

# Accurately estimating breast cancer survival in Spain: cross-matching local cancer registries with the National Death Index

M. Carmen Martos,<sup>1,2</sup> Carme Saurina,<sup>3,4</sup> Cristina Feja,<sup>1,3</sup> Marc Saez,<sup>3,4</sup>  
M. Carmen Burriel,<sup>1</sup> Maria Antònia Barceló,<sup>3,4</sup> Pilar Gómez,<sup>1</sup>  
Gemma Renart,<sup>3,4</sup> Tomás Alcalá,<sup>2,5</sup> and Rafael Marcos-Gragera<sup>6</sup>

## Suggested citation

Martos MC, Saurina C, Feja C, Saez M, Burriel MC, Barceló MA, et al. Accurately estimating breast cancer survival in Spain: cross-matching local cancer registries with the National Death Index. *Rev Panam Salud Publica.* 2009;26(1):51–4.

## ABSTRACT

**Objectives.** To assess the impact of using data from the National Death Index (NDI) of Spain to estimate breast cancer survival rates among residents of Girona and Zaragoza diagnosed in 1995–1999.

**Methods.** This was an observational, longitudinal epidemiologic study, using two population-based cancer registries. Data collected were of female residents of Girona or Zaragoza who had been diagnosed with breast cancer in 1995–1999. Observed and relative 5-year survival rates were estimated, first using the information available from the Girona and Zaragoza cancer registries, and then with the inclusion of NDI data. The 5-year relative survival rate and corresponding 95% Confidence Intervals were estimated using the Hakulinen method. The Kaplan-Maier method and Log Rank test were used to compare survival curves.

**Results.** No statistically significant difference in survival curves was observed in Girona for the data obtained before and after cross-matching with the NDI. However, there was a significant difference in Zaragoza. A comparison of the relative survival rates of each of the two registries before NDI cross-matching showed differences of 3.9% (5-year) and 16.1% (10-year) between the two, whereas after the cross-match, the difference was only 0.5% (5-year) and 1.2% (10-year).

**Conclusions.** In Spain it is imperative that there be systematic use of NDI data to supplement cancer registries, so that comparisons of relative survival rates between registries can be improved.

## Key words

Breast neoplasms, surveillance; medical record linkage; vital statistics; cancer survival rate; Spain.

<sup>1</sup> Dirección General de Salud Pública, Departamento de Salud y Consumo, Gobierno de Aragón, Spain. Send correspondence to: M. Carmen Martos, Dirección General de Salud Pública, Departamento de Salud y Consumo, Gobierno de Aragón, Vía Universitat 36, 50017 Zaragoza, Spain; telephone: +34-976715915; fax: 34-976715016; email: cmartos@aragon.es

<sup>2</sup> Instituto Aragonés de Ciencias de la Salud, Zaragoza, Spain.

<sup>3</sup> Grup de Recerca en Estadística, Economia Aplicada i Salut (GRECS), Universitat de Girona, Spain.

<sup>4</sup> CIBER de Epidemiología y Salud Pública (CIBERESP), Spain.

In Spain, breast cancer is the most common malignancy among women. The world age-adjusted mortality rate for 1998–2002 was 14.7 deaths per 100 000 women/year. Although globally, the es-

timated age-standardized incidence rate for breast cancer in this period was 59.0 cases per 100 000 women-years, it was higher in Girona (66.5), Navarra (66.2), and Basque Country (64.3), and lower in Zaragoza (54.6), Asturias (54.1), Albacete (47.7), and Cuenca (45.2) (1). In addition, Municipal Mapping of Cancer Mortality in Spain shows a medium-high risk of mortality from breast cancer in the city of

<sup>5</sup> Universidad de Zaragoza, Spain.

<sup>6</sup> Unitat d'Epidemiologia y Registre del Càncer de Girona (UERCG), Pla Director d'Oncologia, Departament de Salut, Girona, Spain.

Zaragoza and in many towns in the province of Girona (2).

Survival is one way to measure the overall organizational efficiency of health services that manage cancer patients. Population models of survival based on cancer registries provide a useful tool for evaluating health systems and the effectiveness of interventions (3). In Spain, survival rates for breast cancer have risen in recent years, mortality rates have remained stable or decreased, and overall incidence has increased moderately (4). However, increases in survival rates are not easy to interpret because they could be due, at least in part, to improvements in treatment, more effective treatment due to earlier diagnoses, or over-diagnosis due to screening programs, or simply to lead-time bias due to earlier diagnoses (5–8). Furthermore, because of certain problems, especially in terms of coverage and continued follow-up of patients, the quality of data from cancer registries may compromise the interpretation of the geographical variations of survival rates (7, 9).

