

# Cálculo de la supervivencia relativa. Comparación de métodos de estimación de la supervivencia esperada

Ramón Clèries<sup>a</sup> / Josepa Ribes<sup>a</sup> / Víctor Moreno<sup>a</sup> / Laura Esteban<sup>a</sup> / Laura Pareja<sup>a</sup> / Jordi Gálvez<sup>a</sup> / José Miguel Martínez<sup>b</sup> / Francesc Xavier Bosch<sup>a</sup> / Josep M. Borràs<sup>c</sup>

<sup>a</sup>Servei d'Epidemiologia i Registre del Càncer, Institut Català d'Oncologia, L'Hospitalet de Llobregat, Barcelona, España. <sup>b</sup>Unitat de Recerca en Salut Laboral, Departament de Ciències Experimentals i de la Salut, Universitat Pompeu Fabra, Barcelona, España;

<sup>c</sup>Pla Director d'Oncologia de Catalunya. España.

(Relative survival computation. Comparison of methods for estimating expected survival)

## Resumen

El método más utilizado en el cálculo de la supervivencia de los pacientes diagnosticados de cáncer es la supervivencia relativa. Este método tiene en cuenta la estimación de la supervivencia esperada de dichos pacientes a partir de la mortalidad observada en la zona geográfica de la que proceden. Los métodos utilizados con mayor frecuencia en dicha estimación son los de Ederer (I y II) y el método de Hakulinen. Para dichos cálculos es necesaria la disponibilidad de las tablas de supervivencia estratificadas por edad y por período de calendario. En este trabajo se presenta un ejemplo en el que se muestra cómo realizar dichos cálculos de tal forma que permitan al investigador decidir qué método es el más adecuado según el tipo de análisis que lleve a cabo. Se mostrará que, si el tiempo de seguimiento de la cohorte es inferior a 10 años, cualquiera de los métodos debería dar resultados similares. Sin embargo, el método recomendable es el de Hakulinen, ya que tiene en cuenta la heterogeneidad en el momento de las censuras debidas a posibles pérdidas de seguimiento o abandonos.

**Palabras clave:** Supervivencia relativa. Supervivencia esperada. Retirada del estudio. Supervivencia a largo término.

## Abstract

Relative survival is the most commonly used method to determine survival in patients diagnosed with cancer. This method takes into account estimation of expected survival in cancer patients based on the observed mortality in the geographical area to which they belong. The most frequently used methods for estimation of expected survival are the Ederer (I and II) and Hakulinen methods. Survival tables for the geographical areas stratified by age and calendar year are required for these calculations. The present article presents an example of how to perform these estimations and how to choose the most appropriate method for the type of analysis to be performed. This article shows that if the follow-up of the cohort is less than 10 years, any of these methods should give similar results. However, the Hakulinen method is preferred, since it accounts for heterogeneity due to potential withdrawals.

**Key words:** Relative Survival. Expected Survival. Patient withdrawal. Long-term survival.

## Introducción

En un estudio longitudinal en el que se analiza el riesgo de fallecer por una determinada enfermedad, por ejemplo un tipo de cáncer, los sujetos pueden experimentar en los años siguientes al diagnóstico de la enfermedad una mortalidad más

elevada que la población general<sup>1</sup>. La mortalidad de la cohorte puede deberse a la enfermedad de estudio o a otras causas no relacionadas, lo que implica la necesidad de corregir la supervivencia estimada de estos pacientes<sup>1</sup>.

Un método para determinar la proporción de pacientes que sobreviven a la enfermedad en estudio, en la hipotética situación en que ésta es la única posible causa de defunción, es la utilización de la supervivencia por causa específica<sup>1,2</sup>. En este método se consideran como eventos únicamente las muertes causadas por la enfermedad en estudio. Los tiempos de seguimiento de los pacientes que no han fallecido por dicha causa se consideran como tiempos censurados<sup>2</sup>. Este método no tiene en cuenta otras causas de fallecimiento en competición y puede originar una subestimación de

*Correspondencia:* Dr. Ramón Clèries.  
Servei d'Epidemiologia i Registre del Càncer.  
Institut Català d'Oncologia.  
Gran Via, s/n, km 2,7. 08907 L'Hospitalet de Llobregat.  
Barcelona. España.  
Correo electrónico: r.cleries@iconcologia.net

*Recibido:* 4 de noviembre de 2005.

