

Hablemos de...

Análisis de supervivencia

GUADALUPE GÓMEZ Y ERIK COBO

Departament d'Estadística i Investigació Operativa. Universitat Politècnica de Catalunya. Barcelona. España.

El análisis de supervivencia es una herramienta imprescindible en la investigación clínica y epidemiológica. Gran número de artículos en las revistas más prestigiosas usan alguno de sus métodos. Aunque el origen del nombre se debe a la construcción de tablas de vida en el siglo XVII, sus aplicaciones se extienden al estudio del tiempo transcurrido entre dos eventos de interés, como podría ser la aparición de un acontecimiento adverso tras una intervención terapéutica o el tiempo transcurrido entre el inicio de una infección y su diagnóstico.

En el proceso de obtención de pruebas empíricas en las que basar las decisiones clínicas, el análisis de supervivencia abordará las tres situaciones clásicas que resuelve la inferencia estadística:

- Estudio univariante. Descripción y resumen de los tiempos de vida, usualmente a partir de la estimación e interpretación de las funciones de supervivencia y de riesgo. Estos análisis permiten a su vez predecir el comportamiento futuro de pacientes de características similares. Por ejemplo, habiendo observado que 9 de 10 pacientes intervenidos de una neoplasia

gástrica avanzada superaban los 6 meses de vida, ¿qué se sabe sobre lo que sucederá en futuros pacientes?

- Estudio bivariante. Comparación del patrón de supervivencia de dos poblaciones: ¿es cierto que los miembros de cierto grupo viven más tiempo y tienen mayor esperanza de vida o es simplemente un resultado casual observado en esta muestra, pero que no se repetirá en el futuro?

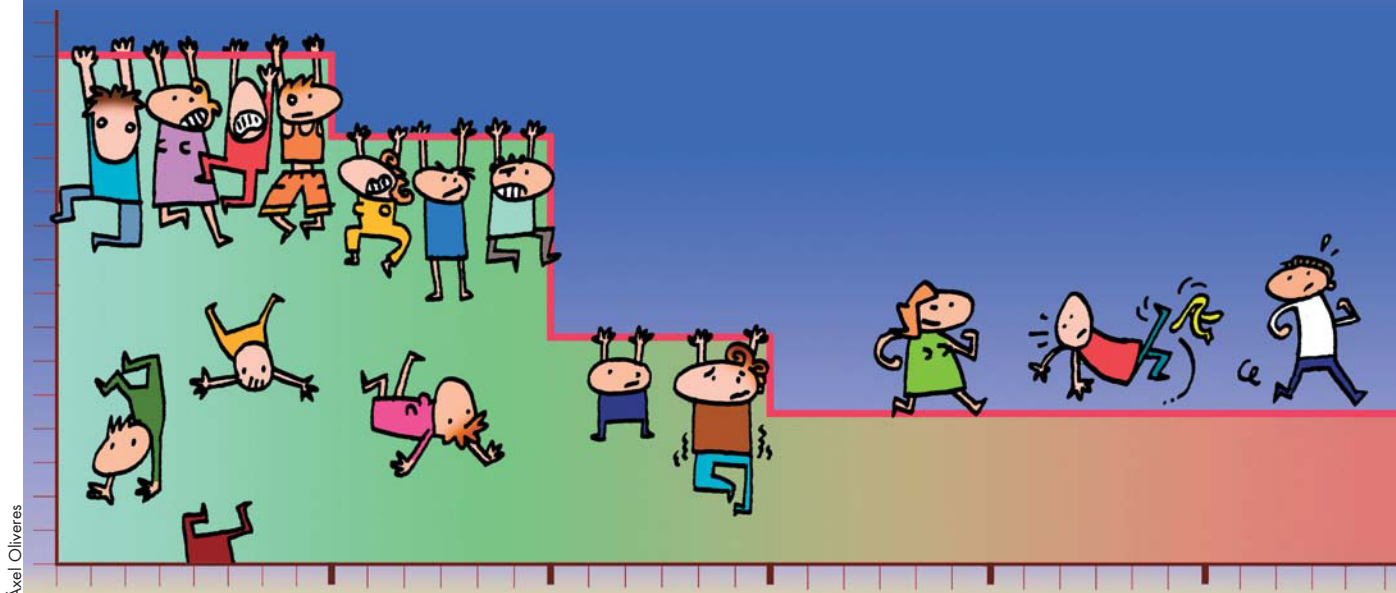
- Estudio multivariante. Construcción de un modelo que, teniendo en cuenta las características de los pacientes, ayude a predecir su tiempo de vida y a seleccionar los factores de riesgo que contribuyen a esta predicción una vez tenida en cuenta la influencia de otros factores más relevantes.

En estas líneas se abordarán, en primer lugar, dos características de las variables del tipo "tiempo entre dos eventos": la asimetría y la censura, ya que requieren un análisis específico. A continuación se presentan las funciones de supervivencia y riesgo, la comparación de curvas y el modelo de riesgos proporcionales de Cox. El lector que desee profundizar puede consultar la literatura específica¹⁻⁸.

