

# ANÁLISIS DE SUPERVIVENCIA DEL REGISTRO POBLACIONAL DE CÁNCER DE MAMA DE LA PROVINCIA DE CASTELLÓN

M. Alacreu<sup>1</sup>, M. A. Martínez-Beneito<sup>1</sup>, A. Torrella<sup>1</sup>, C. Armero<sup>2</sup>, A. Frau<sup>3</sup>, I. Melchor<sup>1</sup>, E. Guallar<sup>1</sup>, J. J. Abellán<sup>2,4</sup> y O. Zurriaga<sup>1</sup>

<sup>1</sup>Àrea d'Epidemiologia. Direcció General de Salut Pública. Conselleria de Sanitat. Generalitat Valenciana

<sup>2</sup>Departament d'Estadística i Investigació Operativa. Universitat de València.

<sup>3</sup>Servei d'Oncologia Mèdica. Hospital Provincial de Castelló.

<sup>4</sup>Institut Valencià d'Estadística.

## Antecedentes

Dentro de la Dirección General de Salud Pública e integrado en el Servicio de Epidemiología (actual Área de Epidemiología), se creó el Registro monográfico de cáncer de mama de la provincia de Castellón en 1998. Dicho registro, es un sistema de información de base poblacional sobre mujeres diagnosticadas de cáncer de mama. La información se recopila mediante la revisión de las historias clínicas en distintos centros hospitalarios de la provincia de Castellón y en dos de la provincia de Valencia. Sobre esta fuente de información desarrollamos una línea de investigación cuyo objetivo es analizar la influencia de las variables, recogidas en el registro, sobre el tiempo de supervivencia. Con este estudio pretendemos detectar y valorar los principales factores que afectan significativamente a la evolución de esta patología tumoral. Así mismo, también estudiamos si las variables pronóstico sufrían cambios cuando eliminábamos el efecto producido por cualquier causa de muerte que no fuera la del cáncer de mama.

## Métodos

El banco de datos inicial está constituido por 1015 observaciones procedentes del Registro Poblacional de Cáncer de Mama de la provincia de Castellón, que corresponden a los casos incidentes en el periodo 1995-1999. Iniciamos el estudio aplicando procedimientos de supervivencia. Escogimos para ello el modelo de regresión de Cox porque permite una modelización multivariante de la supervivencia, con la que se puede cuantificar para cada individuo la evolución temporal del riesgo de fallecer desde la fecha de incidencia, teniendo en cuenta la influencia de las características individuales de cada paciente.

En la primera fase del estudio planteamos regresiones de Cox univariantes, para valorar e identificar aquellas variables que mejor explicaban la mortalidad de las pacientes, así como los tests que evaluaban si se cumplía la proporcionalidad de riesgos.

En esta etapa, dos de las variables que adquirieron especial importancia fueron, la variable que recoge la confirmación de una nueva reproducción de cáncer en la mama (recidiva), y la que describe la confirmación de que la paciente ha desarrollado un cáncer en otra localización tumoral durante el seguimiento (metástasis). Ambas variables pueden dar origen a resultados erróneos si no se les da el tratamiento técnico adecuado, ya que recogen sucesos que se producen durante el seguimiento y marcan un antes y un después tanto en la evolución de la enfermedad como en sus consecuencias. Por ejemplo, si desde que incide una paciente hasta que fallece ha pasado mucho tiempo y poco antes de fallecer ha desarrollado metástasis, los resultados podrían llevarnos a interpretaciones erróneas del tipo “las pacientes con metástasis viven mucho tiempo”. Sin embargo, lo que realmente está pasando es que cuando se desarrolla metástasis, la esperanza de vida es muy corta.

Para obtener un análisis estadístico correcto, consideramos la recidiva, y el desarrollo de metástasis durante el seguimiento como “variables cambiantes en el tiempo”. Esta estrategia consiste básicamente en generar un nuevo banco de datos en el que el seguimiento de una paciente, a quien sucede una recidiva y/o una metástasis, esté “troceado” en varios registros. Por ejemplo, uno que recoja la situación de la paciente desde la incidencia hasta la recidiva (si se produce), otro registro que reúna la información desde la recidiva hasta la metástasis (si procede) y por último, otro que recopile los detalles desde la metástasis hasta el fallecimiento o la fecha fin de estudio, que en este caso se fijó a 31-12-2001. Utilizando este procedimiento pasamos a trabajar del banco de datos original, en el que cada registro recogía la información de una paciente, a un nuevo banco de datos en el que cada paciente puede identificarse en varios registros.

Otra cuestión a señalar, es que las covariables que pasan a formar parte de la regresión de Cox contribuyen de forma lineal en el modelo pero en el caso de las covariables numéricas, existe la posibilidad de valorar si también tiene importancia una componente no lineal, mediante el uso de splines. De esta forma mejoramos el ajuste, permitiendo así una interpretación de los resultados, más acordes con la realidad.

Por último, dado que el cáncer de mama es un tumor de los que mejor pronóstico tienen y que por lo tanto estas pacientes pueden fallecer por otras causas diferentes a esta enfermedad, en la última fase de la investigación planteamos el análisis de supervivencia considerando como evento el fallecimiento por cáncer de mama. El objetivo era eliminar el efecto del fallecimiento por otras causas de muerte para averiguar si las variables de influencia cambiaban.