There are 11 population-based cancer registries in Spain that were included by the International Agency for Research on Cancer in its publication, *Cancer Incidence in Five Continents*. Zaragoza and Girona were among them (1). The use of standardized criteria for choosing sources and collecting information, along with the systematic elaboration of quality indicators, ensures the comparability of incidence estimates among different registries (10). However, since follow-up procedures are not standardized and actually vary widely due to issues with accessing individual mortality data, comparability of survival rates is greatly compromised.

Currently, information on cancer survival rates is scarce. As such, only data provided by registries participating in the EURO-CARE4 study are available. According to this data, 5-year relative survival for female breast cancer diagnosed in 1995–1999 in Spain was 80.8% (7).

Since 2006, Spain has made available its National Death Index (NDI). The index may contribute to improving the geographic comparability of estimated survival rates. The NDI is an information system containing personal data on each and every one of the deaths inscribed in the civil register throughout the country. It has the structure of a database and its variables include: name of the deceased,

national identity card number (since 1996), date and place of birth, sex, marital status, nationality, home address, date of death, and place of death. According to the user's manual (1), it contains data from 1987 onwards. Access is restricted to research centers, government health boards, and health centers and other establishments, all of which are state-controlled.

The objective of this article is to assess the impact using NDI data to estimate survival rates among patients who were residents of Girona and Zaragoza and were diagnosed with breast cancer in 1995–1999.

## MATERIALS AND METHODS

An observational, longitudinal, epidemiologic study was carried out using the population-based the cancer registries of Girona and Zaragoza, two areas in northeastern Spain, to identify female residents diagnosed with breast cancer (*International Classification of Diseases—10th Edition*, code C50) (11) in 1995–1999. In 2001, the Girona registry covered a population of 553 661, and the registry of Zaragoza, 861 855. The sources of information for each of the registries are their respective regional and local hospitals, as well as hematology and pathology departments. However, the patient health status and, if applicable, date of death, were monitored in different ways by each registry. In Girona, which is one of four provinces in Catalonia, data from the Catalan Mortality Registry and the Girona Cancer Registry are cross-matched annually through record-linkage. In Zaragoza, monitoring was done through death certificates in which the cause of death specified “cancer,” “carcinoma,” or “neoplasia,” as well as by the Health Identity Card Registration and hospital discharges due to death.

For the purposes of this study, all cases of breast cancer from the two registries were cross-matched, through record-linkage, with data from the NDI (12). The cross-matching was first done automatically via the website of the Ministry of Health, followed by a manual search of the database for cases not identified previously. The variables used for the automatic cross-matching were: last name(s); first name(s); and date of birth. Patients not found automatically were sought manually by: first name(s), last name(s), date of birth, and National

Identity Number, if available (recorded in the databases since 1996).

For the survival rate analysis, cases obtained only from the death certificates were excluded: 26 in Girona (1.9%) and 68 in Zaragoza (3.5%). Patients were followed up until 31 December 2005. The 5-year observed and relative survival rates in Girona and Zaragoza were subsequently estimated using information available previously from the registries; later, the NDI data (date of death) were included as well. To that end, the Kaplan-Maier method and the Log Rank test were employed for comparison (13). The 5-year relative survival rates and 95% Confidence Intervals (95%CI), estimated using the Hakulinen method, were obtained using WAERS-ICO software, available from Catalan Institute of Oncology website (14).

## RESULTS

In 1995–1999, there were 1 381 new cases of invasive breast cancer registered in Girona; of these, 26 (1.9%) were reported only by means of a death certificate. In Zaragoza, 1 951 cases were identified, of which 68 (3.5%) were reported only by death certificate. When cross-matching the data in the NDI, the date of death was modified or completed in 13 cases in Girona and 210 cases in Zaragoza, representing 0.9% and 10.8% of cases, respectively.

