

EDITORIALES

Curvas de supervivencia y modelos de regresión: errores y aciertos en la metodología de aplicación

M. Fernández, V. Abraira, C. Quereda y J. Ortuño

Servicio de Nefrología y Unidad de Bioestadística Clínica. Hospital Ramón y Cajal. Madrid.

INTRODUCCION

Existen diversos campos en la medicina en los que tiene interés estudiar la variable tiempo hasta que ocurre un determinado evento y su dependencia con otras variables explicatorias¹⁻⁵. Por ejemplo, en el estudio de enfermedades crónicas es importante conocer el tiempo desde el inicio de la enfermedad hasta que ocurre la muerte del enfermo (tiempo de supervivencia) y su dependencia con respecto a la aplicación de distintos procedimientos clínicos o terapéuticos. En otros casos tiene interés comparar diferentes tratamientos preventivos o factores de riesgo, por lo que se estudia el tiempo que tarda en aparecer un determinado evento.

La estimación de la supervivencia puede realizarse por técnicas no paramétricas (curvas de Kaplan-Meier¹) o paramétricas (estas últimas son de mayor complejidad matemática). Ambas permiten comparar curvas, y por tanto, si se generan grupos en función de otra variable, permiten estudiar su dependencia en la supervivencia. Este tipo de análisis es univariable y precisa categorizar la variable. La estimación de supervivencia con modelos de regresión tiene como ventaja que permite usar multivariados y variables continuas. Al igual que en el caso anterior, existen modelos no paramétricos, como el de Cox⁶, y no paramétricos, siendo estos últimos de mayor complejidad.

Curvas de Kaplan-Meier

Es un método para estimar, para cada tiempo, la probabilidad de que un evento ocurra después. E

evento en estudio puede ser muerte (curvas de supervivencia), fallo de un injerto renal, efectividad de un tratamiento, aparición de una complicación clínica, etc. Los métodos más usados son las Tablas de Vida o Curvas Actuariales y las Curvas de Kaplan-Meier¹. En las primeras, la medida de los eventos se realiza a intervalos fijos de tiempo, mientras que en las segundas el tiempo de medida es variable (esto es ventajoso, ya que los resultados no dependen de la elección de los intervalos).

Las **curvas de Kaplan-Meier**¹ estiman la probabilidad de que ocurra el evento para cada tiempo, junto con el correspondiente error estándar. También se representa el número de casos expuestos al riesgo en cada momento. Los programas estadísticos que permiten construir las curvas de supervivencia muestran datos numéricos y una representación gráfica de las curvas (tabla I y fig. 1). Incluyen los siguientes datos numéricos: 1) Tiempo: tiempo transcurrido en ocurrir la muerte, en desarrollar una determinada complicación clínica, en presentar una respuesta a un tratamiento, etc. El tiempo puede representarse en segundos, minutos, horas, días, meses o años. 2) Número de sujetos que continúan expuestos al riesgo al finalizar este tiempo. 3) Sucesos o eventos: número de casos que presentan la respuesta en ese tiempo (número de muertes, número de casos que han presentado una respuesta favorable a un tratamiento, etc.). 4) Pérdidas: número de individuos que han salido del estudio. 5) Función de supervivencia: se muestra la función de probabilidad de supervivencia acumulada para ese tiempo. Como en cada momento se estima la función de supervivencia, se aporta el correspondiente error estándar que permite cuantificar la precisión en dicha estimación. 6) Función de riesgo: estima en cada tiempo la probabilidad de que suceda el evento en ese instante, condicionado a que no haya ocurrido antes (no todos los programas calculan este parámetro).

Correspondencia: Dr. C. Quereda.
S. Nefrología. Hospital Ramón y Cajal.
Ctra. de Colmenar, Km. 9,100
28029 Madrid

Análisis de supervivencia. Kaplan-Meier

Tiempo	Sujetos	Eventos	Pérdidas	F. riesgo	F. superv.	Varianza
413.0	7	1	0	0,1429	0,8571	0,0175
1075.0	6	1	0	0,1667	0,7413	0,0292
1801.0	5	1	0	0,2000	0,5714	0,0350
3044.0	4	1	0	0,2500	0,4286	0,0350
3351.0	3	1	0	0,3333	0,2857	0,0292
5551.0	2	1	0	0,5000	0,1429	0,0175
6277.0	1	1	0	1,0000	0,0000	—

Intervalos de confianza al 95% de la función de supervivencia

Tiempo	Método de Greenwood	Transformación logarítmica
413.0	1,1164	0,5979
1075.0	1,0489	0,3796
1801.0	0,9390	0,2048
3044.0	0,7952	0,620
3351.0	0,6204	0,0489
5551.0	0,4021	0,1164
6277.0	—	—

Comparación de las curvas por la prueba «Long-Rank». J-cuadrado: 3,22; G.L.: 1; p = 0,069

Tabla 1.—Tabla de datos para el cálculo del análisis de supervivencia por el método de Kaplan-Meier.