Puntos clave

- El análisis de supervivencia estudia el tiempo entre dos eventos cualesquiera.
- La función de supervivencia indica la probabilidad de que un paciente supere cierto tiempo de vida. La función de riesgo (*hazard function*) representa la probabilidad, por unidad de tiempo, de presentar el evento en el lapso subsiguiente, condicionado a que éste no se había presentado antes.

- El método estadístico más usual para calcular las funciones de supervivencia y de riesgo es el del producto (Kaplan-Meier y el de tablas de vida (actuarial)).
- La prueba de hipótesis sobre si varias curvas provienen de la misma población puede realizarse con los métodos de *log-rank* y de Gehan.
- El modelo de riesgos proporcionales de Cox es el equivalente a la regresión lineal múltiple en el análisis de supervivencia.



Àxel Oliveres

Asimetría de la variable tiempo

A diferencia de otras variables, el tiempo sólo se desplaza en una dirección: el colesterol puede aumentar o disminuir, pero el tiempo nunca retrocede. Las distribuciones que gobiernan las variables del tipo tiempo presentan unas formas diferentes a las del resto de variables biológicas, que aconsejan planteamientos alternativos de análisis.

La figura 1 muestra la forma típica de esta variable para un ejemplo hipotético de meses de supervivencia tras el diagnóstico de una neoplasia de mal pronóstico: muchos casos tienen valores de unos pocos meses, pero algunos pacientes viven mucho, incluso uno de ellos llega a superar los 8 años. Nótese que la media de esta variable, como resultado de estas observaciones extremas, se sitúa alrededor de 15 meses, valor que no representa en absoluto la supervivencia habitual de esta muestra. La mediana, en cambio, tiene un valor de entre 10 y 11 meses, que resume mejor el tiempo de vida típico de estos pacientes. Lo mismo sucede con la desviación típica, cuyo valor, algo mayor de 15 meses, es demasiado alto para representar la distancia típica de los casos que no alcanzan la supervivencia media y es demasiado bajo para representar esta distancia con el centro para los casos que superan el promedio de vida. Nótese, finalmente, que la superposición de una distribución normal con esta media y desviación típica muestra que el modelo de Gauss-Laplace no se puede utilizar para representar estos casos: ¡el 95% de los pacientes vivirían entre -15 y +45 meses!

Ni la media, ni la desviación típica ni la distribución normal resumen bien el tiempo de supervivencia.

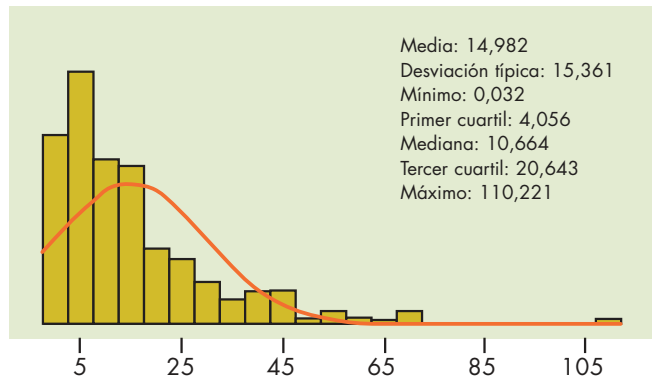


Figura 1. Histograma, con curva normal superpuesta, de los datos completos de supervivencia.

Censura

La segunda característica que presenta la variable tiempo es que requiere un largo período para ser observada en su totalidad. En el ejemplo anterior, a los 24 meses de seguimiento, ya se conocía el tiempo de vida exacto del 80% de los pacientes, pero se debería esperar aún más de 6 años para conocer el otro 20% restante, lo que restaría actualidad y valor científico a los resultados. Por esta razón, suele marcarse un tiempo de seguimiento máximo. Un diseño habitual consistiría en incorporar pacientes durante 12 meses y parar el estudio tras otros 12 meses de seguimiento. Ahora, a los pacientes incluidos el primer día del estudio se les habrá realizado un seguimiento de hasta un máximo de 24 meses. Y a los incluidos al final del primer año, hasta de 12 meses. Pero únicamente se conoce el tiempo exacto de vida de los casos que hayan tenido una supervivencia menor a su tiempo de seguimiento. Para el resto, se sabe que han vivido más que el tiempo de seguimiento observado, por lo que diremos que sus valores están censurados (por la derecha). La figura 2 muestra los tiempos que se habrían observado si se hubiera realizado un seguimiento de entre 12 y 24 meses a los pacientes de la figura 1. Nótese que, al haberse eliminado los mayores tiempos, tanto la media como la desviación tipo ofrecen resultados (10 y 7 en cifras redondas) muy diferentes de los reales de la figura 1 (15 y 15). Ya se ha comentado que estos indicadores no resumían bien a las variables asimétricas y ahora se muestra que, en el caso de datos censurados, proporcionan resultados sesgados, que no informan ni de la auténtica media ni de la desviación típica. En cambio, los datos estadísticos ba-