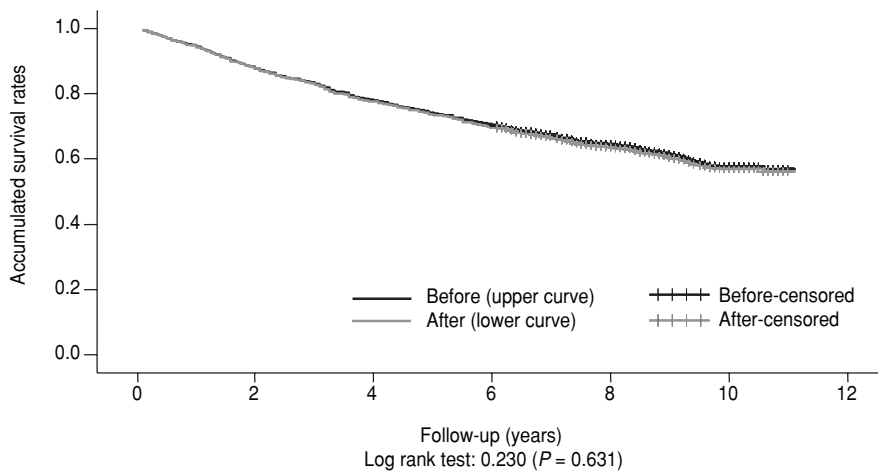
No statistically significant difference in survival curves was observed in Girona for the data obtained before and after the cross-matching with the NDI (Figure 1). However, there was a significant difference ( $P < 0.001$ ) in Zaragoza (Figure 2).

When comparing the relative survival rates of each of the two registries prior to the NDI cross-matching, there was a difference 3.9%, whereas after the cross-matching the difference was only 0.5%. The difference prior to using NDI was much higher at the 10-year relative survival mark, with 16.1% and 1.2% respectively (Table 1).

## DISCUSSION

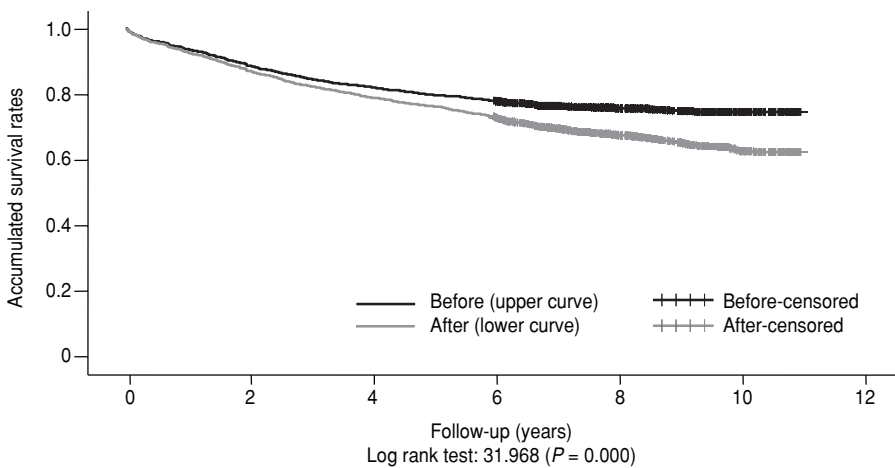
The differences found between the 5- and 10-year relative survival estimates for breast cancer in Girona and Zaragoza are explained, at least in part, by whether or not there is access to individualized mortality data. While the population-based

**FIGURE 1. Survival functions before and after cross-matching with the National Death Index, Girona, 1995–1999**



**Note:** Colored version of this figure is available at <http://journal.paho.org>

**FIGURE 2. Survival functions before and after cross-matching with the National Death Index, Zaragoza, 1995–1999**



**Note:** Colored version of this figure is available at <http://journal.paho.org>

**TABLE 1. Five-year and 10-year relative breast cancer survival rates, before and after cross-matching with the National Death Index (NDI), Girona and Zaragoza, 1995–1999**

Registry	Breast cancer survival rates				
	Before NDI cross-matching		After NDI cross-matching		Difference (%)
	%	95%CI <sup>a</sup>	%	95%CI <sup>a</sup>	
<b>5-year relative survival rates</b>					
Girona	81.0	78.5–83.6	80.5	78.0–83.1	0.5
Zaragoza	85.6	83.6–87.6	81.7	79.6–83.8	3.9
<b>10-year relative survival rates</b>					
Girona	71.4	67.9–75.1	70.2	66.7–73.8	1.2
Zaragoza	85.6	83.0–88.2	69.5	66.2–73.0	16.1

<sup>a</sup>95% Confidence Interval.

cancer registries in Catalonia (which includes Girona) have access to the Catalonia Mortality Registry and systematically cross-match each registry, this is not possible in Aragon (which includes Zaragoza) and other autonomous areas, such

as Andalucía. The main reason for this in Aragon is that legislation regarding health and demographic statistics is not as progressive as it is in Catalonia, making it difficult to access data from information systems outside the central health

administration. The slight overestimation of breast cancer survival produced by data from the Girona Cancer Registry was due to patient emigration following cancer diagnosis. Access to the NDI revealed the true status of these patients.