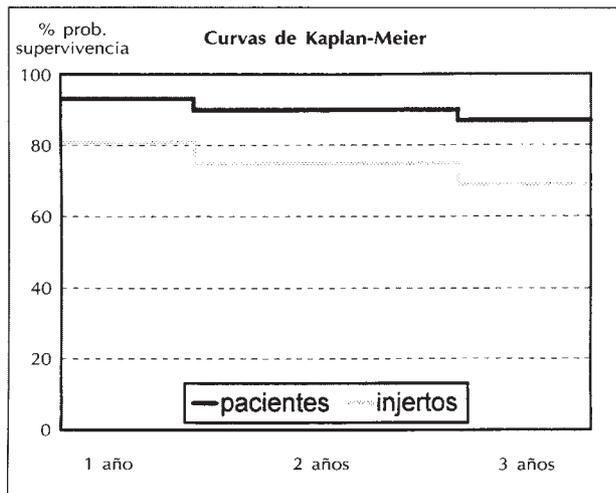


Fig. 1.—Representación gráfica de una curva de Kaplan-Meier.

Para comparar si existen diferencias estadísticamente significativas en las curvas de supervivencia existen distintas pruebas, de las cuales la más usada es la prueba del Logrank.

La **prueba del Logrank** se basa en la diferencia entre los casos observados (ej.: fallecidos reales) y los casos esperados si no existiera diferencia entre los grupos de estudio. Si el valor de la diferencia es

pequeño, no hay razones para pensar que existen diferencias entre los dos grupos; sin embargo, si el valor es elevado, es más lógico pensar, según criterios probabilísticos, que la diferencia entre los dos grupos no es debida al azar.

Es importante considerar que esta prueba compara de forma global una curva de supervivencia con otra, es decir, se aplica en todo el rango de observación. La representación gráfica incluye las dos curvas de supervivencia de los grupos a comparar (fig. 1). Si *a priori* se quiere saber si existen diferencias de supervivencia en un tiempo dado (por ejemplo, conocer la mortalidad precoz al año de tratamiento), esto se podrá hacer a partir de la función de supervivencia en ese punto de los grupos a estudiar y la varianza (error estándar elevado al cuadrado).

FACTORES QUE INFLUYEN Y CONDICIONAN LOS RESULTADOS DE LAS CURVAS DE SUPERVIVENCIA

1. Tiempo de seguimiento de los individuos

Hay que considerar que la variable dependiente a estudiar es el tiempo hasta que ocurre un suceso y está definida por el intervalo que transcurre entre el momento en el que empieza la observación y el tiempo real en el que ocurre el evento. En la mayoría de los estudios, la observación no comienza en el mismo tiempo para todos los sujetos, y si el estudio se ha diseñado para que acabe en un tiempo determinado, los casos que entraron más tarde tienen un menor tiempo de observación. Por ejemplo, si se analiza la supervivencia de los enfermos que entran en hemodiálisis, la observación de cada sujeto empezaría con la entrada en diálisis de cada paciente, y obviamente, como éstos van entrando en fechas diferentes, los últimos en entrar tendrán un período más corto de observación. En estos últimos casos si no ocurre un evento, no se sabrá si no ocurre porque existe poca probabilidad de que suceda o porque no se ha observado el suficiente tiempo para que ocurra.

Actitudes a seguir ante este problema

A) Diseñar el estudio para que el período de observación acabe no en la misma fecha para todos los casos, sino, para cada individuo, un tiempo determinado después del inicio de la observación. De esta forma todos los sujetos estudiados tendrán el mismo tiempo de seguimiento.

B) Si el diseño del estudio se ha hecho para que acabe en un tiempo determinado, dependiendo de criterios clínicos podrían eliminarse del estudio los casos que tengan una evolución muy corta. Un efecto parecido se consigue si se acorta el tiempo de observación, consiguiendo así aumentar la precisión del estudio al final del mismo. Por ejemplo, si el seguimiento medio de un grupo de enfermos en estudio es de 10 años y existe un número muy pequeño de casos que se les ha seguido durante 15 años, la precisión al estimar supervivencia a los 15 años será pequeña, debido al tamaño de la muestra. Si el número de casos es pequeño, se detecta la falta de precisión en esta parte de la curva al encontrar errores estándar grandes.

