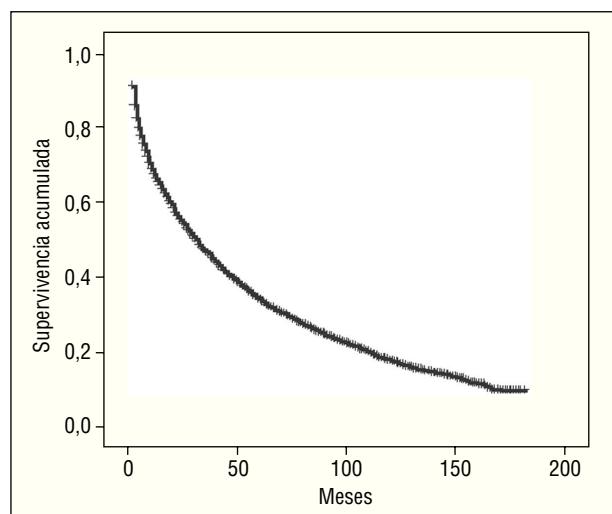


Fig. 5. Probabilidades de supervivencia en el tiempo.

Si lo que se desea estudiar es si la variable dependiente está relacionada con la supervivencia de los individuos y conocer simultáneamente el efecto independiente de una serie de factores sobre esta supervivencia, se utilizará la regresión de Cox (modelo de riesgos proporcionales)<sup>38-40</sup>. Este método debe utilizarse cuando se estudie un suceso que sólo ocurra una vez como máximo, puesto que para sucesos o complicaciones que se repitan más de una vez en algunos pacientes durante el seguimiento debería recurrirse a la regresión de Poisson<sup>41</sup>. La regresión de Cox es la extensión multivariable del análisis de supervivencia y usa modelos próximos a la regresión logística (fig. 6). Esta técnica permite identificar y evaluar la relación entre un conjunto de variables explicativas y la tasa de aparición del suceso de interés, y predecir las probabilidades de supervivencia (o de permanencia sin el evento) para un determinado sujeto a partir del patrón de valores que presenten sus variables pronóstico. La regresión de Cox asume algunos de los mismos supuestos que el método de Kaplan-Meier: *a)* el suceso debe ser irreversible; *b)* debe poder ocurrir una sola vez, y *c)* la censura no debe ser informativa. En la interpretación de una regresión de Cox, el parámetro de asociación que se obtiene es una *hazard ratio* (razón de riesgo), que corresponde a una tasa instantánea, que conceptualmente sólo requiere una duración de tiempo infinitesimal para que ocurra el suceso. Así, el *hazard* a tiempo *t* es el cociente entre quienes mueren en ese tiempo y quienes estaban en riesgo de morir. El valor mínimo de la razón de riesgo es 0, el máximo infinito. El modelo de Cox promedia de manera ponderada las razones de riesgo de los diversos momentos en que se produce alguna muerte.

En el análisis de supervivencia también hay opciones paramétricas para la estimación de funciones de supervivencia y para la realización de contrastes de hipótesis sobre éstas. A tal efecto será necesario suponer ciertas funciones de probabilidad para la variable tiempo de espera y comprobarlas mediante ciertas pruebas de ajuste.

Algunas distribuciones paramétricas se han empleado repetidamente en la literatura médica para llevar a cabo

análisis de supervivencia. Probablemente la distribución exponencial es la más simple de todas ellas. La característica más importante de esta distribución es que su función de riesgo se mantiene constante, caracterizando la exponencial como una distribución sin memoria de tiempo (un paciente tiene la misma probabilidad de morir en cualquier momento). La distribución de Weibull es una generalización de la exponencial que ha recibido una atención especial durante los últimos años debido a su simplicidad y flexibilidad<sup>42,43</sup>. Además de la exponencial, se han utilizado otras distribuciones como la lognormal, gamma, Rayleigh y Gompertz, aunque en menor medida<sup>44</sup>.

En cuanto a las limitaciones del análisis de supervivencia en estudios de EPOC, cabe destacar que en los últimos 5 años han aparecido varios estudios epidemiológicos de base poblacional que han señalado un posible impacto del tratamiento con corticoides inhalados en la mortalidad observada en pacientes con EPOC<sup>45</sup>, y cuyas conclusiones han suscitado varias críticas basadas fundamentalmente en el denominado “sesgo de tiempo inmortal”, referido a la ventaja que supone para el grupo de tratamiento la necesidad de que los pacientes incluidos en él tengan forzosamente que sobrevivir durante el período posterior a la inclusión y anterior al tratamiento. Este sesgo se superpone con el denominado de “tiempo variable”, que hace referencia al hecho de que al computar el tiempo de supervivencia en ocasiones se incluyen períodos de no tratamiento. Ambos sesgos operan en la dirección de demostrar el efecto. Sin embargo, cabe destacar que pueden cuantificarse y, por lo tanto, corregirse.