*Aceptado:* 8 de marzo de 2006.

la supervivencia, siendo entonces una posible alternativa el método de decrementos múltiples que corresponde a un análisis multiestado<sup>3</sup>. Otra limitación es que la disponibilidad y la calidad de los certificados de defunción no siempre son idóneas y pueden originar un sesgo en la estimación de la supervivencia por causa específica<sup>2</sup>.

Una posible solución a estos problemas es la utilización de la supervivencia relativa (SR)<sup>2,4</sup>, definida como el cociente entre la supervivencia observada (SO) y la supervivencia esperada (SE)<sup>4</sup>. En este método, la SO se calcula asumiendo que los eventos son todas las defunciones independientemente de la causa de muerte. La SE debe estimarse a partir de la mortalidad de la población residente en el área geográfica de la cual proceden los miembros de la cohorte en estudio<sup>1,4,5</sup>, y la SE es una corrección del sesgo en la estimación de la SO. Hay, al menos, 3 aplicaciones que permiten el cálculo de la SR: a) RELSURV<sup>2,6</sup>, un programa que permite estimar un modelo de regresión de SR a partir del método de Estève<sup>2</sup>, que además permite estimar la SE con los métodos de Ederer<sup>4,7</sup>; b) SURV3<sup>8</sup>, programa que permite utilizar los métodos de Hakulinen<sup>9</sup> y Ederer en la estimación de la SE, y c) WAERS<sup>10</sup>, una aplicación web para el cálculo de la SR mediante el método de Hakulinen y que permite al usuario seleccionar la población de referencia.

En este trabajo se describe el procedimiento para calcular la SR en función del método escogido para estimar la SE, comentando las ventajas e inconvenientes de cada uno de ellos a partir de un ejemplo hipotético.

### Procedimiento de cálculo de la supervivencia relativa

Para el cálculo de la supervivencia relativa el primer paso que hay que realizar es definir el período de estudio. Supongamos que definimos este tiempo en años. El tiempo de supervivencia se define como el intervalo entre el diagnóstico de la enfermedad y el evento de interés, que en este caso es la defunción del paciente. Si el paciente no ha podido ser seguido hasta el final del período de estudio, estableceremos que su tiempo de supervivencia es censurado<sup>1</sup>. En este caso, el tiempo de seguimiento se calculará desde la fecha de diagnóstico hasta la de último contacto con el paciente. De esta forma tendremos, para cada paciente, 2 variables respuesta: tiempo de seguimiento y censura/evento. Además, de cada paciente debemos conocer la edad y el año de entrada en el estudio (diagnóstico de la enfermedad).

En el cálculo de la SO se pueden utilizar 2 estimadores: el del método actuarial (tablas de vida) o el de Kaplan-Meier (KM)<sup>1</sup>. El primer método no es tan utili-

zado como el segundo, debido a que consiste en dividir el período de observación en intervalos indexados como  $i = 1, 2, 3, 4, \dots$  dando lugar a lo que a menudo se conoce como datos de supervivencia agrupados<sup>1,11</sup>. Este método suele utilizarse cuando sólo se conoce el año de diagnóstico y el año de fallecimiento del paciente, debido a que los pacientes tienen seguimientos de año en año sin tener en cuenta el día y el mes de fallecimiento. Si se conoce el tiempo de supervivencia exacto, es decir, la fecha exacta de diagnóstico y defunción del paciente, se utiliza el estimador de KM.

La SE se estima a partir de las tasas de supervivencia poblacionales (TSP) de la región de residencia de los miembros de la cohorte. Definamos  $\lambda$  como la tasa de mortalidad poblacional –por todas las causas– para los individuos de un grupo de edad y en un año en concreto en el área de residencia de los pacientes. Esta tasa puede calcularse como el cociente entre el número de fallecimientos y el número de individuos (residentes) de riesgo para ese grupo de edad y ese año en la región de procedencia de los pacientes<sup>1,12</sup>. Sea  $\mu_{ij}$  la TSP en el año  $j$ -ésimo para los pacientes del grupo de edad  $i$ -ésimo. A partir de  $\lambda_{ij}$  podemos estimar  $\mu_{ij}$  mediante:

$$\hat{\mu}_{ij} = e^{-\lambda_{ij}} \quad (1)$$

denominándose estimador de la TSP<sup>12</sup>. Una vez determinadas las tasas para todos los años del período de estudio y para todas las edades, se aplican dichas tasas a la cohorte de pacientes en estudio.

