

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

**Statistical analysis involving mortality prediction factors for the identification of prostate cancer**

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

<sup>I</sup> Facultad de Medicina No. 1, 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> Hospital Provincial Docente Clínicoquirúrgico "Saturnino Lora Torres", Santiago de Cuba, Cuba.

**RESUMEN**

Se efectuó una investigación observacional y analítica, desde septiembre de 2013 hasta igual mes de 2014, con el fin de evaluar la utilidad de una nueva técnica de análisis estadístico implicativo para la identificación de los factores pronósticos de la mortalidad por cáncer de próstata en la provincia de Santiago de Cuba. Según la regresión logística, los factores que empeoraron el pronóstico en los pacientes con cáncer de próstata fueron la afectación ganglionar y el grado III de la diferenciación histológica, y según el análisis estadístico implicativo, lo hicieron la afectación ganglionar, las metástasis y las complicaciones. El análisis estadístico implicativo complementó a la regresión logística en la identificación de los factores pronósticos, con lo cual se logró una mejor comprensión de la causalidad y se elevó la calidad de este tipo de estudio.

**Palabras clave:** análisis estadístico implicativo, cáncer de próstata, factores pronósticos.

**ABSTRACT**

An observational and analytic investigation was carried out, from September, 2013 to the same month in 2014, with the purpose of evaluating the usefulness of a new technique of involving statistical analysis for the identification of mortality prediction factors for prostate cancer in Santiago de Cuba. According to the logistical regression, the factors that worsened the prediction in the patients with prostate cancer were the ganglionic disorder and the grade III of the histological differentiation, and according to the involving statistical analysis, the ganglionic disorder, the metastasis and complications. The involving statistical analysis supplemented the logistical regression in the identification of the prediction factors, with which a better understanding of the causation was achieved and there was an increase in the quality of this type of study.

**Key words:** involving statistical analysis, prostate cancer, prediction factors.

## INTRODUCCIÓN

La identificación de los factores pronósticos es de gran importancia para la prevención secundaria y terciaria de las enfermedades, pues en muchos casos indicarán el tratamiento.<sup>1</sup>

En otro orden de ideas, el cáncer de próstata es la neoplasia maligna más frecuente en varones, y la segunda causa de muerte por cáncer después de las neoplasias broncopulmonares. Actualmente se observa un aumento exponencial de la mortalidad por esta causa a partir de los 64 años.<sup>2,3</sup>

Los factores pronósticos de primera magnitud en el cáncer de próstata son el estadio y el grado de diferenciación.<sup>4</sup>

Al respecto, en la valoración del estadio se requiere un tratamiento adecuado del fragmento hístico extraído, para poder determinar la extensión del tumor, el volumen del tumor o porcentaje de invasión en la próstata, la situación de los márgenes de resección, la invasión de las vesículas seminales o no, el estudio de los ganglios linfáticos con remisión y la clasificación tumor-nódulo-metástasis (TNM) o de Whitmore-Jewett.

El grado de diferenciación del tumor se realiza con el índice de Gleason, puesto que 90 % de los carcinomas son de tipo acinar. Con este índice se valora la diferenciación estructural y la heterogeneidad del tumor en los grados bajo, medio y alto.

Actualmente no existen marcadores predictivos con evidencia aplicables a la biopsia de próstata; es decir, al contrario de lo que ocurre en otros carcinomas, como el de mama, en que la expresión cuantitativa de receptores hormonales en la biopsia pronostica la respuesta al tamoxifeno, en el cáncer de próstata no ocurre lo mismo con los receptores de andrógenos y la respuesta a hormonoterapia.<sup>5,6</sup>

Para el estudio de los factores pronósticos se emplean herramientas de la estadística, ciencia que se ocupa de los métodos y procedimientos para recoger, clasificar, resumir, hallar regularidades y analizar los datos --siempre y cuando la variabilidad e incertidumbre sean una causa intrínseca de estos--, así como realizar inferencias a partir de ellos, con la finalidad de ayudar a tomar decisiones y, en su caso, formular predicciones. Al respecto, tradicionalmente se utilizan las técnicas estadísticas multivariadas, y de ellas, la regresión logística o regresión de Cox.<sup>7</sup>

El análisis estadístico implicativo es un método de análisis no simétrico de datos que permite, partiendo de un conjunto de datos que interrelaciona una población de sujetos u objetos con un conjunto de variables, la extracción y estructuración del conocimiento en forma de normas y reglas generalizadas y, a partir de la contingencia de estas, la explicación y, consecuentemente, una determinada previsión en distintas ramas del saber.<sup>8,9</sup>

Su origen se encuentra en la modelización estadística de la cuasi-implicación: cuando la variable o la conjunción de variables "a" se observa en la población, entonces generalmente la variable "b" también lo está.