C) Del apartado anterior se deduce que al estimar la probabilidad de supervivencia en un tiempo dado no se puede afirmar nada con precisión si no se incluye en los resultados el «error estándar».

El error estándar nos informa de los márgenes de seguridad en los que se mueve la estimación de la probabilidad de supervivencia dada. Así, si la probabilidad de supervivencia de un paciente en diálisis es de 0,70 (70%) a los 5 años con un error estándar de 0,03, esto implica que la supervivencia del 70% puede oscilar, con una confianza del 95%, entre 76% ($0,70 + 0,03 \times 1,96$) y 64% ($0,70 - 0,03 \times 1,96$), es decir, es una estimación bastante precisa (si los rangos oscilaran entre 40 y 90% sería una estimación muy imprecisa).

D) Cuando se va a estudiar la posibilidad de que ocurra una muerte o un evento clínico en una enfermedad, el tiempo de observación debe estar en función del tiempo que con criterios clínicos se sabe que evoluciona dicha enfermedad. Por ejemplo, si se quieren estudiar las complicaciones médicas que acontecen en una enfermedad crónica que evoluciona en años, no debería diseñarse un estudio de supervivencia cuyo tiempo de observación fueran meses.

Por lo tanto, el conocimiento clínico de la enfermedad a estudiar es el que condiciona el tiempo necesario de seguimiento en el estudio. Además, como las variables tiempo de supervivencia suelen presentar una distribución asimétrica, como medida de centralización es más útil la mediana que la media.

2. Pérdidas de casos durante el seguimiento

Cuando se realiza un seguimiento de una cohorte de pacientes existen casos en los que se pierde su evolución durante el estudio. El método de Kaplan-Meier tiene la ventaja de que permite utilizar la información de las pérdidas, siendo el período de seguimiento el tiempo que transcurre desde que el

sujeto entra en el estudio hasta que se pierde su evolución.

3. Selección de los pacientes en estudio y homogeneidad de las muestras

El método de Kaplan-Meier permite calcular la probabilidad de supervivencia de una población estratificando en función de que cumpla o no ciertas condiciones. Por ejemplo, estimar la probabilidad de supervivencia de pacientes en diálisis según sean o no hipertensos. Pues bien, el método de Kaplan-Meier asume que el muestreo aleatorio hace que los grupos a comparar sean homogéneos con respecto a todas las demás variables, a excepción de la variable a estudiar (en nuestro ejemplo, las muestras serían homogéneas a excepción de la variable hipertensión arterial). Sin embargo, esta asunción no siempre es así, pues el muestreo aleatorio sólo garantiza que las muestras homogéneas son las más probables.

4. Variables no medibles que condicionan supervivencia: «Efecto centro» y «Efecto aprendizaje»

Efecto centro

Es bien conocido que existen factores relacionados con el centro hospitalario y los profesionales sanitarios que pueden condicionar de manera más decisiva la supervivencia de los enfermos que las propias variables del enfermo (por ejemplo, la selección individualizada del tipo de enfermos que llega a un centro, el excelente cuidado de los pacientes, la mayor experiencia en ciertas pautas terapéuticas, etc.). Además, en el mismo sentido es posible que una variable que tiene valor pronóstico para la población global pueda tener un significado pequeño en una situación local. ¿Cómo se puede entonces generalizar las variables que condicionan la supervivencia para la población general y que no sean exclusivas de la muestra analizada? Básicamente existen dos formas:

1. *Diseño de estudios multicéntricos y prospectivos:*

Al analizar muestras de distintos centros hospitalarios, si se encuentran variables con valor pronóstico, podremos tener certeza de que las variables predictivas encontradas no son específicas de un determinado centro.

2. *Validación externa del modelo:*

Si el estudio se ha realizado en un único centro hospitalario, se debe validar el modelo con muestras de enfermos de otros centros para tener seguridad de que dicho modelo se cumple en la pobla-

ción general (por ejemplo, si se encuentra que la cifra sérica de colesterol condiciona mortalidad en la población en estudio de un centro, se debería contrastar si en poblaciones homogéneas, pero de centros distintos, se cumple dicho hallazgo).

