



# Aplicación de métodos y herramientas de la ciencia de datos para el análisis del coste de la asistencia sanitaria de los pacientes con cáncer de pulmón en un hospital universitario de tercer nivel en España

JUAN LUIS CRUZ-BERMÚDEZ<sup>1\*</sup>, CONSUELO PAREJO<sup>2</sup>, FRANCISCO MARTÍNEZ-RUIZ<sup>3</sup>, JUAN CRISTÓBAL SÁNCHEZ-GONZÁLEZ<sup>2</sup>, RAQUEL GÓMEZ-BRAVO<sup>2</sup>, ARTURO RAMOS MARTÍN-VEGUE<sup>4</sup>, ANA ROYUELA<sup>5</sup>, ALEJANDRO RODRÍGUEZ-GONZÁLEZ<sup>6</sup>, ERNESTINA MENASALVAS-RUIZ<sup>6\*</sup>, MARIANO PROVENCIO<sup>2\*</sup>.

<sup>1</sup>Servicio de Informática. Hospital Universitario Puerta de Hierro-Majadahonda. Universidad Politécnica de Madrid. <sup>2</sup>Servicio de Oncología Médica. Hospital Universitario Puerta de Hierro-Majadahonda. <sup>3</sup>Facultad de Matemáticas. Universidad de Valencia. <sup>4</sup>Servicio de Admisión y Documentación Clínica. Hospital Universitario Puerta de Hierro-Majadahonda. <sup>5</sup>Unidad de Bioestadística. Instituto de Investigación Sanitaria Puerta de Hierro - Segovia de Arana. <sup>6</sup>Centro de Tecnología Biomédica. Universidad Politécnica de Madrid.

## Resumen

**Objetivos:** Estudiar el coste de la asistencia sanitaria de los pacientes con cáncer de pulmón en el Hospital Universitario Puerta de Hierro Majadahonda (HUPHM), aprovechando nuestra Historia Clínica Electrónica (HCE) y el registro clínico estructurado de casos de cáncer de pulmón.

**Métodos:** Aplicamos Cross Industry Standard Process for Data Mining (CRISP-DM) para integrar y analizar datos de actividad extraídos de la HCE y datos clínicos de pacientes con cáncer de pulmón procedentes del registro de Oncología Médica. Los costes de la actividad asistencial fueron obtenidos de la definición de precios oficiales de la Comunidad de Madrid, recogidos en la Orden 731/2013 de precios públicos en el Servicio Madrileño de Salud. Hemos descrito y cuantificado factores con influencia en los costes de la asistencia sanitaria desde un punto de vista univariante.

**Resultados:** Los costes de hospitalización supusieron un 87% de los costes totales, excluyendo los costes de los tratamientos ambulatorios y la radioterapia. El estadio y ECOG-PS al diagnóstico y el índice de comorbilidad de Charlson mostraron diferencias en los costes asistenciales de los pacientes.

**Conclusiones:** La aplicación de CRISP-DM nos ha permitido describir el coste de la actividad hospitalaria del cáncer de pulmón mediante equipos multidisciplinares y la integración de datos de la HCE y de registros clínicos estructurados. Es un primer paso para la aplicación de algoritmos de Machine Learning predictivos y de soporte a la decisión. Los factores descritos relacionados con el coste son coherentes con el sentido común y el conocimiento clínico actual, si bien hemos podido cuantificarlos y pueden ser de utilidad para planificar recursos y mejorar la gestión de estos pacientes.

**Palabras clave:** Cáncer de pulmón, costes de la asistencia sanitaria, Data Science, Data Mining, CRISP-DM.

## Abstract

**Objectives:** To study the health care cost of lung cancer patients in Hospital Universitario Puerta de Hierro Majadahonda (HUPHM) by leveraging our Electronic Patient Record (EPR) and structured clinical registry of lung cancer cases.

**Methods:** We applied the Cross Industry Standard Process for Data Mining (CRISP-DM) to integrate and analyze activity data extracted from the EPR and clinical data of lung cancer patients from a Medical Oncology registry. Activity costs were obtained from the official definition published in the law "Orden 731/2013 de precios públicos en el Servicio Madrileño de Salud". We have described and quantified factors with an influence in health care costs from a univariate point of view.

**Results:** Hospitalization costs accounted for the 87% of overall costs, excluding ambulatory treatments and radiotherapy costs. Stage and ECOG-PS at diagnosis and Charlson Comorbidity Index showed differences in health care costs for lung cancer patients.

**Conclusions:** Applying CRISP-DM has allowed us to describe hospital lung cancer health care costs through a multidisciplinary team and integrating activity data from EPR and clinical structured data. This is a first step towards the application of predictive and decision support Machine Learning algorithms. The factors described related with costs are consistent with common sense and current clinical knowledge, although we have been able to quantify them and could be used to plan resources and improve quality of care.

**Key words:** Lung cancer, health care cost, Data Science, Data Mining, CRISP-DM.

## INTRODUCCIÓN

El cáncer continúa siendo un problema de salud y económico de primera magnitud. De acuerdo con Eurostat el cáncer causa una de cada cuatro muertes en Europa<sup>1</sup>. Específicamente el cáncer de pulmón es el que más fallecimientos provoca, un 20,8% de todas las muertes por cáncer, con más de 266.000 fallecidos en Europa en 2011. Desde un punto de vista económico, el cáncer costó a la UE 126.000 millones de euros en 2009, siendo el coste de la asistencia sanitaria para estos pacientes de 51.000 millones de euros (40%). La pérdida de productividad derivada de la muerte prematura del paciente ascendió a 42.600 millones y la pérdida de días de trabajo a 9.430 millones. De todos los tipos de cáncer, el cáncer de pulmón tuvo el impacto económico más alto (18.800 millones, un 15% del total). En España, se estima que el coste total del cáncer fue de 9.016 millones de euros (un 0,86% del PIB) en 2009, siendo el coste del cáncer de pulmón de 1.258 millones de euros (un 13,95%)<sup>2</sup>. El cáncer de pulmón es el que mayor incidencia presenta a nivel mundial (1.824.701 nuevos casos en 2012 de acuerdo al estudio GLOBOCAN 2012<sup>3</sup>) y el que mayor número de muertes produce (una tasa de mortalidad de 33,3/100.000 habitantes en 2017, lo que supone 183.400 fallecimientos). En España, de acuerdo a los datos publicados en 2018 por la Sociedad Española de Oncología Médica (SEOM)<sup>4</sup>, se estiman 28.645 nuevos casos en 2017. El número de fallecimientos en 2016 por cáncer de pulmón ascendió a 22.187, siendo el tumor que más muertes produjo. Estos datos, sin embargo, no proceden de un registro de casos nacional, sino de la extrapolación a los datos de población española de los datos de GLOBOCAN 2012.

### \*Correspondence to:

Juan Luis Cruz-Bermúdez  
 email: jlcruz@idiphim.org  
 ORCID: 0000000156501160

Ernestina Menasalvas  
 email: ernestina.menasalvas@ctb.upm.es

Mariano Provencio  
 email: mariano.provencio@salud.madrid.org



Estamos convencidos de que la gestión de la asistencia hospitalaria del cáncer de pulmón puede beneficiarse de la aplicación de técnicas informáticas de procesamiento de la información clínica. Si bien el Hospital Universitario Puerta de Hierro Majadahonda (HUPHM) dispone de Historia Clínica Electrónica (HCE) desde finales de 2008, ésta no está pensada para realizar investigación clínica. Con la excepción de los datos demográficos y de los datos de actividad asistencial, la mayor parte de los datos clínicos se registran de forma no estructurada (en texto libre), lo que impide su análisis directo y su aprovechamiento secundario. Para suplir esta carencia, los servicios clínicos del HUPHM elaboran registros de casos asistenciales que integran datos clínicos estructurados de interés para la práctica asistencial y para la realización de investigación clínica.

Sin embargo, son pocos los estudios realizados hasta la fecha que integran una información clínica detallada para una patología concreta y relevante, como el cáncer de pulmón, con datos masivos de actividad asistencial disponibles en la HCE para estudiar los costes de la asistencia sanitaria. El análisis de estos datos integrados mediante técnicas estadísticas y de ciencia de datos (*Data Science*) podría aportar un conocimiento muy valioso que mejore la gestión de estos pacientes.

Los costes de la asistencia sanitaria se han calculado históricamente mediante sistemas de contabilidad analítica<sup>5</sup>, bien basados primero en el reparto del coste real observado entre los distintos servicios (centros de coste o grupos funcionales homogéneos, GFH), bien posteriormente a través del cálculo del coste por paciente y por proceso (proyecto Geclif del Ministerio de Sanidad y Consumo<sup>6</sup>). Mediante Geclif se realiza la asignación de recursos y la imputación de productos generados y de costes de los GFH implicados a los pacientes, para posteriormente realizar el cálculo del coste por Grupo Relacionado por el Diagnóstico (GRD). Sin embargo, el coste por paciente se ha considerado un producto intermedio para obtener un coste por GRD, mucho más manejable. Por otra parte, no ha sido objeto de estos sistemas de costes el estudio de los fac-

tores clínicos con influencia en el coste asistencial de patologías (diagnósticos) concretas, como el cáncer de pulmón.

En este contexto, tratamos en este estudio de aplicar de forma práctica la metodología *Cross Industry Standard Process for Data Mining* (CRISP-DM)<sup>7</sup> y las herramientas necesarias para la integración, depuración y análisis de los datos clínicos y asistenciales disponibles para mejorar el conocimiento existente acerca del coste de la asistencia sanitaria de los pacientes de cáncer de pulmón en el HUPHM.

