

Enfoque: Métodos contemporáneos en bioestadística (II)

Análisis de riesgos competitivos

Melania Pintilie*

Department of Biostatistics, Ontario Cancer Institute/Princess Margaret Hospital, University Health Network, Dalla Lana School of Public Health, University of Toronto, Canadá

Historia del artículo:

On-line el 31 de mayo de 2011

Palabras clave:

Análisis de supervivencia
Análisis de riesgos competitivos
Modelo de Fine y Gray

Keywords:

Survival analysis
Competing risks analysis
Fine and Gray model

RESUMEN

La necesidad de desarrollar tratamientos o programas específicos para una enfermedad requiere un análisis de los resultados que sea específico para dicha enfermedad. Criterios de valoración como la insuficiencia cardíaca, la muerte debida a una enfermedad específica o el control de la enfermedad local en el cáncer pueden ser imposibles de observar debido a la aparición previa de un tipo de evento diferente (como la muerte por otra causa). El evento que dificulta o modifica la posibilidad de observar el evento de interés se denomina riesgo competitivo.

Las técnicas habituales de análisis del tiempo hasta el evento aplicadas en presencia de riesgos competitivos producen unos resultados sesgados o no interpretables. La estimación de la probabilidad del evento debe calcularse, pues, con el empleo de técnicas específicas como la función acumulativa de incidencia introducida por Kalbfleisch y Prentice. El modelo introducido por Fine y Gray puede aplicarse para evaluar una covariable cuando hay riesgos competitivos. Con el empleo de técnicas específicas para el análisis de los riesgos competitivos se asegurará que los resultados no estén sesgados y puedan interpretarse correctamente.

© 2011 Sociedad Española de Cardiología. Publicado por Elsevier España, S.L. Todos los derechos reservados.

An Introduction to Competing Risks Analysis

ABSTRACT

The need to develop treatments and/or programs specific to a disease requires the analysis of outcomes to be specific to that disease. Such endpoints as heart failure, death due to a specific disease, or control of local disease in cancer may become impossible to observe due to a prior occurrence of a different type of event (such as death from another cause). The event which hinders or changes the possibility of observing the event of interest is called a competing risk.

The usual techniques for time-to-event analysis applied in the presence of competing risks give biased or uninterpretable results. The estimation of the probability of the event therefore needs to be calculated using specific techniques such as the cumulative incidence function introduced by Kalbfleisch and Prentice. The model introduced by Fine and Gray can be applied to test a covariate when competing risks are present. Using specific techniques for the analysis of competing risks will ensure that the results are unbiased and can be correctly interpreted.

Full English text available from: www.revespcardiol.org

© 2011 Sociedad Española de Cardiología. Published by Elsevier España, S.L. All rights reserved.

Abreviaturas

CIF: función acumulativa de incidencia
F&G: modelo de Fine y Gray
KM: Kaplan-Meier
PH Cox: modelo de riesgos proporcionales de Cox
RC: riesgos competitivos

INTRODUCCIÓN

Los riesgos competitivos (RC) han sido reconocidos como un caso especial de análisis de tiempo hasta el evento desde el siglo

xviii. Se han publicado trabajos en el campo de la estadística o la matemática que han incorporado nuevos avances, entre ellos la monografía de David y Moeschberger¹. Con la disponibilidad de datos más amplios, claros y precisos respecto a los diferentes tipos de resultados, los RC volvieron a surgir como un tipo crucial de análisis dentro de los análisis de tiempo hasta el evento que es necesario para comprender mejor una enfermedad. Era preciso establecer una conexión entre los resultados matemáticos y el campo aplicado. Varios autores han contribuido a mejorar el conocimiento de las situaciones de RC^{2,3}. Otros han mejorado y desarrollado técnicas y en algunos casos han proporcionado códigos de programas informáticos preparados para el uso en el análisis estadístico aplicado⁴⁻⁶.

INTRODUCCIÓN AL ANÁLISIS DE TIEMPO HASTA EL EVENTO

En muchos estudios, el resultado evaluado se observa de forma longitudinal. De esta forma, se observa a cada sujeto de la cohorte

* Autor para correspondencia: Biostatistics Department, Ontario Cancer Institute/ UHN, 610 University Ave., Toronto, M5G 2M9, Canadá.

Correo electrónico: Pintilie@uhnres.utoronto.ca

durante un tiempo hasta que se produce el evento. Por ejemplo, el evento de interés puede ser la muerte, un infarto de miocardio o la recurrencia del cáncer. Los objetivos del estudio pueden ser estimar la probabilidad de aparición del evento o su asociación con covariables de interés como el tratamiento o las características de los sujetos. El análisis estadístico utilizado se denomina *análisis de tiempo hasta el evento* o, a veces, *análisis de supervivencia*. El método más frecuente para estimar la probabilidad de un evento es un enfoque no paramétrico, generalmente denominado método de Kaplan-Meier⁷ (KM) o *método de límite de producto*. El supuesto principal de la estimación de KM de la supervivencia es que las observaciones censuradas acabarían presentando el evento si el seguimiento fuera lo bastante largo.

En el resto de este artículo, se indicarán las probabilidades del evento en vez de la probabilidad de ausencia del evento. Por ejemplo, en vez de la probabilidad de supervivencia, se presentará la probabilidad de muerte, que puede estimarse mediante el complemento del estimador de KM: 1-KM.

