

# Cálculo automatizado de la supervivencia relativa vía web. El proyecto WAERS del Instituto Catalán de Oncología

Ramon Clèries / Josepa Ribes / Jordi Gálvez / Àngels Melià / Víctor Moreno / Francesc Xavier Bosch  
Servei d'Epidemiologia i Registre del Càncer. Institut Català d'Oncologia. Barcelona. España.

(Automatic calculation of relative survival through the web.  
The WAERS project of the Catalan Institute of Oncology)

## Resumen

La medida utilizada habitualmente para estimar la supervivencia del cáncer es la supervivencia relativa, definida como el cociente entre la supervivencia observada y la esperada. La supervivencia esperada se calcula a partir de la mortalidad de una población de referencia. La disponibilidad y la preparación de tablas de mortalidad de la población general no es siempre posible y requiere software específico para su cálculo. A tal efecto, el Instituto Catalán de Oncología (ICO) ha desarrollado la aplicación WAERS, una aplicación web que proporciona la estimación de la supervivencia relativa para una cohorte de pacientes. El usuario debe preparar los datos en un formato específico y enviarlos a un servidor remoto que se encuentra en el ICO. Este servidor calcula la supervivencia relativa y devuelve los resultados en un fichero a una dirección que ha indicado el usuario.

Mediante esta aplicación, los registros de cáncer de base hospitalaria y poblacional y los registros de otras enfermedades pueden estimar la supervivencia relativa de sus cohortes seleccionando a la población de referencia que consideren (provincia o comunidad autónoma). También puede ser útil para estudios de mortalidad en cohortes.

**Palabras clave:** Cáncer. Supervivencia relativa. Aplicación web. Registro de cáncer.

## Abstract

The most commonly used measure to estimate cancer survival is relative survival, defined as the ratio between observed and expected survival. Expected survival is computed on the basis of the mortality of a reference population. Mortality tables for the general population are not always available and their calculation requires specific software. For that purpose, the Catalan Institute of Oncology developed WAERS (Web-Assisted Estimation of Relative Survival), a web-based application that estimates the relative survival for a cohort of patients. The user prepares data in a specific format and sends them to a remote server located at the Catalan Institute of Oncology. This server computes relative survival and returns a file with the results to the electronic address supplied by the user.

By means of this application, hospital- and population-based Spanish cancer registries and registries of other diseases can estimate relative survival of their cohorts using their reference population (province or autonomous community). This application could also be useful for cohort mortality studies.

**Key words:** Cancer. Relative survival. Web application. Cancer registry.

## Fundamento y objetivos

El estudio de la supervivencia de un grupo de la población puede estar motivado por la posibilidad de que dicho grupo esté sujeto a un riesgo de muerte diferente del de su población de referencia. Por ejemplo, los individuos con cáncer ge-

neralmente experimentan una mortalidad más elevada que la población general en los años siguientes al diagnóstico de cáncer. De esta forma, la mortalidad debería ser considerada como resultado de 2 efectos: el de la causa en estudio –cáncer– y el conjunto de las otras causas<sup>1</sup>.

Con la finalidad de estimar el efecto de la causa en estudio, cuando se calcula la supervivencia por cáncer se puede emplear la supervivencia neta, en la que sólo se consideran como eventos las muertes por cáncer. La supervivencia neta nos da una medida de la supervivencia del paciente corregida por otras posibles causas de fallecimiento<sup>1</sup>. En estudios de cáncer realizados desde una perspectiva clínica, ya sean ensayos clínicos aleatorizados o estudios de tipo observacional, es habitual utilizar la supervivencia por causa específica para estimar la supervivencia neta. En la supervi-

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

*Recibido:* 23 de julio de 2004.

*Aceptado:* 22 de septiembre de 2004.

vencia específica se consideran los tiempos de seguimiento de los pacientes que han fallecido por otra causa distinta del cáncer como tiempos censurados<sup>2</sup>. Este procedimiento no tiene en cuenta correctamente otras causas de fallecimiento en competición y puede originar una subestimación de la supervivencia. Como alternativa se puede emplear el método de decrementos múltiples que corresponde a un análisis multiestado<sup>3</sup>. Otra limitación del cálculo de la supervivencia específica es que se necesita disponer de la causa de fallecimiento. No siempre está disponible la información de la mortalidad o no es suficientemente precisa, debido a los errores de codificación en algunos tipos de cáncer, como el de hígado, pulmón o huesos, donde asientan metástasis de otras localizaciones tumorales. Todo ello puede originar un sesgo en la estimación de la supervivencia neta. Para solventar estos problemas y poder comparar la supervivencia entre poblaciones diferentes, se ha propuesto como estimador de la supervivencia neta la supervivencia relativa (SR)<sup>2</sup>.

