

## ARTÍCULO ORIGINAL

### **Análisis estadístico implicativo para la identificación de factores pronósticos de la mortalidad por cáncer de pulmón**

### **Statistical analysis involving the identification of mortality prediction factors due to lung cancer**

**Dra. Alina Moraga Rodríguez,<sup>I</sup> Dra. C. Larisa Zamora Matamoras,<sup>II</sup> MsC. Nelsa María Sagaró del Campo,<sup>III</sup> Dra. Annia Moraga Rodríguez<sup>IV</sup> y Dra. Alina Rodríguez Griñán<sup>V</sup>**

<sup>I</sup> Facultad No. 2 de Ciencias Médicas, Universidad de Ciencias Médicas, Santiago de Cuba, Cuba.

<sup>II</sup> Universidad de Oriente, Santiago de Cuba, Cuba.

<sup>III</sup> Universidad de Ciencias Médicas, Santiago de Cuba, Cuba.

<sup>IV</sup> Facultad No. 1 de Ciencias Médicas, Universidad de Ciencias Médicas, Santiago de Cuba, Cuba.

<sup>V</sup> Hospital Provincial Docente Clínicoquirúrgico "Saturnino Lora Torres", Universidad de Ciencias Médicas, Santiago de Cuba, Cuba.

### **RESUMEN**

Se realizó un estudio observacional y analítico para evaluar la utilidad de una nueva técnica de análisis estadístico implicativo en la identificación de los factores pronósticos de una entidad clínica en la provincia de Santiago de Cuba, de septiembre de 2013 a igual mes de 2014, en cuyo caso fue seleccionado el cáncer de pulmón por ser la neoplasia maligna más frecuente en el territorio. Como factores pronósticos de tumores pulmonares, según la regresión logística fueron identificados la afectación ganglionar, la presencia de metástasis y la localización central del tumor, y según el análisis estadístico implicativo se detectaron la afectación ganglionar, la presencia de metástasis, el grado III de diferenciación histológica y el estadio avanzado. El análisis estadístico implicativo resultó ser una técnica apropiada que complementa a la regresión logística en la identificación de factores pronósticos, pues se logra una mejor comprensión de la causalidad y eleva la calidad de este tipo de investigación.

**Palabras clave:** análisis estadístico implicativo, regresión logística, cáncer del pulmón, factores pronósticos.

### **ABSTRACT**

An observational and analytic study to evaluate the usefulness of a new technique of statistical analysis involving the prediction factors identification of a clinical entity in Santiago de Cuba, was carried out from September, 2013 to the same month in 2014, in this case lung cancer was selected for being the most frequent malignancy in the territory. As prediction factors of lung tumors, according to the logistical regression the nodular involvement, the metastasis and the central localization of the tumor were identified, and according to the statistical involving analysis, the nodular involvement, and metastasis, histologic differentiation stage III and advanced stage were detected. The statistical involving analysis turned out to be an appropriate technique that

supplements logistical regression in the prediction factors identification, because a better understanding of the causation is achieved and the quality of this investigation increases.

**Key words:** statistical involving analysis, logistical regression, lung cancer, prediction factors.

## INTRODUCCIÓN

Los factores pronósticos son aquellos que señalan la evolución clínica de un padecimiento una vez que está presente; por lo que la identificación de estos es de gran interés para la prevención secundaria y terciaria. Determinar los factores pronósticos en las neoplasias resulta muy necesario para definir el tratamiento y posible evolución, pues el cáncer es uno de los procesos morbosos más complejos y resulta el más complicado en la esfera de la medicina por la amplia multicausalidad que incide en su origen, la complejidad de sus mecanismos patogénicos y biomoleculares, y la variedad de formas y tipos que pueden producirse en los seres vivos. En la población de Santiago de Cuba existe una alta incidencia de cáncer como causa de muerte y los 3 tipos más frecuentes son el de pulmón, el de próstata y el de mama, en ese orden.<sup>1,2</sup>