INTRODUCCIÓN A LOS RIESGOS COMPETITIVOS

No es infrecuente que un participante de un estudio experimentalmente más de un tipo de evento. Se produce una situación de RC cuando la aparición de un tipo de evento modifica la capacidad de observar el evento de interés. Miyasaka et al⁸ llevaron a cabo un estudio en una cohorte de base comunitaria formada por pacientes a los que se diagnosticó fibrilación auricular entre 1986 y 2000 en el *Olmsted County*, Minnesota, Estados Unidos. La variable de valoración primaria fue la aparición de demencia. La mediana de seguimiento fue de 4,6 años. Otros tipos de eventos evaluados fueron el ictus y la muerte. De los 2.837 individuos con fibrilación auricular, 299 presentaron demencia y 1.638 fallecieron hasta el momento en el que se realizó el análisis. Los números de pacientes con ictus no se indican y se censuran en el análisis. La conclusión del estudio fue que la incidencia de demencia en los individuos con fibrilación auricular es elevada (el 10,5% a los 5 años con el método de KM). La aparición de un ictus antes de la demencia no afecta a la observación de esta y, por lo tanto, no es un evento que constituya un RC. Para esta argumentación, no tendremos en cuenta que los ictus múltiples pueden causar demencia. En cambio, la muerte sin que se haya presentado demencia previamente hace que la observación de la demencia sea imposible. Así pues, la muerte sin demencia es un evento de RC para la variable de valoración de demencia. También un traumatismo craneal grave podría considerarse un evento de RC, puesto que los cambios conductuales del paciente podrían hacer que el diagnóstico de demencia resultara imposible.

Una situación de RC más sutil es la que se produce en el estudio realizado por Whalley et al⁹ sobre la importancia de la ecocardiografía. Esta cohorte de 228 pacientes ancianos sintomáticos fue examinada mediante ecocardiografía y fue objeto de un seguimiento para identificar la hospitalización cardiovascular o la muerte cardiovascular. La hipótesis era que las características ecocardiográficas podrían predecir los eventos cardiovasculares. El criterio de valoración principal se definió como una variable combinada formada por muerte y/u hospitalización cardiovascular. Para este tipo de variable, una muerte debida a causas distintas de la enfermedad cardiovascular es un evento de RC y, como tal, un paciente deja de estar en riesgo de sufrir cualquiera de los eventos de interés.

Se llevó a cabo un ensayo doble ciego, aleatorizado, de tres grupos, en un total de 931 centros de 24 países, para evaluar el efecto de valsartán frente a valsartán+captopril frente a captopril solo (ensayo VALIANT)¹⁰ en la mortalidad por todas las causas. En total, se incluyó a 14.703 pacientes tras infarto de miocardio con

disfunción ventricular izquierda y/o insuficiencia cardiaca en una proporción 1:1:1 en los tres grupos. Dado que se consideraba evento cualquier muerte, este tipo de variable de valoración no tiene RC. El estudio confirmó la hipótesis de que la supervivencia observada en los tres grupos era diferente. Sin embargo, se identificó la hemorragia gastrointestinal (GI) como un efecto secundario grave en los tres grupos. Moukarbel et al¹¹ estudiaron los posibles factores que podían predecir la hemorragia GI. Para esta variable de valoración, la muerte sin hemorragia GI constituye un RC claro.

Un número creciente de investigadores reconoce la presencia de RC y la necesidad de aplicar técnicas adecuadas para ello. Núñez et al¹² estudiaron una cohorte de 972 pacientes con síndromes coronarios agudos sin elevación del segmento ST entre 2001 y 2005. Uno de los objetivos del estudio fue identificar factores asociados a la rehospitalización por insuficiencia cardiaca aguda. Entre los factores estudiados estaban los de diabetes, antecedentes previos de cardiopatía isquémica, insuficiencia renal crónica, antecedentes de tabaquismo y antecedentes de tratamiento. Los autores reconocieron la posibilidad de RC como la muerte antes de la rehospitalización y aplicaron correctamente técnicas específicas para tener en cuenta la situación de RC.

Melberg et al¹³ estudiaron una cohorte de 1.234 pacientes con enfermedad coronaria sintomática que recibieron 2 tipos de tratamientos: cirugía de revascularización arterial coronaria (n = 594) o intervención coronaria percutánea (n = 640). De las 301 muertes observadas durante el seguimiento, el 42,5% fueron muertes cardiacas y el resto, muertes no cardiacas. Los autores presentan los resultados correspondientes a la mortalidad por todas las causas, así como los de mortalidad cardiaca y mortalidad no cardiaca. Señalan que el porcentaje de mortalidad por todas las causas es la suma de los porcentajes de mortalidad cardiaca y de mortalidad no cardiaca, estimados correctamente teniendo en cuenta los RC. Los autores resaltan la importancia de analizar cada uno de los eventos de interés en vez de combinarlos en la mortalidad total. Esta cuestión ha sido expuesta también a nivel más general por Mell y Jeong¹⁴.

Como podría inferirse de los ejemplos citados, la principal cuestión que se plantea cuando hay RC es si no tener en cuenta los RC y censurar las observaciones que correspondan a un RC o tenerlos en cuenta. Cuando no se tiene en cuenta los RC y se censura las observaciones de RC, el análisis se reduce a un escenario de tiempo hasta el evento «habitual». Dado que este tipo de análisis es muy conocido y teniendo en cuenta la disponibilidad de programas informáticos para realizarlo, muchos investigadores recurren a este enfoque, como se ha visto en los ejemplos citados. Sin embargo, hay acuerdo unánime, y no sólo entre los estadísticos^{2,15-18}, en cuanto a que la estimación de la probabilidad del evento sobrestima en este caso la probabilidad real. La siguiente cuestión que surge de manera natural es si pueden realizarse modelos dentro de estos límites (ignorar/censurar los RC). Esto resulta más ambiguo y más difícil de apreciar. Aunque este tipo de análisis puede no estar exento de valor, su interpretación tiene casi siempre importantes dificultades. La principal exigencia es que el evento de RC (cuyas observaciones se han censurado y se han mezclado con las observaciones censuradas verdaderas) debe ser independiente del evento de interés. Si es así, los resultados podrían interpretarse como el efecto de covariables cuando no existían eventos de RC. Sin embargo, este supuesto no suele ser válido y no puede verificarse ni ponerse a prueba. En conclusión, cada vez que las observaciones de RC son censuradas, la estimación de la probabilidad del evento es incorrecta y la interpretación del efecto de las covariables no está clara debido a la falta de conocimiento de la independencia entre el evento de interés y el evento de RC.