Además del análisis de la supervivencia, hay otros resultados distintos de la mortalidad más difíciles de medir pero de gran importancia, como: estado funcional, calidad de vida, ausencia de dolor y otros síntomas, satisfacción con los cuidados médicos y calidad de la muerte. Todo ello es especialmente cierto para pacientes con enfermedades crónicas como la EPOC, para los cuales la calidad de vida puede ser de gran relevancia.

En los últimos años, en los estudios de coste-beneficio ha aparecido otro concepto que permite ponderar la supervivencia respecto a la calidad de vida: los años de

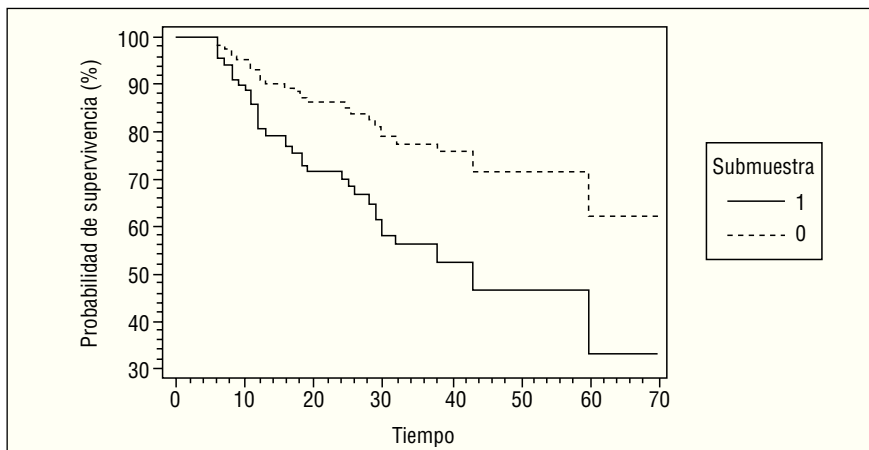


Fig. 6. Probabilidades de supervivencia estimadas según la regresión de Cox.



vida ajustados por calidad. Éstos pueden ser útiles en este tipo de estudios, ya que permiten incorporar la calidad y la cantidad de vida ganada tras la aplicación de un procedimiento terapéutico por medio de sencillos métodos de cálculo.

## Conclusiones

Diversas investigaciones llevadas a cabo recientemente arrojan ciertas dudas respecto a la eficacia del tratamiento farmacológico en la progresión y mortalidad de la EPOC, debidas, en primer lugar, a la débil significación estadística observada. No obstante, para valorar en su justa medida los resultados de dichos estudios también merecen consideración diversos problemas metodológicos, muchos de ellos derivados de la larga duración de la enfermedad y de las consiguientes dificultades que entraña el seguimiento de los pacientes.

De hecho, la elevada tasa de abandonos afecta con frecuencia la validez interna, tanto más cuando la distribución no es homogénea entre los grupos de intervención y de control. Asimismo, es preciso valorar el eventual efecto del sesgo de Berkson –derivado de incluir exclusivamente a pacientes hospitalarios–, corregir los sesgos de tiempo inmortal y de tiempo variable, y evitar el empleo de pruebas estadísticas basadas en distribuciones gaussianas.

Por otra parte, también conviene tener en cuenta que los rigurosos criterios de inclusión y de exclusión necesariamente empleados en los ECA para optimizar la comparabilidad entre los grupos investigados suelen restar representatividad a las muestras examinadas respecto a los pacientes atendidos en la práctica clínica habitual.

En suma, en el momento actual no se dispone de datos concluyentes. Por ello es necesario realizar nuevos ECA de mayor duración –tanta como sea posible– y con menores tasas de abandonos, al objeto de contrastar el efecto de los tratamientos farmacológicos en condiciones experimentales, sin que ello sea obstáculo para ampliar y completar a posteriori el conocimiento objetivo mediante la realización de estudios cuasi experimentales y/o de cohortes en poblaciones naturales.

## Declaración de conflicto de intereses

Los autores han declarado no tener ningún conflicto de intereses.