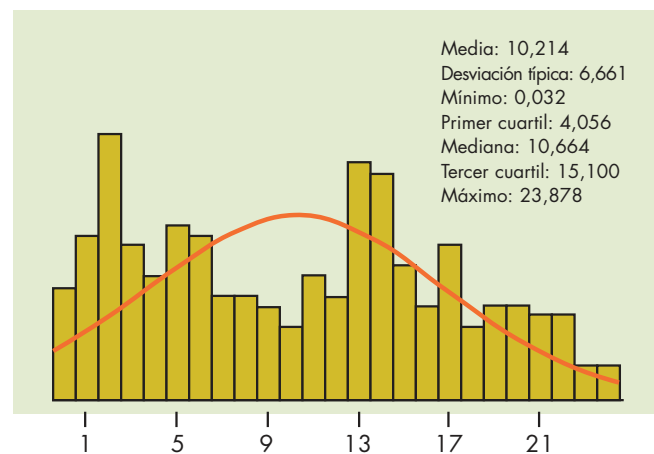


Figura 2. Histograma, con curva normal superpuesta, de los datos censurados de supervivencia.

sados en rangos proporcionan resultados fiables siempre que el porcentaje de datos no censurados supere el percentil que se desea estimar, en este ejemplo puede usarse la mediana o percentil 50, ya que hay un 69% de datos no censurados. Nótese que el valor estimado de la mediana coincide con el de la figura 1.

Un problema adicional acaece cuando la censura es el resultado de una pérdida (*lost to follow-up*) previa al final del período de seguimiento, especialmente si se desconoce la razón. Si, por ejemplo, al ver próximo su final los pacientes deciden “desaparecer” –quizá volver al

La censura se produce cuando el tiempo de supervivencia supera el período de seguimiento.

pueblo o cambiar de centro de soporte–, se desconocería su defunción y se obtendrían cifras optimistas sobre la supervivencia. Y si el individuo fallece por una causa diferente a la de interés (*competing risks analysis*) surge el problema de si es consecuencia o no de la enfermedad en estudio. La mayoría de los métodos usados en análisis de supervivencia presuponen que los individuos censurados están sujetos a la misma probabilidad de fallecer que los que permanecen en el estudio. Ésta es la llamada condición de no informatividad.

No normalidad de la variable tiempo

Gran parte de los análisis estadísticos se basan en la hipótesis de que la variable tiempo se distribuye según una ley normal. En este caso, la inferencia se reduce a la estimación de los dos parámetros que rigen esta distribución, la media y la varianza. Como este método no se adapta suficientemente bien a muestras formadas por tiempos de vida, hay que recurrir a otras distribuciones o a otras técnicas. Los procedimientos no paramétricos, es decir, aquellos que no imponen a los datos una distribución concreta, son la alternativa más usada en estudios de supervivencia. Además, las técnicas no paramétricas permiten tener en

cuenta el carácter secuencial de los datos, de forma que cada individuo sólo contribuye al estudio mientras está bajo observación, es decir, mientras no aparece la censura.

En resumen, que la variable de interés sea el tiempo y que éste se mida secuencialmente tiene como consecuencia una distribución asimétrica y la presencia de censura. Estas circunstancias desaconsejan el uso de la distribución normal, que tan bien caracterizan la media y la desviación típica. En el siguiente punto se explica cuál es la forma correcta de resumir y analizar este tipo de datos.

Función de supervivencia

Para resumir y analizar tiempos de vida se usa la denominada función de supervivencia. Esta función puede calcularse en cada instante y da cuenta de la probabilidad de que un paciente sobreviva el mismo.

La figura 3 representa cómo van variando, en el tiempo, la proporción de unidades que están libres del episodio de interés. Puede verse, por ejemplo, que en el mes 10, el 60% de los pacientes del estudio aún no lo había presentado. Y que el intervalo de confianza (IC) de lo que sucede en la población origen de la muestra es muy estrecho, indicando que se dispone de buena información (de hecho, este cálculo se basa en 1.000 pacientes). En estas líneas no se especifica cómo se calcula la función de supervivencia a partir de los datos pero informamos de que se ha obtenido mediante el estimador no paramétrico de Kaplan-Meier.

Nótese en esta función un descenso más marcado en los primeros meses que en los últimos, lo que podría indicar una mayor intensidad de presentación del episodio al inicio que al final. Sin embargo, este mayor descenso es explicable por el simple hecho de que al final se presentan menos episodios, ya que quedan menos casos: si, por ejemplo, el episodio es el fallecimiento, ¡los pacientes ya fallecidos no son susceptibles de volver a morir!

La función de supervivencia indica la probabilidad de que un paciente supere cierto tiempo de vida.