Se plantea que el pronóstico en estos casos está determinado por numerosos factores clínicos, por ejemplo: la edad. En relación con la neoplasia es importante el tipo histológico, el grado de diferenciación histológica, el estadio (clasificación tumor-nódulo-metástasis), el tumor primario, la localización, la presencia de metástasis, la afectación ganglionar o el tipo de tratamiento, ya sea quirúrgico o con agentes químicos o radiaciones.<sup>3,4</sup>

Como resultado del crecimiento anormal de células cancerosas en el tejido pulmonar, se produce el cáncer del pulmón, el cual puede derivar en metástasis e infiltración a otros tejidos del cuerpo; este constituye el más mortífero tanto para hombres como para mujeres. Cuanto más temprano se detecte, mayores serán las probabilidades de supervivencia.<sup>5</sup>

El factor pronóstico más importante en el grupo de los carcinomas de células no pequeñas, es la extensión de la enfermedad. Actualmente se define siguiendo la clasificación tumor-nódulo-metástasis (TNM). Solo los tumores pequeños (menor o igual a 3 cm), con o sin afectación pleural, bien delimitados y a 2 cm o más de la carina, presentan posibilidad de supervivencia.<sup>6,7</sup>

Para el estudio de los factores pronósticos tradicionalmente se utilizan las técnicas estadísticas multivariadas y de estas, la regresión logística o regresión de Cox.<sup>8</sup>

Ahora bien, el análisis estadístico implicativo es un método de estudio no simétrico de datos que permite, a partir de un conjunto de datos que interrelaciona una población de sujetos u objetos con un grupo de variables, la extracción y estructuración del conocimiento en forma de normas y reglas generalizadas y, a partir de la contingencia de estas reglas, la explicación y, en consecuencia, una determinada previsión en distintas ramas del saber.<sup>9,10</sup>

Así, con el objetivo de identificar estos factores en el cáncer más frecuente en la provincia de Santiago de Cuba, y de mostrar además la utilidad del análisis estadístico implicativo, se desarrolló un estudio de casos y controles en los 3 hospitales donde se tratan los pacientes con dicha afección.

## MÉTODOS

Se realizó un estudio observacional y analítico para evaluar la utilidad de una nueva técnica de análisis estadístico implicativo (ASI), en la identificación de los factores pronósticos de una entidad en un territorio y momento determinados.

Al respecto, la entidad seleccionada fue el cáncer de pulmón, debido a su elevada incidencia, prevalencia y letalidad en el mundo, el país y específicamente en el territorio escogido para efectuar el estudio: la provincia de Santiago de Cuba, donde constituye un problema de salud prioritario, por ser la primera causa de muerte de manera sostenida y con tendencia ascendente desde el 2008. El período de estudio fue de septiembre de 2013 a igual mes de 2014.

Tomando como variable dependiente el estado vivo o fallecido y como explicativas los factores pronósticos indicados en la bibliografía y por expertos consultados, se llevó a cabo el análisis mediante la regresión logística y la nueva técnica propuesta, y se compararon ambas en cuanto a su utilidad en la identificación de factores pronósticos.

El universo de los casos estuvo constituido por todas las personas fallecidas con el diagnóstico clínico e histológico de cáncer de pulmón, procedentes de dicha provincia y mayores de 18 años, que asistían a consulta especializada a tal efecto en los hospitales Oncológico Docente Provincial "Conrado Benítez García", Provincial Docente Clínicoquirúrgico "Saturnino Lora Torres" y General Docente "Dr. Juan Bruno Zayas Alfonso", donde se encontraba su historia clínica en el período de estudio; en tanto, el universo de los controles estuvo conformado por las personas vivas con los mismos criterios anteriores.

Para la selección de los casos y los controles se empleó un muestreo estratificado no proporcional, donde los estratos fueron los 3 hospitales mencionados en la definición del universo, y en cada estrato se realizó un muestreo simple aleatorio, con el programa MS Excel para los números aleatorios -- dada la existencia de un listado de las historias clínicas en los departamentos de Registros Médicos de cada institución hospitalaria --, lo que facilitó la localización de las historias tanto de pacientes vivos (controles) como fallecidos (casos).

En el procesamiento y la presentación de los resultados se emplearon los sistemas informáticos SPSS (*Statistical Product and Service Solutions*) versión 11.5; CHIC (*Classification Hiérarchique Implicative et Cohésitive*) versión 4.1 y EPIDAT versión 3.0. Se confeccionó una base de datos en MS Excel, en formato .csv como condición necesaria para la aplicación del CHIC, la cual se exportó al SPSS.