Ahora bien, desde la década de los 70, y gracias al trabajo del francés Régis Gras en el campo del análisis estadístico implicativo (*analyse statistique implicative* -- ASI --), se viene estudiando el fenómeno de la creación de reglas de inferencia imprecisas (o parciales) entre variables binarias (atributos o caracteres) que describían a una población

de individuos (sujetos u objetos), o sea, se trata de descubrir reglas inductivas no simétricas para modelar relaciones del tipo "si a, entonces casi b".

A diferencia de los métodos de análisis simétricos basados, por ejemplo, en una distancia o en una correlación, los conjuntos de reglas obtenidos pueden conducir a hipótesis de causalidad.

El sistema informático CHIC (*Classification Hiérarchique Implicative et Cohésitive*) es un software de análisis de datos, inicialmente concebido por Regis Gras en relación con las técnicas del Análisis Estadístico Implicativo, el cual ha sido sucesivamente desarrollado para las computadoras personales por Saddo Ag Almouloud, Harrisson Ratsimba-Rajohn y, en su versión actual, por Raphaël Couturie. Se caracteriza por su sencillez de manejo, ser innovador en el tratamiento de la información y proporcionar de forma rápida gran cantidad de cálculos estadísticos y gráficos-árboles, y grafo de clasificación e implicación.

## MÉTODOS

Se efectuó una investigación observacional y analítica, desde septiembre de 2013 hasta igual mes de 2014, con el fin de evaluar la utilidad de una nueva técnica de análisis estadístico implicativo para la identificación de los factores pronósticos de la mortalidad por cáncer de próstata en la provincia de Santiago de Cuba.

El universo de los casos estuvo constituido por todas las personas fallecidas con el diagnóstico clínico e histológico de cáncer de próstata, 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 versión 4.1 y EPIDAT versión 3.0. Se confeccionó una base de datos en MS Excel, en formato .cvs 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 y el grado de diferenciación histológica. Los que presentaban afectación ganglionar poseían entre 3,652 y 46,484 más probabilidades de fallecer, con un nivel de confiabilidad de 95 % (tabla 1), incluso este hecho resultó significativo con 99 % de confianza.

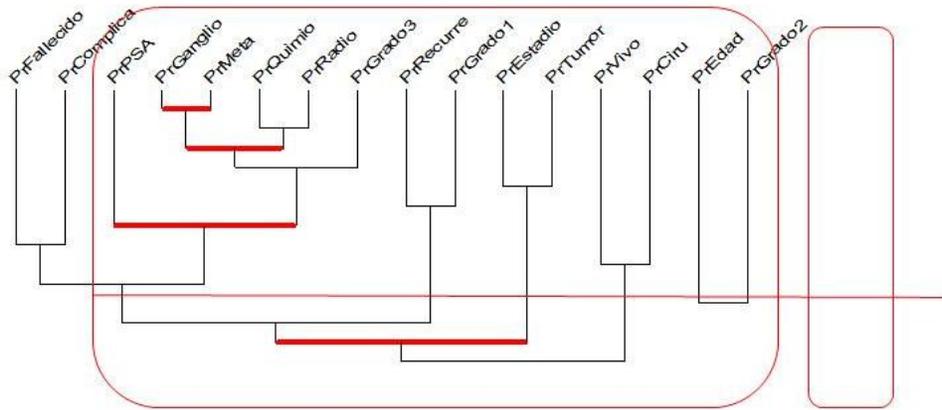
**Tabla 1.** Variables incluidas en el 10mo. paso de la regresión logística (modelo atrás condicional) y sus indicadores

Variables	B	DE	Wald	Gl.	p	Exp. ( $\beta$ )	IC 95 % para Exp. ( $\beta$ )	
							Inferior	Superior
Antígeno prostático específico	-1,177	0,641	3,371	1	0,066	0,308	0,088	1,083
Afectación ganglionar	2,567	0,649	15,652	1	0,000	13,030	3,652	46,484
Grado de diferenciación histológica	-1,087	0,412	6,975	1	0,008	0,337	0,151	0,756
Cirugía	-0,960	0,525	3,348	1	0,067	0,383	0,137	1,071
Constante	1,573	0,751	4,389	1	0,036	4,821		

La probabilidad de fallecer, con un nivel de confiabilidad de 95 %, se incrementó entre 0,151 y 0,756 veces respecto a la de no fallecer, a medida que se pasó de un estadio de diferenciación a otro. El valor de la constante en el modelo de regresión logística indica que, aún cuando no esté presente ninguno de los factores analizados, la probabilidad de fallecer es 4,82 veces mayor que la de no fallecer.