By using NDI data in conjunction with both cancer registries to obtain the vital status of breast cancer patients and, if applicable, the date of death, 5- and 10-year relative survival rates were found to be similar: Girona, 80.5% and 70.2%; and Zaragoza, 81.7% and 69.5%. Five-year relative survival figures are slightly higher than the average for Europe (79.2%) described by the EURO CARE-4 project for the same period (15), although France and all of the Scandinavian countries except Denmark have better 5-year relative breast cancer survival rates than do Girona and Zaragoza (8). Since Scandinavian cancer registries have national coverage and systematic record-linkage with mortality registries and population databases, differences in survival rates were presumed to be reliable. However, in France, just as in Spain, there is no national cancer registry, and survival is estimated from regional registers; furthermore, there are difficulties in accessing individualized mortality data. Therefore, survival rate estimates based on these records must be interpreted with caution.

Data on survival rates collected from the population-based cancer registries allow global assessment of prevention and control activities (16). Moreover, the health authorities need cancer survival statistics in order to plan and execute local cancer control programs and to optimize the allocation of resources. However, differences in the quality of data may jeopardize geographic and temporal comparability. This may also influence the interpretation of the results obtained (17).

Most of the population-based cancer registries in Spain depend on the governments of autonomous regions. Legislative differences and variations in their health systems, priorities, and resources may compromise the comparability of data obtained from their information systems. It is therefore necessary to develop initiatives at the national level that will improve the comparability of health indicators across autonomous communities.

In summary, these results show that systematic use NDI data to supplement cancer registries is imperative to improving comparisons of relative survival rates between registries.

**Acknowledgments.** This study received funding from the following: the Fondo de Investigación (Health Research Fund) of the Ministry of Health and Consumption, projects ETS-06/90553 and ETS-07/90453; Agencia de Evaluación de Tecnología e Investigación Médicas

(Agency for Technology Evaluation and Medical Research); Catalan Health Service, Generalitat de Catalunya (Autonomous Catalan Government), project AATRM-006/01/2006; and the University of Girona as part of the IdIBGi 2006 project. We appreciate the comments of

the three anonymous peer reviewers, the editor, and the attendees of the XXV Scientific Conference of the Epidemiology Association of Spain, in Córdoba, Spain, 17–19 October 2007, where a preliminary version of this work was presented.