A partir de (1), la estimación de la SE se puede llevar a cabo por 3 métodos, conocidos como Ederer I<sup>4,7</sup>, Ederer II<sup>7</sup> y Hakulinen<sup>1,2,9,11,12</sup>.

### Cálculo de la supervivencia esperada mediante los métodos de Ederer

Los métodos de Ederer se basan en calcular una supervivencia poblacional media en la cohorte para cada momento en el que se ha producido un evento. Sea  $t_i$  un momento en el que se ha producido un evento y  $SP(t_i)$ , la supervivencia poblacional media en la cohorte en el intervalo  $(t_{i-1}, t_i)$ . Si definimos  $SE(t_i)$  como la supervivencia esperada a tiempo  $t_i$ , la  $SE(t_i)$  se calcula mediante la fórmula:

$$SE(t_i) = \prod_{\forall j | t_j \leq t_i} SP(t_j) \quad (2)$$

es decir,  $SE(t_i) = SP(t_1) \times \dots \times SP(t_2) \times \dots \times SP(t_i)$ .

La diferencia entre los 2 métodos de Ederer se encuentra en el número de pacientes que se considera en cada momento. Ederer I considera para cada inter-

valo a todos los pacientes, mientras que Ederer II sólo considera a los que están en riesgo durante ese intervalo<sup>4,7</sup>.

Para calcular la  $SE(t_i)$  mediante el método de Ederer I se sigue el siguiente esquema:

1. Determinar el tiempo mínimo y máximo de supervivencia en la cohorte, dividiendo en intervalos de igual longitud dicho tiempo.
2. Estimar para cada paciente la supervivencia poblacional que tendría en cada tiempo, aunque no esté en riesgo en dicho tiempo.
3. Calcular la media de las supervivencias poblacionales de cada individuo en  $t_i$ , siendo ésta la  $SP(t_i)$ .
4. Aplicar (2).

A partir de este esquema se puede comprobar que el método de Ederer II consiste en modificar el paso 2<sup>4,7</sup>. Veamos un ejemplo de cálculo mediante los métodos de Ederer. Supongamos una cohorte formada por 5 individuos diagnosticados de un tipo determinado de cáncer durante el período 1985-1989. La cohorte ha sido seguida durante un período de 5 años, y con edades comprendidas entre los 35 y 59 años en el momento del diagnóstico. Este seguimiento implica: a) que la cohorte habrá sido seguida, como máximo, hasta 1994, y b) que los miembros de la cohorte habrán cumplido, como máximo, 64 años al cierre del estudio. Por tanto, para la estimación de la SE nos hará falta disponer de las tasas de SP para los grupos de edad comprendidos entre 35 y 64 años durante el período 1985-1994. En la tabla 1 se muestra un ejemplo de tabla de supervivencia poblacional. Cada año de calendario queda representado por 3 filas. Las 10 columnas de la cada fila contienen las SP de los individuos por edades, siendo la primera fila las SP de los individuos de edades comprendidas entre 35 y 44 años, la segunda fila las SP de los individuos de 45-54 años y la tercera las SP de los individuos de 55-64 años. El procedimiento de cálculo de la SE de nuestro ejemplo estará basado en el seguimiento de la cohorte aplicado a los datos de esta tabla.

La tabla 2 contiene los datos de los individuos de la cohorte y un ejemplo de cálculo de la SE mediante los métodos de Ederer. Para cada individuo conocemos, además de un identificador (ID), la edad y el año de calendario de la entrada en la cohorte, su estado vital al final del seguimiento y el tiempo (en años) que éste ha sido seguido. Por ejemplo, fijémonos en el individuo número 3: este paciente entra en la cohorte en 1988 con 35 años de edad y es seguido durante 2 años. A partir de los datos de la tabla 1, la TSP para dicho paciente durante 1988 es de 0,9981. En el segundo año de seguimiento de 36 años en el año 1989, siendo de 0,9982. En el tercer año se le aplica la TSP de los individuos de 37 años en 1990, siendo 0,9980. Hasta aquí

tendríamos las estimaciones necesarias para el método de Ederer II. En cambio, la utilización del método de Ederer I conlleva aplicar la TSP a los hipotéticos cuarto y quinto años de seguimiento para ese paciente. Estas TSP son las de los individuos de 38 años en 1991 y 39 años en 1992. Es decir, son los valores 0,9978 y 0,9972, respectivamente. En cada año en el que se ha producido un evento o censura, este procedimiento implica determinar una estimación de la SP para la cohorte. Por ejemplo, la SP de la cohorte calculada mediante el método de Ederer II para el segundo año de seguimiento se obtiene como la media de las SP de los individuos 2 al 5 en ese año (4 individuos). En cambio, el método de Ederer I tiene en cuenta a los 5 individuos de la cohorte. Finalmente, la SE se determina a partir de las SP que correspondan a cada método y aplicando (2).