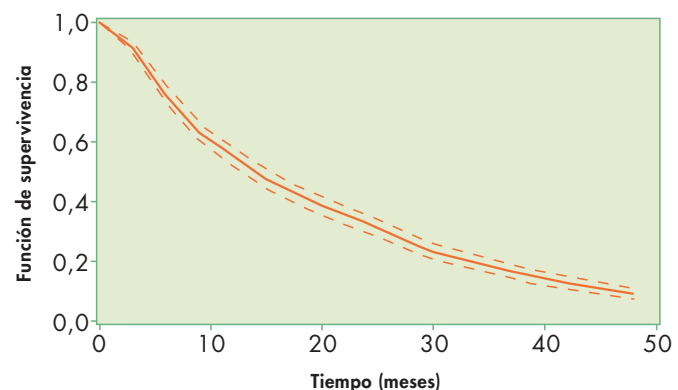


Figura 3. Función de supervivencia con su correspondiente intervalo de confianza del 95%.

Función de riesgo

Para hacer comparables los descensos del inicio (que se obtienen de una población mayor) con los descensos finales (provenientes de una población restante más pequeña) se calcula la proporción de casos que presentan el episodio en un momento determinado sobre la proporción de casos que llegan a ese momento, lo que se conoce como función de riesgo, fuerza de mortalidad o tasa condicional de fallo (*hazard rate*). Otra forma de interpretar la función de riesgo en el instante t es entenderla como la distribución (densidad) residual del tiempo de vida después de t .

En la figura 4 puede verse que la función de riesgo correspondiente a los datos presentados en la figura 3 es aproximadamente constante durante el período de seguimiento, con una tasa de fallo cercana al 0,05 mensual. Cuando la función de riesgo es constante se denomina tasa de riesgo. Un ejemplo de tasa de riesgo constante sería la de las hemorragias digestivas tras la administración de antiinflamatorios, de los que se dice que producen el mismo riesgo sea la primera o la enésima vez que se toman.

La forma de la función de riesgo dependerá del fenómeno estudiado. Un riesgo creciente corresponde a una población

La función de riesgo (*hazard function*) representa la probabilidad, por unidad de tiempo, de presentar el episodio en el lapso subsiguiente, condicionado a que éste no se había presentado antes.

Cuando la función de riesgo es constante se denomina tasa de riesgo.

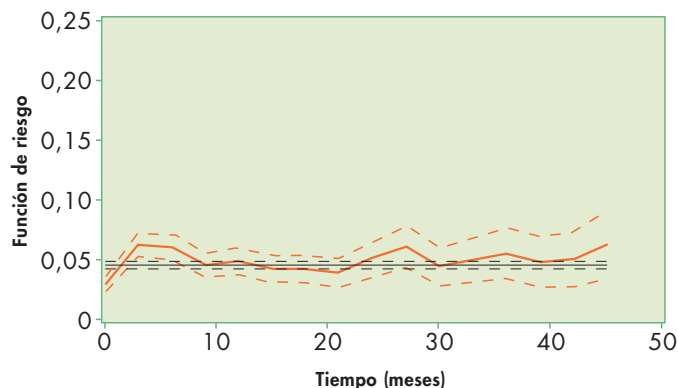


Figura 4. Función de riesgo (línea discontinua) y tasa de riesgo (línea continua) con sus correspondientes intervalos de confianza del 95%.

que envejece y se encuentra, por ejemplo, en el análisis del tiempo de vida de pacientes con leucemia que no responden al tratamiento (fig. 5a). Un riesgo decreciente corresponde a poblaciones en las que los individuos se fortalecen con el paso del tiempo, por ejemplo, después de una operación quirúrgica (fig. 5b). La función de riesgo puede adoptar otras formas, como la de bañera o la de joroba. La forma de bañera (fig. 5c) es apropiada como modelo para poblaciones que se controlan desde el nacimiento: al principio están las enfermedades infantiles, después se estabiliza y luego sigue un proceso creciente.