Efecto aprendizaje

Las innovaciones en tratamientos médicos y el inicio de nuevas técnicas quirúrgicas suelen acompañarse del abandono de protocolos estándar, y por un período de tiempo los resultados inmediatos pueden no ser tan favorables como los que se alcanzan al adquirir mayor experiencia. Este exceso de riesgo transitorio, que aparece asociado al inicio de un nuevo tratamiento, puede dar información falsa y mostrar como deletéreo un efecto que sería beneficioso al ganar experiencia. Obviamente, los datos del período de aprendizaje y consolidación de la experiencia deben ser recogidos como control de calidad interno del equipo; sin embargo, si lo que se pretende es comparar supervivencia entre el tratamiento antiguo y nuevo, deberán compararse éstos cuando ambos estén consolidados, excluyendo del estudio los casos tratados durante el período de aprendizaje.

5. Análisis retrospectivos y prospectivos

Los estudios prospectivos garantizan una mejor recogida de los datos; sin embargo, en muchos estudios retrospectivos, la recogida de la información es incompleta, dando lugar a bases de datos inconsistentes. Además, en muchos de estos estudios es difícil controlar cuándo empieza el proceso, salvo que sean episodios muy significativos en la historia clínica de la enfermedad (así, por ejemplo, en la evolución de la insuficiencia renal crónica, la fecha de entrada en diálisis, fecha de realización de un trasplante renal, etc., son eventos muy significativos que siempre quedan recogidos en la historia clínica y que se pueden recoger con seguridad en un estudio retrospectivo).

En general, cuando se diseña un estudio que pretende demostrar causalidad (por ejemplo, efecto beneficioso de un determinado fármaco en la evolución de una enfermedad), se deben diseñar estudios prospectivos. Cuando se realizan estudios de supervivencia de enfermedades crónicas que cursan en años, como el tratamiento sustitutivo de la insuficiencia renal, puede estar justificado el uso de análisis retrospectivos, ya que sería difícil seguir de forma prospectiva una evolución de 10 ó 15 años. En estos casos, sólo una metódica recogida de la información podrá garantizar la calidad de los resultados.

6. Categorizar variables biológicas

Cuando se generan curvas de supervivencia, las variables a analizar deben ser categóricas. Por ejemplo, si se pretende analizar cómo influye la edad en la supervivencia, se puede categorizar la variable edad en décadas, en mayores y menores de 65 años, etcétera. Si se analiza hipertensión arterial, se puede categorizar como ausencia o presencia de hipertensión o con un número mayor de categorías, como sería: ausencia de hipertensión, presencia de hipertensión controlada con medicación y presencia de hipertensión no controlada con medicación.

Como es evidente que la mayor parte de las variables biológicas son continuas, categorizar estas variables de forma arbitraria puede no ajustarse a la realidad y, por tanto, la estratificación debe realizarse con sensatez y guiados siempre por un criterio clínico. Se debe buscar, con criterios clínicos, el punto de corte que haga más evidente las diferencias entre los grupos. Después se puede contrastar estadísticamente qué diferencias son más significativas. Por ejemplo, si se quiere conocer la supervivencia en función de la edad, se realizarán curvas de Kaplan-Meier utilizando la variable edad categorizada en décadas; en menores de 40 años, entre 40-65 y mayores de 65 años, etc., y al final se elegirá el modelo que más se ajuste a la realidad y que muestre diferencias más evidentes. Las diferencias entre los grupos quedarán contrastadas por el grado de significación estadística. A pesar de ello, es posible que al hacer discreta una variable continua se esté perdiendo información, conduciendo a resultados poco significativos.

Los modelos de regresión permiten analizar supervivencia con variables continuas, si bien no garantizan que se ajuste más a la realidad que cuando se categoriza la variable.

7. Se elegirá el tipo de análisis en función del objetivo del estudio

Esta norma básica, aunque evidente, no siempre es llevada a cabo. Matas y cols.³ utilizan varios ejemplos donde se infringe esta norma. Por ejemplo, si se pretende demostrar que un mayor número de identidades en el tipaje HLA donante-receptor se asocia con un menor número de pérdidas de injertos renales de causa inmunológica, no deberían incluirse en este estudio los injertos renales perdidos por recidiva de la enfermedad de base, los injertos nunca funcionantes, los injertos perdidos por problemas técnicos, etc., ya que crean confusión en la interpretación de la hipótesis inicial.

En un sentido parecido, Matas y cols.³ encuentran que los factores de riesgo para perder un injerto renal varían dependiendo de que se considere la muerte del enfermo como causa de pérdida del injerto o no. Este aspecto de enorme interés refleja cómo dependiendo de la selección *a priori* de los pacientes, de la aceptación o no de casos con mayor morbilidad, de los criterios utilizados para definir la pérdida del injerto, etc., se está influyendo en las representaciones gráficas de las curvas de supervivencia y en los factores de riesgo encontrados.