Ambos métodos de Ederer pueden dar estimaciones similares de la SE si el tiempo máximo de seguimiento de los pacientes se encuentra entre los 5 y 10 años<sup>7,9</sup>. Pero si el seguimiento medio de la cohorte es superior a 10 años, el estimador recomendado es el obtenido por el método de Hakulinen, también denominado *long-term*<sup>5,9</sup>.

#### Cálculo de la supervivencia esperada mediante el método de Hakulinen

En la aplicación del método de Hakulinen, la estimación de la SE se realiza mediante la utilización de un estimador similar al de KM que tiene en cuenta las censuras. Es decir, el número de pacientes de riesgo de cada intervalo se calcula teniendo en cuenta un número de abandonos esperados en dicho tiempo. El procedimiento se basa en determinar para cada intervalo el número de pacientes de riesgo. Sea  $r_i$  el número de pacientes de riesgo a tiempo  $t_i$ , la suma de las SP de cada uno de los  $r_i$  pacientes de riesgo es un estimador del número de pacientes de riesgo esperados, que definimos como  $r_i^*$ . De esta forma tenemos que  $r_i - r_i^*$  nos da el número de defunciones esperadas,  $d_i^*$ . Finalmente, el número de abandonos esperados en el intervalo  $w_i$  se calcula aplicando una corrección a la suma de las SP de los individuos en riesgo en dicho intervalo y cuyo tiempo de seguimiento no exceda ese intervalo. Esta corrección se denomina «anticipación al tiempo de censura»<sup>9</sup>, habitualmente 1/2. De esta forma, el estimador de la  $SE(t_i)$  mediante el método de Hakulinen es:

$$SE(t_i) = \prod_{\forall j|t_j \leq t_i} \left( 1 - \frac{d_j^*}{(r_j^* - \frac{1}{2} w_j)} \right) \quad (3)$$

**Tabla 1. Supervivencia poblacional de una determinada área geográfica, en el período 1985-1994, para los individuos de edades comprendidas entre 35 y 64 años**

Año	Edades									
	35	36	37	38	39	40	41	42	43	44
	45	46	47	48	49	50	51	52	53	54
	55	56	57	58	59	60	61	62	63	64
1985	0,9978	0,9977	0,9974	0,9973	0,9971	0,9969	0,9963	0,9963	0,9962	0,9960
	0,9954	0,9951	0,9945	0,9944	0,9938	0,9934	0,9922	0,9922	0,9920	0,9916
	0,9903	0,9897	0,9885	0,9883	0,9871	0,9862	0,9837	0,9836	0,9832	0,9824
1986	0,9979	0,9978	0,9975	0,9975	0,9971	0,9968	0,9963	0,9963	0,9961	0,9959
	0,9955	0,9953	0,9948	0,9947	0,9940	0,9933	0,9923	0,9922	0,9918	0,9914
	0,9905	0,9901	0,9890	0,9888	0,9873	0,9859	0,9838	0,9836	0,9828	0,9820
1987	0,9974	0,9974	0,9971	0,9971	0,9968	0,9964	0,9960	0,9960	0,9959	0,9958
	0,9946	0,9945	0,9940	0,9938	0,9932	0,9925	0,9916	0,9915	0,9913	0,9912
	0,9887	0,9885	0,9873	0,9870	0,9857	0,9843	0,9823	0,9821	0,9817	0,9814
1988	0,9981	0,9980	0,9978	0,9978	0,9975	0,9974	0,9969	0,9969	0,9968	0,9966
	0,9961	0,9959	0,9954	0,9953	0,9948	0,9945	0,9935	0,9934	0,9933	0,9929
	0,9918	0,9913	0,9904	0,9902	0,9892	0,9884	0,9863	0,9862	0,9859	0,9852
1989	0,9983	0,9982	0,9980	0,9979	0,9977	0,9976	0,9971	0,9971	0,9970	0,9970
	0,9964	0,9962	0,9958	0,9957	0,9952	0,9949	0,9940	0,9939	0,9938	0,9936
	0,9924	0,9920	0,9911	0,9909	0,9899	0,9893	0,9873	0,9873	0,9869	0,9865
1990	0,9981	0,9980	0,9980	0,9979	0,9977	0,9974	0,9971	0,9970	0,9969	0,9969
	0,9960	0,9958	0,9957	0,9955	0,9952	0,9946	0,9940	0,9936	0,9936	0,9935
	0,9915	0,9911	0,9910	0,9905	0,9899	0,9887	0,9873	0,9865	0,9865	0,9864
1991	0,9980	0,9979	0,9979	0,9978	0,9976	0,9974	0,9970	0,9969	0,9969	0,9969
	0,9958	0,9957	0,9956	0,9954	0,9950	0,9946	0,9937	0,9936	0,9935	0,9934
	0,9911	0,9909	0,9908	0,9903	0,9895	0,9886	0,9869	0,9865	0,9864	0,9861
1992	0,9979	0,9978	0,9975	0,9975	0,9972	0,9970	0,9965	0,9965	0,9964	0,9962
	0,9956	0,9953	0,9948	0,9947	0,9942	0,9938	0,9926	0,9926	0,9924	0,9920
	0,9907	0,9902	0,9891	0,9889	0,9877	0,9869	0,9845	0,9845	0,9841	0,9832
1993	0,9980	0,9979	0,9976	0,9976	0,9973	0,9970	0,9965	0,9965	0,9963	0,9961
	0,9957	0,9955	0,9950	0,9949	0,9943	0,9936	0,9927	0,9926	0,9922	0,9918
	0,9910	0,9906	0,9895	0,9894	0,9880	0,9866	0,9846	0,9844	0,9837	0,9829
1994	0,9976	0,9975	0,9973	0,9972	0,9969	0,9966	0,9962	0,9962	0,9961	0,9960
	0,9949	0,9948	0,9943	0,9941	0,9935	0,9929	0,9920	0,9919	0,9917	0,9916
	0,9893	0,9891	0,9879	0,9877	0,9864	0,9851	0,9832	0,9830	0,9826	0,9824