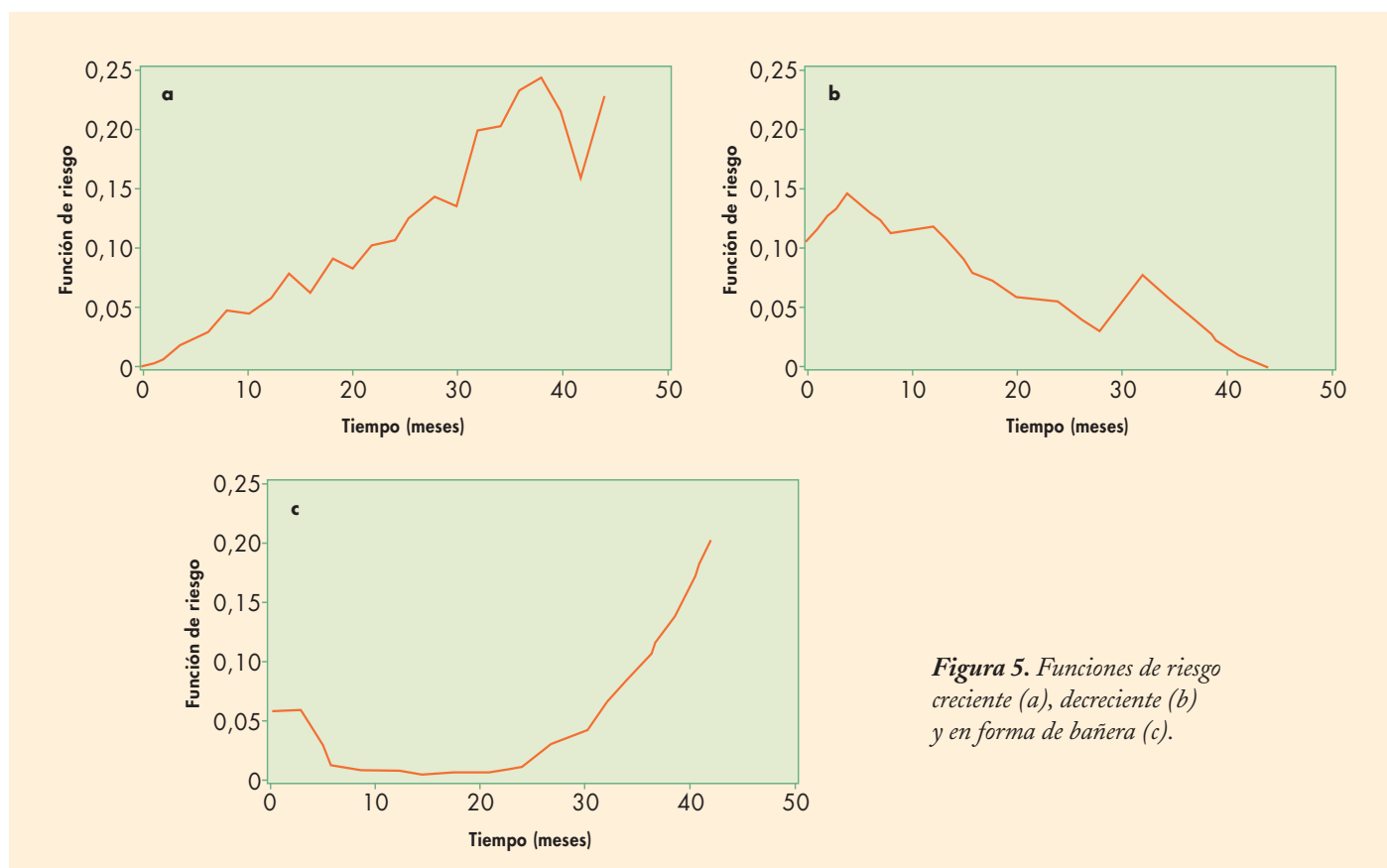


Figura 5. Funciones de riesgo creciente (a), decreciente (b) y en forma de bañera (c).

Estimación instantánea (Kaplan-Meier) y por intervalo (actuarial)

De la misma manera que se pueden calcular la velocidad instantánea o la velocidad “media” durante un período, la función de riesgo también puede corresponder a un instante o a un intervalo. El método de estimación empleado en la figura 4 es el llamado método actuarial (*life tables*), que calcula, en un intervalo, la proporción de casos que lo superan respecto al total de casos que lo inician. Una dificultad de este método la presentan los casos perdidos durante cada intervalo (están vivos al inicio pero no se sabe si lo están al final). La solución habitual considera que los individuos perdidos o que abandonan durante el estudio lo hacen, en término medio, a la mitad del intervalo y por tanto, que estuvieron en riesgo durante la primera mitad. Un planteamiento alternativo (estimador de Kaplan-Meier) consiste en “actualizar” la estimación de la función de supervivencia en cada momento en que aparece un episodio, y no al final de intervalos predefinidos. Este procedimiento es el más utilizado para estimar la función de supervivencia (fig. 3). El estimador para la función de riesgo

Los dos métodos estadísticos más usuales para calcular las funciones de supervivencia y de riesgo son el del producto (Kaplan-Meier) y el de tablas de vida (actuarial). Difieren en que el segundo predefine intervalos de tiempo.

Como en toda inferencia estadística, conviene acompañar las estimaciones de intervalos de confianza.

se calcula bien a partir del estimador de la función de supervivencia, bien a partir de la proporción de episodios entre los individuos en riesgo en cada instante (fig. 4).

Así, aun en el caso de que la tasa de fallo fuera constante, las fluctuaciones del muestreo originarán que las estimaciones difieran de un intervalo a otro. Por ello, para visualizar correctamente la información contenida en la muestra conviene representar también la banda de confianza que informa sobre los auténticos valores poblacionales de la tasa de fallo que son compatibles con los resultados muestrales observados. Nótese en la figura 4 cómo éste se va ampliando con el tiempo, y muestra la mayor incertidumbre originada por el menor número de casos que alcanzan las mayores cifras del seguimiento. Nótese también que, si se decide asumir que la tasa es constante durante todo el período de seguimiento, se podrán combinar todas las tasas mensuales en una sola, con el doble beneficio de parsimonia científica y eficiencia estadística: parsimonia porque se explica el mismo fenómeno de forma más simple; y eficiencia porque se obtienen estimadores más precisos, con un IC más estrecho.