La SR se define como el cociente entre la supervivencia observada y la supervivencia esperada<sup>4</sup>. La supervivencia observada es la supervivencia de la cohorte teniendo en cuenta el tiempo de seguimiento del individuo y si éste ha fallecido o no, independientemente de la causa. La supervivencia esperada determina la supervivencia que se esperaría en la misma cohorte según la mortalidad por todas las causas de la población de la cual proceden los individuos<sup>5</sup>. La SR se interpreta como la proporción de pacientes que habrá sobrevivido en la hipotética situación de que la enfermedad en estudio fuese la única y posible causa de defunción. Debido a que la SR es un cociente, se pueden obtener valores > 1. Tal situación indicaría que la supervivencia de los pacientes en estudio es superior a la de la población general.

Hay al menos 2 aplicaciones que permiten la estimación de la supervivencia relativa: *a)* RELSURV2<sup>6</sup>, que emplea el método de Estève et al<sup>2</sup> a partir de un modelo de Cox para la estimación de la SR, y *b)* SURV3<sup>7</sup>, que utiliza el método de Hakulinen<sup>8</sup>. Ambas aplicaciones tienen incorporados tests de comparación de curvas de SR, pero el inconveniente se encuentra en que el usuario debe preparar las tablas de mortalidad de la población de referencia y suministrarlas en el formato requerido.

El Instituto Catalán de Oncología (ICO) ha iniciado el proyecto WAERS (Web Assisted Estimation of the Relative Survival) con el objetivo de proporcionar al usuario una aplicación informática que permita el cálculo de la SR de una cohorte de pacientes sin la necesidad de disponer de las tablas de mortalidad de la población de referencia, y sin instalar en su ordenador el software necesario para su cálculo.

---

## Material y método

El cálculo de la SR mediante WAERS se realiza a partir de la estimación de la supervivencia observada, calculada mediante el método de Kaplan-Meier<sup>9</sup>, y de la supervivencia esperada, calculada a partir del método de Hakulinen<sup>8</sup>. Dichos cálculos se realizan de forma remota en un servidor instalado en el ICO. Los cálculos estadísticos se han programado utilizando el paquete estadístico R<sup>10</sup>. La aplicación web se ha programado mediante el lenguaje PHP<sup>11</sup>.

Para el cálculo de la supervivencia relativa de una cohorte mediante WAERS es necesario proporcionar las siguientes variables para cada individuo de la cohorte: *a)* edad en el diagnóstico; *b)* año de diagnóstico; *c)* tiempo de seguimiento en años; *d)* sexo; *e)* área de residencia, y *f)* estado vital del individuo al final del período de seguimiento. Cabe señalar que la aplicación se ha diseñado para tener en cuenta como variable tiempo los años de seguimiento del paciente. Por ejemplo, para un paciente que ha entrado en enero de 1985 y ha sido seguido hasta marzo de 1986, el tiempo de seguimiento ha sido de 1 año y 2 meses. La variable tiempo de seguimiento para este paciente tomará el valor  $1 + 2/12 = 1,16$ .

El usuario debe preparar un fichero de texto (formato ASCII) con la información de estas 6 variables dispuestas en columnas separadas por tabuladores. La información de cada individuo se dispone en una fila. El fichero se envía al servidor mediante una página web y el usuario deberá especificar una dirección de correo electrónico para recibir los resultados mediante un fichero texto.

Para cada paciente se deberán codificar las variables. Por ejemplo, la variable área de residencia hace referencia a la provincia o comunidad autónoma de residencia del paciente, es decir, su población de referencia. Supongamos que un paciente reside en Cataluña; entonces, su código de residencia es 1 según la codificación indicada en la página web, pero si sabemos que reside en Tarragona entonces dicho código es 5. La aplicación WAERS está disponible en: <http://rht.iconcologia.catsalut.net/cas/surv.htm> (fig. 1).