MODELOS DE REGRESION

Como ya se ha comentado, cuando se quiere comparar supervivencia entre dos grupos resulta difícil obtener dos muestras homogéneas en todos los factores que puedan afectar al resultado en estudio, salvo el factor de interés, ya que ello conduciría a muestras tan pequeñas que sería imposible demostrar nada con un nivel de probabilidad suficiente.

Por otra parte, el estudio de los factores que condicionan supervivencia pueden conducir a interpretaciones erróneas cuando dichos factores están interrelacionados entre sí. Por ejemplo, la variable nefropatía de base puede estar relacionada con el desarrollo de hipertensión arterial, y si se pretende analizar si la nefropatía de base condiciona supervivencia tendremos que medir el grado de interrelación que existe entre estas dos variables para saber con certeza si lo que condiciona supervivencia es la nefropatía de base o su asociación con hipertensión arterial.

Por ello resulta interesante disponer de algún método de análisis que nos permita manejar toda la muestra disponible, considerando el efecto de todas las posibles variables explicativas en el cálculo de la supervivencia. Los modelos de regresión permiten estudiar una variable dependiente (supervivencia, muerte, aparición de cardiopatía isquémica, etc.) en función del tiempo y otras posibles variables explicativas (edad, enfermedad de base, hipertensión arterial, etc.). En otras palabras, permite relacionar covariables predictivas y probabilidad para que ocurra un determinado evento.

Un modelo de regresión ampliamente usado es el desarrollado por Cox⁶, que se denomina también modelo proporcional. El modelo permite calcular el riesgo relativo en función de posibles variables explicativas y asume que, a medida que pasa el tiempo, el riesgo de los dos grupos a comparar varía, pero de forma proporcional el uno con el otro. Esto es una limitación del modelo, ya que no siempre se cumple. Por tanto, antes de aplicar el modelo se

debe verificar que efectivamente el riesgo varía de forma proporcional (al hacer las curvas de Kaplan-Meier para cada uno de los grupos, si el riesgo varía de forma proporcional, las curvas deben ser paralelas y no cruzarse).

En la actualidad está en fase de investigación el desarrollo de otros modelos de regresión que no precisan que se cumpla dicha asunción, pero aún no están disponibles en los paquetes informáticos.

Los modelos de regresión pueden usarse con dos objetivos⁷:

1) **Predictivo**: en el que el interés del investigador es predecir lo mejor posible la variable dependiente (muerte, aparición de infarto agudo de miocardio), en función de un conjunto de variables independientes (hipertensión arterial, cardiopatía isquémica, etc.).

2) **Estimativo**: en el que el interés se centra en estimar la relación de una o más variables independientes con la variable dependiente. El objetivo estimativo se utiliza con mayor frecuencia en los estudios etiológicos en los que se pretende encontrar factores determinantes de una enfermedad o proceso.

Estrategias de modelización

Existen 3 estrategias de modelización o formas de desarrollar el modelo de regresión⁷:

1. Estrategia «hacia detrás» o Backward (en terminología anglosajona).

Es la más habitual y la recomendada en la mayoría de los textos. Se aplica introduciendo el número máximo de variables que interese analizar (covariables) y se van quitando del modelo de una en una siguiendo los siguientes criterios: a) si el objetivo es predictivo, se quitan las variables dependiendo del grado de significación estadística, es decir, se quitan aquellas variables que tengan un menor grado de significación estadística hasta conseguir un modelo en el que todas las variables sean estadísticamente significativas. Como el objetivo es predictivo, este conjunto de variables finales serán las que mejor predigan la variable dependiente (supervivencia, probabilidad de que ocurra un IAM, etc.).

b) Si el objetivo es estimativo, se van quitando las variables en función de que sean o no de confusión, y en el modelo final quedaría la variable de estudio y todas las variables de confusión para la misma. Aquellas variables estadísticamente significativas, pero que no sean de confusión, desaparecerán del modelo.

2. Estrategia «hacia delante» o Forward

A diferencia del anterior, se parte de una única variable y se van introduciendo todas las demás va-

riables hasta conseguir el modelo definitivo. La inclusión de variables se hace como en el supuesto anterior, dependiendo de que se utilice un objetivo predictivo o estimativo. Hay que considerar que el modelo final puede quedar con variables estadística y no estadísticamente significativas.