**Cálculo del intervalo de confianza para la supervivencia relativa**

Para el cálculo de la varianza de la  $SR(t_i)$ , se considera una transformación log-log complementaria a la  $SO(t_i)^{11}$ . Sea  $LSO(t_i) = \log(-\log[SO(t_i)])$ , cuya varianza puede ser aproximada mediante el método delta<sup>11</sup>:

$$V[LSO(t_i)] \approx \frac{1}{(SO(t_i)\log[SO(t_i)])^2} \text{Var}[SO(t_i)] \quad (4)$$

A partir de (4), y considerando  $SE(t_i)$  como una constante, se pueden obtener los límites inferior y

superior de un intervalo de confianza a nivel  $\alpha$  para la  $SR(t_i)$ . Sea  $X_1(j_i) =$

$$e^{[LSO(t_i)] - Z_{\frac{\alpha}{2}} \sqrt{V[LSO(t_i)]}} \text{ y } X_2(t_i) = e^{[LSO(t_i)] + Z_{\frac{\alpha}{2}} \sqrt{V[LSO(t_i)]}},$$

siendo  $Z_{\frac{\alpha}{2}}$  el valor de la distribución normal estándar

que deja un área a su derecha de  $\frac{\alpha}{2}$ , los límites del

intervalo de confianza se obtienen mediante:

$$\left[ \frac{e^{-X_1(t_i)}}{SE(t_i)}, \frac{e^{-X_2(t_i)}}{SE(t_i)} \right] \quad (5)$$

**Tabla 2. Métodos de Ederer: datos de la cohorte y cálculo de la supervivencia esperada utilizando la supervivencia poblacional de la tabla 1**

ID	Edad entrada	Año entrada	Estado	Tiempo	SE Año 1	SE Año 2	SE Año 3	SE Año 4	SE Año 5
1	45	1989	0	1	0,9964	0,9958	0,9956	0,9947	0,9943
2	59	1986	1	2	0,9873	0,9843	0,9863	0,9873	0,9865
3	35	1988	0	3	0,9981	0,9982	0,9980	0,9978	0,9972
4	54	1988	0	4	0,9929	0,9924	0,9911	0,9908	0,9889
5	55	1985	1	5	0,9903	0,9901	0,9873	0,9902	0,9899
Media de la SP en la cohorte (Ederer I):					0,9930	0,9922	0,9917	0,9922	0,9914
SR(t) (Ederer I):					0,9930	0,9852	0,9770	0,9693	0,9610
Media de la SP en la cohorte (Ederer II):					0,9930	0,9913	0,9921	0,9905	0,9899
SR(t) (Ederer II):					0,9930	0,9843	0,9766	0,9673	0,9575

Estado: estado vital (0, vivo; 1, fallecido). Tiempo: años de seguimiento desde la entrada en la cohorte. ID: identificador para cada miembro de la cohorte. SP: supervivencia poblacional. SR: supervivencia relativa. SE: supervivencia esperada.