Con vistas a describir los factores pronósticos presentes en la población de estudio se utilizaron como medidas de resumen el número y el porcentaje. Para determinar en qué medida estos factores actúan como causantes y agravantes del cáncer, se emplearon 2 metodologías de análisis estadístico.

En primer lugar se aplicó la regresión logística binaria luego de verificar el cumplimiento de los supuestos de la prueba, con la identificación de los factores que tuvieron significación estadística, según el intervalo de confianza de la exponencial de beta. Se empleó el método por pasos hacia adelante (razón de verosimilitud), puesto que fue el que demostró un mejor ajuste al modelo, según la prueba de Hosmer y Lemeshow.

Posteriormente, se aplicó el ASI para identificar las variables que poseían relación causal significativa y, por tanto, constituían factores pronósticos del empeoramiento de la enfermedad.

Por último, se compararon ambas técnicas tomándolas como “medios diagnósticos” para la identificación de factores pronósticos, a través del empleo de la regresión logística como estándar de oro, y se consideró cada variable como factor pronóstico o no, según los resultados de ambas técnicas. Se calcularon la sensibilidad, la especificidad, los valores predictivos y las razones de verosimilitud del ASI, así como los índices de Youden, de validez y de Kappa para medir la concordancia. También se identificaron las similitudes y diferencias entre estas.

## RESULTADOS

- Con la regresión logística

Resultaron ser factores de mal pronóstico: la afectación ganglionar, la presencia de metástasis y la localización central. Los que presentaban afectación ganglionar poseían entre 1,135 y 14,243 veces más probabilidades de fallecer que los que no la mostraban, con un nivel de confiabilidad del 95 % (tabla 1). Asimismo, con un nivel de confiabilidad de 95 %, los que presentaban metástasis o el tumor localizado centralmente, poseían respectivamente, 2,954-51,488 y 1,135-13,035 más probabilidades de fallecer.

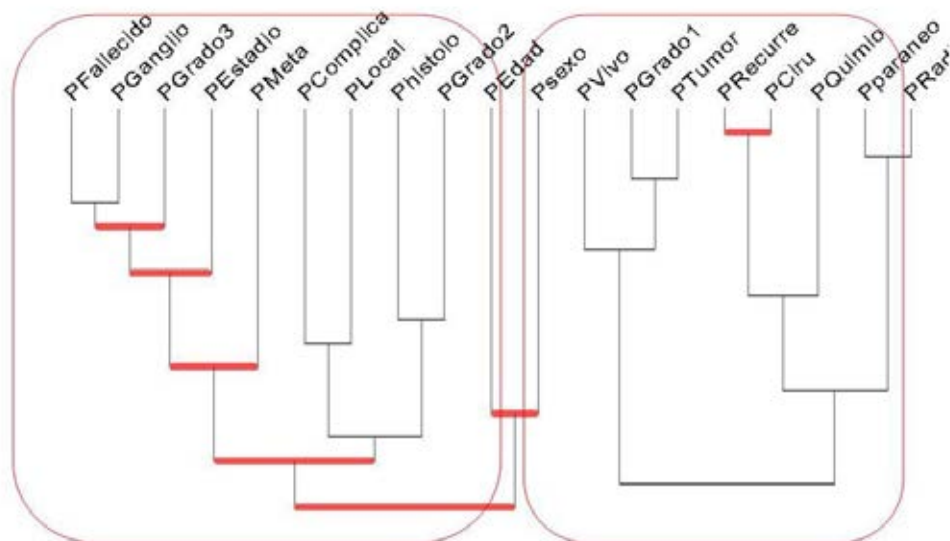
El valor de la constante en el modelo de regresión logística indica que, de no estar presente ninguno de los factores analizados, la probabilidad de fallecer es 0,131 veces menor que la de no fallecer.