3. Estrategia «Stepwise»

Es una modificación de la anterior para evitar que el modelo final quede con variables no estadísticamente significativas. Cada vez que se añade una variable se verifica si alguna de las variables introducidas es no significativa, en cuyo caso se quitaría.

Veamos qué resultados da un determinado paquete estadístico para el cálculo de un modelo de Cox (Programa Presta, V. Abaira y cols. 1992) (tabla II).

Regresión de Cox

Variable	Beta	Exp (beta) (RR)	EE beta	J ²	p
HTA	-1,12317	0,322331	0,60086	0,50017	0,51351
Edad	0,147141	0,15852	0,092482	0,53127	0,1072
Log. máx. verosimilitud sin variables =				-18,27550	
Log. máx. verosimilitud modelo completo =				-14,77048	
Intervalos de confianza al 95% de los riesgos relativos					
HTA	0,01398	7,42989			
Edad	0,96645	1,38876			

Tabla II.-Tabla de datos para el cálculo de un modelo de regresión de Cox.

El programa muestra para cada variable independiente a analizar: 1) coeficiente de la variable (si este coeficiente es negativo indica un efecto protector, y si es positivo implica un factor de riesgo); 2) exponencial del coeficiente (es el riesgo relativo (RR) asociado a la modificación en una unidad de la variable. Si el RR es menor que 1 implica que la variable tiene un efecto protector, y si es mayor que 1 que es factor de riesgo. Por ejemplo, si el RR para la variable hipertensión arterial es 2, implicaría que los hipertensos tienen doble riesgo que los no hipertensos); 3) error estándar del coeficiente; 4) valor de la J-cuadrado y de la p asociado al mismo para contrastar la hipótesis nula (indica si la variable que se estudia como variable predictiva es estadísticamente significativa).

El programa también especifica la J-cuadrado del modelo que informa sobre la significación estadística del modelo completo.

La J-cuadrado score contrasta la J-cuadrado del modelo y ambas deben mostrar valores similares. Si

la muestra es pequeña puede observarse disparidad entre estos dos valores.

El número de variables independientes que se pueden estudiar depende del programa estadístico utilizado (por ejemplo, Presta admite hasta 20). El programa también presenta los intervalos de confianza al 95 % de los riesgos relativos y la matriz de covarianzas de los coeficientes (esta última es útil si se estudian variables Dummy y/o interacción).

CONSIDERACIONES QUE HAY QUE TENER AL APLICAR EL MODELO MULTIPLE DE REGRESION DE COX

1. Antes de realizar el modelo de regresión múltiple se deberá realizar un análisis simple de cada una de las variables a estudiar mediante una curva de Kaplan-Meier para observar el comportamiento de esa variable de forma individual. Es interesante comprobar cómo una variable, que condiciona mortalidad al analizarla de forma individual, pierde su significación estadística al hacerla interrelacional o al ajustarla con otras variables.

2. **Interacción y confusión en la regresión logística.**

Estos dos conceptos tienen que ver con la interferencia que una o dos variables pueden realizar en la asociación entre otras. Existe confusión cuando la asociación entre dos variables difiere significativamente según que se considere o no otra variable, y a esta última variable se la denomina variable de confusión. Existe interacción cuando la asociación entre dos variables varía según los diferentes niveles de otra u otras variables.

3. **Colinealidad en regresión logística.**

Si se va a aplicar un modelo de regresión en un determinado paquete estadístico para estudiar variables que condicionan supervivencia, el programa solicita introducir la variable dependiente (supervivencia) y el número de variables independientes a estudiar (HTA, cifras de colesterol, etc.).

Existen problemas de colinealidad cuando una variable independiente es función de otras variables o están relacionadas entre sí. Por ejemplo, esto puede ocurrir al introducir variables que expresen el mismo fenómeno clínico. Analicemos el siguiente caso: se quiere estudiar la cardiopatía isquémica, claudicación intermitente, accidente cerebrovascular agudo y arteriosclerosis como variables que condicionan mortalidad. Al realizar el modelo de regresión podrían producirse problemas de colinealidad, ya que la variable arteriosclerosis engloba a las otras. En estos casos el paquete estadístico informa que la matriz de covarianzas es singular y el modelo es irresoluble. La

solución de este problema es eliminar las variables que sean función de otras (en el ejemplo se eliminaría la arteriosclerosis). Hay que considerar que quitar estas variables colineales no supone una pérdida de información, puesto que esta información está englobada en las otras variables.

Si la colinealidad es por interacción de dos variables, se debe quitar la variable interacción. Existen también problemas de colinealidad cuando el número de variables a estudiar es mayor que el número de casos.