## MÉTODOS

Se ha realizado un estudio descriptivo, longitudinal, observacional y retrospectivo sobre el registro asistencial de pacientes con cáncer de pulmón y los datos de actividad asistencial correspondientes, en los ámbitos ambulatorio, de urgencias y de hospitalización, con las características descritas a continuación. El estudio se ha realizado de acuerdo a los principios de la Declaración de Helsinki y su protocolo fue previamente aprobado por el Comité Ético de Investigación con medicamentos (CEIm) del Hospital.

### CRISP-DM

Se ha utilizado como marco metodológico CRISP-DM. CRISP-DM es la metodología más usada para el desarrollo de proyectos de descubrimiento de conocimiento y minería de datos y el estándar de facto de la industria<sup>8</sup>. Durante nuestro estudio hemos aplicado esta metodología, con un alcance superior al contenido en este estudio, adaptándola a las necesidades del entorno sanitario y del problema de análisis planteado, tomando como referencia estudios previos que aplican esta metodología al análisis de datos clínicos y asistenciales<sup>9,10</sup>.

### Fuentes de datos

Para la realización del estudio se parte de los datos del registro asistencial de pacientes con cáncer de pulmón, con una cohorte de 778 pacientes cuyo diagnóstico se produjo desde el 01/01/2003 hasta 31/12/2016. Dicho registro dispone de 56 variables, entre las que se encuentran variables demográficas (fecha de

nacimiento, sexo, raza), hábito tabáquico y su cuantificación, antecedentes de cáncer personales o en familiares de primer grado, situación de seguimiento y fallecimiento, y variables del diagnóstico, tales como fecha, histología, grado histológico, estadio, realización de estudios moleculares o el valor de la escala Eastern Cooperative Oncology Group Performance Status<sup>11</sup> (ECOG-PS). Para identificar adecuadamente a los pacientes fallecidos y actualizar sus fechas de fallecimiento, se realizaron cruces de información con el Índice Nacional de Defunciones (INDEF)<sup>12</sup>. El HUPHM dispone de un sistema de Historia Clínica Electrónica (Selene® de Cerner Corporation) que integra información de gestión asistencial y clínica para todos los pacientes. Dicho sistema, junto con los sistemas departamentales del hospital, integran su información mediante procesos de extracción, transformación y carga (ETL) en un sistema data *warehouse* en el marco de un mapa complejo de sistemas informacionales y descargas oficiales de información. Se obtuvo la siguiente información de actividad del sistema data *warehouse* y del sistema agrupador AP-GRD del centro, y se cruzó con el registro de casos de pacientes con cáncer de pulmón:

- Datos de actividad citada desde el 30/07/2008 al 01/08/2017: 7 844 820 registros y 20 columnas.
- Datos de asistencias en urgencias desde el 01/09/2008 hasta el 25/10/2017: 1 317 145 registros y 23 columnas.
- Datos de episodios de hospitalización del 01/01/2009 al 31/12/2014 procedentes del registro hospitalario codificado (CMBD) mediante CIE-9-MC y agrupado mediante AP-GRD v27: 149 951 registros y 149 columnas.

Los datos del registro de pacientes y de actividad asistencial fueron cruzados y exportados en formato Microsoft Excel, anonimizados y procesados adicionalmente para obtener un total de 133 variables derivadas, necesarias para los distintos análisis realizados durante los ciclos de la metodología CRISP-DM. Se calculó el índice de comorbilidad de Charlson (CCI o *Charlson comorbidity index*<sup>13</sup>) a partir de las 17 comorbilidades que define y que son obtenidas a partir del CMBD (Conjunto Mínimo Básico de

Datos) empleando los diagnósticos codificados mediante CIE-9-MC en todos los episodios de hospitalización posteriores al diagnóstico de cáncer de pulmón de los pacientes. Para ello se utilizó una herramienta desarrollada en Stata<sup>14</sup> que implementa una versión mejorada (Enhanced ICD-9-CM Coding Algorithm) del algoritmo de Deyo<sup>15</sup> con un mejor resultado<sup>16</sup>.

Para el cálculo de costes se siguió una aproximación bottom-up, obteniendo los costes de las asistencias unitarias y multiplicando por su número, de acuerdo a lo realizado en otros estudios de referencia<sup>17</sup>. El coste de las asistencias unitarias fue obtenido de la definición de precios oficiales de la Comunidad de Madrid, recogidos en la Orden 731/2013 de precios públicos en el Servicio Madrileño de Salud<sup>18</sup> que se basan en los resultados de métodos de contabilidad analítica como los descritos previamente.

### Análisis de datos

La obtención de los datos de actividad asistencial, filtrados con las variables de interés, se realizó mediante Oracle Business Intelligence Discoverer 10g Versión 2 (10.1.2.1). Para las tareas de preparación, cruce y depuración de datos se utilizó Microsoft Excel 2016, como una herramienta de uso sencillo y ampliamente disponible en el entorno hospitalario.

Los análisis exploratorios iniciales fueron realizados mediante Tableau 10.3, herramienta de análisis visual que permite de forma rápida la identificación de relaciones y tendencias entre los datos, y mediante Stata 12.0. Los análisis estadísticos fueron realizados mediante Stata 12.0, y Epidat 4.2<sup>19</sup>.

Las variables cualitativas (categóricas) se expresan mediante frecuencias, en términos absolutos y en porcentajes. Las variables cuantitativas (numéricas) como media, mediana, percentiles 25 (P25) y 75 (P75), desviación estándar (DT) y rango según el caso. Para comparar proporciones de variables categóricas se aplica el test ji-cuadrado de Pearson o el test exacto de Fisher. Para las variables cuantitativas, en la comparación de medidas centrales entre grupos se utilizan pruebas paramétricas (t de student y ANOVA) o pruebas no paramétricas (U-Mann-Whitney y Kruskal-Wallis) utili-



zando las más adecuadas en cada caso en función del cumplimiento de la asunción de la normalidad de los datos y del número total de pacientes de cada grupo. La supervivencia de los pacientes se estimó utilizando el método de Kaplan-Meier.

Las pruebas con un valor de  $p < 0,05$  fueron consideradas estadísticamente significativas, aplicando según el caso la corrección de Bonferroni en comparaciones entre múltiples grupos. Se utilizaron en todos los casos pruebas bilaterales.

## RESULTADOS

### Descripción de la cohorte de pacientes

Conformamos una cohorte de 522 pacientes, cuyo diagnóstico se produce desde el 01/01/2009 hasta 31/12/2016, para los cuales determinamos que sus datos han sido registrados con suficiente homogeneidad y calidad, y nuestra Historia Clínica Electrónica, implantada en septiembre de 2008, contiene todas sus asistencias sanitarias. En la Tabla 1 se describen las principales características al diagnóstico recogidas para esta cohorte de pacientes.

**TABLA 1**

PRINCIPALES CARACTERÍSTICAS AL DIAGNÓSTICO DE LA COHORTE DE PACIENTES ANALIZADA

Total	Total	Total	Total
	522 (100%)		522 (100%)
<b>Sexo</b>		<b>Estadio al diagnóstico</b>	
Hombre	392 (75,1%)	IA	59 (11,3%)
Mujer	130 (24,9%)	IB	51 (9,8%)
<b>Edad al diagnóstico</b>		IIA	31 (5,9%)
Media (DT)	65,9 (9,9)	IIB	30 (5,7%)
Mediana (Mín.-Máx.)	66 (33-92)	IIIA	129 (24,7%)
<b>Raza</b>		IIIB	67 (12,8%)
Caucásico	499 (95,6%)	IV	155 (29,7%)
Latino	14 (2,7%)	<b>Antecedentes familiares de 1º grado</b>	
Africano	9 (1,7%)	Sí	131 (25,1%)
<b>Hábito tabáquico</b>		No	124 (23,8%)
Ex fumador (>1 año)	322 (61,7%)	Desconocido	267 (51,1%)
Paquetes año. Mediana (P25-P75)	50 (40-65)	<b>Antecedentes personales</b>	
Cigarrillos día. Mediana (P25-P75)	20 (20-40)	Sí	104 (19,9%)
Fumador activo	129 (24,7%)	No	331 (63,4%)
Paquetes año. Mediana (P25-P75)	50 (37-70)	Desconocido	87 (16,7%)
Cigarrillos día. Mediana (P25-P75)	20 (20-40)	<b>ECOG-PS al diagnóstico</b>	
Nunca fumador	63 (12,1%)	0	289 (55,4%)
Desconocido	8 (1,5%)	1	181 (34,7%)
<b>Histología</b>		2	37 (7,1%)
Adenocarcinoma	268 (51,3%)	3	5 (1,0%)
Escamoso	190 (36,4%)	4	1 (0,2%)
Carcinoma de células grandes	32 (6,1%)	Desconocido	9 (1,7%)
NOS/Indiferenciado	9 (1,7%)	<b>ECOG-PS al diagnóstico</b>	
Carc. neuroendocrino de cél. grande	6 (1,1%)	1-2	37 (7,1%)
Carc. de célula pequeña	6 (1,1%)	3-4	59 (11,3%)
Adenoescamoso	4 (0,8%)	>=5	192 (36,8%)
Sarcomatoide	1 (0,2%)	Desconocido	234 (44,8%)
Tumor carcinoide	1 (0,2%)		
Otros	5 (1,0%)		

Fuente: Elaboración propia.

Dichos pacientes fueron seguidos desde el diagnóstico hasta cierre del estudio (01/08/2017), pérdida de seguimiento o fallecimiento, con una mediana de seguimiento de 21,1 meses (rango 0,1-103,6 meses). Se registraron 226 muertes (43,3%) durante el seguimiento, con una mediana de supervivencia estimada de 51,6 meses (IC95% 32,7-70,5 meses). La mediana de la supervivencia para los pacientes fallecidos fue de 326 días, y la media de 441 días (IC95% 383 – 500).