### Con el análisis estadístico implicative

En el árbol de similaridad se aprecia la formación de dos grupos o clases que engloban a todos los factores pronósticos (figura 1).



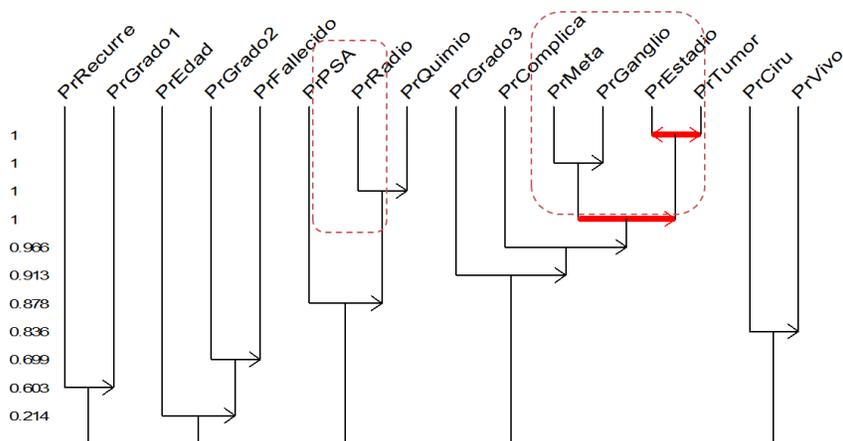
**Fig. 1.** Árbol de similaridad

Al efectuar un corte con un índice de similaridad por encima de 90 %, se apreciaron los factores más similares, los que formaron 4 grupos:

- 1) unión de los fallecidos con los que presentaron complicaciones,
- 2) mayor cantidad de variables similares (afección ganglionar, metástasis, radioterapia, quimioterapia, grado III de diferenciación histológica y valores elevados del antígeno – PSA--) con tres nodos significativos en los niveles 1, 3 y 6,
- 3) similitud entre los que poseen un grado I de diferenciación histológica y la recurrencia, lo que es lógico porque el tumor menos diferenciado tiene más tendencia a recidivas.
- 4) similitud de los que se encuentran en estadio avanzado con los que presentan tumor de mayor tamaño, lo que es lógico porque el tamaño incide en la clasificación en estadios.

Los fallecidos tuvieron factores similares a los que presentaron complicaciones y los vivos a los operados, con índices de similaridad de 9,72 y 89,43, respectivamente.

El nodo más significativo en el árbol cohesivo de la figura 2 está en el nivel 1, también hay un nodo significativo en el nivel 4.



**Fig. 2.** Relaciones causales entre factores pronósticos (árbol cohesivo)

La conjunción de estas variables definieron las reglas (figura 3). Con la máxima intensidad se apreciaron las siguientes 4 reglas:

$R_1$ : ( $PrEstadio \Leftrightarrow PrTumor$ ), la cual significa que un estadio avanzado de la enfermedad implicó siempre una tumoración mayor de 5 cm. Esto es lógico porque el tamaño del tumor influye en su estadificación clínica; igualmente los que tuvieron mayor tamaño tumoral siempre clasificaron en estadio avanzado.

$R_2$ : ( $PrMeta \Rightarrow PrGanglio$ ), la cual significa que la presencia de metástasis implicó siempre afectación ganglionar.

$R_3$ : ( $PrRadio \Rightarrow PrQuimio$ ), que significa que haber recibido radioterapia implicó haber recibido quimioterapia; signo de que en todos estos pacientes se asoció la combinación terapéutica.

$R_4$ : ( $R_2 \Rightarrow R_1$ ), significa que un estadio avanzado de la enfermedad implicó siempre una tumoración mayor de 5 cm y viceversa, y esto está determinado siempre por otra regla que es que la presencia de metástasis indicó siempre afectación ganglionar.