4. *Casi colinealidad en regresión logística.*

Si existe este problema conduce a estimadores poco precisos. Se puede detectar al observar riesgos relativos relevantes no estadísticamente significativos y con errores estándar grandes. En este caso se pueden originar modelos matemáticos inestables muy sensibles a pequeños cambios (por ejemplo, modelos en los que cambian de forma significativa los riesgos relativos con cambios pequeños en el número de casos).

5. *Uso inadecuado de variables.*

Un problema frecuente que se plantea cuando se utiliza el modelo de Cox es el manejo de las variables tiempo dependientes, ya que en estos casos el riesgo puede no variar de forma proporcional con el tiempo. Por ejemplo, supongamos que la cifra de albúmina al inicio de una enfermedad condiciona mortalidad, pero a su vez, a medida que la enfermedad avanza, las cifras de albúmina son cada vez menores. La cifra basal de albúmina podría utilizarse como factor predictivo de mortalidad (planteamiento correcto). Sin embargo, si pretendemos aplicar el modelo con las cifras de albúmina que se obtienen a lo largo de la evolución, una albúmina baja puede ser la consecuencia de una peor evolución de la enfermedad y, por tanto, no sería marcador predictivo, sino consecuencia de la propia enfermedad.

Por tanto, si se analiza este tipo de variables realizando un uso inadecuado del modelo se puede incurrir en errores de inferencia en el análisis de los datos^{8, 9}.

Vamos a analizar con mayor detalle algunos de los errores que se cometen con mayor frecuencia:

1. Utilización de variables tiempo-dependientes para predecir eventos futuros.

Como ya se ha explicado, el modelo de Cox permite calcular el riesgo relativo para que ocurra un evento en cada instante de tiempo después de la entrada en un estudio. Debido a que la probabilidad de que ocurra dicho evento en un tiempo determinado es una consecuencia, su valor tiene que depender sólo de información pasada. En este tipo de estudios es fundamental controlar qué información futura conocida no se está utilizando para definir cri-

terios de inclusión y que no se utiliza información futura para estimar la probabilidad de que ocurra el evento. Veamos un ejemplo para clarificar este concepto. Consideremos un estudio de pacientes con deterioro progresivo de la función renal. La medida de la función renal residual al inicio del seguimiento podría ser un factor de riesgo que condicionara supervivencia, en el sentido de que una menor función residual se asociara con mayor morbilidad (esto sería una forma correcta de plantear el modelo). Sin embargo, supongamos ahora que el modelo se plantea de la siguiente forma: considerar como predictores de mortalidad la media de los valores de función renal residual que se tienen a lo largo del seguimiento. Debido a que la función renal residual típicamente desciende hasta cero entre los pacientes con enfermedad renal progresiva, los pacientes que sobreviven más tiempo tienden a tener una media de función renal inferior, y podría concluirse erróneamente que una menor función renal residual se asocia con mejor supervivencia.

Este ejemplo demuestra cómo el uso de información futura para construir una variable inapropiada (en el ejemplo, la media de todas las funciones renales residuales de un paciente) puede conducir a conclusiones erróneas.

En resumen, los eventos clínicos que aparecen a lo largo de la evolución no deben ser usados para construir una variable que, incluida en un modelo proporcional, sirva para predecir eventos futuros.

2. No se debe ajustar simultáneamente por variables basales y variables tiempo-dependientes en el mismo modelo de regresión. En general, es seguro ajustar por características basales, pero se debe tener mayor precaución cuando se incorporan variables tiempo-dependientes al análisis.

Esto no significa que estas variables no sean útiles y se puedan utilizar en ciertas ocasiones. Por ejemplo, son de gran utilidad al comparar distintos tratamientos. En estos casos hay que considerar que durante el proceso de aprendizaje puede aumentar de forma transitoria el riesgo en el grupo donde se aplica el nuevo tratamiento conduciendo a modelos no proporcionales, por lo que si se aplicara el modelo de forma puntual reflejaría el exceso de riesgo puntual. Si se realiza un seguimiento a más largo plazo se evidencia cómo los riesgos relativos cambian de forma no proporcional con el tiempo.