### Análisis del coste de consultas externas y pruebas

De la cohorte de 522 pacientes, 516 (98,9%) requirieron alguna asistencia ambulatoria en el hospital posterior al diagnóstico, con un total de 33 611 asistencias ambulatorias realizadas y una media de 40,2 asistencias por año.

La Orden 731/2013 de precios públicos en el Servicio Madrileño de Salud establece que las primeras consultas médicas tienen un coste de 130 euros, las consultas médicas sucesivas de 78 euros, y las consultas de enfermería 18 euros (aplicando el dato definido para las consultas de enfermería de Atención Primaria, dado que no está especificada la consulta en el ámbito hospitalario). De esta forma el coste total de las consultas para los 516 pacientes asciende a 1.396.524€.

Utilizando el periodo de seguimiento de los pacientes, obtenemos que el coste medio anual por paciente, para los 516 pacientes que presentan al menos una consulta ambulatoria, es de 1.714 euros (IC95% 1.591€ - 1.838€). En la Tabla 2 se compara el número de asistencias ambulatorias anual y el coste anual de las consultas según diferentes características al diagnóstico. En lo relativo al coste medio anual, no observamos

**TABLA 2**

NÚMERO Y COSTE ANUAL DE LAS ASISTENCIAS AMBULATORIAS, DE URGENCIAS Y DE HOSPITALIZACIÓN SEGÚN SEXO, CHARLSON, ECOG-PS Y ESTADIO

	Nº de pacientes	Asistencias ambulatorias				Asistencias en urgencias				Asistencias en hospitalización			
		Número de asistencias por año		Coste estimado por año (sólo consultas)		Número de asistencias por año		Coste estimado por año		Número de ingresos por año		Coste estimado por año	
		Media (DT)	p-valor	Media (DT)	p-valor	Media (DT)	p-valor	Media (DT)	p-valor	Media (DT)	p-valor	Media (DT)	p-valor
<b>Total</b>	<b>516</b>	<b>40,2 (30,4)</b>		<b>1714 (1428)</b>		<b>2,4 (3,3)</b>		<b>441 (596)</b>		<b>2,0 (6,5)</b>		<b>14251 (54440)</b>	
<b>Sexo del paciente</b>			<b>0,482</b>		<b>0,297</b>		<b>0,246</b>		<b>0,246</b>		<b>&lt;0,001</b>		<b>&lt;0,001</b>
Mujer	126	36,7 (24,5)		1.508 (970)		2,0 (2,7)		356 (493)		1,1 (3,2)		5.981 (16.144)	
Hombre	390	41,3 (32,1)		1.781 (1543)		2,6 (3,5)		468 (623)		2,2 (7,2)		16.923 (61.731)	
<b>Índice de Charlson</b>			<b>&lt;0,001</b>		<b>&lt;0,001</b>		<b>&lt;0,001</b>		<b>&lt;0,001</b>		<b>&lt;0,001</b>		<b>&lt;0,001</b>
1-2	37	25,4 (28,6)		1.078 (1037)		0,7 (1,2)		130 (212)		0,7 (0,7)		5.078 (6.377)	
3-4	58	29,1 (26,7)		1.287 (1044)		2,0 (2,7)		356 (490)		1,8 (3,6)		26.391 (78.409)	
5-8	83	48,6 (26,1)		2.094 (1.820)		2,9 (3,2)		520 (578)		5,0 (13,5)		32.015 (103.920)	
>8	107	50,0 (30,0)		2.221 (1501)		3,9 (4,2)		700 (754)		4,4 (5,8)		27.829 (41.707)	
<b>ECOG-PS</b>			<b>0,060</b>		<b>0,019</b>		<b>&lt;0,001</b>		<b>&lt;0,001</b>		<b>0,021</b>		<b>&lt;0,001</b>
0	286	38,0 (30,4)		1.588 (1.209)		1,9 (2,9)		351 (531)		1,2 (2,8)		9.695 (32.751)	
1	178	41,9 (27,5)		1.719 (1182)		2,9 (3,5)		531 (626)		2,0 (4,2)		11.220 (23.958)	
>=2	43	49,3 (38,4)		2.648 (2.788)		3,6 (3,7)		651 (665)		7,1 (19,0)		59.465 (155.997)	
<b>Estadio</b>			<b>&lt;0,001</b>		<b>&lt;0,001</b>		<b>&lt;0,001</b>		<b>&lt;0,001</b>		<b>0,009</b>		<b>0,045</b>
I	107	18,9 (19,1)		885 (831)		1,1 (2,5)		206 (443)		0,5 (1,4)		6.625 (31.807)	
II	60	30,0 (22,0)		1.333 (910)		0,9 (1,1)		162 (208)		0,9 (1,6)		5.690 (9.109)	
III	196	45,6 (30,2)		1.865 (1.195)		2,3 (2,6)		416 (471)		1,7 (3,2)		12.665 (40.307)	
IV	153	52,0 (31,5)		2.252 (1.848)		4,1 (4,3)		747 (774)		3,7 (11,1)		24.973 (83.850)	

Fuente: Elaboración propia.



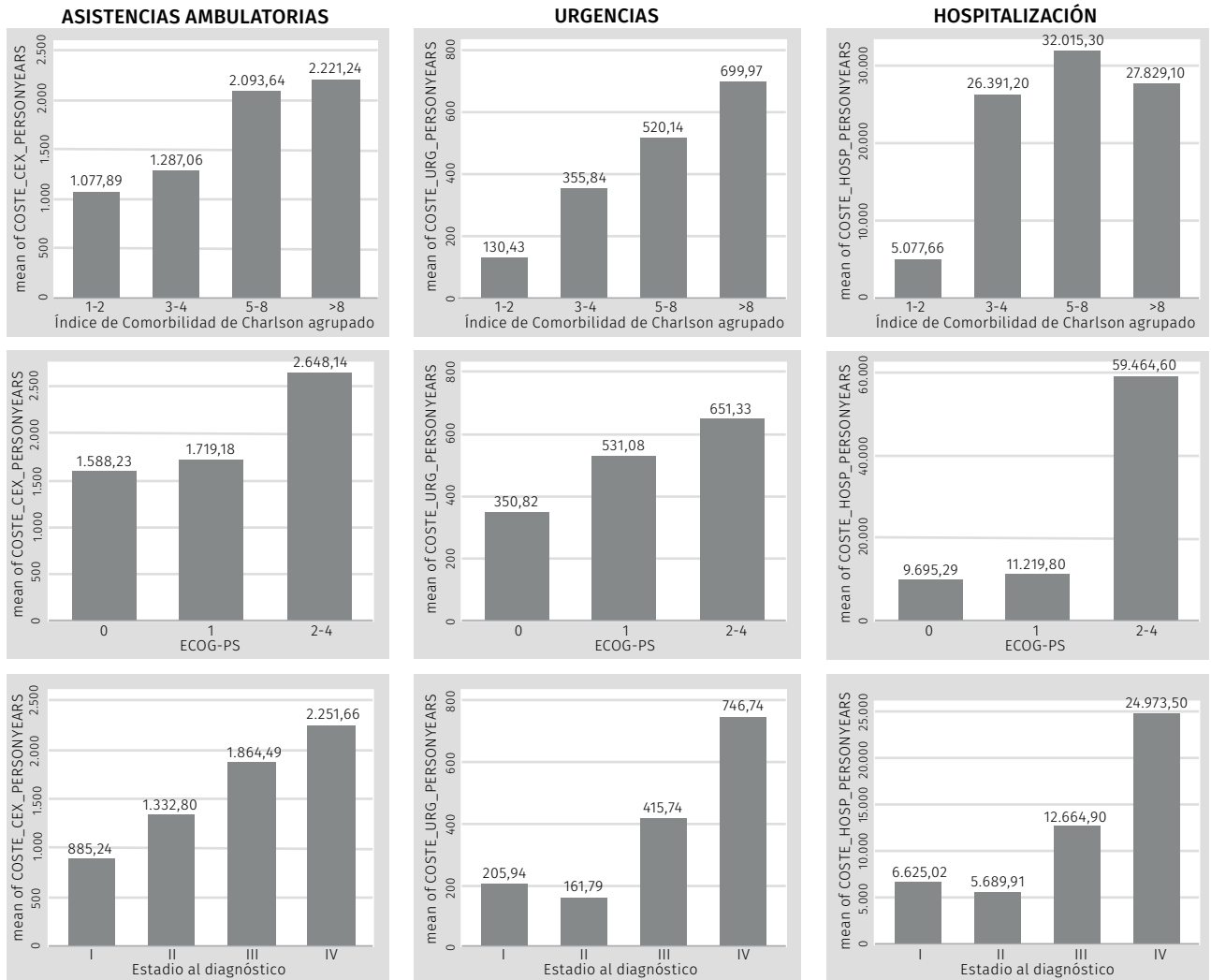
diferencias significativas por sexo ( $p = 0,297$ ), mientras que apreciamos diferencias significativas según el índice Charlson de comorbilidad ( $p < 0,001$ ), el ECOG ( $p = 0,019$ ) y el estadio al diagnóstico ( $p < 0,001$ ). En la Figura 1 se muestran los costes medios anuales en función de estos tres factores al diagnóstico.