## **Resultados**

El banco de datos final estuvo constituido por 1173 registros, de los cuales 99 fueron desechados por contener valores perdidos en la variable que recoge el número

de ganglios afectos y 109 fueron eliminados en la selección de los casos invasivos, con objeto de obtener resultados comparables con la literatura científica. Con esta fuente empezamos el proceso de selección de entre más de cien variables. Cabe destacar que variables como el año de incidencia, el hospital, los tratamientos de radioterapia y quimioterapia, el tipo histológico, etc., no resultaron significativas. Después de varios ensayos, obtuvimos una regresión de Cox definitiva, en la que resultaban variables significativas la edad, el método de detección, los antecedentes familiares, el tamaño del tumor, el número de ganglios afectos, la metástasis al inicio, el grado de diferenciación, los receptores hormonales estrogénicos, el tratamiento con hormonoterapia y la metástasis en el seguimiento. Esta regresión se evaluó considerando como evento el fallecimiento de la paciente sin distinguir la causa de muerte, obteniendo como resultados que la edad entra en el modelo de forma lineal y que por cada año de edad, se incrementa el riesgo de fallecer en 1.8 %.

Respecto al método de detección, no existen diferencias significativas entre la detección por el programa de cribado o por cáncer de intervalo pero sí que existen diferencias entre haber sido detectada por el programa de cribado y otros métodos, incrementándose el riesgo en este caso un 127.2 %. Por otra parte, tener antecedentes familiares de primer grado aumenta el riesgo un 111 %, no se detectan diferencias significativas entre no tener antecedentes familiares y tenerlos de segundo grado o en otras localizaciones tumorales. En referencia al tamaño del tumor, vemos que el aumento de riesgo es significativo a partir de 5 centímetros de diámetro o si el tumor presenta afección de la piel o carcinoma inflamatorio, siendo el aumento un 110.4 %. El número de ganglios afectos, es una variable cuantitativa que entra en el modelo de forma lineal y con la que obtenemos que por cada ganglio aumenta el riesgo un 6.7 %. Tener metástasis al inicio, aumenta el riesgo de fallecer en un 875.6 %. El grado de diferenciación es un factor que en el peor de los casos (grado de diferenciación poco diferenciado) aumenta el riesgo de fallecer hasta un 128 %. Tener receptores hormonales estrogénicos positivos, se comporta como factor protector, disminuyendo el riesgo un 43 %. El tratamiento con hormonoterapia también disminuye el riesgo un 53 %. Por último, desarrollar metástasis durante el seguimiento, aumenta el riesgo en un 3521 %.

En la última etapa del estudio, ajustamos el modelo de Cox, considerando como evento el fallecimiento por cáncer de mama como única causa de muerte. Curiosamente, tanto la edad como el grado de diferenciación dejaron de ser variables significativas y reproducir de nuevo el cáncer en la mama (recidiva) pasó a tener un efecto significativo. En cuanto al método de detección, observamos un aumento de riesgo a fallecer por cáncer de mama de un 73.4 %, si la detección se hace por un método diferente al programa de cribado. Respecto a los antecedentes familiares, no se detectan diferencias entre no tener antecedentes y tenerlos de primer o de segundo grado, sin embargo existe un aumento de riesgo, del 55.7 %, entre no tener antecedentes y tenerlos en

otras localizaciones tumorales. Referente al tamaño del tumor, se aprecian diferencias significativas en todas las categorías. Se observa que cuanto mayor es el tamaño del tumor, mayor riesgo de fallecer por cáncer de mama. Por otra parte, al introducir en el modelo el número de ganglios afectos comprobamos que no sólo contribuía significativamente de forma lineal. Mediante la función spline, verificamos que había una componente no lineal significativa, que explicaba que la diferencia de riesgo, entre tener  $n$  ganglios afectos y tener  $n + 1$ , no es siempre el mismo. La metástasis al inicio supone un aumento de riesgo del 259.4%. Tener receptores hormonales estrogénicos positivos disminuye el riesgo, de fallecer por cáncer de mama, en un 61%. Recibir tratamiento con hormonoterapia supone reducir el riesgo en un 37%. Reproducir de nuevo el cáncer en la mama durante el seguimiento aumenta el riesgo un 60% y por último, el factor que peor pronóstico tiene es generar una metástasis durante el seguimiento, que produce un aumento de riesgo de un 993.2%.

## **Conclusiones**

Comparando los dos ajustes planteados con el modelo de Cox, observamos que tanto la edad de una paciente como el grado de diferenciación, dejan de tener carácter pronóstico de la supervivencia cuando eliminamos el efecto por cualquier causa de muerte distinta al cáncer de mama y ajustamos por el resto de variables. Los antecedentes familiares, el tamaño tumoral y el número de ganglios afectos son variables que, aun manteniéndose significativas en las dos regresiones cambian su matiz cuando seleccionamos sólo las muertes por cáncer de mama. Estas diferencias señalan la importancia de conocer la causa de muerte, para obtener resultados más precisos que contribuyan a identificar políticas de intervención en el proceso asistencial, que mejoren la supervivencia del cáncer.

## **Agradecimientos**

Este trabajo ha recibido financiación del fondo de investigación sanitaria (FIS 01/0557E) y ha sido parcialmente financiado por una subvención de la Conselleria de Sanitat de la Generalitat Valenciana al Departament d'Estadística e Investigació Operativa de la Universitat de València.