DISCUSION

Las técnicas estadísticas aplicadas a la investigación clínica constituyen un instrumento de gran utilidad en las ciencias biológicas. Las curvas de su-

pervivencia y los modelos de regresión son técnicas ampliamente utilizadas, y en nefrología han sido una herramienta útil y única para estimar la supervivencia de los enfermos en diálisis o con trasplante renal, así como la supervivencia del injerto renal funcional. Sin embargo, como toda técnica matemática, el uso inadecuado de la misma, así como el desconocimiento en algunos casos de los principios básicos de su aplicación, ha conducido a resultados que, siendo estadísticamente significativos, se alejan de forma sustancial de la realidad clínica. Por otra parte, y a pesar del uso creciente por parte de la comunidad nefrológica de estos métodos de análisis, las aportaciones científicas referentes a estas técnicas en la revista de nuestra Sociedad han sido prácticamente inexistentes¹⁰.

Matas y cols.³ han demostrado que los factores de riesgo que influyen en la supervivencia a largo plazo del injerto renal varían dependiendo de que se considere la muerte del enfermo como causa de pérdida del injerto, y recomiendan a los centros trasplantadores reconsiderar la forma de abordar el análisis de los datos. Además, hay que tener presente que estos hallazgos estadísticos permiten comparar pronósticos entre centros trasplantadores y, más aún, ayudan a tomar decisiones clínicas y a elegir pautas de tratamiento. El objetivo de esta reflexión ha sido poner de manifiesto que, dependiendo del tipo de diseño que se haga *a priori*, se puede influir en los resultados de las curvas de supervivencia, y, por tanto, cuando se trate de comparar análisis de distintos equipos hay que tener presente si las variables de estudio han sido objeto de la misma consideración.

En el mismo sentido hay que tener presente que la interpretación de las curvas de supervivencia elaboradas por distintos equipos requieren ser observadas con un análisis crítico. En este capítulo se han descrito los aspectos más importantes que influyen en la generación de las curvas, y son estos factores los que hay que buscar para una correcta interpretación de los datos. Se debe conocer cómo se ha hecho la selección de los pacientes, cuál es el seguimiento medio de los casos, cómo se ha definido el suceso o evento clínico, cuántos casos se han perdido en la evolución, qué sesgos se han cometido en la distribución de los grupos, cuál es el error estándar de la estimación de la curva, cómo se han

categorizado las variables continuas, cuál ha sido el rigor en la recogida de los datos en los estudios retrospectivos, si existen variables de confusión, etc.

También hay que analizar de forma crítica las curvas de supervivencia elaboradas por nosotros. ¿Quién mejor que el que realiza el estudio es capaz de saber el rigor del método utilizado? Las limitaciones del estudio deben ser expuestas en la metodología de trabajo, porque éstas no quitan brillantez al estudio y garantizan transparencia y control de calidad. En otras palabras, no se trata de demostrar mejores supervivencias, sino de acercarnos con el método estadístico a la realidad de los enfermos.

Por último, se debe identificar con exactitud y prudencia qué factores encontrados con las curvas de supervivencia están influyendo en la toma de decisiones clínicas y si los parámetros encontrados se ajustan a la experiencia y realidad clínica. En este sentido no hay que olvidar que los factores de riesgo encontrados en una muestra de enfermos pueden ser consecuencias del efecto centro y, por tanto, precisan de una validación externa para dar consistencia al modelo.

Bibliografía

1. Kaplan E, Meir P: *J Am Stat Assoc* 53: 457, 1958.
2. Guttman RD: The graft survival curve: ideology and rhetoric. *Transplant Proc* 24 (6): 2407-2410, 1992.
3. Matas AJ, Kristen J y cols.: Half-life and risk factors for kidney transplant outcome. Importance of death with function. *Transplantation* 55: 757-761, 1993.
4. Appleton DR: Analysis of data in nephrology. Survival analysis and analysis of variance. *Nephrol Dial Transplant* 3: 340-344, 1988.
5. Port FK, Wolfe RA y cols.: Comparison of survival probabilities for dialysis patients versus cadaveric renal transplant recipients. *JAMA* 270: 1339-1343, 1993.
6. Cox DR: Regression models and life-tables. *J R Stat Soc* 34: 187-202, 1972.
7. Abaira V, Pérez de Vargas A: *Métodos multivariantes en Bioestadística*. Ed. Centro de Estudios Ramón Areces, 1996.
8. Wolfe RA, Strawderman RL: Logical and statistical fallacies in the use of Cox Regression Models. *Am J Kidney Dis* 27 (1): 124-129, 1996.
9. Gjertson DW, Terasaki PI: The large center variation in half-lives of kidney transplants. *Transplantation* 53: 357, 1992.
10. Barrio V, Madero R: Estadística en la investigación en medicina. *Nefrología* vol IX, n.º 3: 236-242, 1989.