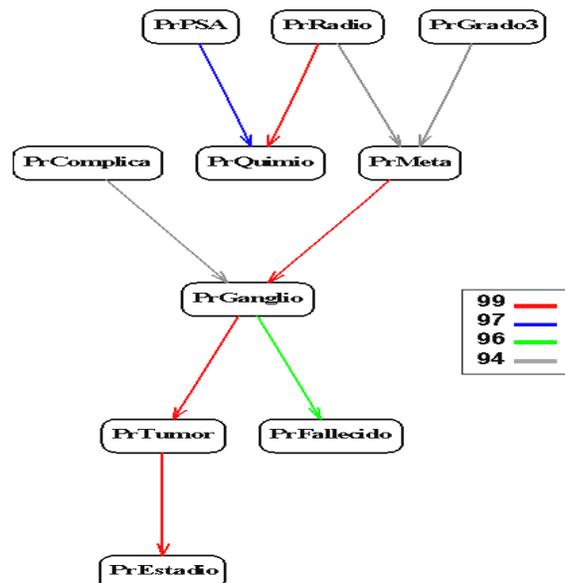
Con un alto nivel de cohesión también se observaron las siguientes reglas:

$R_5$ : ( $PrComplica \Rightarrow R_4$ ); esto significa que la regla 4 es una consecuencia de que el paciente haya presentado complicaciones, con un índice de cohesión de 0,966.

$R_6$ : ( $PrGrado3 \Rightarrow R_5$ ), con una confianza de 91,3 %, donde toda la regla  $R_5$  es implicada a su vez por un grado II de diferenciación histológica.

$R_7$ : [ $PrPSA \Rightarrow (PrRadio \Rightarrow PrQuimio)$ ], la regla 7 plantea, con 0,878 de cohesión, que todos los pacientes que recibieron radioterapia necesitaron quimioterapia y esto estuvo determinado por tener un valor de PSA elevado.

$R_8$ : ( $PrCiru \Rightarrow PrVivo$ ); en esta se plantea que con una cohesión de 0,836, los individuos operados siguieron con vida.



**Fig. 3.** Grafo implicativo de las relaciones causales entre los factores pronóstico con 95 % y más de intensidad implícita.

## Análisis comparativo entre ambas técnicas

Mediante la regresión logística se identificaron como factores de mal pronóstico las variables afectación ganglionar, con 99 % de confianza, y el grado de diferenciación histológica III, con 95 %, mientras que con el ASI fueron la presencia de metástasis, con 93 % de confianza, la presencia de complicaciones, con 91 %, y la afectación ganglionar (tabla 2).

**Tabla 2.** Variables identificadas como factores pronóstico de cáncer de próstata según el ASI y la regresión logística (RL)

ASI	RL		Total
	Sí	No	
Sí	1	2	3
No	1	8	9
Total	2	10	12

Con ambas técnicas se logró identificar la afectación ganglionar como factor de mal pronóstico. A través de la regresión logística se identificó un factor que no se determinó con el ASI y con el ASI fueron identificados 2 que no se definieron con la regresión logística.

Al evaluar la técnica del ASI como identificador de factores pronósticos en el cáncer de próstata, con 95 % de confianza se obtuvo de 24,58 a 75,42 % de capacidad para detectar las variables significativamente asociadas al cáncer, ya sea como factor de buen o de mal pronóstico (sensibilidad). También se demostró entre 74,85 y 85,15 % de capacidad para delimitar las variables que no constituyen factor pronóstico (especificidad).

Un factor pronóstico identificado por el ASI tiene una probabilidad de 33,33 % de serlo realmente (valor predictivo positivo), así mismo una variable que no constituye un factor pronóstico tienen una probabilidad de 88,89 % de no serlo en realidad (valor predictivo negativo).

Con esta técnica es 2,5 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 modo, fue 0,63 veces más probable que una variable que no era factor pronóstico fuera clasificada como tal (razón de verosimilitud negativa), y la concordancia entre ambas técnicas fue de 0,25 (índice de Kappa).

La diferencia entre la tasa de positivos verdaderos y de positivos falsos fue de 30 % (índice de Youden). Con el ASI se clasificó correctamente a 75 % de las variables (índice de validez).