Cuando se realiza el análisis teniendo en cuenta el RC (y codificándolo de manera diferente del evento de interés o la censura), la probabilidad se estima correctamente y el modelo tiene una interpretación sencilla. No hay ningún supuesto de independencia que dificulte la interpretación. El coeficiente de una covariable estimada de esta forma constituye el efecto de esa covariable en las probabilidades observadas.

Varios autores^{19,20} han intentado comparar mediante simulaciones los dos enfoques en cuanto a la potencia de las pruebas. Sin embargo, el investigador tiene que tener presente que el principal problema es la interpretación de los resultados. Con independencia de lo potentes que sean las pruebas, el análisis debe dar respuesta a la pregunta del estudio.

ESTIMACIÓN DE LA PROBABILIDAD DEL EVENTO

Es una práctica común aplicar el método de KM para estimar la probabilidad de un evento. La fórmula típica para la estimación de KM es $\hat{S}(t) = \prod_{t_i \leq t} \frac{n_i - d_i}{n_i}$, en donde $t_1 < t_2 < t_3 < \dots$ son los puntos temporales ordenados en los que se ha observado el evento, n_i representa el número de pacientes en riesgo en ese momento t_i y d_i es el número de eventos en ese momento t_i .

Esta fórmula puede transformarse mediante manipulación algebraica para expresar la probabilidad del evento, de la siguiente forma:

$$\hat{P}(t) = 1 - \hat{S}(t) = \sum_{t_i \leq t} \frac{d_i}{n_i} \hat{S}(t_{i-1}). \quad (1)$$

En presencia de RC existen al menos dos tipos de eventos: el evento de interés, que se identifica con el subíndice e , y el evento del riesgo competitivo, que se identifica con subíndice c . Kalbfleisch y Prentice introdujeron la fórmula de la probabilidad de un evento de interés en presencia de RC:

$$\hat{P}_e(t) = \sum_{t_i \leq t} \frac{d_{ei}}{n_i} \hat{S}(t_{i-1}) \quad (2)$$

Es interesante señalar la relación entre (1) y (2). Dado que d_i es el número de todos los eventos en t_i , puede interpretarse como la suma del número de eventos de interés d_{ei} y el número de eventos de RC d_{ci} en ese momento t_i . En consecuencia, la probabilidad de cualquier tipo de evento puede descomponerse de la siguiente forma:

Probabilidad de todos los eventos

$$\begin{aligned} &= \sum_{t_i \leq t} \frac{d_i}{n_i} \hat{S}(t_{i-1}) = \sum_{t_i \leq t} \frac{d_{ei} + d_{ci}}{n_i} \hat{S}(t_{i-1}) \\ &= \sum_{t_i \leq t} \frac{d_{ei}}{n_i} \hat{S}(t_{i-1}) + \sum_{t_i \leq t} \frac{d_{ci}}{n_i} \hat{S}(t_{i-1}) = \hat{P}_e(t) + \hat{P}_c(t) \end{aligned} \quad (3)$$

Así pues, la probabilidad de todos los eventos puede descomponerse en las probabilidades de cada tipo de ellos.

Si se utiliza 1-KM para calcular la probabilidad de un evento de interés en presencia de RC, la supervivencia de todos los eventos en la fórmula (2) se sustituye por la estimación de KM basada en los eventos de interés únicamente. Esto comportará un sesgo en los resultados, como se indicará más adelante. El principal supuesto que se acepta para el uso del método de KM es que los pacientes censurados, si fueran seguidos durante un tiempo suficiente, acabarían sufriendo el evento. Sin embargo, cuando el método de KM se utiliza en presencia de RC, los pacientes que sufren tipos de eventos distintos del evento de interés suelen ser censurados a pesar de que dejen de estar en riesgo de sufrir el evento de interés.

Además, la elegante descomposición presentada en (3) no puede hacerse para la fórmula de 1-KM.

En situaciones aplicadas, es posible que haya otros varios tipos de eventos que no sean el de interés. En este caso, pueden agruparse todos bajo el paraguas de eventos de RC.

Ilustraremos con un ejemplo que el uso del método de KM no es apropiado en presencia de RC.

Descripción del ejemplo

Utilizaremos para ilustrarlo un conjunto de datos obtenidos para estudiar los efectos tardíos del tratamiento del linfoma de Hodgkin. El resultado principal es la hospitalización por causas cardíacas. El linfoma de Hodgkin es un tipo de cáncer que aparece sobre todo en adultos jóvenes. En sus fases iniciales es casi curable, con una supervivencia general a 10 años del 70%. Así pues, una cohorte de este tipo de pacientes es ideal para el estudio de los efectos secundarios a largo plazo del tratamiento. La serie de datos utilizada aquí es un subgrupo de una cohorte más amplia que se presentará en otra publicación. Además, los datos se han modificado para nuestros fines. Por ejemplo, en aras de una mayor simplicidad, conservamos en la serie de datos sólo a los pacientes que fueron tratados con quimioterapia o radioterapia y excluimos a los que recibieron un tratamiento combinado. Para aumentar la tasa de RC (muerte sin hospitalización cardíaca), incluimos a pacientes de todos los estadios. Se imputaron algunas fechas de seguimiento y de muerte. Dadas las modificaciones que se hicieron en los datos, no se puede extraer conclusiones clínicas de este análisis. Los datos presentados aquí se refieren a 689 registros con 93 hospitalizaciones cardíacas y 467 muertes.