### Ejemplo

En el siguiente ejemplo se calcula la SR de los pacientes diagnosticados de cáncer de próstata del registro hospitalario de tumores del Instituto Catalán de Oncología y del Hospital Universitario de Bellvitge<sup>13</sup>. Estos pacientes han entrado en el período 1991-2000 y han sido seguidos hasta 2001. De éstos, se ha seleccionado a los residentes en Cataluña, que son un total de 1.354. La edad media de los pacientes es de 65 años (error estándar = 4,5; rango, 46-73). Para el cálculo de la SE ha sido necesario disponer de las tablas de mortalidad de Cataluña en el período 1991-2001 y para los grupos de edad de 46-84 años. El seguimiento máximo para estos pacientes se observa en los que entraron en 1991 y están vivos al final del seguimiento, es decir, 11 años. En la tabla 3 se muestran las estimaciones de las SE según los 3 métodos. La primera columna indica el intervalo de tiempo de seguimiento, la segunda el número de individuos de riesgo, la tercera el número de individuos que fallecen a final de cada intervalo, la cuarta la SO, la quinta el error estándar de la SO y, finalmente, en las 3 últimas se encuentran las estimaciones de la SE según los tres métodos. Como puede observarse, los métodos de Ederer dan estimaciones muy similares de la SE con valores superiores a la SE estimada mediante el método de Hakulinen, tal como se puede observar en la figura 1. En la tabla 4 se muestran las estimaciones de la SR a partir de los 3 métodos. Nótese que a partir del séptimo año de seguimiento la SR calculada mediante la estimación de la SE utilizando el método de Hakulinen es superior a la estimada por los métodos de Ederer. Por otro lado, a partir del séptimo año, la curva de SR calculada a partir de la estimación de la SE por el método de Ederer II se asemeja a la curva de SR que utiliza la SE estimada mediante el método de Hakulinen (fig. 2).

**Tabla 3. Cálculo de la supervivencia observada y estimaciones de la supervivencia esperada mediante los métodos de Ederer I, Ederer II y Hakulinen**

t	n	Fallecen	SO(t)	EE[SO(t)]	SE(t) Ederer I	SE(t) Ederer II	SE(t) Hakulinen
(0-1)	1.354	206	0,85	0,009	0,96	0,96	0,96
(1-2)	1.138	120	0,76	0,011	0,92	0,91	0,91
(2-3)	788	80	0,68	0,013	0,87	0,87	0,87
(3-4)	535	43	0,63	0,014	0,83	0,83	0,82
(4-5)	381	19	0,60	0,015	0,79	0,78	0,77
(5-6)	274	23	0,55	0,017	0,75	0,74	0,72
(6-7)	177	11	0,51	0,019	0,71	0,70	0,67
(7-8)	104	8	0,47	0,022	0,67	0,66	0,62
(8-9)	64	3	0,45	0,024	0,63	0,62	0,57
(9-10)	36	1	0,44	0,026	0,59	0,58	0,52
(10-11)	17	1	0,41	0,035	0,54	0,52	0,46

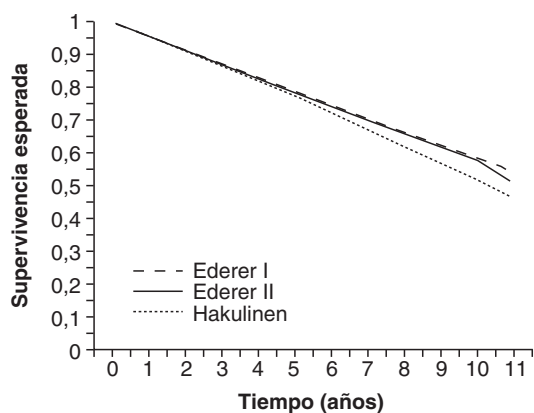
t: tiempo en el que se produce un evento/censura; n: número de individuos de riesgo a tiempo t; SO(t): supervivencia observada (Kaplan-Meier), calculada a tiempo t; EE(SO(t)): error estándar de la supervivencia observada a tiempo t; SE(t): supervivencia esperada.