**Tabla 1.** Variables incluidas en el 10mo paso de la regresión logística (modelo atrás condicional) y sus indicadores

Variables	B	DE	Wald	Gl.	p	Exp. ( $\beta$ )	IC: 95 % para Exp. ( $\beta$ )	
							Inferior	Superior
Afectación ganglionar	1,391	645	4,650	1	031	4,021	1,135	14,243
Presencia de metástasis	2,512	729	11,872	1	001	12,333	2,954	51,488
Localización central	1,347	623	4,683	1	030	3,847	1,135	13,035
Quimioterapia	1,197	622	3,699	1	054	302	089	1,023
Constante	2,036	630	10,431	1	001	131		

- Con el análisis estadístico implicative

En el árbol de similaridad se aprecia la formación de 2 grandes grupos o clases que engloban a todos los factores (figura 1). En una de las clases se encuentra la categoría “fallecido” y en la otra, la categoría “vivo”, lo cual significa que ambas tienen comportamientos diferentes y cada una se asemeja a un grupo de factores distintos. En la clase a la que pertenecen los vivos están las variables que se asocian con un mejor pronóstico, supuestamente como los tratamientos quirúrgicos y radiaciones, el grado I de diferenciación histológica, la recurrencia que evidentemente se presentará cuando el paciente alarga su vida y la presencia del síndrome paraneoplásico.

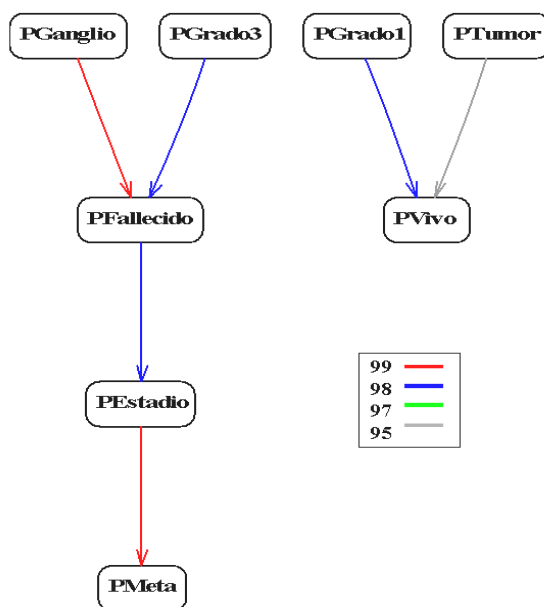


**Fig. 1.** Agrupación de factores pronósticos según índice de similaridad (árbol de similaridad)

Respecto al grafo implicativo se aprecia la formación de 2 grafos (figura 2); el primero se generó a partir de un conjunto de factores que incluye al fallecido, y está compuesto por varios caminos: poseer afectación ganglionar o un grado III de diferenciación histológica implica que el paciente va a fallecer (C1); con 98 % de intensidad se pudo plantear que el número de pacientes que falleció y que no estaba en un estadio III era ínfimo (C2), de igual forma el número de pacientes que falleció y no presentaba metástasis fue muy pequeño (C3), lo que se aseguró con 99 % de intensidad de la implicación.

El segundo grafo está compuesto por 2 caminos que implican estar vivo al tener un grado I de diferenciación histológica (98 %) y un tumor de tamaño mayor de 5 cm (95 %), con implicación directa, lo cual coincidió con el grafo implicativo en modo de cono para las relaciones causales de los factores pronósticos que influyeron en una mejor evolución del afectado. Los grafos en modo de cono para las relaciones causales de los factores pronósticos que influyeron en el desenlace mejor o fatal, reafirmaron lo que se aprecia en el grafo implicativo principal.

Del análisis implicativo se obtuvo que constituyeron factores de mal pronóstico en el cáncer de pulmón: la afectación ganglionar, con una intensidad igual o superior a 99 %, el estadio y el grado III de diferenciación histológica avanzado, con una confianza entre 98 y 99 %, y la presencia de metástasis, con una confianza entre 94 y 96 %.