Las tasas de hospitalización cardíaca y de muerte sin un evento cardíaco se calcularán con el empleo tanto del método de KM (1) como de la función acumulativa de incidencia (CIF) introducida por Kalbfleisch y Prentice²¹ para este fin (2).

El método de Kaplan-Meier aplicado a una situación de riesgos competitivos sobrestima la tasa real del evento

En la figura 1 se presentan las estimaciones de CIF y 1-KM para la hospitalización cardíaca del grupo tratado con quimioterapia solamente. La línea discontinua correspondiente a las estimaciones de 1-KM está por encima de la línea continua que representa las estimaciones de CIF. Puede demostrarse matemáticamente que 1-KM siempre sobrestima la probabilidad del evento. Un error de concepto frecuente es que las estimaciones de 1-KM son correctas

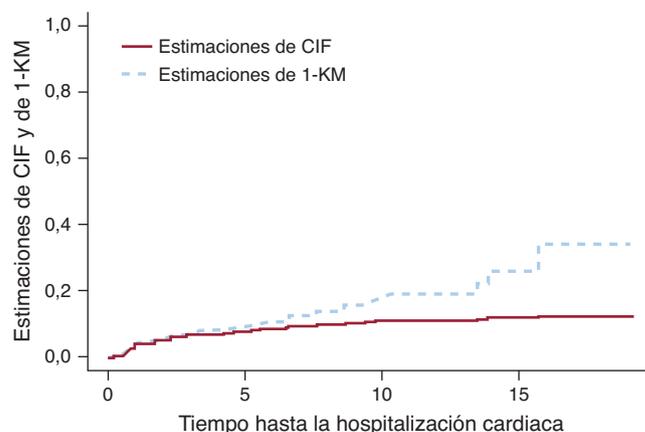


Figura 1. Estimaciones de la función acumulativa de incidencia frente a las de 1-Kaplan-Meier. CIF: función acumulativa de incidencia; KM: método de Kaplan-Meier.

si los dos eventos son independientes. La independencia de los eventos es siempre, en el mejor de los casos, cuestionable, pero incluso cuando los datos se simulan como eventos independientes, hay diferencia entre las estimaciones de CIF y 1-KM. La magnitud de la diferencia depende del número de eventos, tanto para los eventos de interés como para los eventos de RC. En el estudio de Miyasaka et al⁸, la incidencia de demencia a los 5 años con el empleo del método de KM fue del 10,5%. El número de RC (muertes) fue de aproximadamente tres cuartos del número total de eventos, lo que indica que su estimación puede ser muy superior a lo que se observa.

La función acumulativa de incidencia realiza una partición de la probabilidad de algún acontecimiento adverso (hospitalización cardíaca o muerte) en sus probabilidades componentes

Algebraicamente, esto se demuestra en (3). Sin embargo, para una mejor comprensión de la forma en que esto funciona, se mostrará gráficamente que, de la probabilidad de todos los eventos, una parte participa en la CIF para uno de los eventos y la otra, en la CIF para el otro evento. La figura 2A muestra la probabilidad de algún acontecimiento adverso: hospitalización cardíaca o muerte sin hospitalización cardíaca. La figura 2B incluye sólo la curva entre los 10,7 y 10,85 años, de manera que los pasos son visibles. En cada paso hay un círculo. Los círculos claros aparecen en los pasos en que se

observó una muerte, mientras que los círculos negros están en los pasos en que se produjo una hospitalización cardíaca. Los pasos con círculos negros participan en la CIF para la hospitalización cardíaca en la figura 2C y los pasos con círculos claros participan en la curva correspondiente a la muerte en la figura 2D. Así pues, cada paso contribuirá a la probabilidad del evento que causa. De esta forma, en cualquier momento concreto, la probabilidad de todos los eventos es la suma de la probabilidad del evento de interés y la probabilidad de RC. Obsérvese que los últimos tres paneles (figs. 2B-D) muestran la misma ventana temporal y tienen la misma longitud del eje y, de tal manera que puede compararse el tamaño de los pasos de cada uno de ellos. En la tabla 1 se indican estas probabilidades a 1, 2, 3, 4 y 5 años.

Dado que 1-KM sobrestima la probabilidad de un evento, si intentáramos añadir las estimaciones de 1-KM para la hospitalización cardíaca a la 1-KM para la muerte, obtendríamos una tasa muy superior a la probabilidad de aparición de algún acontecimiento adverso. En algunos casos, el número obtenido es incluso superior a 1, lo cual demuestra que, en presencia de RC, las estimaciones de 1-KM no son probabilidades adecuadas.

¿Realmente los métodos de función acumulativa de incidencia estiman la probabilidad correcta del evento?