### Observaciones y comparación de los métodos de cálculo de la supervivencia esperada

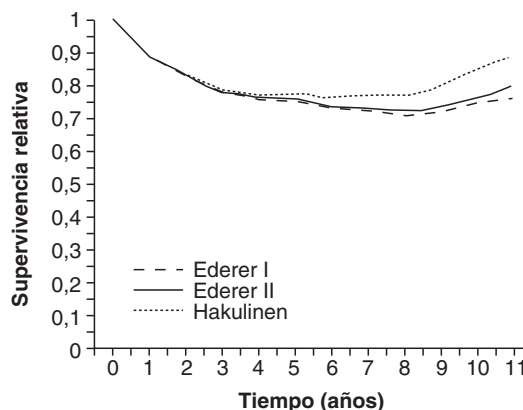
En el cálculo de la SR quedan definidos claramente 3 procedimientos: disposición de las tablas de mortalidad, estimación de la supervivencia esperada y determinación de la SR.

El primer paso presenta 2 inconvenientes: a) la disponibilidad de las TSP de la población de referencia en el período en el que se realiza el estudio, y b) el coste de preparación de los datos. Supongamos, por ejemplo, que no disponemos de las TSP del período 1999-2000, pero sí las del año 1998. En tal caso podemos utilizar en dicho período las TSP de 1998 o incluso re-

**Figura 1. Estimación de la supervivencia esperada. Comparación de los 3 métodos: Ederer I, Ederer II y Hakulinen.**



**Figura 2. Supervivencia relativa en función del método de estimación de la supervivencia esperada (SE) (Ederer I, Ederer II y Hakulinen).**



**Tabla 4. Cálculo de la de la supervivencia relativa e intervalos de confianza (IC) del 95% a partir de las estimaciones de la supervivencia esperada**

t	SR(t)		IC del 95%		SR(t)		IC del 95%	
	Ederer I	Ederer I	Ederer II	Ederer II	Hakulinen	Hakulinen	Hakulinen	Hakulinen
(0-1)	0,89	(0,87-0,90)	0,89	(0,87-0,90)	0,89	(0,87-0,90)		
(1-2)	0,83	(0,80-0,85)	0,83	(0,81-0,85)	0,83	(0,81-0,85)		
(2-3)	0,78	(0,75-0,81)	0,78	(0,75-0,81)	0,79	(0,75-0,81)		
(3-4)	0,75	(0,72-0,79)	0,76	(0,72-0,79)	0,77	(0,73-0,80)		
(4-5)	0,75	(0,71-0,79)	0,76	(0,72-0,80)	0,77	(0,73-0,81)		
(5-6)	0,73	(0,68-0,77)	0,73	(0,69-0,78)	0,75	(0,71-0,80)		
(6-7)	0,72	(0,67-0,77)	0,73	(0,68-0,78)	0,76	(0,71-0,82)		
(7-8)	0,71	(0,64-0,77)	0,72	(0,65-0,78)	0,76	(0,69-0,83)		
(8-9)	0,72	(0,64-0,79)	0,73	(0,65-0,81)	0,79	(0,71-0,87)		
(9-10)	0,75	(0,66-0,83)	0,76	(0,67-0,85)	0,85	(0,75-0,94)		
(10-11)	0,76	(0,63-0,88)	0,80	(0,67-0,93)	0,89	(0,74-1,03)		

SR(t): supervivencia relativa.

alizar una extrapolación. Ambas posibilidades pueden conllevar un sesgo en las estimaciones de la SR<sup>11</sup> sólo evaluable a posteriori, cuando se disponga de las tasas para el período.

La elección del método de estimación de la SE es el apartado más discutible acerca del cálculo de la SR. Tanto el método de Ederer II como el de Hakulinen son los recomendables, ya que minimizan el sesgo en la estimación de la SE que se puede dar por la heterogeneidad de las edades de los pacientes de la cohorte<sup>4,7,9</sup>. En un análisis de SR a largo plazo y cuando hay pocos individuos de riesgo, el método de Ederer I puede sobrestimar la SE y, por lo tanto, subestimar la SR<sup>4,9,11</sup>. En el ejemplo utilizado anteriormente, se observa la tendencia a la sobrestimación de la SE mediante el método de Ederer I (tabla 3 y fig. 1). Se ha descrito que

sólo en el caso de estratificación por grupos de edad es indiferente la utilización de cualquiera de los 3 métodos, aunque se recomienda el método de Ederer I únicamente en modelos de curabilidad<sup>9,11</sup>.