Los pacientes que presentan al menos una consulta ambulatoria con índice Charlson mayor o igual a 5 (coste medio 2.165€, N=190) presentan un 100%

más coste medio anual que los pacientes con Charlson 1-2 (coste medio 1.078, N=37) y un 68% más coste que los pacientes con Charlson 3-4 (coste medio 1.287€, N=58). Los pacientes con ECOG-PS 2-4 (coste medio 2.648€, N=43) presentan un coste medio anual un 54% mayor que los pacientes con ECOG-PS 1 (coste medio 1.719€, N=178) y un 67% mayor que los pacientes con ECOG-PS 0 (coste medio 1.588€, N=286). Los pacientes con estadio IV (coste medio 2.252€, N=153) presentan un coste medio anual un

**FIGURA 1**

DISTRIBUCIONES UNIVARIANTES DEL COSTE DE LAS ASISTENCIAS AMBULATORIAS, DE URGENCIAS E INGRESOS EN HOSPITALIZACIÓN SEGÚN EL ÍNDICE DE COMORBILIDAD DE CHARLSON, ECOG-PS Y ESTADIO AL DIAGNÓSTICO.



Fuente: Elaboración propia.

69% mayor que los pacientes con estadio II (coste medio 1.333€, N=60) y un 154% mayor (2,5 veces más) que los pacientes con estadio I (coste medio 885€, N=107).

### **Análisis del coste de las asistencias en urgencias**

Tomando los 516 pacientes para los que hubo un seguimiento ambulatorio, 385 pacientes (74,6%) presentaron actividad de urgencias posterior al diagnóstico, con un total de 1.536 registros de asistencias en urgencias realizadas.

De acuerdo a la Orden 731/2013, obtenemos como precios para las urgencias generales 180 euros. De esta forma el coste total de las urgencias para los 385 pacientes asciende a 276.480 euros. La media para los 516 pacientes fue de 441€, mientras que fue de 536€ euros para los 385 que tuvieron al menos una asistencia en urgencias.

En la Tabla 2 se compara el número de asistencias en urgencias anuales y el coste estimado anual de los 516 pacientes para los que hubo un seguimiento ambulatorio según diferentes factores al diagnóstico. Podemos ver que existen diferencias significativas en el coste medio anual de los pacientes que requieren asistencia de urgencias según índice de Charlson ( $p < 0,001$ ), ECOG-PS ( $p < 0,001$ ) o estadio ( $p < 0,001$ ). En la Figura 1 se muestran los costes medios de las asistencias en urgencias según estos factores.

Los pacientes con Charlson  $> 8$  (coste medio 700€, N=107) presentan un coste un 97% superior que los pacientes con Charlson 3-4 (coste medio 356€, N=58) y un 438% superior (5,38 veces más) que los pacientes con Charlson 1-2 (coste medio 130€, N=37). Frente a los pacientes con Charlson 1-2, los pacientes con Charlson 5-8 (coste medio 520€, N=83) presentan un coste 4 veces mayor, y los pacientes con Charlson 3-4 un coste 2,7 veces mayor. Los pacientes con ECOG-PS 2-4 (coste medio 651€, N=43) presentan un coste un 85% superior a los ECOG-PS 0 (coste medio 351€, N=286), y los pacientes con ECOG-PS 1 (coste medio 531€, N=178) un coste un 51% superior a los ECOG-PS 0. Los pacientes con estadio IV (coste medio 747€, N=153) presentan un coste un 80% mayor que los pacientes con estadio III (coste medio 416€, N=196), 4,6 veces más coste que los

pacientes estadio II (162€, N=60) y 3,6 veces más que los pacientes estadio I (206€, N=107).

### **Análisis de la actividad de hospitalización**

Considerando los 516 pacientes que tuvieron seguimiento ambulatorio, 285 pacientes (55,2%) generaron 749 episodios de hospitalización posteriores al diagnóstico de cáncer de pulmón (o en los 30 días inmediatamente anteriores a la fecha de diagnóstico) entre los años 2009 y 2014. Los pacientes que ingresan lo hacen una media de 3,57 veces al año con una estancia media anual de 40,1 días, 1,97 veces en media para el conjunto de los 516 pacientes, con una estancia media anual de 22,1 días.

El coste medio del episodio de hospitalización, según la Orden 731/2013, fue 6.693€. Se muestran en la Figura 2 los GRDs correspondientes a los pacientes de la cohorte completa de 522 pacientes (753 ingresos) según su estadio al diagnóstico. El coste medio anual de los ingresos en hospitalización es de 14.251€, que asciende a 25.802€ considerando sólo los 285 pacientes con algún ingreso registrado.

En la Tabla 2 observamos diferencias significativas en el coste medio anual de hospitalización de los pacientes según su índice de Charlson ( $p < 0,001$ ), ECOG ( $p < 0,001$ ) o estadio ( $p = 0,045$ ). En la Figura 1 se muestran los costes medios anuales de los ingresos en hospitalización según estos tres factores.

Los pacientes con Charlson  $> 8$  (coste medio 27.829€, N=107) presentan un coste medio anual un 5% mayor que los pacientes con Charlson 3-4 (coste medio 26.391€, N=58) y 5,5 veces más coste que los pacientes con Charlson 1-2 (coste medio 5.078€, N=37). Los pacientes con ECOG-PS 2-4 (coste medio 59.465€, N=43) presentan un coste 5,3 veces superior a los pacientes con ECOG-PS 1 (coste medio 11.220€, N=178) y 6,1 veces superior a los pacientes con ECOG-PS 0 (coste medio 9.695€, N=286). Los pacientes con estadio III (coste medio 12.665€, N=196) presentan un coste un 91% mayor que los pacientes estadio I (6.625€, N=107).

Observamos también, a diferencia de los ámbitos ambulatorios y de urgencias, una diferencia estadísticamente significativa según el sexo del paciente ( $p < 0,001$ ). Así, los pacientes hombres tienen un coste 2,8 veces superior a las mujeres.





**FIGURA 2**

GRDS DE LOS EPISODIOS DE HOSPITALIZACIÓN PARA CADA ESTADIO

GRD Estadio I

GRD	Tipo Grd	Descripción del GRD	% del Número de episodios	Número de episodios	Estancia promedio
75	Q	PROCEDIMIENTOS TORÁCICOS MAYORES	32,77%	39,0	7,5
541	M	NEUMONÍA SIMPLE Y OTROS TRAST. RESPIRATORIOS EXC. BRONQUITIS & ASMA CON CC MAYOR	11,76%	14,0	7,6
538	Q	PROCEDIMIENTOS TORÁCICOS MAYORES CON CC MAYOR	9,24%	11,0	11,8
410	M	QUIMIOTERAPIA	3,36%	4,0	12,5
10	M	NEOPLASIAS DE SISTEMA NERVIOSO CON CC	2,52%	3,0	6,3
77	Q	OTROS PROCEDIMIENTOS QUIRÚRGICOS DE APARATO RESPIRATORIO SIN CC	2,52%	3,0	1,7
89	M	NEUMONÍA SIMPLE & PLEURITIS EDAD>17 CON CC	2,52%	3,0	10,7
127	M	INSUFICIENCIA CARDIACA & SHOCK	1,68%	2,0	12,3
148	Q	PROCEDIMIENTOS MAYORES DE INTESTINO DELGADO & GRUESO CON CC	1,68%	2,0	20,0
203	M	NEOPLASIA MALIGNA DE SISTEMA HEPATOBILIAR O DE PÁNCREAS	1,68%	2,0	26,4
404	M	LINFOMA & LEUCEMIA NO AGUDA SIN CC	1,68%	2,0	8,5
463	M	SIGNOS & SÍNTOMAS CON CC	1,68%	2,0	10,1
82	M	NEOPLASIAS RESPIRATORIAS	1,68%	2,0	5,5
100	M	SIGNOS & SÍNTOMAS RESPIRATORIOS SIN CC	0,84%	1,0	6,0
101	M	OTROS DIAGNÓSTICOS DE APARATO RESPIRATORIO CON CC	0,84%	1,0	6,0
116	Q	OTRAS IMPLANTACIONES DE MARCAPASOS CARDIACO PERMANENTE	0,84%	1,0	2,7
145	M	OTROS DIAGNÓSTICOS DE APARATO CIRCULATORIO SIN CC	0,84%	1,0	105,0
188	M	OTROS DIAGNÓSTICOS DE APARATO DIGESTIVO EDAD>17 CON CC	0,84%	1,0	2,0
272	M	TRASTORNOS MAYORES DE PIEL CON CC	0,84%	1,0	9,0
303	Q	PROCEDIMIENTOS S. RIÑÓN Y URÉTER POR NEOPLASIA	0,84%	1,0	6,0
315	Q	OTROS PROCEDIMIENTOS SOBRE RIÑÓN & TRACTO URINARIO	0,84%	1,0	2,0
36	Q	PROCEDIMIENTOS SOBRE RETINA	0,84%	1,0	1,3
394	Q	OTROS PROC. QUIRÚRGICOS HEMATOLÓGICOS Y DE ÓRGANOS HEMOPOYÉTICOS	0,84%	1,0	7,0
406	Q	TRAST. MIELOPROLIFERATIVO O NEO. MAL DIFER. CON PROC. QUIRÚRGICO MAYOR CON CC	0,84%	1,0	6,8
407	Q	TRAST. MIELOPROLIFERATIVO O NEO. MAL DIFER. CON PROC. QUIRÚRGICO MAYOR SIN CC	0,84%	1,0	4,3
409	M	RADIOTERAPIA	0,84%	1,0	47,0
533	M	OTROS TRAST. SISTEMA NERVIOSO EXC. AIT, CONVULSIONES & CEFALEA CON CC MAYOR	0,84%	1,0	4,0
539	Q	PROCEDIMIENTOS RESPIRATORIOS EXCEPTO PROC. TORÁCICOS MAYORES CON CC MAYOR	0,84%	1,0	2,2
540	M	INFECCIONES & INFLAMACIONES RESPIRATORIAS EXCEPTO NEUMONÍA SIMPLE CON CC MAYOR	0,84%	1,0	15,9
543	M	TRAST. CIRCULATORIOS EXCEPTO IAM, ENDOCARDITIS, ICC & ARRITMIA CON CC MAYOR	0,84%	1,0	14,0
553	Q	PROC. AP. DIGEST. EXCEPTO HERNIA & PROC. MAYOR ESTÓMAGO O INTEST. CON CC MAYOR	0,84%	1,0	8,7
557	M	TRASTORNOS HEPATOBILIARES Y DE PÁNCREAS CON CC MAYOR	0,84%	1,0	10,0
561	M	OSTEOMIELITIS, ARTRITIS SÉPTICA & TRAST. T. CONECT. CON CC MAYOR	0,84%	1,0	8,0
580	M	INFECCIONES & PARASITOSIS SISTÉMICAS EXCEPTO SEPTICEMIA CON CC MAYOR	0,84%	1,0	20,8
64	M	NEOPLASIA MALIGNA DE OIDO, NARIZ, BOCA & GARGANTA	0,84%	1,0	2,0
76	Q	OTROS PROCEDIMIENTOS QUIRÚRGICOS DE APARATO RESPIRATORIO CON CC	0,84%	1,0	1,0
867	Q	ESCISIÓN LOCAL Y EXTRACCIÓN DE DISPOSITIVO FIJACIÓN INTERNA EXCEPTO CADERA Y FÉMUR, SIN CC	0,84%	1,0	0,7
87	M	EDEMA PULMONAR & INSUFICIENCIA RESPIRATORIA	0,84%	1,0	2,0
874	Q	LINFOMA Y LEUCEMIA CON PROC. QUIRÚRGICO MAYOR, CON CC	0,84%	1,0	56,0
877	Q	OXIG. MEMB. EXTRAC. O TRAQUEOSTOMIA CON VENT. MEC. +96 hrs O SIN DIAG. PRINC. TRAST. ORL CON PROC.	0,84%	1,0	118,0
88	M	ENFERMEDAD PULMONAR OBSTRUCTIVA CRÓNICA	0,84%	1,0	1,9
901	M	SEPTICEMIA SIN VENT. MEC. +96 HORAS EDAD >17	0,84%	1,0	5,0
94	M	NEUMOTORAX CON CC	0,84%	1,0	2,8
Total general			100,00%	119,0	10,7