Comparación de curvas

Para comparar dos curvas de supervivencia, una primera aproximación estudiaría un único tiempo de vida, por ejemplo, los 3 años. Sin embargo, este procedimiento desprecia el resto de información, por lo que necesita muestras amplias para encontrar diferencias. Además, la misma elección de este punto en el tiempo puede ser controvertida. Por todo esto, es más apropiado combinar la información de todos los tiempos para decidir si las curvas en comparación provienen de la misma población. Esta combinación puede hacerse otorgando la misma ponderación a todos los tiempos (prueba del *log-rank*, equivalente a la de Mantel y Haenszel para combinar tablas 2 x 2), o dando un peso

La prueba de hipótesis sobre si varias curvas provienen de la misma población puede realizarse con los métodos de *log-rank* y de Gehan.

mayor a los tiempos iniciales que tienen más observaciones (prueba de Gehan, generalización de la de Wilcoxon).

Estos dos procedimientos proporcionan un grado de significación (valor de p) que permite detectar y establecer la existencia de diferencias significativas y la supremacía de la supervivencia de una población con respecto a la otra. Sin embargo, la ausencia de un parámetro que cuantifique las diferencias entre ambas poblaciones, dificulta la evaluación de la relevancia clínica. En el siguiente apartado se expone cómo el modelo de Cox propone un parámetro que permite resumir las diferencias entre las poblaciones en comparación.

Modelo de riesgos proporcionales (regresión de Cox)

El modelo de Cox es el equivalente en análisis de la supervivencia al modelo de regresión lineal. Se trata de un modelo semiparamétrico, ya que no exige ninguna forma para la función de riesgo (puede ser creciente o decreciente, etc.) pero, en cambio, define un parámetro que es la razón entre ambas funciones de riesgo (*hazard ratio*). Este planteamiento responde de forma natural a la pregunta del clínico, que no ambiciona conocer cuál es el tiempo exacto de vida de un paciente determinado (parte del modelo no parametrizada), y centra sus esfuerzos en identificar las intervenciones que puedan aumentarlo (parte del modelo parametrizada). El modelo de Cox no impone como premisa una forma para la función de riesgo, pero sí asume que la razón de riesgo toma el mismo valor durante todo el período de seguimiento. Como ejemplo, se puede pensar que si se desarrolla un nuevo tipo de antiinflamatorio, es deseable que la reducción del riesgo de hemorragia digestiva se mantenga mientras se toma la medicación. En la figura 6 puede verse la forma que adopta la función de supervivencia y la función de riesgo cuando este efecto, es decir, esta reducción que mide la razón de riesgo es constante. Las funciones de supervivencia de ambos grupos se van separando progresivamente, resultado del mayor riesgo en uno de ellos. Por su parte, las funciones de riesgo son paralelas, con una cierta oscilación por las fluctuaciones del muestreo. La razón de riesgo de este ejemplo es 0,45 (IC del 95%, 0,38-0,52), lo que indica que uno de los grupos tiene un riesgo que es algo menor de la mitad que el correspondiente al otro grupo (la reducción es del 55%; IC del 95%, 48 al 62%). Obsérvese que el IC es razonablemente preciso. Una vez más, adoptar la premisa de riesgos proporcionales hará el modelo más parsimonioso y simple de comunicar, al mismo tiempo que aumenta su eficiencia estadística. Así, el médico estaría en condiciones de decir a su paciente: "mire,

La razón de riesgo (*hazard ratio*) estima el cociente entre las funciones de riesgo de dos curvas durante el período de seguimiento.

El modelo de riesgos proporcionales de Cox es el equivalente a la regresión lineal múltiple en el análisis de supervivencia.

yo no puedo decirle cuántos cumpleaños le quedan, pero puedo garantizarle que, si realiza la intervención que le propongo, su riesgo queda reducido entre un 48 y un 62%".

Por supuesto que esta premisa de proporcionalidad de los riesgos no tiene por qué ser siempre correcta y es obligación del investigador y del estadístico analizar hasta qué punto los datos permiten seguir empleándola. La figura 7 muestra una situación en la que el episodio tiene una tendencia diferente a presentarse en ambos grupos: mientras que en uno de ellos se presenta con más frecuencia en la primera mitad del seguimiento, en el otro grupo el episodio acontece más a menudo en la segunda mitad. Esta situación, sin duda más rebuscada, podría darse en un tratamiento con más de un mecanismo de acción, de forma que en una parte del seguimiento se observaran unos resultados contrarios a los de la otra parte. Por supuesto, en esta situación no sería correcto asumir un efecto constante del tratamiento, entre otras cosas porque ambos efectos se anularían mutuamente y no se observaría nada (razón de riesgo = 1,095; IC del 95%, 0,940-1,227).