## BIBLIOGRAFÍA

1. Fletcher C, Peto R. The natural history of chronic airflow obstruction. *Br Med J*. 1977;1:1645-8.
2. Mortalidad por causa de muerte, España y comunidades autónomas (1981-2004). Disponible en: <http://www.msces/estadEstudios/estadisticas/estadisticas/estMinisterio/mortalidad/seriesTablas.htm>
3. GOLD executive committee. Global Initiative for Chronic Obstructive Lung Disease. Disponible en: <http://www.goldcopd.com>
4. Anthonisen NR, Connett JE, Kiley JP, Altose MD, Bailey WC, Buist AS, et al. Effects of smoking intervention and the use of an

inhaled anticholinergic bronchodilator on the rate of decline of FEV1. *The Lung Health Study*. *JAMA*. 1994;272:1497-505.

5. The Lung Health Study Research Group. Effect of inhaled triamcinolone on the decline in pulmonary function in chronic obstructive pulmonary disease. *N Engl J Med*. 2000;343:1902-9.
6. Burge PS, Calverley PM, Jones PW, Spencer S, Anderson JA, Maslen TK. Randomised, double blind, placebo controlled study of fluticasone propionate in patients with moderate to severe chronic obstructive pulmonary disease: the ISOLDE trial. *BMJ*. 2000;320:1297-303.
7. Pauwels RA, Lofdahl CG, Laitinen LA, Schouten JP, Postma DS, Pride NB, et al. Long-term treatment with inhaled budesonide in persons with mild chronic obstructive pulmonary disease who continue smoking. *European Respiratory Society Study on Chronic Obstructive Pulmonary Disease*. *N Engl J Med*. 1999;340:1948-53.
8. Jensen HH, Godtfredsen NS, Lange P, Vestbo J. Potential misclassification of causes of death from COPD. *Eur Respir J*. 2006;28:781-5.
9. McGarvey LP, John M, Anderson JA, Zvarich M, Wise RA. Ascertainment of cause-specific mortality in COPD: operations of the TORCH Clinical Endpoint Committee. *Thorax*. 2007;62:411-5.
10. Anthonisen NR, Wright EC, Hodgkin JE. Prognosis in chronic obstructive pulmonary disease. *Am Rev Respir Dis*. 1986;133:14-20.
11. Mannino DM, Buist AS, Petty TL, Enright PL, Redd SC. Lung function and mortality in the United States: data from the First National Health and Nutrition Examination Survey follow up study. *Thorax*. 2003;58:388-93.
12. Gunen H, Hacievliyagil SS, Kosar F, Mutlu LC, Gulbas G, Pehlivan E, et al. Factors affecting survival of hospitalised patients with COPD. *Eur Respir J*. 2005;26:234-41.
13. Casanova C, Cote C, De Torres JP, Aguirre-Jaime A, Marín JM, Pinto-Plata V, et al. Inspiratory-to-total lung capacity ratio predicts mortality in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2005;171:591-7.
14. Nishimura K, Izumi T, Tsukino M, Oga T. Dyspnea is a better predictor of 5-year survival than airway obstruction in patients with COPD. *Chest*. 2002;121:1434-40.
15. Schols AM, Slangen J, Volovics L, Wouters EF. Weight loss is a reversible factor in the prognosis of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 1998;157:1791-7.
16. Landbo C, Prescott E, Lange P, Vestbo J, Almdal TP. Prognostic value of nutritional status in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 1999;160:1856-61.
17. Oga T, Nishimura K, Tsukino M, Sato S, Hajiro T. Analysis of the factors related to mortality in chronic obstructive pulmonary disease: role of exercise capacity and health status. *Am J Respir Crit Care Med*. 2003;167:544-9.
18. Gerardi DA, Lovett L, Benoit-Connors ML, Reardon JZ, ZuWallack RL. Variables related to increased mortality following outpatient pulmonary rehabilitation. *Eur Respir J*. 1996;9:431-5.
19. Bowen JB, Votto JJ, Thrall RS, Haggerty MC, Stockdale-Woolley R, Bandyopadhyay T, et al. Functional status and survival following pulmonary rehabilitation. *Chest*. 2000;118:697-703.
20. Domingo-Salvany A, Lamarca R, Ferrer M, García-Aymerich J, Alonso J, Félez M, et al. Health-related quality of life and mortality in male patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2002;166:680-5.
21. Decramer M, Celli B, Tashkin DP, Pauwels RA, Burkhart D, Cassino C, et al. Clinical trial design considerations in assessing long-term functional impacts of tiotropium in COPD: the UPLIFT trial. *COPD*. 2004;1:303-12.
22. Seemungal T, Stockley R, Calverley P, Hagan G, Wedzicha JA. Investigating new standards for prophylaxis in reduction of exacerbations – the INSPIRE study methodology. *COPD*. 2007;4:177-83.
23. Lung Health Study Research Group. Effect of inhaled triamcinolone on the decline in pulmonary function in chronic obstructive pulmonary disease. *N Engl J Med*. 2000;343:1902-9.
24. Pinto-Plata VM, Cote C, Cabral H, Taylor J, Celli BR. The 6-min walk distance: change over time and value as a predictor of survival in severe COPD. *Eur Respir J*. 2004;23:28-33.
25. Spencer S, Calverley PM, Burge PS, Jones PW. Impact of preventing exacerbations on deterioration of health status in COPD. *Eur Respir J*. 2004;23:698-702.