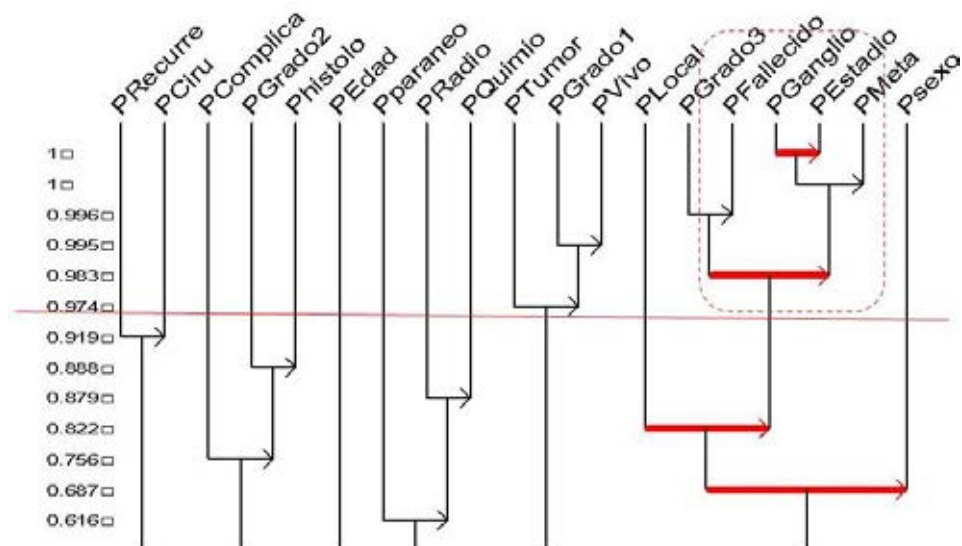
**Fig. 2.** Grafo implicativo de las relaciones causales entre los factores pronósticos con 95 % y más de intensidad implicativa

El árbol cohesitivo en la figura 3 muestra nodos significativos en los niveles 1, 5, 10 y 12, además de una jerarquía orientada, la cual estructura los factores pronósticos en varias clases que definen R-reglas:

- $R_1$  ( $PGanglio \Rightarrow PEstadio$ ), lo cual significa que el hecho de tener afectación ganglionar implica estar en un estadio avanzado de la enfermedad.
- $R_2$  ( $PGanglio \Rightarrow PEstadio \Rightarrow PMeta$ ), lo que significa que la afectación ganglionar implicó siempre estar en el estadio más avanzado, y denota siempre el hecho de presentar metástasis.

Las 2 reglas anteriores se forman con el mayor índice de cohesión 1. Igualmente, entre las reglas de mayor grado de intensidad se encuentran:

- $R_3$  ( $PGrado3 \Rightarrow PFallecido$ ), la cual plantea que tener un grado III de diferenciación histológica implica que se va a fallecer, con una cohesión de 0,996.
- $R_4$  ( $PGrado1 \Rightarrow PVivo$ ); esta regla plantea que presentar un grado I de diferenciación histológica implica que se va a permanecer vivo, con una cohesión de 99,5 %.
- $R_5$ : ( $R_3 \Rightarrow R_2$ ), la cual significa que la regla  $R_3$  es una condición necesaria para que se cumpla la regla  $R_2$ , con una cohesión de 98,3 %.
- $R_7$ : ( $PTumor \Rightarrow (PGrado1 \Rightarrow PVivo)$ ), con 97,4 % de riesgo, poseer un grado I de diferenciación histológica implica que el paciente se mantenga vivo y esto está implicado por el tamaño tumoral.
- $R_8$ : ( $PRecurre \Rightarrow PCiru$ ); una recurrencia implica someterse a una operación con 91,9 % de riesgo.



**Fig. 3.** Relaciones causales entre factores pronósticos (árbol cohesitivo)

- Análisis comparativo entre ambas técnicas

La regresión logística (RL) identificó como factores de mal pronóstico la afectación ganglionar, la presencia de metástasis y la localización central del tumor; mientras que los factores de mal pronóstico identificados por el ASI, fueron la afectación ganglionar, el grado III de diferenciación histológica, la presencia de metástasis y el estadio avanzado.

Ambas técnicas identificaron la afectación ganglionar, con 99 % de confiabilidad en el caso del ASI y 95 % en la regresión logística; también ambas identificaron la presencia de metástasis como factor de mal pronóstico, con una confianza de 94 y 99 % en el ASI y la regresión logística, respectivamente. Con la regresión se reconoció un factor no determinado por el ASI y este, a su vez, identificó 2 factores no determinados por la regresión logística (tabla 2).