El modelo de Cox puede considerar simultáneamente la hipotética relación de varias variables con la supervivencia y estudiar, de esta manera, si la relación de algunas persiste una vez se ha tenido en cuenta al resto. Para hacerlo, extiende la premisa de riesgos proporcionales a cada variable considerada, tanto variables cuantitativas como categóricas. La construcción del modelo es una tarea muy delicada, ya que se debe especificar la relación que cumple con esta premisa de proporcionalidad, al mismo tiempo que se evita introducir variables que tengan entre ellas una elevada colinealidad. Si el modelo está bien elaborado, su interpretación es más simple, ya que la razón de riesgo establecida para una variable puede interpretarse "independientemente", es decir, de manera fija, del resto de variables⁹⁻¹⁰. La decisión final sobre si una relación puede interpretarse como causal debe basarse en la existencia de la correspondiente hipótesis antes de la recogida de los datos, así como de su procedencia, observacional o experimental.

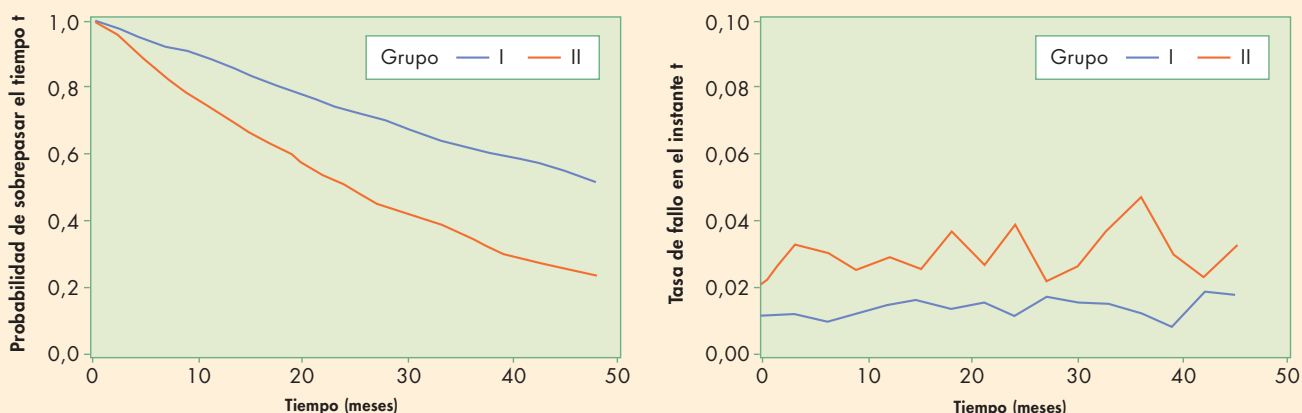


Figura 6. Forma de las funciones de supervivencia y de riesgo cuando la razón de riesgo es constante.