Para este fin, se simuló un conjunto de datos de 500 registros, de tal manera que no hubiera ninguna censura antes de los 5 años y se

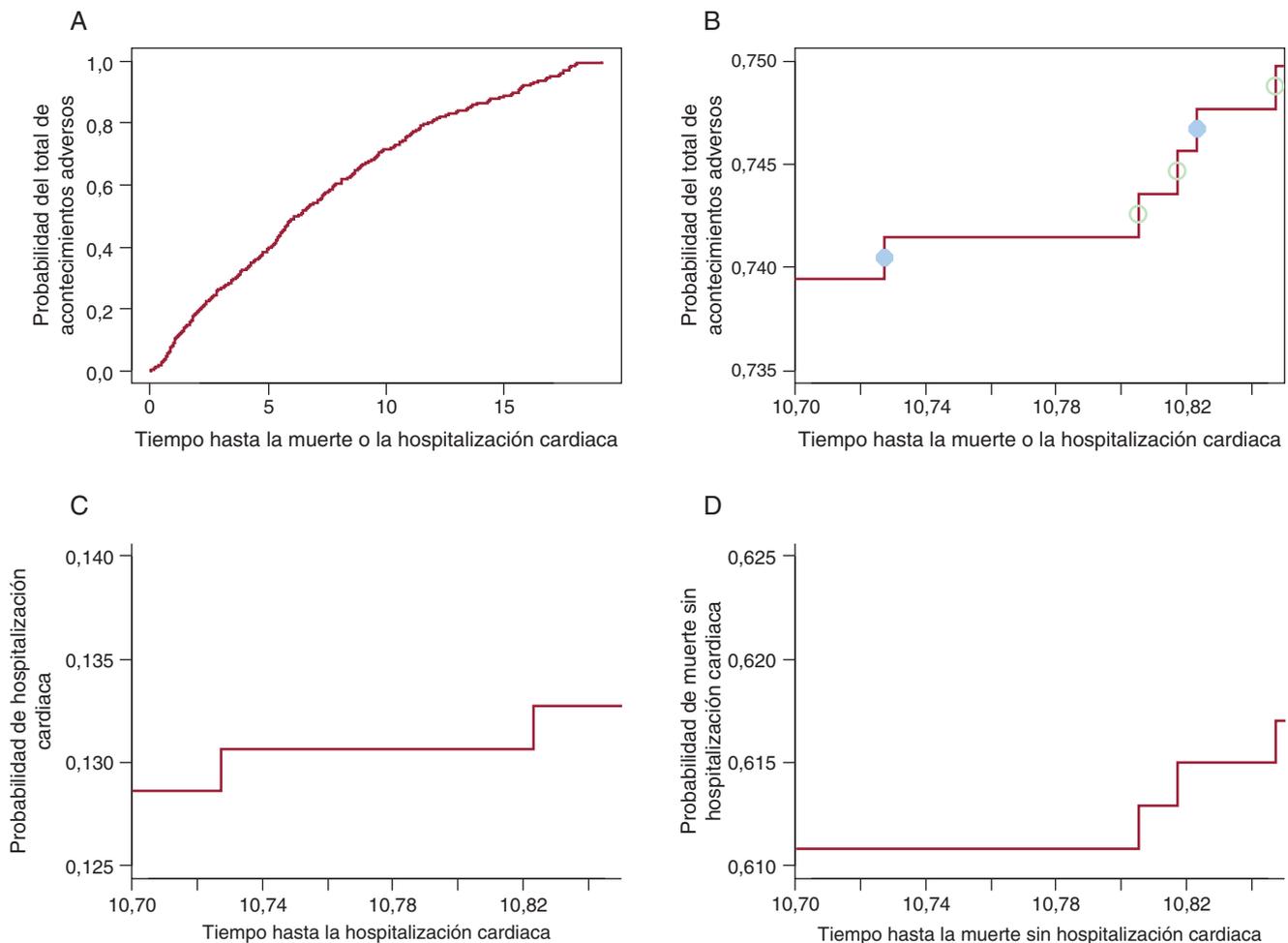


Figura 2. Partición de la probabilidad del total de acontecimientos adversos en sus probabilidades componentes. A. Probabilidad de hospitalización cardíaca o muerte. B. Probabilidad de hospitalización cardíaca o muerte únicamente para la ventana temporal de 10,7- 10,85 años. Los círculos negros indican la hospitalización cardíaca y los círculos claros indican las muertes sin hospitalización cardíaca. C. Probabilidad de hospitalización cardíaca en la ventana temporal de 10,7-10,85 años. D. Probabilidad de muerte sin hospitalización cardíaca en la ventana temporal de 10,7-10,85 años.

Tabla 1

La probabilidad de algún acontecimiento adverso es la suma de sus probabilidades componentes

Año de notificación	Probabilidad de hospitalización cardiaca	Probabilidad de muerte	Probabilidad de hospitalización cardiaca o muerte
1	0,038	0,054	0,092
2	0,054	0,139	0,193
3	0,072	0,193	0,265
4	0,076	0,25	0,327
5	0,087	0,305	0,392

consideran dos tipos de eventos: tipo 1 y 2. En la [tabla 2](#) se indica, para cada tipo de evento, el número observado hasta ese momento concreto, la tasa bruta y la estimación de la CIF, que son exactamente iguales. La igualdad sólo se produce cuando no hay observaciones censuradas hasta ese momento de valoración. En presencia de observaciones censuradas dentro de los años presentados, la igualdad no se da y la forma correcta de estimar la probabilidad es la CIF, y no la tasa bruta.

En conclusión, para calcular la probabilidad del evento en presencia de RC, es preciso utilizar el método introducido por Kalbfleisch y Prentice, al que se suele denominar curva de incidencia acumulativa.

MODELIZACIÓN

Un aspecto importante en un análisis es verificar la asociación entre una covariable y el evento de interés, considerado solo o introduciendo un ajuste para otros factores. En ausencia de RC, esto suele realizarse con el empleo de un modelo de riesgos proporcionales de Cox (PH Cox)²².