---

## Ejemplo de aplicación

Supongamos que disponemos de una cohorte de 20 varones residentes en Cataluña y diagnosticados de un determinado tipo de cáncer. Estos pacientes han sido diagnosticados durante el período 1985-1989 y han sido seguidos hasta diciembre de 1989. El primer paso es la preparación de los datos. En la tabla 1 se muestra el fichero con el formato de los datos de los pacientes. La primera columna de la tabla corresponde a un iden-

Figura 1. Página web donde se realiza la solicitud de cálculo de la SR accesible desde la dirección <http://rht.iconcologia.catsalut.net/cas/surv.htm>

**Cálculo automatizado de la Supervivencia Relativa**

Esta herramienta permite el cálculo automático de la Supervivencia Relativa a partir de los datos proporcionados por el usuario. La supervivencia relativa (SR) se interpreta como la proporción de pacientes vivos después de T años de seguimiento en la situación hipotética de que el cáncer en cuestión es la única causa posible de mortalidad. La SR se calcula como el cociente entre la supervivencia observada en un grupo de pacientes dividida por la supervivencia esperada que tendrían estos pacientes teniendo en cuenta la mortalidad de una población de referencia. Una ventaja de este dato es que la información de la causa de muerte específica no es necesaria. Por otra parte, la SR da una medida del exceso de mortalidad experimentado por los pacientes diagnosticados de cáncer.

Las variables tienen que venir en el siguiente orden:

- Número de Registro
- Edad a la entrada
- Sexo del individuo  
1=Hombre, 2=Mujer
- Tiempo de seguimiento (en años)
- Evento  
1=SI, 0=NO
- Año de entrada
- Población de referencia de la que se utilizan las tasas de mortalidad:  
1=Catalunya, 2=Barcelona, 3=Girona, 4=Lleida, 5=Tarragona

Se tiene que pasar el fichero en formato ASCII con el siguiente formato:

```

1 87.00 1 0.08 1 1992 1
2 70.00 1 10.33 0 1993 1
3 75.00 1 9.83 0 1991 1
4 84.00 1 2.67 1 1992 1
5 85.00 2 5.00 1 1983 1
6 79.00 2 9.17 0 1988 1
7 74.00 2 2.58 1 1992 1
    
```

**ATENCIÓN!!: ESTE PROCESO PUEDE TARDAR ALGUNOS MINUTOS!!**

E-Mail donde se enviarán los resultados:

Nombre del fichero:

tificador del paciente definido por el usuario. La segunda columna corresponde a la edad de diagnóstico del paciente (en años); la tercera, al sexo (valor 1 para los varones y 2 para las mujeres); la cuarta, al período de seguimiento de cada paciente (en años); la quinta, al estado vital del paciente (valor 1 si el individuo ha fallecido y 0 si el paciente se encuentra vivo al final del seguimiento); la sexta columna, al año de diagnóstico y, finalmente, en la última columna, a la población de referencia escogida. Actualmente sólo consta en la web la población de Cataluña y, por ello, en el ejemplo está codificada como 1. Aunque en la tabla 1, como ejemplo de fichero, aparecen en la primera fila los nombres de las variables para facilitar la explicación, el fichero no debe incluir dichos nombres.

Una vez enviados los datos, el servidor remitirá el fichero de resultados como se observa en la tabla 2. Este

fichero contiene las variables Risk (número de pacientes en riesgo a tiempo T), T (tiempo en que se ha producido el evento o la censura), SRel (supervivencia relativa), ICinf y ICSup (límites, inferior y superior del intervalo de confianza del 95% de la supervivencia relativa) y la variable SObs con la supervivencia observada. Por ejemplo, en la primera fila se indica que hay 20 pacientes en riesgo hasta el primer año de seguimiento, la SR es del 96% con un intervalo de confianza (IC) del 95% de 86-101% y una supervivencia observada del 95%. Tal como se observa en la tabla 1, hay 5 pacientes que han sido seguidos durante 1 año. De esta forma, para el segundo año de seguimiento nos quedan 15 pacientes en riesgo y una SR del 90%, con un IC del 95% de 76-101% y una supervivencia observada del 89% (fila 2 de la tabla 2). Y así sucesivamente hasta el último tiempo de seguimiento. El formato en que

**Tabla 1. Formato del fichero de texto que se enviará a la aplicación WAERS**

ID	Edad	Sexo	Seguimiento	Fallecimiento	Año	Área
1	35	1	1	0	1989	1
2	45	1	1	0	1989	1
3	55	1	1	0	1989	1
4	44	1	1	0	1989	1
5	54	1	1	1	1985	1
6	61	1	2	0	1988	1
7	35	1	2	0	1988	1
8	44	1	2	0	1988	1
9	45	1	2	0	1988	1
10	54	1	2	0	1988	1
11	55	1	2	1	1985	1
12	60	1	3	1	1985	1
13	35	1	3	0	1987	1
14	45	1	3	0	1987	1
15	55	1	3	1	1985	1
16	44	1	4	0	1986	1
17	54	1	4	0	1986	1
18	60	1	4	1	1986	1
19	35	1	5	1	1985	1
20	44	1	5	0	1985	1

Nota: El fichero enviado no debe contener la primera fila con los nombres de las variables.