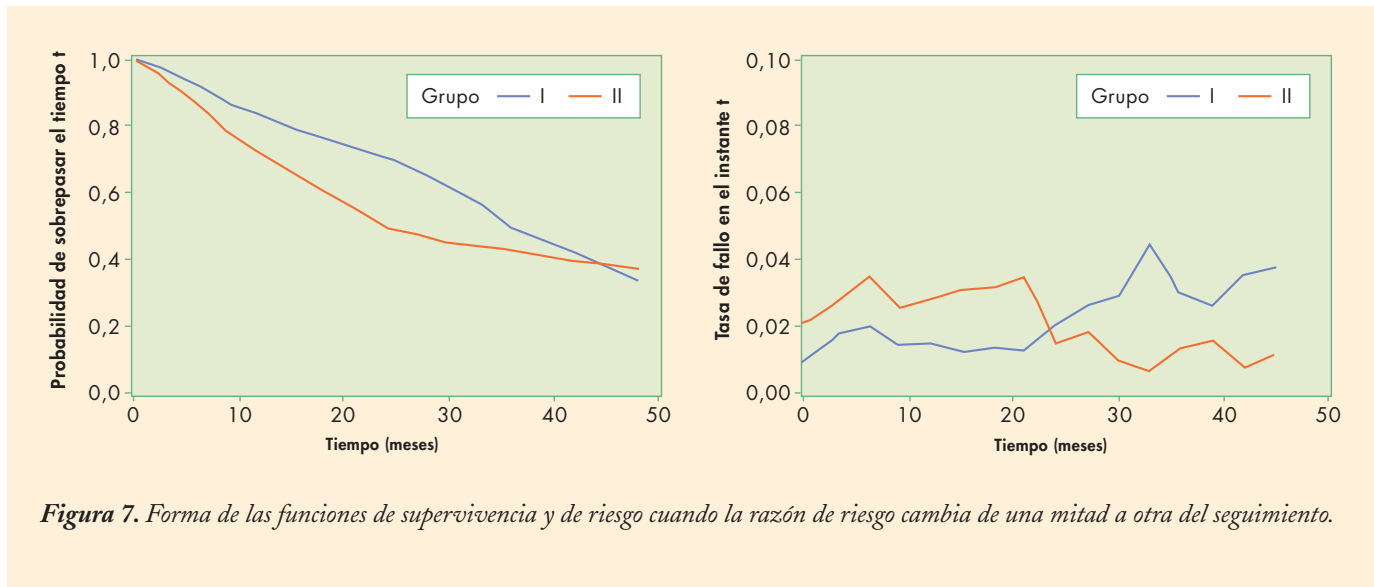


Figura 7. Forma de las funciones de supervivencia y de riesgo cuando la razón de riesgo cambia de una mitad a otra del seguimiento.

Conclusión

El análisis de supervivencia debe basarse en el estudio de las funciones de supervivencia y de las funciones de riesgo. El modelo de riesgos proporcionales de Cox, sin imponer condiciones sobre la forma de la supervivencia, permite al investigador disponer de un estimador del efecto de un tratamiento, junto con su IC, a partir del cual puede evaluar tanto la significación estadística como la relevancia clínica. Sin embargo, se trata de una metodología sofisticada, con premisas que deben ser evaluadas por un estadístico, y que requiere

cálculos intensivos. Afortunadamente, la incorporación al mercado de profesionales estadísticos, expertos en el uso de las modernas herramientas estadísticas e informáticas, ha facilitado que el análisis de supervivencia se haya convertido en una metodología habitual en gran parte de artículos de revistas médicas de prestigio. El clínico, convenientemente asesorado por un estadístico experto en ellas, tiene ahora la oportunidad de incorporar este potencial en su trabajo clínico e investigador.

Bibliografía



1. Aalen OO, Gjessing HK. Understanding the shape of the hazard rate: A process point of view. *Statist Sci* 2001;1:1-22.
2. Collet DS. *Modelling survival data in medical research*. Londres: Chapman and Hall, 1994.
3. Gómez, G. Técnicas estadísticas en análisis de supervivencia. *ICB digital 5* [consultado 17/06/2002]. Disponible en: <http://www.icf.uab.es/icbdigital/>
4. Gómez, G. *Análisis de supervivencia* Barcelona. Universidad Politècnica de Catalunya, 2004. ISBN 84-688-5607-X.
5. Hougaard P. *Fundamentals of survival data*. *Biometrics* 1999;55:13-22.
6. Kalbfleisch JD, Prentice RL. *The statistical analysis of failure time data*. New York: John Wiley and Sons, 1980.
7. Klein JP, Moeschberger ML. *Survival analysis. Techniques for censored and truncated data*. New York: Springer-Verlag, 1997.
8. Kleinbaum DG. *Survival analysis. A self-learning text*. New York: Springer-Verlag, 1996.
9. Cobo E, Corchero C. Ajuste: qué variables, cómo y cuándo. *Formación Médica Continuada en Atención Primaria* 2003;10:741-2.
10. Cobo E. Análisis multivariante en investigación médica: criterios para la inclusión de variables. *Med Clin (Barc)* 2002;119:230-7.

Actualización Enfermedades metabólicas

Enfermedad de Wilson

José Antonio Solís

Hemocromatosis

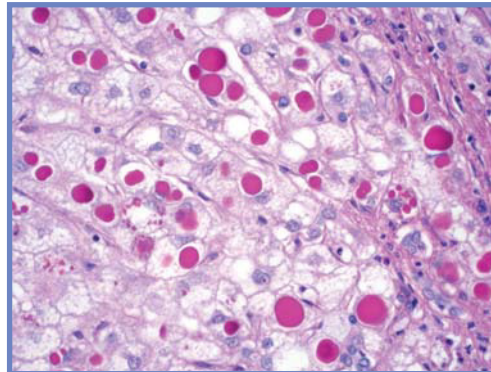
Alberto Pardo y Enrique Quintero

Déficit de alfa-1-antitripsina

J. Ignacio Herrero

Porfirias hepáticas

Raúl Andrade



El lugar en terapéutica de...

Inhibidores de la angiogénesis en el tratamiento del cáncer

Constantino Fontdevila

Revisión técnica diagnóstica

Pruebas de función pancreática exocrina

Salvador Navarro

Ensayos clínicos y práctica clínica

Cromoendoscopia para la detección de la displasia
en la enfermedad inflamatoria intestinal

Margarita Menacho y Adolfo Parra

Prevención de...

Programas de cribado en el cáncer colorrectal hereditario (I)

Francisco-Rodríguez-Moranta

Programas de cribado en el cáncer colorrectal familiar (y II)

Francisco Rodríguez-Moranta y Francesc Balaguer



Ventana a otras especialidades

Tromboembolia pulmonar

Manuel Monreal

Hablemos de...

El síndrome de *burnout*

Gemma Parramon y Joan de Pablo