En presencia de RC, el modelo PH Cox no tiene una interpretación sencilla. Si el tiempo transcurrido hasta los dos tipos de eventos puede considerarse independiente, los resultados pueden interpretarse como indicativos del efecto en la situación en la que no existen RC. Sin embargo, el supuesto de independencia rara vez puede asumirse o verificarse y, por lo tanto, los resultados del modelo PH Cox no suelen ser interpretables.

Modelo de Fine y Gray

Fine y Gray⁶ (F&G) modificaron el modelo PH Cox para tener en cuenta la presencia de RC. La modificación técnica consiste en mantener las observaciones de RC en el conjunto de riesgo con una ponderación decreciente. De esta forma, el método F&G proporciona un modelo de los riesgos de subdistribución. El efecto estimado con el empleo del modelo F&G muestra las diferencias actuales y reales entre los grupos de tratamiento en términos de cocientes de riesgos de subdistribución. El supuesto de proporcionalidad de los riesgos continúa siendo una exigencia, pero naturalmente se refiere a los riesgos de subdistribución. El modelo F&G puede acomodar coeficientes dependientes del tiempo al modelo de no proporcionalidad de los riesgos. Este modelo puede aplicarse tanto al evento de interés (hospitalización cardiaca) como al RC (muerte).

Se aplicaron los modelos PH Cox y F&G al conjunto de datos de linfoma de Hodgkin con objeto de poner a prueba la opción de tratamiento de quimioterapia frente a la de radioterapia. En este ejemplo ([tabla 3](#)), los resultados de los modelos PH Cox y F&G difieren sustancialmente (primeras dos filas). Como se ha mencionado antes, los resultados del PH Cox no son interpretables y no pueden utilizarse. La segunda fila muestra que hay más

hospitalizaciones cardiacas en el grupo de radioterapia y la tercera fila indica más muertes en el grupo de quimioterapia. En la [figura 3](#) se muestran estos resultados gráficamente. Es posible que la quimioterapia sola se administrara a pacientes con enfermedad avanzada y que tuvieran también mayor probabilidad de muerte a causa del cáncer que padecían. En cambio, el tratamiento de radioterapia sola se aplicó probablemente a pacientes que se encontraban en un estadio inicial y que vivieron más tiempo tras el diagnóstico del linfoma de Hodgkin. Estos pacientes tenían una mayor probabilidad de sufrir efectos secundarios tardíos, como la enfermedad cardiaca.

Como puede apreciarse en este ejemplo, la interpretación de los resultados es un trabajo de colaboración entre el estadístico y el clínico, que tiene un conocimiento profundo de la enfermedad.

La presencia de RC complica tanto el análisis como la interpretación de los datos. Para permitir al lector una interpretación correcta de los resultados, los autores tienen que incluir información sobre los eventos observados aunque puedan no parecer importantes a primera vista. Así pues, cuando se observa la variable de valoración a lo largo del tiempo, los autores tienen que incluir el evento de interés, si es posible que haya RC, cuántos pacientes sufren cualquiera de estos tipos de eventos y la duración del seguimiento. En presencia de RC, resulta informativo incluir el análisis para el evento de interés, así como el análisis para los RC, ya que se complementan entre sí y podrían ser útiles en la interpretación de los resultados.

El enfoque logístico

Supongamos en primer lugar que estamos en un marco de referencia en el que no hay RC. Cuando se prevé que el resultado se produzca en corto tiempo (p. ej., 1 año), el instrumento de elección para muchos investigadores es la regresión logística. Esto es apropiado si cada individuo de la cohorte dispone del seguimiento mínimo, en este caso 1 año. De hecho, la estimación de la mortalidad a 1 año coincidirá con la estimación de 1-KM. El punto de corte temporal debe ser el mismo para cada individuo de la cohorte. Así pues, si el resultado de interés es la mortalidad a 1 año y un individuo de la cohorte fallece al año y 2 días, esa persona debe clasificarse como «ausencia de eventos a 1 año». Esto puede reducir el número de eventos, lo cual se traduce en un análisis que no es ideal cuando se producen muchos eventos observados después del momento de corte.

Las mismas reglas básicas son de aplicación cuando hay RC. Todos los individuos de la cohorte deben disponer del seguimiento mínimo elegido como momento de corte, y ese momento de corte debe aplicarse a todos los individuos de la cohorte. En general, los coeficientes y valores de *p* darán el mismo mensaje, pero no serán exactamente los mismos en la regresión logística y en el modelo F&G. En primer lugar, en la regresión logística el coeficiente representa el logaritmo de la *odds ratio*, mientras que en el modelo F&G es el logaritmo del cociente de las subdistribuciones de riesgos. Además, en el análisis logístico, se utilizan todos los eventos y, naturalmente, se utiliza un modelo diferente.

CÁLCULO DE LA POTENCIA ESTADÍSTICA

Cuando la medida utilizada es el tiempo hasta el evento, el cálculo de la potencia tiene dos etapas. El primer paso consiste en calcular el número de eventos necesarios para detectar una magnitud del efecto específica. A continuación, se calcula el número de pacientes necesario para observar ese número de eventos. En los apartados anteriores se ha hecho hincapié en que, cuando hay RC, no es posible observar todos los eventos de interés

Tabla 2

Probabilidad de los dos tipos de evento cuando no se censura ninguna observación hasta los 5 años

Año de notificación	Número de eventos de tipo 1	Tasa bruta de eventos de tipo 1 (%)	CIF para el evento de tipo 1 (%)	Número de eventos de tipo 2	Tasa bruta de eventos de tipo 2 (%)	CIF para el evento de tipo 2 (%)
1 año	31	6,2	6,2	39	7,8	7,8
2 años	49	9,8	9,8	74	14,8	14,8
3 años	62	12,4	12,4	105	21	21
4 años	76	15,2	15,2	140	28	28
5 años	87	17,4	17,4	170	34	34

CIF: función acumulativa de incidencia.