**Tabla 2.** Factores pronósticos de cáncer de pulmón identificados con ambas técnicas

ASI	RL		Total
	Sí	No	
Sí	2	2	4
No	1	10	11
Total	3	12	15

## DISCUSIÓN

El tipo histológico no resultó un factor pronóstico, según lo obtenido con las 2 técnicas aplicadas, a pesar de lo que se registra en la bibliografía sobre el carcinoma microcítico o indiferenciado de célula pequeña, de que presenta manifestación clínica y respuesta al tratamiento distintas del resto de los carcinomas. Este tumor se caracteriza por su rápido crecimiento y diseminación precoz, que hacen que más 60 % de los pacientes

posea metástasis en el momento del diagnóstico, y aunque es muy sensible a la quimioterapia y radioterapia, la supervivencia media no supera los 15 meses.<sup>6</sup> No obstante, se pudo observar un predominio de este tipo celular en los fallecidos.

Sánchez Hernández *et al*<sup>11</sup> aseveran en su investigación, que uno de los factores que mejora el pronóstico es la intervención quirúrgica; sin embargo, como el diagnóstico se suele realizar en fases avanzadas de la enfermedad, cuando solo 20 % de los casos puede ser intervenido, este no resulta un factor que mejore el pronóstico, y con los tratamientos no quirúrgicos, basados en quimioterapia y radioterapia, tampoco se han logrado desenlaces más favorables.

En un estudio efectuado en China,<sup>12</sup> con el procedimiento quirúrgico se halló un buen pronóstico en pacientes que presentaban la neoplasia en los estadios I y II, siempre que se empleó la técnica de recepción y reconstrucción de la arteria pulmonar. El análisis multivariado desarrollado en otra investigación china,<sup>13</sup> donde se siguieron 600 pacientes por 5 años, indicó que aquellos con metástasis ósea, sin otro órgano afectado, tuvieron buen pronóstico.

Al evaluar la técnica del ASI como identificador de factores pronósticos relacionados con el cáncer de pulmón, de acuerdo a los indicadores estimados con 95 % de confianza, se encontró una capacidad para detectar las variables significativamente asociadas a dicho cáncer (sensibilidad) entre 49,68 y 83,66 %, ya sea como factor de buen o de mal pronóstico. También demostró entre 79,04 y 87,63 % de capacidad para identificar las variables que no constituían factores pronósticos (especificidad).

Un factor pronóstico identificado por el ASI tiene una probabilidad de 50 % de serlo realmente (valor pronóstico positivo); así mismo, una variable que no constituye un factor pronóstico según esta técnica, tiene una probabilidad de 90,91 % de no serlo en realidad (valor pronóstico negativo).

Con esta técnica es 4 veces más probable detectar un factor pronóstico entre los verdaderos factores pronósticos que entre las variables que no lo son (razón de verosimilitud positiva). De igual manera, es poco probable (0,4 %) que una variable que no sea factor pronóstico salga clasificada como tal (razón de verosimilitud negativa) y la concordancia entre ambas técnicas fue de 0,44 (índice de Kappa).<sup>14</sup>

La incertidumbre sobre el futuro de una enfermedad puede estimarse sobre la experiencia clínica y la información epidemiológica. En la práctica clínica diaria es preciso tener un conocimiento del pronóstico para tomar decisiones en cuanto a diagnóstico y tratamiento, pero además para informar al paciente de su futuro.

Cabe señalar que la diferencia entre la tasa de positivos verdaderos y la de positivos falsos fue de 50 % (índice de Youden). El ASI clasificó correctamente 80 % de las variables (índice de validez).

Los factores asociados al cáncer de pulmón, identificados en este estudio fueron:

- Según la regresión logística, factores pronósticos de empeoramiento: afectación ganglionar, presencia de metástasis y localización central del tumor. No se determinó ningún factor de mejoramiento.
- Según el análisis estadístico implicativo, factores pronósticos de empeoramiento: afectación ganglionar, presencia de metástasis, grado III de diferenciación histológica y estadio avanzado. No se determinó ningún factor de mejoramiento.