Este sesgo también puede darse en la comparación de curvas de SR mediante modelos de regresión. Estos modelos permiten determinar, por ejemplo, el efecto de la edad y el estadio de la enfermedad en la supervivencia relativa. Entre ellos cabe destacar el modelo de Estève de riesgos proporcionales<sup>2,14</sup> y los modelos de Hakulinen-Tenkanen<sup>11,14</sup> y Dickman<sup>14</sup>, basados en modelos lineales generalizados. Para estos modelos es preferible la estimación de la SE mediante los métodos de Ederer II y de Hakulinen<sup>8,11</sup>. A pesar de ello, el método de Ederer II no tiene en cuenta una posible heterogeneidad entre abandonos, por lo que se recomienda el método de Hakulinen como método estándar.

Cabe destacar que hay una posible metodología alternativa al método de Hakulinen, denominada método de Verheul<sup>15</sup>. Dicha metodología modifica los tiempos potenciales de censuras, pero se ha comprobado que las estimaciones de la SE con dicho método no presentan diferencias con las obtenidas mediante el método de Hakulinen<sup>11,15</sup>.

### Conclusiones

Para evitar los sesgos en la estimación de la SE, es recomendable el cálculo de la SR subdividiendo, cuando sea posible, la cohorte en grupos de edad. Si esta subdivisión no es posible o el tiempo de seguimiento medio es superior a 10 años, el método de Hakulinen es el recomendable.

## Agradecimientos

Este trabajo se ha financiado parcialmente por las Redes de Centros de Investigación Cooperativa en Epidemiología y Salud Pública (RCESP, C03/09) y de Cáncer (C03/10) y por la beca DURSI (núm. expediente 2003XT 00023).

## Bibliografía

1. Estève J, Benhamou E, Raymond L. Statistical Methods in Cancer Research. Descriptive Epidemiology. IARC Scientific Publications, N.º 128. Lyon: IARC; 1994.
2. Estève J, Benhamou E, Croasdale M, Raymond L. Relative survival and the estimation of net survival: elements for further discussion. *Stat Med.* 1990;9:529-38.
3. Llorca J, Delgado-Rodríguez. Análisis de supervivencia en presencia de riesgos competitivos: estimadores de la probabilidad de suceso. *Gac Sanit.* 2004;18:391-7.
4. Ederer F, Axtell LM, Cutler SJ. The relative survival rate: a statistical methodology. *Natl Cancer Inst Monogr.* 1961;6:101-21.
5. Hakulinen T. On long-term relative survival rates. *J Chronic Dis.* 1977;30:431-43.
6. Hédelin G. RELSURV 2.0: a program for relative survival analysis. Estrasburgo: Department of Epidemiology and public health, Faculty of Medicine, Louis Pasteur University; 1997.
7. Ederer F, Heise H. The effect of eliminating deaths from cancer on general population survival rates, methodological note 11. End results evaluation section. National Cancer Institute; 1959.
8. Dickman PW, Hakulinen T, Voutilainen ET. SURV3, Relative Survival Analysis. Helsinki: Finnish Cancer Registry; 2000.
9. Hakulinen, T. Cancer survival corrected for heterogeneity in patient withdrawal. *Biometrics.* 1982;38:933-42.
10. Clèries R, Ribes J, Gálvez J, Melià A, Moreno V, Bosch FX. Cálculo automatizado de la supervivencia relativa vía web. El proyecto WAERS del Instituto Catalán de Oncología. *Gac Sanit.* 2005;19:71-5.
11. Dickman P. Teaching and seminars [citado 10 Jun 2006]. Disponible en: <http://www.pauldickman.com/teaching/index.php>
12. Therneau TM, Grambsch PM. Modeling survival data: extending the Cox model. New York: Springer-Verlag; 2000.
13. Registre Hospitalari de Tumors ICO/CSUB [citado 2 Oct 2005]. Disponible en: <http://rht.iconcologia.net>
14. Dickman PW, Sloggett A, Hills M, Hakulinen T. Regression models for relative survival. *Stat Med.* 2004;23:51-64.
15. Verheul HA, Dekker E, Bossuyt P, Moulijn AC, Dunning AJ. Background mortality in clinical survival studies. *Lancet.* 1993;341:872-5.