Tabla 3

Efecto del tratamiento cuando se utilizan los modelos de riesgos proporcionales de Cox y de Fine y Gray

Variable de valoración	Modelo	HR	IC del 95%	p
Hospitalización cardiaca	PH Cox	1,07	0,71-1,63	0,74
Hospitalización cardiaca	F&G	1,63	1,1-2,445	0,02
Muerte sin hospitalización cardiaca	F&G	0,38	0,31-0,47	< 0,0001

F&G: modelo de Fine y Gray; HR: razón de riesgos; IC: intervalo de confianza; PH Cox: modelo de riesgos proporcionales de Cox.

Las razones de riesgos muestran el aumento de los riesgos en el grupo de radioterapia en comparación con el grupo de quimioterapia.

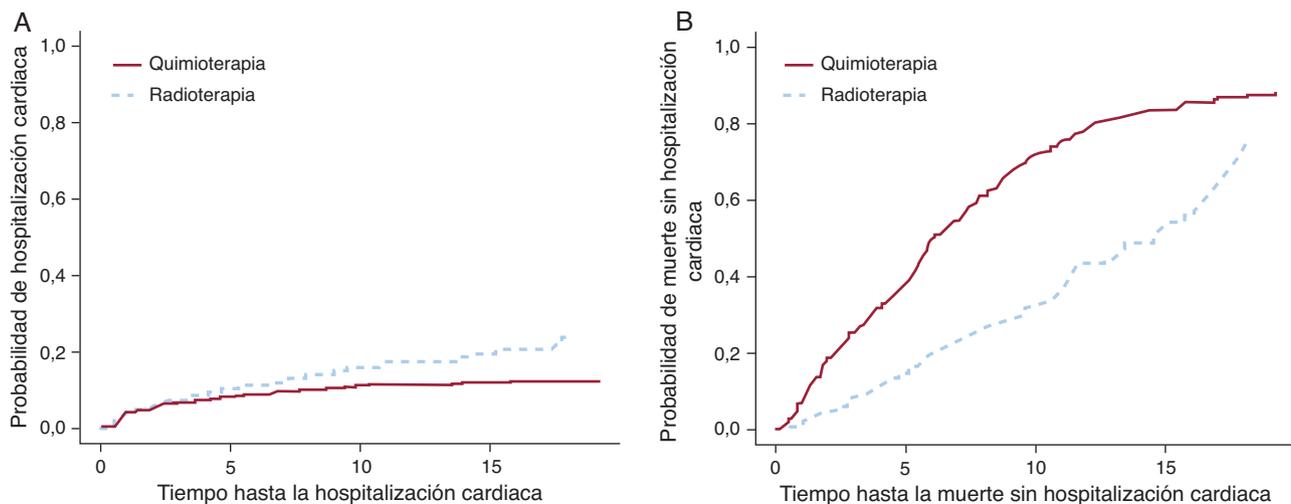
debido a la aparición de RC. Dado que el número de eventos ocupa un papel central en el cálculo de la potencia, es necesaria una precaución adicional para asegurar que se tiene en cuenta los RC. Si no se consideran los RC, el estudio tendrá una potencia estadística insuficiente y, por lo tanto, es probable que no dé resultado (y posiblemente no sea ético).

SOFTWARE

El programa informático R con código fuente abierto CRAN (*the Comprehensive R Archive Network*) (disponible en: <http://cran.r-project.org/>) ofrece un paquete (cmprsk) desarrollado por el Dr. Robert Gray que contiene los instrumentos necesarios para un análisis completo que tenga en cuenta los RC. Así, se podría obtener gráficos de probabilidad para el evento de interés y un valor de p basado en la prueba de Gray, que es un *log-rank test* modificado para verificar la situación de RC. En el paquete se incluye también una función para la modelización con el empleo del enfoque de

F&G. Luca Scruca mejoró la presentación de los resultados de la función de modelización para una lectura más fácil al incorporar al paquete una función de tipo resumen. El modelo aporta la posibilidad de verificar la proporcionalidad de los riesgos, y pueden incluirse términos para coeficientes dependientes del tiempo. El código no permite acomodar una truncación por la izquierda ni datos de agrupaciones. La truncación por la izquierda sería útil para el análisis de eventos múltiples/recurrentes por paciente o para el análisis de la cohorte de casos. Se ha desarrollado un código para estudios de cohortes de casos (Pintilie et al²³) que puede solicitarse a los autores. Zhou et al²⁴ ampliaron el modelo F&G para acomodar datos estratificados y habrá también una versión para datos con agrupamiento. En este momento, puede obtenerse el código solicitándolo a los autores en ambos casos, pero es probable que se envíe a CRAN.

El STATA 11 ha implementado recientemente el modelo F&G. Es preciso tener presente que los gráficos obtenidos con el empleo del STATA son predictivos, en vez de ser gráficos de probabilidad observada. Hay dos advertencias que tener en cuenta cuando se

**Figura 3.** Efecto del tratamiento en la hospitalización cardiaca y la muerte.

utilizan curvas de predicción: *a*) las líneas se mostrarán siempre como si se cumpliera la proporcionalidad de riesgos, y *b*) el número de pasos en cada curva será mayor que el número de eventos en cada subgrupo, lo cual creará la impresión de que hay más eventos de los que realmente hay.