ANDRÉS MARTÍNEZ J ET AL. PROBLEMAS METODOLÓGICOS MÁS RELEVANTES EN LOS ESTUDIOS DE MORTALIDAD Y PROGRESIÓN DE LA ENFERMEDAD EN PACIENTES CON EPOC

26. Hilleman DE, Dewan N, Malesker M, Friedman M. Pharmacoeconomic evaluation of COPD. *Chest*. 2000;118:1278-85.
27. Miravittles M, Murio C, Guerrero T, Gisbert R. Costs of chronic bronchitis and COPD: a 1-year follow-up study. *Chest*. 2003;123: 784-91.
28. Aaron SD, Fergusson D, Marks GB, Suissa S, Vandemheen KL, Doucette S, et al. Counting, analysing and reporting exacerbations of COPD in randomised controlled trials. *Thorax*. 2008;63:122-8.
29. Sin DD, McAlister FA, Man SF, Anthonisen NR. Contemporary management of chronic obstructive pulmonary disease: scientific review. *JAMA*. 2003;290:2301-12.
30. Suissa S, Ernst P, Vandemheen KL, Aaron SD. Methodological issues in therapeutic trials of chronic obstructive pulmonary disease. *Eur Respir J*. 2008;31:927-33.
31. Celli BR, Cote CG, Marín JM, Casanova C, Montes de Oca M, Méndez RA, et al. The body-mass index, airflow obstruction, dyspnea, and exercise capacity index in chronic obstructive pulmonary disease. *N Engl J Med*. 2004;350:1005-12.
32. Cote CG. Surrogates of mortality in chronic obstructive pulmonary disease. *Am J Med*. 2006;119:54-62.
33. Puhan MA, Guyatt GH, Goldstein R, Mador J, McKim D, Stahl E, et al. Relative responsiveness of the Chronic Respiratory Questionnaire, St. Georges Respiratory Questionnaire and four other health-related quality of life instruments for patients with chronic lung disease. *Respir Med*. 2007;101:308-16.
34. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc*. 1958;53:457-81.
35. Klein JP, Moeschberger ML. *Survival analysis: techniques for censored and truncated data*. New York: Springer; 1997.
36. Hosmer DW, Lemeshow S. *Applied survival analysis: regression modeling of time to event data*. New York: Wiley & Sons; 1999.
37. Martínez González MA, Sánchez Villegas A, Faulín Fajardo FJ. *Bioestadística amigable*. 2.ª ed. Madrid: Díaz de Santos; 2006.
38. Cox DR. Regression models and life-tables. *J Roy Statist Soc B*. 1972;34:187-220.
39. Collet D. *Modelling survival data in medical research*. New York: Chapman and Hall; 1974.
40. Cox DR, Oakes D. *Analysis of survival data*. London: Chapman and Hall; 1984.
41. Katz MH. *Multivariable analysis: a practical guide for clinicians*. Cambridge: Cambridge University Press; 1999.
42. Ducrocq V, Quaas R, Pollak E, Casella G. Length of productive life in dairy cows. 2. Variance component estimation and sire evaluation. *J Dairy Sci*. 1988;71:3071-9.
43. Ducrocq V, Quaas R, Pollak E, Casella G. Length of productive life of dairy cows. 1. Justification of a Weibull model. *J Dairy Sci*. 1988;71:3061-70.
44. Ducrocq V. *An analysis of length of productive life in dairy cattle*. Ithaca, NY: Cornell University; 1987.
45. De Lucas Ramos P, Ferreiro A, Rodríguez González-Moro J. Agonistas adrenérgicos b2 y corticoides. *Arch Bronconeumol*. 2008;43:11-7.