**FIGURA 2**

## GRDS DE LOS EPISODIOS DE HOSPITALIZACIÓN PARA CADA ESTADIO

## GRD Estadio II

GRD	Tipo Grd	Descripción del GRD	% del Número de episodios	Número de episodios	Estancia promedio
75	Q	PROCEDIMIENTOS TORÁCICOS MAYORES	26,26%	26,00	8,95
410	M	QUIMIOTERAPIA	11,11%	11,00	1,64
82	M	NEOPLASIAS RESPIRATORIAS	10,10%	10,00	10,64
541	M	NEUMONÍA SIMPLE Y OTROS TRAST. RESPIRATORIOS EXC. BRONQUITIS & ASMA CON CC MAYOR	9,09%	9,00	11,44
77	Q	OTROS PROCEDIMIENTOS QUIRÚRGICOS DE APARATO RESPIRATORIO SIN CC	8,08%	8,00	2,03
538	Q	PROCEDIMIENTOS TORÁCICOS MAYORES CON CC MAYOR	5,05%	5,00	14,20
89	M	NEUMONÍA SIMPLE & PLEURITIS EDAD>17 CON CC	3,03%	3,00	2,67
102	M	OTROS DIAGNÓSTICOS DE APARATO RESPIRATORIO SIN CC	2,02%	2,00	5,00
144	M	OTROS DIAGNÓSTICOS DE APARATO CIRCULATORIO CON CC	2,02%	2,00	9,00
394	Q	OTROS PROC. QUIRÚRGICOS HEMATOLÓGICOS Y DE ÓRGANOS HEMOPOYÉTICOS	2,02%	2,00	3,00
76	Q	OTROS PROCEDIMIENTOS QUIRÚRGICOS DE APARATO RESPIRATORIO CON CC	2,02%	2,00	0,99
101	M	OTROS DIAGNÓSTICOS DE APARATO RESPIRATORIO CON CC	1,01%	1,00	5,00
15	M	ACCIDENTE CEREBROVASCULAR NO ESPECÍFICO & OCLUSIÓN PRECEREBRAL SIN INFARTO	1,01%	1,00	7,60
239	M	FRACTURAS PATOLÓGICAS & NEOPLASIA MALIGNA MUSCULOESQUELÉTICA & T. CONECTIVO	1,01%	1,00	27,00
240	M	TRASTORNOS DE T. CONECTIVO CON CC	1,01%	1,00	8,00
277	M	CELULITIS EDAD>17 CON CC	1,01%	1,00	2,00
320	M	INFECCIONES DE RIÑÓN & TRACTO URINARIO EDAD>17 CON CC	1,01%	1,00	6,00
398	M	TRASTORNO DE S. RETICULOENDOTELIAL & INMUNITARIOS CON CC	1,01%	1,00	14,00
430	M	PSICOSIS	1,01%	1,00	6,87
463	M	SIGNOS & SÍNTOMAS CON CC	1,01%	1,00	20,00
468	Q	Procedimiento quirúrgico extensivo sin relación con diagnóstico principal	1,01%	1,00	9,00
544	M	ICC & ARRITMIA CARDIACA CON CC MAYOR	1,01%	1,00	12,05
557	M	TRASTORNOS HEPATOBILIARES Y DE PÁNCREAS CON CC MAYOR	1,01%	1,00	9,00
561	M	OSTEOMIELITIS, ARTRITIS SÉPTICA & TRAST. T. CONECT. CON CC MAYOR	1,01%	1,00	2,00
574	M	TRASTORNOS DE SANGRE, ÓRGANOS HEMOPOYÉTICOS & INMUNOLÓGICOS CON CC MAYOR	1,01%	1,00	4,00
588	M	BRONQUITIS & ASMA EDAD >17 CON CC MAYOR	1,01%	1,00	5,00
78	M	EMBOLISMO PULMONAR	1,01%	1,00	12,00
87	M	EDEMA PULMONAR & INSUFICIENCIA RESPIRATORIA	1,01%	1,00	5,00
88	M	ENFERMEDAD PULMONAR OBSTRUCTIVA CRÓNICA	1,01%	1,00	0,51
95	M	NEUMOTÓRAX SIN CC	1,01%	1,00	2,00
Total general			100,00%	99,00	7,56