**Tabla 2. Fichero resultante de la utilización de WAERS con los datos de la cohorte descrita en la tabla 1 y utilizando como población de referencia Cataluña**

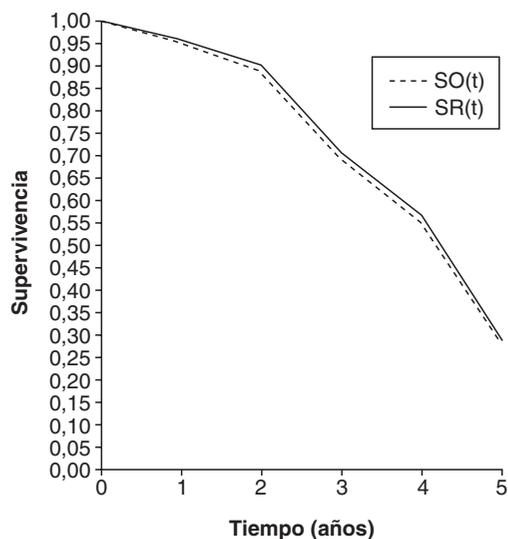
Risk	T	SRel	ICinf	ICSUp	SObs
20	1	0,96	0,86	1,01	0,95
15	2	0,90	0,76	1,01	0,89
9	3	0,70	0,48	1,02	0,69
5	4	0,57	0,32	1,02	0,55
2	5	0,28	0,06	1,03	0,28

son devueltos los datos permite que se pueda realizar fácilmente una gráfica conjunta de la supervivencia relativa y observada, tal como muestra la figura 2.

### Observaciones y trabajo futuro

Esta aplicación supone para el usuario la ventaja de preparar el fichero de los pacientes a estudiar sin tener que preparar las tablas de mortalidad de su población de referencia. No es necesario disponer del software instalado. Entre las limitaciones actuales de esta aplicación, que se mejorarán en el futuro, destaca la imposibilidad de calcular la supervivencia relativa estratificada por una variable de control. Por ejemplo, si

**Figura 2. Gráfico de la supervivencia observada y relativa (SR) de los 20 pacientes de la cohorte. (SO)**



se desea calcular la supervivencia relativa en función de la variable sexo, el usuario deberá preparar 2 ficheros, uno con los varones y otro con las mujeres, y realizar el proceso descrito anteriormente para cada uno de ellos. Por el momento tampoco es posible ajustar modelos multivariados a la SR. Próximamente se incorporará el modelo de Hakulinen y Tenkanen<sup>12</sup> que permitirá evaluar el efecto de covariables sobre la SR.

Finalmente, en la actualidad sólo están disponibles las tablas de mortalidad de la Comunidad Autónoma de Cataluña y sus provincias para los períodos 1985-2001. En breve se prepararán las tablas para todas las provincias españolas y comunidades autónomas para el período 1975-2001, a partir de los datos cedidos por el Instituto Nacional de Estadística (INE). Esta aplicación puede ser de utilidad para registros de cáncer, tanto de base hospitalaria como poblacional, y registros de otras enfermedades (infartos de miocardio, diabetes, sida, etc.). También puede ser útil para estudios de mortalidad en cohortes.

### Agradecimientos

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

## Bibliografía

1. Estève J, Benhamou E, Raymond L. Statistical methods in cancer research (vol. IV): descriptive epidemiology. (IARC Scientific Publications, No. 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 M. 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. Department of Epidemiology and public health, Faculty of Medicien, Louis Pasteur University, Strasbourg, France; 1997.
7. Dickman PW, Hakulinen T, Voutilainen ET. SURV3, relative survival analysis. Finnish cancer registry. Helsinki; 2000.
8. Hakulinen T. Cancer survival corrected for heterogeneity in patient withdrawal. *Biometrics*. 1982;38:933-42.
9. Dickman P, Hakulinen T. Notes on the population-based cancer survival analysis course. Disponible en: [www.pauldickman.com](http://www.pauldickman.com)
10. R Development Core Team (2004). R: a language and environment for statistical computing. Foundation for Statistical Computing, Vienna, Austria, 2004. Disponible en: <http://www.R-project.org>
11. PHP Professional Projects. Ashish Daniel Wilfred, Meeta Gupta & Kartik Bhatnagar Premier Press; 2002.
12. Dickman PW, Sloggett A, Hills M, Hakulinen T. Regression models for relative survival. *Stat Med*. 2004;23:51-64.