## REFERENCIAS BIBLIOGRÁFICAS

1. Pita Fernández S, Valdés Cañedo FA. Determinación de factores pronósticos. Cad Aten Primaria. 1997; 4: 26-29 [citado 23 Ene 2014].
2. Cuba. Ministerio de Salud Pública. Dirección Nacional de Registros Médicos y Estadísticas de Salud. Anuario Estadístico de Salud 2013. La Habana: MINSAP; 2014. p. 65 [citado 25 Ago 2014].
3. Álvarez Sintés R. Afecciones respiratorias. Neoplasia de pulmón. En: Álvarez Sintés R, Hernández Cabrera G, Báster Moro JC, García Núñez RD, Louro Bernal I, Céspedes Lantigua LA, et al. Medicina general integral. 2 ed. Vol 2. La Habana: Editorial Ciencias Médicas; 2008. p. 71-7.
4. Sereno Moyano M, González-Barón M, Redondo Sánchez A, Martínez Jáñez N. Factores pronósticos del cáncer de pulmón. Revisiones en cáncer. 2003; 17(2): 43-55.
5. Larrínaga B, Córdoba A, Gómez M. Factores pronósticos relacionados con la histología de los tumores. ANALES Sis San Navarra. 2000 [citado 23 Ene 2014]; 24(Sup 1).
6. American Cancer Society. Cáncer de pulmón microcítico (cáncer de pulmón de células pequeñas) [citado 23 Ene 2014].
7. Olasolo JJ, Redondo EA, León Jiménez A, Rueda Ramos A. Carcinoma no microcítico de pulmón. Supervivencia y factores pronósticos del tratamiento radioterápico. Archivos de Bronconeumología. 2003; 39(2): 81-6.
8. Zamora Matamoros L, Orús Bánega P, Díaz Silvera J. El Análisis Estadístico Implicativo, instrumento común de investigación en una experiencia de cooperación multidisciplinar: "Visualizar" una expresión de discontinuidad del rendimiento académico en estudiantes universitarios de Matemática y Computación usando análisis estadístico implicativo. Quaderni di Ricerca di Ricerca in Didattica (Mathematics). 2010; 20(Suppl 1): 451-75.
9. Orús Báguena P, Zamora Matamoros L, Gregori Huerta P. Trabajos teóricos. Conceptos fundamentales del Análisis Estadístico Implicativo (ASI) y su soporte computacional CHIC. En: Teoría y aplicaciones del Análisis Estadístico Implicativo: primera aproximación en lengua hispana. Castellón: Universitat Jaume I de Castellón; 2009. p. 65-103 [citado 23 Ene 2014].
10. Edo Gual PI, Huerta Palau. Estudios sobre los problemas ternarios de probabilidad condicional de nivel NO [tesis doctoral]. Valencia: Universitat de València; 2014.
11. Sánchez Hernández I, Izquierdo Alonso JL, Almonacid Sánchez C. Situación epidemiológica y pronóstica del cáncer de pulmón en nuestro medio. Arch Bronconeumol. 2006; 42(11): 594-99.
12. Ma Q, Liu D, Guo Y, Shi B, Tian Y, Song Z, et al. Surgical techniques and results of the pulmonary artery reconstruction for patients with central non-small cell lung cancer. J Cardiothorac Surg. 2013; 8: 219.

13. Guo Q, Wu M, An T, Zhao J, Duan J, Wang Z, et al. Characteristics and treatment of bone metastases in 322 cases non-small cell lung cancer: a retrospective study. Zhongguo Fei Ai Za Zhi. 2014; 17(9): 656-62.
14. Lastra Aras E. Cáncer de pulmón. Factores pronósticos. El Médico Interactivo [citado 23 Ene 2014].

Recibido: 16 de diciembre de 2015.

Aprobado: 29 de enero de 2016.

**Alina Moraga Rodríguez.** Facultad No. 2 de Ciencias Médicas, avenida Cebreco, km 1 ½, reparto Pastorita, Santiago de Cuba, Cuba. Correo electrónico:  
[annia@medired.scu.sld.cu](mailto:annia@medired.scu.sld.cu)