## FIGURA 2

### GRDS DE LOS EPISODIOS DE HOSPITALIZACIÓN PARA CADA ESTADIO

#### GRD Estadio III

GRD	Tipo Grd	Descripción del GRD	% del Número de episodios	Número de episodios	Estancia promedio
541	M	NEUMONÍA SIMPLE Y OTROS TRAST. RESPIRATORIOS EXC. BRONQUITIS & ASMA CON CC MAYOR	13,01%	48,0	9,3
410	M	QUIMIOTERAPIA	11,38%	42,0	1,7
75	Q	PROCEDIMIENTOS TORÁCICOS MAYORES	11,38%	42,0	7,5
82	M	NEOPLASIAS RESPIRATORIAS	7,86%	29,0	8,3
101	M	OTROS DIAGNÓSTICOS DE APARATO RESPIRATORIO CON CC	6,23%	23,0	7,0
76	Q	OTROS PROCEDIMIENTOS QUIRÚRGICOS DE APARATO RESPIRATORIO CON CC	4,88%	18,0	3,3
89	M	NEUMONÍA SIMPLE & PLEURITIS EDAD>17 CON CC	3,25%	12,0	9,4
538	Q	PROCEDIMIENTOS TORÁCICOS MAYORES CON CC MAYOR	2,98%	11,0	14,9
87	M	EDEMA PULMONAR & INSUFICIENCIA RESPIRATORIA	2,98%	11,0	4,9
77	Q	OTROS PROCEDIMIENTOS QUIRÚRGICOS DE APARATO RESPIRATORIO SIN CC	2,71%	10,0	1,2
577	M	TRAST. MIELOPROLIFERATIVO & NEO. MAL DIFERENCIADA CON CC MAYOR	2,44%	9,0	2,0
10	M	NEOPLASIAS DE SISTEMA NERVIOSO CON CC	2,17%	8,0	9,8
394	Q	OTROS PROC. QUIRÚRGICOS HEMATOLÓGICOS Y DE ÓRGANOS HEMOPOYÉTICOS	1,63%	6,0	1,7
874	Q	LINFOMA Y LEUCEMIA CON PROC. QUIRÚRGICO MAYOR, CON CC	1,08%	4,0	1,5
127	M	INSUFICIENCIA CARDIACA & SHOCK	0,81%	3,0	5,7
14	M	ICTUS CON INFARTO	0,81%	3,0	21,9
182	M	ESOFAGITIS, GASTROENTERITIS & TRAST. DIGEST. MISCELANEOS EDAD>17 CON CC	0,81%	3,0	4,3
398	M	TRASTORNO DE S. RETICULOENDOTELIAL & INMUNITARIOS CON CC	0,81%	3,0	6,7
533	M	OTROS TRAST. SISTEMA NERVIOSO EXC. AIT, CONVULSIONES & CEFALEA CON CC MAYOR	0,81%	3,0	8,0
539	Q	PROCEDIMIENTOS RESPIRATORIOS EXCEPTO PROC. TORÁCICOS MAYORES CON CC MAYOR	0,81%	3,0	7,7
574	M	TRASTORNOS DE SANGRE, ÓRGANOS HEMOPOYÉTICOS & INMUNOLÓGICOS CON CC MAYOR	0,81%	3,0	14,0
96	M	BRONQUITIS & ASMA EDAD>17 CON CC	0,81%	3,0	3,1
99	M	SIGNOS & SÍNTOMAS RESPIRATORIOS CON CC	0,81%	3,0	7,3
11	M	NEOPLASIAS DE SISTEMA NERVIOSO SIN CC	0,54%	2,0	8,5
138	M	ARRITMIAS CARDIACAS & TRASTORNOS DE CONDUCCIÓN CON CC	0,54%	2,0	7,0
239	M	FRACTURAS PATOLÓGICAS & NEOPLASIA MALIGNA MUSCULOESQUELÉTICA & T. CONECTIVO	0,54%	2,0	19,5
463	M	SIGNOS & SÍNTOMAS CON CC	0,54%	2,0	4,5
78	M	EMBOLISMO PULMONAR	0,54%	2,0	6,0
813	M	GASTROENTERITIS NO BACTERIANA Y DOLOR ABDOMINAL EDAD > 17 CON CC	0,54%	2,0	18,5
85	M	DERRAME PLEURAL CON CC	0,54%	2,0	15,0
88	M	ENFERMEDAD PULMONAR OBSTRUCTIVA CRÓNICA	0,54%	2,0	6,0
882	M	DIAGNÓSTICO DE SISTEMA RESPIRATORIO CON VENTILACIÓN MECÁNICA <96 hrs	0,54%	2,0	9,5
90	M	NEUMONÍA SIMPLE & PLEURITIS EDAD>17 SIN CC	0,54%	2,0	9,5
1	Q	CRANEOTOMÍA EDAD>17 CON CC	0,27%	1,0	16,1
100	M	SIGNOS & SÍNTOMAS RESPIRATORIOS SIN CC	0,27%	1,0	6,7
102	M	OTROS DIAGNÓSTICOS DE APARATO RESPIRATORIO SIN CC	0,27%	1,0	1,0
110	Q	PROCEDIMIENTOS CARDIOVASCULARES MAYORES CON CC	0,27%	1,0	12,0
121	M	TRAST. CIRCULATORIOS CON IAM & COMPL. MAYORES, ALTA CON VIDA	0,27%	1,0	3,9
130	M	TRASTORNOS VASCULARES PERIFÉRICOS CON CC	0,27%	1,0	23,0
172	M	NEOPLASIA MALIGNA DIGESTIVA CON CC	0,27%	1,0	7,0
174	M	HEMORRAGIA GASTROINTESTINAL CON CC	0,27%	1,0	2,0
18	M	TRASTORNOS DE NERVIOS CRANEALES & PERIFÉRICOS CON CC	0,27%	1,0	5,0
180	M	OBSTRUCCIÓN GASTROINTESTINAL CON CC	0,27%	1,0	13,0
2	Q	CRANEOTOMÍA EDAD>17 SIN CC	0,27%	1,0	14,0
202	M	CIRROSIS & HEPATITIS ALCOHÓLICA	0,27%	1,0	11,0
207	M	TRASTORNOS DEL TRACTO BILIAR CON CC	0,27%	1,0	4,0
211	Q	PROC. DE CADERA & FÉMUR EXCEPTO ARTICULACIÓN MAYOR EDAD>17 SIN CC	0,27%	1,0	6,0

GRD	Tipo Grd	Descripción del GRD	% del Número de episodios	Número de episodios	Estancia promedio
216	Q	BIOPSIAS DE SISTEMA MUSCULOESQUELETICO & TEJIDO CONECTIVO	0,27%	1,0	19,0
233	Q	OTROS PROC. QUIRURGICOS DE S.MUSCULOESQUELETICO & T.CONECTIVO CON CC	0,27%	1,0	3,0
240	M	TRASTORNOS DE T.CONECTIVO CON CC	0,27%	1,0	7,9
303	Q	PROCEDIMIENTOS S.RIÑON Y URETER POR NEOPLASIA	0,27%	1,0	7,0
308	Q	PROCEDIMIENTOS MENORES SOBRE VEJIGA CON CC	0,27%	1,0	1,2
309	Q	PROCEDIMIENTOS MENORES SOBRE VEJIGA SIN CC	0,27%	1,0	1,3
311	Q	PROCEDIMIENTOS TRANSURETRALES SIN CC	0,27%	1,0	3,0
320	M	INFECCIONES DE RIÑON & TRACTO URINARIO EDAD>17 CON CC	0,27%	1,0	6,0
406	Q	TRAST.MIELOPROLIFERATIVO O NEO. MAL DIFER. CON PROC. QUIRURGICO MAYOR CON CC	0,27%	1,0	0,4
423	M	OTROS DIAGNOSTICOS DE ENFERMEDAD INFECCIOSA & PARASITARIA	0,27%	1,0	26,0
425	M	REACCION DE ADAPTACION AGUDA & DISFUNCION PSICOSOCIAL	0,27%	1,0	15,0
429	M	ALTERACIONES ORGANICAS & RETRASO MENTAL	0,27%	1,0	7,0
449	M	ENVENENAMIENTO & EFECTO TOXICO DE DROGAS EDAD>17 CON CC	0,27%	1,0	4,0
452	M	COMPLICACIONES DE TRATAMIENTO CON CC	0,27%	1,0	10,0
468	Q	Procedimiento quirúrgico extensivo sin relación con diagnóstico principal	0,27%	1,0	3,8
543	M	TRAST.CIRCULATORIOS EXCEPTO IAM, ENDOCARDITIS, ICC & ARRITMIA CON CC MAYOR	0,27%	1,0	37,0
544	M	ICC & ARRITMIA CARDIACA CON CC MAYOR	0,27%	1,0	6,0
55	Q	PROCEDIMIENTOS MISCELANEOS SOBRE OIDO, NARIZ, BOCA & GARGANTA	0,27%	1,0	2,0
550	Q	OTROS PROCEDIMIENTOS VASCULARES CON CC MAYOR	0,27%	1,0	8,0
552	M	TRAST. AP. DIGESTIVO EXCEPTO ESOF., GASTROENT. & ULC.NO COMPL. CON CC MAYOR	0,27%	1,0	5,0
554	Q	PROCEDIMIENTOS SOBRE HERNIA CON CC MAYOR	0,27%	1,0	7,0
575	Q	PROCEDIMIENTOS S. SANGRE, ORGANOS HEMOPOYETICOS & INMUNOLOGICOS CON CC MAYOR	0,27%	1,0	4,0
582	M	LESIONES, ENVENENAMIENTOS O EFECTO TÓXICO DROGAS EXC. TRAUMA MULTIPLE CON CC MAYOR	0,27%	1,0	1,0
584	M	SEPTICEMIA CON CC MAYOR	0,27%	1,0	40,2
585	Q	PROCEDIMIENTO MAYOR ESTOMAGO,ESOFAGO,DUODENO,I.DELGADO & GRUESO CON CC MAYOR	0,27%	1,0	41,0
66	M	EPISTAXIS	0,27%	1,0	3,0
69	M	OTITIS MEDIA & ITRS EDAD>17 SIN CC	0,27%	1,0	16,0
710	M	HIV CON DIAG MAYORES RELACIONADOS CON DIAG MAYORES MULT. O SIGNIF. SIN TBC	0,27%	1,0	8,0
820	M	MALFUNCION, REACCION O COMPL. DE DISPOSITIVO, INJERTO O TRANSPLANTE GENITOURINARIO	0,27%	1,0	8,8
853	Q	PROCEDIMIENTO CARDIOVASCULAR PERCUTANEO CON STENT LIBERADOR DE FARMACO, CON IAM	0,27%	1,0	11,0
875	Q	LINFOMA Y LEUCEMIA CON PROC. QUIRURGICO MAYOR, SIN CC	0,27%	1,0	2,0
877	Q	OXIG. MEMB. EXTRAC. O TRAQUEOSTOMIA CON VENT. MEC. + 96 hrs O SIN DIAG. PRINC. TRAST. ORL CON PROC.	0,27%	1,0	21,0
878	Q	TRAQUEOSTOMIA CON VENT. MEC. + 96 hrs O SIN DIAG. PRINC. TRASTORNOS ORL SIN PROC. QUIR. MAYOR	0,27%	1,0	16,6
889	M	CONVULSIONES EDAD > 17 CON CC	0,27%	1,0	9,0
897	M	DIAGNOSTICO HEMATOLOGICO O INMUNOLOGICO MAYOR EXC CRISIS ENF.CEL.FALCIFORMES Y COAGUL	0,27%	1,0	11,0
Total general			100,00%	369,0	7,3