## REFERENCES

- Curado M P, Edwards B, Shin HR, Storm H, Ferlay J, Heanue M, Boyle P eds. Cancer incidence in five continents, Vol. IX. Lyon: International Agency for Cancer Research; 2007. (Scientific Publication No. 160).
- Lopez-Abente G, Ramis R, Pollán M, Aragonés N, Pérez-Gómez B, Gómez-Barroso D, et al. Atlas de municipal de mortalidad por cáncer en España, 1989–1998. Madrid: Instituto de Salud Carlos III; 2006.
- Stracci F, La Rosa F, Falsetti E, Ricci E, Aristei C, Bellezza G, Bolis GB, Fenocchio D, Gori S, Rulli A, Mastrandrea V. A population survival model for breast cancer. *Breast*. 2005; 14(2):94–102.
- Izquierdo A, González JR, Moreno V, Galcerán J, Marcos-Gragera R, Clèries R, Borràs J. Incidencia de cáncer de mama en Cataluña: Análisis de tendencias. *Med Clin (Barc)*. 2006;126(8):289.
- Sant M, Francisci S, Capocaccia R, Verdecchia A, Allemani C, Berrino F. Time trends of breast cancer survival in Europe in relation to incidence and mortality. *Int J Cancer*. 2006; 119: 2417–22.
- Zahl PH, Strand BH, Maehlen J. Incidence of breast cancer in Norway and Sweden during introduction of nationwide screening: prospective cohort study. *Br Med J*. 2004;328(7445):921–24.
- Berrino F. The EURO CARE study: strengths, limitations and perspectives of population-based, comparative survival studies. *Ann Oncol*. 2003;14(Suppl 5):v9-13.
- Berrino F, Capocaccia R. Survival of European cancer patients. In: Coleman MP, Alexe D-M, Albret T, McKee M. Responding to the challenge of cancer in Europe. Ljubljana: Institute of Public Health of the Republic of Slovenia; 2008. Pp. 151–76.
- Sant M, Francisci S, Capocaccia R, Verdecchia A, Allemani C, Berrino F. Should we use incidence, survival or mortality to assess breast cancer trends in European women? *Nature*. 2006;3:228–9.
- Tyczynski JE, Démaret E, Parkin DM. Standards and guidelines for cancer registration in Europe 2003, Vol. I. Lyon: International Agency for Cancer Research; 2003. (Scientific Publication No. 40).
- International Statistical Classification of Diseases and Related Health Problems, 10th revision. Available at: <http://www.who.int/classifications/apps/icd/icd10online/>. Accessed 23 February 2008.
- Navarro C. El índice nacional de defunciones: un avance en la accesibilidad de los datos de mortalidad largamente esperado. *Gac Sanit*. 2006;20:421–3.
- Pérez-Hoyos S. Análisis de supervivencia. *Quaderns de salut pública administració de serveis de salut*. Valencia: Institut Valencià d'Estudis en Salut Pública; 1997.
- 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.
- Berrino F, De Angelis R, Sant M, Rosso S, La-sota MB, Coebergh JW, Santaquilani M. Survival for eight major cancers and all cancers combined for European adults diagnosed in 1995–99: results of the EURO CARE-4 study. *Lancet Oncol*. 2007;8(9):773–83.
- Martinez-Garcia C, Barlow L, Moller T, Lutz JM, Andre M, Steward JA. European health systems and cancer care. *Ann Oncol*. 2003; 14(Suppl 5):v41–60.
- Rachet B, Riga M, Mity E, Romanengo M, Quinn M, Cooper N, Coleman M. Geographical comparisons of cancer survival indicators. *Health Stat Q*. 2004;(22):5–13.

Manuscript received on 15 June 2008. Revised version accepted for publication on 7 December 2008.

## RESUMEN

### Estimación precisa de la supervivencia de cáncer de mama en España: emparejamiento de los registros locales de cáncer con el Índice Nacional de Defunciones

**Objetivo.** Evaluar el efecto de utilizar los datos del Índice Nacional de Defunciones (IND) de España para estimar las tasas de supervivencia de cáncer de mama en las mujeres residentes en Girona y Zaragoza que recibieron el diagnóstico de cáncer de mama en 1995–1999.

**Métodos.** Se realizó un estudio epidemiológico observacional y longitudinal basado en el empleo de los registros de cáncer de mujeres residentes en Girona y Zaragoza que habían recibido el diagnóstico de cáncer de mama en 1995–1999. Se estimaron las tasas de supervivencias observada y relativa a 5 años, primero según la información disponible en los registros de cáncer de Girona y Zaragoza y luego con la inclusión de los datos del IND. Se calcularon las tasas de supervivencia relativa a 5 años y sus correspondientes intervalos de confianza de 95% por el método de Hakulinen. Las curvas de supervivencia se compararon por el método de Kaplan-Maier y la prueba de rangos logarítmicos.

**Resultados.** No se encontraron diferencias estadísticamente significativas entre las curvas de supervivencia de Girona antes y después de emparejar los datos locales con los del IND; sin embargo, hubo diferencias significativas entre las curvas de Zaragoza. Al comparar las tasas de supervivencia relativa de cada uno de los registros antes de emparejarlos con los datos del IND se encontraron diferencias de 3,9% (a 5 años) y 16,1% (a 10 años), mientras que después del emparejamiento, la diferencia entre ellas fue solamente de 0,5% (a 5 años) y 1,2% (a 10 años).

**Conclusiones.** En España es imperativo el empleo sistemático de los datos del IND para complementar los registros de cáncer de manera de mejorar las comparaciones de las tasas de supervivencia relativa cuando se utilizan diferentes registros.

## Palabras clave

Neoplasias de la mama, vigilancia; registro médico coordinado; estadísticas vitales; tasa de supervivencia; España.