CONCLUSIONES

La disponibilidad de series amplias de datos con un seguimiento completo para varias variables de valoración está aumentando continuamente. También aumenta la necesidad de análisis que se centren en una variable de valoración precisa, como la muerte por insuficiencia cardíaca o el control de la enfermedad o el control de la enfermedad local. Con todas estas variables de valoración, es posible que haya RC. Así pues, es esencial tener en cuenta los RC desde la fase de diseño hasta la interpretación de los resultados. Aunque el modelo PH Cox puede tener escaso valor cuando se considera una independencia, las estimaciones de KM no son correctas y no pueden interpretarse. Por consiguiente, es preciso aplicar técnicas específicas, como los modelos de CIF y F&G que proporcionan el programa R y parcialmente el programa STATA.

CONFLICTO DE INTERESES

Ninguno.

BIBLIOGRAFÍA

- David HA, Moeschberger ML. *The theory of competing risks*. London: Griffin; 1978.
- Gooley TA, Leisenring W, Crowley J, Storer BE. Estimation of failure probabilities in the presence of competing risks: New representations of old estimators. *Stat Med*. 1999;18:695-706.
- Pintilie M. *Competing risks: A practical perspective*. 1.^a ed. Chichester: Wiley; 2006.
- Pepe MS, Mori M. Kaplan-Meier, marginal or conditional-probability curves in summarizing competing risks failure time data? *Stat Med*. 1993;12:737-51.
- Gray RJ. A class of k-sample tests for comparing the cumulative incidence of a competing risk. *Ann Stat*. 1988;16:1141-54.
- Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. *J Am Stat Assoc*. 1999;94:496-509.
- Kaplan EL, Meier P. Nonparametric-estimation from incomplete observations. *J Am Stat Assoc*. 1958;53:457-81.
- Miyasaka Y, Barnes ME, Petersen RC, Cha SS, Bailey KR, Gersh BJ, et al. Risk of dementia in stroke-free patients diagnosed with atrial fibrillation: data from a community-based cohort. *Eur Heart J*. 2007;28:1962-7.
- Whalley GA, Wright SP, Pearl A, Gamble GD, Walsh HJ, Richards M, et al. Prognostic role of echocardiography and brain natriuretic peptide in symptomatic breathless patients in the community. *Eur Heart J*. 2008;29:509-16.
- Pfeffer MA, McMurray JJV, Velazquez EJ, Rouleau JL, Kober L, Maggioni AP, et al. Valsartan, captopril, or both in myocardial infarction complicated by heart failure, left ventricular dysfunction, or both. *N Engl J Med*. 2003;349:1893-906.
- Moukarbel GV, Signorovitch JE, Pfeffer MA, McMurray JJV, White HD, Maggioni AP, et al. Gastrointestinal bleeding in high risk survivors of myocardial infarction: the VALIANT Trial. *Eur Heart J*. 2009;30:2226-32.
- Núñez J, Sanchis J, Núñez E, Bodí V, Mainar L, Miñana G, et al. Insuficiencia cardíaca aguda post-alta hospitalaria tras un síndrome coronario agudo sin elevación del segmento-ST y riesgo de muerte e infarto agudo de miocardio subsiguiente. *Rev Esp Cardiol*. 2010;63:1035-44.
- Melberg T, Nygard OK, Kuiper KK, Nordrehaug JE. Competing risk analysis of events 10 years after revascularization. *Scan Cardiovasc J*. 2010;44:279-88.
- Mell LK, Jeong JH. Pitfalls of using composite primary end points in the presence of competing risks. *J Clin Oncol*. 2010;28:4297-9.
- Southern DA, Faris PD, Brant R, Galbraith PD, Norris CM, Knudtson ML, et al. Kaplan-Meier methods yielded misleading results in competing risk scenarios. *J Clin Epidemiol*. 2006;59:1110-4.
- Gelman R, Gelber R, Henderson IC, Coleman CN, Harris JR. Improved methodology for analyzing local and distant recurrence. *J Clin Oncol*. 1990;8:548-55.
- Satagopan JM, Ben-Porat L, Berwick M, Robson M, Kutler D, Auerbach AD. A note on competing risks in survival data analysis. *Br J Cancer*. 2004;91:1229-35.
- Putter H, Sasako M, Hartgrink HH, Van de Velde CJH, Van Houwelingen JC. Long-term survival with non-proportional hazards: results from the Dutch Gastric cancer Trial. *Stat Med*. 2005;24:2807-21.
- Varadhan R, Weiss CO, Segal JB, Wu AW, Scharfstein D, Boyd C. Evaluating health outcomes in the presence of competing risks a review of statistical methods and clinical applications. *Med Care*. 2010;48:S96-105.
- Freidlin B, Korn EL. Testing treatment effects in the presence of competing risks. *Stat Med*. 2005;24:1703-12.
- Kalbfleisch JD, Prentice RL. *The statistical analysis of failure time data*. 1.^a ed. New York: Wiley; 1980.
- Cox DR. Regression models and life-tables. *J R Statist Soc B*. 1972;34:187-220.
- Pintilie M, Bai Y, Yun LS, Hodgson DC. The analysis of case cohort design in the presence of competing risks with application to estimate the risk of delayed cardiac toxicity among Hodgkin lymphoma survivors. *Stat Med*. 2010;29:2802-10.
- Zhou B, Latouche A, Rocha V, Fine JP. Competing risks regression for stratified data. *Biometrics*. 2010. doi: 10.1111/j.1541-0420.2010.01493.x [Epub ahead of print].