**FIGURA 2**

GRDS DE LOS EPISODIOS DE HOSPITALIZACIÓN PARA CADA ESTADIO

GRD Estadio IV

GRD	Tipo Grd	Descripción del GRD	% del Número de episodios	Número de episodios	Estancia promedio
541	M	NEUMONÍA SIMPLE Y OTROS TRAST. RESPIRATORIOS EXC. BRONQUITIS & ASMA CON CC MAYOR	16,27%	27,0	13,9
82	M	NEOPLASIAS RESPIRATORIAS	15,66%	26,0	11,3
410	M	QUIMIOTERAPIA	8,43%	14,0	2,0
76	Q	OTROS PROCEDIMIENTOS QUIRÚRGICOS DE APARATO RESPIRATORIO CON CC	6,63%	11,0	15,4
75	Q	PROCEDIMIENTOS TORÁCICOS MAYORES	6,02%	10,0	5,7
10	M	NEOPLASIAS DE SISTEMA NERVIOSO CON CC	5,42%	9,0	13,3
87	M	EDEMA PULMONAR & INSUFICIENCIA RESPIRATORIA	4,82%	8,0	6,3
101	M	OTROS DIAGNÓSTICOS DE APARATO RESPIRATORIO CON CC	2,41%	4,0	11,8
239	M	FRACTURAS PATOLÓGICAS & NEOPLASIA MALIGNA MUSCULOESQUELÉTICA & T. CONECTIVO	2,41%	4,0	9,7
552	M	TRAST. AP. DIGESTIVO EXCEPTO ESOF., GASTROENT. & ULC. NO COMPL. CON CC MAYOR	2,41%	4,0	8,0
539	Q	PROCEDIMIENTOS RESPIRATORIOS EXCEPTO PROC. TORÁCICOS MAYORES CON CC MAYOR	1,81%	3,0	26,7
540	M	INFECCIONES & INFLAMACIONES RESPIRATORIAS EXCEPTO NEUMONÍA SIMPLE CON CC MAYOR	1,81%	3,0	17,0
89	M	NEUMONÍA SIMPLE & PLEURITIS EDAD>17 CON CC	1,81%	3,0	10,7
130	M	TRASTORNOS VASCULARES PERIFÉRICOS CON CC	1,20%	2,0	6,0
413	M	OTROS TRASTORNOS MIELOPROLIFERATIVOS O NEOPLASIAS MAL DIFER. CON CC	1,20%	2,0	15,0
414	M	OTROS TRASTORNOS MIELOPROLIFERATIVOS O NEOPLASIAS MAL DIFER. SIN CC	1,20%	2,0	22,0
468	Q	Procedimiento quirúrgico extensivo sin relación con diagnóstico principal	1,20%	2,0	20,5
543	M	TRAST. CIRCULATORIOS EXCEPTO IAM, ENDOCARDITIS, ICC & ARRITMIA CON CC MAYOR	1,20%	2,0	10,0
561	M	OSTEOMIELITIS, ARTRITIS SÉPTICA & TRAST. T. CONECT. CON CC MAYOR	1,20%	2,0	15,0
77	Q	OTROS PROCEDIMIENTOS QUIRÚRGICOS DE APARATO RESPIRATORIO SIN CC	1,20%	2,0	2,2
100	M	SIGNOS & SÍNTOMAS RESPIRATORIOS SIN CC	0,60%	1,0	6,0
110	Q	PROCEDIMIENTOS CARDIOVASCULARES MAYORES CON CC	0,60%	1,0	9,0
148	Q	PROCEDIMIENTOS MAYORES DE INTESTINO DELGADO & GRUESO CON CC	0,60%	1,0	41,0
161	Q	PROCEDIMIENTOS SOBRE HERNIA INGUINAL & FEMORAL EDAD>17 CON CC	0,60%	1,0	4,0
172	M	NEOPLASIA MALIGNA DIGESTIVA CON CC	0,60%	1,0	20,0
2	Q	CRANEOTOMÍA EDAD>17 SIN CC	0,60%	1,0	8,0
210	Q	PROC. DE CADERA & FÉMUR EXCEPTO ARTICULACIÓN MAYOR EDAD>17 CON CC	0,60%	1,0	28,0
216	Q	BIOPSIAS DE SISTEMA MUSCULOESQUELÉTICO & TEJIDO CONECTIVO	0,60%	1,0	30,0
23	M	ESTUPOR & COMA NO TRAUMÁTICOS	0,60%	1,0	17,0
395	M	TRASTORNOS DE LOS HEMATÍES EDAD>17	0,60%	1,0	10,0
398	M	TRASTORNO DE S. RETICULOENDOTELIAL & INMUNITARIOS CON CC	0,60%	1,0	4,0
419	M	IEBRE DE ORIGEN DESCONOCIDO EDAD>17 CON CC	0,60%	1,0	11
452	M	COMPLICACIONES DE TRATAMIENTO CON CC	0,60%	1,0	5,0
463	M	SIGNOS & SÍNTOMAS CON CC	0,60%	1,0	13,0
482	Q	TRAQUEOSTOMIA CON TRASTORNOS DE BOCA, LARINGE O FARINGE	0,60%	1,0	3,0
544	M	ICC & ARRITMIA CARDIACA CON CC MAYOR	0,60%	1,0	13,0
551	M	ESOFAGITIS, GASTROENTERITIS & ULCERA NO COMPLICADA CON CC MAYOR	0,60%	1,0	4,0
559	Q	PROCEDIMIENTOS MUSCULOESQUELÉTICOS NO MAYORES CON CC MAYOR	0,60%	1,0	24,0
569	M	TRAST. DE RIÑÓN & TRACTO URINARIO EXCEPTO INSUFICIENCIA RENAL CON CC MAYOR	0,60%	1,0	21,0
574	M	TRASTORNOS DE SANGRE, ÓRGANOS HEMOPOYÉTICOS & INMUNOLÓGICOS CON CC MAYOR	0,60%	1,0	3,2
582	M	LESIONES, ENVENENAMIENTOS O EFECTO TÓXICO DROGAS EXC. TRAUMA MÚLTIPLE CON CC MAYOR	0,60%	1,0	2,0
584	M	SEPTICEMIA CON CC MAYOR	0,60%	1,0	7,0
818	Q	SUSTITUCIÓN DE CADERA EXCEPTO POR COMPLICACIONES	0,60%	1,0	9,0
832	M	ISQUEMIA TRANSITORIA	0,60%	1,0	7,0
878	Q	TRAQUEOSTOMIA CON VENT. MEC. + 96 hrs O SIN DIAG. PRINC. TRASTORNOS ORL SIN PROC. QUIR. MAYOR	0,60%	1,0	14,0
898	Q	INFECCIONES Y PARASITOSIS CON PROC. QUIRÚRGICO	0,60%	1,0	9,9
Total general			100,00%	166,0	11,3

## DISCUSIÓN

Nuestro estudio demuestra que la integración de datos clínicos con datos de actividad asistencial, y su análisis mediante metodologías como CRISP-DM, permite obtener nuevo conocimiento de interés para la gestión. Hasta dónde conocemos, es el primer estudio acerca del coste de la asistencia sanitaria en el cáncer de pulmón, en todos sus estadios, en España.

En lo relativo al método aplicado, la aplicación de CRISP-DM ha permitido llegar a resultados de interés sin necesidad de establecer una hipótesis conocida previamente y favorece la participación de perfiles técnicos, sin un conocimiento profundo del dominio de negocio, en la realización de estudios en el ámbito médico. Consideramos que las herramientas y métodos empleados podrían ser aplicados a otras patologías para obtener un conocimiento de interés para la mejora de la gestión asistencial de los pacientes. Así mismo, consideramos que sería interesante en futuros trabajos aplicar métodos multivariantes para establecer la influencia real de cada variable clínica sobre el consumo de asistencias sanitarias y por tanto sobre el coste, teniendo en cuenta la posible interrelación entre variables (factores de confusión), así como aplicar otras técnicas en el ámbito del *Machine Learning* para la implementación de algoritmos predictivos y de soporte a la decisión que permitieran, por ejemplo, identificar pacientes con un perfil de alto consumo de asistencias sanitarias (y por tanto de alto coste) para ofrecerles proactivamente otras alternativas asistenciales.

Para la realización del estudio ha sido clave la integración de distintos orígenes de datos, fundamentalmente procedentes de la HCE, de la codificación de episodios hospitalaria (CMBD) y de los registros de casos cumplimentados manualmente por el Servicio de Oncología. Consideramos fundamental avanzar en la integración del resto de fuentes disponibles, en la estructuración y codificación de la información en los distintos ámbitos asistenciales y en la automatización de la captura de datos en el ámbito hospitalario. En este sentido, el proceso manual de captura de datos en registros

asistenciales paralelos a la HCE provoca errores que inciden en la calidad de los mismos para la realización de posteriores investigaciones. La integración de fuentes adicionales (como en nuestro caso los datos de prescripción y administración de medicación ambulatoria) y la mayor disponibilidad de datos estructurados nos permitiría disponer de más pacientes y más variables para su estudio mediante las distintas técnicas estadísticas y de *Data Mining* y *Machine Learning* existentes. Dado que un alto porcentaje del contenido clínico en la HCE se captura en texto libre por los profesionales, la aplicación de herramientas de procesamiento de lenguaje natural podría ser clave para avanzar en este campo<sup>20</sup>.

En lo relativo a los resultados obtenidos, el coste medio total por paciente al año ascendió a 16.406 euros, siendo un 10,4% (1.714€) atribuible a las asistencias ambulatorias, un 2,7% (441€) a las asistencias en urgencias y un 86,9% (14.251€) a los ingresos en hospitalización. No se consideran, por falta de datos durante la realización del estudio, los costes de la medicación administrada ambulatoriamente y de los tratamientos radioterápicos. En un estudio realizado en Holanda sobre 907 pacientes, seguidos una mediana de 20,7 meses, Van Der Linden et al.<sup>21</sup> obtuvo un coste medio total de 28.468€ incluyendo entre otros conceptos la quimioterapia ambulatoria, lo que supondría un coste anual medio de 16.503€. Langton et al.<sup>22</sup> indicó en un estudio realizado con 4.271 pacientes con cáncer en Australia que más del 90% de los pacientes fallecidos a causa del cáncer ingresaron al menos una vez en el hospital, y que dichos ingresos en hospitalización supusieron más del 80% del coste total en los últimos 6 meses de vida. En otro estudio de pacientes con cáncer de pulmón en estadio IIIB y IV, Lang et al.<sup>23</sup> indicó que la hospitalización y las consultas médicas supusieron un 85% de los costes durante todo el seguimiento del paciente, excluyendo los costes del tratamiento ambulatorio. De acuerdo a un amplio estudio realizado por Yabroff et al.<sup>24</sup> los costes de hospitalización en el último año de vida en cáncer de pulmón son cercanos al 70% del conjunto de costes, sin contemplar los costes de la



medicación administrada ambulatoriamente. Las diferencias en la composición de las cohortes de pacientes, en la configuración de los sistemas sanitarios y en la metodología de los estudios impiden una comparación directa de estos resultados.