## DISCUSIÓN

A pesar de que en este estudio el tratamiento quirúrgico mostró cierta similitud en los pacientes vivos, no se puede considerar de buen pronóstico. Un seguimiento de 8 años en un hospital de Francia a 1 943 prostatectomizados reveló un pobre pronóstico en más de 40 % de ellos, con recurrencia de la enfermedad y muerte.<sup>10</sup>

Otro estudio retrospectivo en España a 925 prostatectomizados mostró el valor pronóstico de esta al curarse una tercera parte de los más jóvenes, independiente de los casos con metástasis y los valores elevados del PSA.<sup>11</sup>

También en una investigación efectuada en Italia se confirmaron los beneficios del tratamiento quirúrgico, teniendo en cuenta además la afectación ganglionar.<sup>12</sup>

La quimioterapia ha sido un factor discutido, que puede mejorar mucho el pronóstico por los nuevos avances moleculares, sobre todo si se emplea antes de que aparezcan las metástasis, como ha sido notificado por algunos autores.<sup>13,14</sup>

En esta serie se identificaron los siguientes factores relacionados con el cáncer de próstata:

- Según la regresión logística, los factores pronósticos de empeoramiento fueron la afectación ganglionar y el grado III de diferenciación histológica. No se determinó ningún factor pronóstico de mejoría.
- Según el análisis estadístico implicativo, los factores de empeoramiento fueron la afectación ganglionar, las metástasis y las complicaciones. No se obtuvo factor pronóstico alguno de mejoría.

## REFERENCIAS BIBLIOGRÁFICAS

1. Pita Fernández S, Valdés Cañedo FA. Determinación de factores pronósticos. Cad Aten Primaria. 1997 [citado 23 Ene 2014]; 4: 26-29.
2. De Torres Ramírez I. Factores pronósticos y predictivos del carcinoma de próstata en la biopsia prostática. Actas Urol Esp. 2007 [citado 23 Ene 2014]; 31(9).
3. Gimeno F, Piñango L. Cáncer de próstata-Diagnóstico precoz. Madrid: Schering-Plough; 1994. p. 11-69
4. 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 (Supl. 1): 7-15.
5. Segarra Tomás J, Millán Rodríguez F, Palou Redorta J, Villavicencio Mavrich H. Factores pronósticos y tablas predictivas del cáncer de próstata clínicamente localizado Servicio de Urología. Actas Urol Esp 2006; 30(6): 567-73.
6. Lastra Aras E. Cáncer de pulmón. Factores pronósticos. El Médico Interactivo [citado 23 Ene 2014].
7. Flores Luna L, Salazar Martínez E, Duarte Torres RM, Torres Mejía G, Alonso Ruiz P, Lazcano Ponce E. Factores pronósticos relacionados con la supervivencia del cáncer de mama. Salud Pública Mex. 2008; 50: 119-25.
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. Ploussard G1, Agamy MA, Alenda O, Allory Y, Mouracade P, Vordos D, et al. Impact of positive surgical margins on prostate-specific antigen failure after radical prostatectomy in adjuvant treatment-naïve patients. BJU Int. 2011; 107(11): 1748-54.
11. Rubio-Briones J, Iborra I, Trassierra M, Collado A, Casanova J, Gómez-Ferrer A, et al. Metastatic progression, cancer-specific mortality and need for secondary treatments in patients with clinically high-risk prostate cancer treated initially with radical prostatectomy. Actas Urol Esp. 2010; 34(7): 610-7.
12. Novara G, Ficarra V, Mocellin S, Ahlering TE, Carroll PR, Graefen M, et al. Systematic review and meta-analysis of studies reporting oncologic outcome after robot-assisted radical prostatectomy. Eur Urol. 2012; 62(3): 382-404.
13. Drouin SJ1, Rouprêt M, Wallerand H, Houédé N. Chemotherapy in early stage of hormone-resistant metastatic prostate cancer: what are the indications? Prog Urol. 2010; 20 Suppl 3: 192-7.
14. Edo Gual PI, Huerta Palau MP. Estudios sobre los problemas ternarios de probabilidad condicional de nivel N0 [tesis doctoral]. Valencia: Universitat de Valencia; 2014.

Recibido: 16 de diciembre de 2015.

Aprobado: 29 de enero de 2016.

*Alina Moraga Rodríguez*. Facultad de Medicina No. 1, avenida de las Américas, entre calles I y E, reparto Sueño, Santiago de Cuba, Cuba. Correo electrónico: [a.rodriguez@infomed.sld.cu](mailto:a.rodriguez@infomed.sld.cu)