De acuerdo a nuestro análisis, se observan diferencias en el coste de la asistencia ambulatoria y de las asistencias en urgencias según el índice de comorbilidad de Charlson, el ECOG-PS y el estadio al diagnóstico (mayor cuanto mayor comorbilidad, mayor ECOG-PS y más avanzado el estadio). En lo relativo a la hospitalización, además de confirmar las diferencias observadas en el resto de ámbitos asistenciales, apreciamos diferencias según el sexo del paciente (un coste 2,8 veces mayor en los hombres que en las mujeres). Esto puede ser debido a las distintas características clínicas de ambos grupos: un 28% de las mujeres de la cohorte tienen ECOG-PS > 0 frente a un 43,5% en el caso de los hombres y un 24,6% de las mujeres tienen un índice Charlson >8 frente a un 40,5% en el caso de los hombres.

Observamos también en hospitalización cómo la distribución de GRDs para los 753 ingresos analizados presenta diferencias según el estadio al diagnóstico del paciente, predominando el procedimiento quirúrgico (GRD 75) en los pacientes estadio I y la neumonía (GRD 541) en los estadios más avanzados.

A tenor de estos resultados, podemos decir que todas estas relaciones son coherentes con el sentido común y el conocimiento clínico actual, si bien hemos podido cuantificarlas en nuestro estudio desde un punto de vista univariante, y pueden ser de utilidad para planificar recursos y mejorar la gestión de estos pacientes. ■

# REFERENCIAS

- 1 in 4 deaths caused by cancer in the EU28. Eurostat Press Releases. 25 Nov 2014. <http://ec.europa.eu/eurostat/en/web/products-press-releases/-/3-25112014-BP>
- 2 Luengo-Fernandez, R., Leal, J., Gray, A., & Sullivan, R. (2013). Economic burden of cancer across the European Union: a population-based cost analysis. *The Lancet Oncology*, 14(12), 1165–1174.
- 3 GLOBOCAN 2012: Estimated Cancer Incidence, Mortality and Prevalence Worldwide in 2012. ARCI: OMS; [último acceso el 13/01/18]. Accesible en: <http://globocan.iarc.fr/Default.aspx>
- 4 Las Cifras del Cáncer en España 2018. Sociedad Española de Oncología Médica (SEOM). [último acceso el 22/03/18]. Accesible en: [https://seom.org/seomcms/images/stories/recursos/Las\\_Cifras\\_del\\_cancer\\_en\\_Espana2018.pdf](https://seom.org/seomcms/images/stories/recursos/Las_Cifras_del_cancer_en_Espana2018.pdf)
- 5 García Cornejo, B. (2004). La experiencia en contabilidad de costes y en benchmarking del Sistema Nacional de Salud español. Balance de una década. *Rev. Adm. Sanit. Siglo XXI* 2, 103–123.
- 6 Instituto Nacional de la Salud. GE- CLIF. Gestión Clínico-Financiera y Coste por Proceso. Madrid: Ministerio de Sanidad y Consumo, Subdirección General de Coordinación Administrativa, 2001.
- 7 Shearer, C. (2000) The CRISP-DM Model: The New Blueprint for Data Mining. *Journal of Data Warehousing*, 5, 13–22.
- 8 Óscar Marbán, Gonzalo Mariscal and Javier Segovia (2009). A Data Mining & Knowledge Discovery Process Model, *Data Mining and Knowledge Discovery in Real Life Applications*, Julio Ponce and Adem Karahoca (Ed.), InTech, DOI: 10.5772/6438. [Último acceso el 09/04/18] Available from: [https://mts.intechopen.com/books/data\\_mining\\_and\\_knowledge\\_discovery\\_in\\_real\\_life\\_applications/a\\_data\\_mining\\_amp\\_knowledge\\_discovery\\_process\\_model](https://mts.intechopen.com/books/data_mining_and_knowledge_discovery_in_real_life_applications/a_data_mining_amp_knowledge_discovery_process_model)
- 9 Rivo, E., de la Fuente, J., Rivo, Á., García-Fontán, E., Cañizares, M.-Á., & Gil, P. (2012). Cross-industry standard process for data mining is applicable to the lung cancer surgery domain, improving decision making as well as knowledge and quality management. *Clinical & Translational Oncology* : Official Publication of the Federation of Spanish

- Oncology Societies and of the National Cancer Institute of Mexico, 14(1), 73–9. <http://doi.org/10.1007/s12094-012-0764-8>
- 10 Pérez, J., Iturbide, E., Olivares, V., Hidalgo, M., Martínez, A., & Almanza, N. (2015). A Data Preparation Methodology in Data Mining Applied to Mortality Population Databases. *Journal of Medical Systems*, 39(11), 152. <http://doi.org/10.1007/s10916-015-0312-5>
  - 11 Oken, M. M., Creech, R. H., Tormey, D. C., Horton, J., Davis, T. E., McFadden, E. T., & Carbone, P. P. (1982). Toxicity and response criteria of the Eastern Cooperative Oncology Group. *American Journal of Clinical Oncology*, 5(6), 649–55. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/7165009>
  - 12 Índice Nacional de Defunciones. [Último acceso el 09/04/18]. Accesible en: [https://www.mssi.gob.es/estadEstudios/estadisticas/estadisticas/estMinisterio/IND\\_TipoDifusion.htm](https://www.mssi.gob.es/estadEstudios/estadisticas/estadisticas/estMinisterio/IND_TipoDifusion.htm)
  - 13 Charlson ME, Pompei P, Ales KL, McKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chron Dis* 1987 May; 40(5): 373-383.
  - 14 Stagg, Vicki, (2017), CHARLSON: Stata module to calculate Charlson index of comorbidity, <https://EconPapers.repec.org/RePEc:boc:bocode:s456719>. [Accessed: 27/02/2018]
  - 15 Deyo, R. A., Cherkin, D. C., & Ciol, M. A. (1992). Adapting a clinical comorbidity index for use with ICD-9 administrative databases. *J Clin Epidemiol*, 45(10), 613–619. [http://doi.org/10.1016/0895-4356\(93\)90104-9](http://doi.org/10.1016/0895-4356(93)90104-9)
  - 16 Quan, H., Sundararajan, V., Halfon, P., & Fong, A. (2005). Coding algorithms for defining comorbidities in. *Medical Care*, 43(11), 1130–1139.
  - 17 Knust, R. E., Portela, M. C., Pereira, C. C. de A., & Fortes, G. B. (2017). Estimated costs of advanced lung cancer care in a public reference hospital. *Revista de Saude Publica*, 51, 53. <https://doi.org/10.1590/S1518-8787.2017051006665>
  - 18 ORDEN 731/2013, de 6 de septiembre, del Consejero de Sanidad, por la que se fijan los precios públicos por la prestación de los servicios y actividades de naturaleza sanitaria de la Red de Centros de la Comunidad de Madrid. [último acceso el 22/03/18]. Accesible en: <http://www.madrid.org/cs/Satellite?blobcol=urldata&blobheader=application%2Fpdf&blobheadername1=Content-Disposition&blobheadervalue1=filename%3DBOCCM-20130910-1.pdf&blobkey=id&blobtable=MungoBlobs&blobwhere=1352916926357&ssbinary=true>
  - 19 Epidat: programa para análisis epidemiológico de datos. Versión 4.2, julio 2016. Consellería de Sanidade, Xunta de Galicia, España; Organización Panamericana de la salud (OPS-OMS); Universidad CES, Colombia. <https://www.sergas.es/Saude-publica/EPIDAT-4-2?idioma=es>
  - 20 Ruiz, E.M., Tuñas, J.M., Bermejo, G., Martín, C.G., Rodríguez-González, A., Zanin, M., Pedro, C.G. De, Méndez, M., Zaretskaia, O., Rey, J., et al. (2018). Profiling Lung Cancer Patients Using Electronic Health Records. *J. Med. Syst.* 42, 1–10.
  - 21 van der Linden, N., Bongers, M.L., Coupé, V.M.H., Smit, E.F., Groen, H.J.M., Welling, A., Schramel, F.M.N.H., and Uyl-de Groot, C.A. (2016). Costs of non-small cell lung cancer in the Netherlands. *Lung Cancer* 91, 79–88.
  - 22 Langton, J. M., Reeve, R., Srasuebkul, P., Haas, M., Viney, R., Currow, D., & Pearson, S. A. (2016, May 24). Health service use and costs in the last 6 months of life in elderly decedents with a history of cancer: A comprehensive analysis from a health payer perspective. *British Journal of Cancer*. Nature Publishing Group. <http://doi.org/10.1038/bjc.2016.75>
  - 23 Lang, K., Marciniak, M. D., Faries, D., Stokes, M., Buesching, D., Earle, C., ... Thompson, D. (2009). Costs of first-line doublet chemotherapy and lifetime medical care in advanced non-small-cell lung cancer in the United States. *Value in Health*, 12(4), 481–488. <https://doi.org/10.1111/j.1524-4733.2008.00472.x>
  - 24 Yabroff, K. R., Lamont, E. B., Mariotto, A., Warren, J. L., Topor, M., Meekins, A., & Brown, M. L. (2008). Cost of Care for Elderly Cancer Patients in the United States. *JNCI: Journal of the National Cancer Institute*, 100(9), 630–641. <http://doi.org/10.1093/jnci/djn103>