

ALGUNOS CONCEPTOS DE ANALISIS DE SOBREVIVENCIA: UNA APLICACIÓN A DATOS DE PACIENTES CON CANCER DE CERVIX

Jenny-Carolina Serrato-Rincón & Constanza Quintero-Guzmán
Universidad Nacional de Colombia

Resumen

Con el fin de ilustrar el uso de los análisis de sobrevivencia, se presenta un estudio con 58 mujeres con cáncer de cerviz para analizar ciertos anticuerpos a la proteína E7 del virus del papiloma humano (HPV), como posibles factores de pronóstico al tiempo de sobrevivencia, medido desde el comienzo de un tratamiento de radioterapia hasta la muerte. Para el análisis se hicieron comparaciones de curvas de sobrevivencia, en el caso de determinar la influencia de una única variable y para el caso multivariado se utilizaron modelos de regresión. Se concluyó que este grupo de pacientes no refleja ninguna de las variables consideradas como factor pronóstico al tiempo de sobrevivencia mencionado.

Palabras claves:

Abstract

In order to illustrate the use of survival analysis, we present a study with 58 women with cervix cancer to analyze certain antibodies to the E7 protein of the humane papiloma virus (HPV), considered as possible prognosis factors for the survival time which is measured from the beginning of an x-ray treatment to death. For the analysis, in the case of determining the influence of a single variable some comparisons of survival curves in the respective category variables were made and for the multivariable case regression models were used. It was concluded that this group of patients does not reflect any of the considered variables as being prognosis factors for the above mentioned survival time.

Key words:

Introducción

El cáncer de cérvix constituye un problema importante de salud pública, aproximadamente el 80% de los 500000 casos que se calcula se presentan anualmente en el mundo corresponde a países en vía de desarrollo, en nuestro país es la primera causa de muerte por cáncer en la mujer en edad reproductiva. En el Instituto Nacional de Cancerología uno de cada cinco pacientes es atendido por esta causa.

El virus de papiloma humano, o papilomavirus humano (*HPV*) es un grupo de más de 80 tipos de virus. Se llaman papiloma virus porque ciertos tipos pueden causar verrugas, o papilomas, que son tumores benignos, no cancerosos. Diferentes tipos de virus de papiloma humano causan las verrugas comunes que crecen en las manos y en los pies y aquellas que se desarrollan en la boca y en el área genital.

El genoma del virus del papiloma humano (*HPV*) ha sido detectado en más de un 90% de los carcinomas cervicales; los tipos virales más frecuentes son el 16 y el 18, los cuales también se presentan con frecuencia en las lesiones premalignas. La exposición a un agente viral usualmente estimula anticuerpos específicos en el suero del huésped. La medición de la respuesta de estos anticuerpos puede ser de utilidad en el diagnóstico de la infección.

En varios estudios del cáncer de cérvix se encuentra una alta frecuencia de anticuerpos a la proteína E7-H *PV16* en el grupo de pacientes con la enfermedad, en comparación con los individuos del grupo control, pero aún no se ha establecido con certeza si la presencia de estos anticuerpos actúa como un factor pronóstico con respecto al tiempo de sobrevivencia de estos pacientes.

Esta aplicación tuvo como objetivo identificar algunos factores de pronóstico para predecir el tiempo de sobrevivencia, tomando como punto inicial el comienzo del tratamiento de radioterapia. En particular, mediante el uso de análisis de sobrevivencia, se buscaba establecer si la presencia de ciertos anticuerpos a la proteína E7 del virus del papiloma humano (*HPV*) puede usarse como factor pronóstico del tiempo de sobrevivencia desde el comienzo de la aplicación del tratamiento hasta la muerte, en pacientes con carcinoma de cérvix tratadas en el Instituto Nacional de Cancerología entre agosto 3 de 1995 hasta agosto 31 de 2001.

Modelos de sobrevivencia

En investigación en salud son frecuentes los estudios cuya variable de interés es el tiempo de sobrevivencia, entendiendo éste como el intervalo de tiempo desde un punto inicial hasta un punto final, fecha de desenlace del evento, por ejemplo, desde la fecha de diagnóstico de una enfermedad hasta la muerte, desde la fecha de remisión hasta la primera recaída, desde el final de un tratamiento hasta a muerte o recaída.

Los datos de sobrevivencia tienen dos características que han dado lugar a desarrollar modelos específicos. Una de éstas es la ocurrencia de datos incompletos que corresponden a los individuos con fecha de desenlace del evento, desconocida, bien sea porque abandonan el estudio o porque éste termina y no ha ocurrido el evento; por lo tanto su dato de tiempo de sobrevivencia es incompleto y se llama dato censurado, se mide desde el punto inicial hasta la última fecha en que se observó al individuo. Los datos completos se llaman datos no censurados.

Si notamos T la variable tiempo de sobrevivencia, la distribución de T se caracteriza por tres funciones, a saber:

1. función de sobrevivencia. Es la probabilidad de que un individuo sobreviva un tiempo mayor que T : $S(t) = \text{Pr ob}(T > t)$

2. Función de densidad de probabilidad. Es el límite de la probabilidad de que un individuo tenga el evento en el intervalo corto, t a $t + \Delta t$, por unidad de tiempo:

$$f(t) = \lim_{\Delta t \rightarrow 0} \frac{\text{Prob}(\text{un individuo tenga el evento en el intervalo}(t, t + \Delta t))}{\Delta t}.$$

3. Función riesgo. Está definida como la probabilidad de ocurrencia del evento, durante un intervalo pequeño de tiempo, dado que el individuo ha sobrevivido el comienzo del intervalo o como el límite de la probabilidad de que un individuo tenga el evento en un intervalo muy corto, t a $t + \Delta t$, por unidad de tiempo, dado que el individuo ha sobrevivido en el tiempo t :

$$f(t) = \lim_{\Delta t \rightarrow 0} \frac{\text{Prob}(\text{un individuo de edad } t \text{ tenga el evento en el intervalo}(t, t + \Delta t))}{\Delta t}.$$

Método de Kaplan-Meier para estimar $S(t)$

Este método también es llamado método del producto límite pues la estimación de la función de sobrevivencia $S(t)$ en cada t , se obtiene del producto acumulado de la proporción de pacientes que sobreviven al tiempo k dado que sobrevivieron al tiempo $k - 1$, y se nota p_k ". El método es ventajoso pues no requiere ninguna suposición acerca de la forma de la función $S(t)$, se puede aplicar a pequeñas, medianas y grandes muestras, y además, tiene en cuenta las observaciones censuradas¹, pues éstas aportan a la proporción.

La estimación de $S(k)$ se obtiene de la siguiente forma, $\hat{S}(k) = p_1 * p_2 * \dots * p_k$, donde $p_k = \frac{n - r}{n - r + 1}$, n es el número total de pacientes, p_k es la proporción de individuos que sobreviven el k -ésimo año después de haber sobrevivido $k-1$ años y r es el mayor número de observación no censurada menor o igual que k , después de enumerar todas las observaciones de menor a mayor incluyendo los tiempos censurados. Si existen varias observaciones iguales, $S(k)$ se estima para la última de éstas.

Aplicando logaritmos y métodos gráficos a la curva de sobrevivencia podemos saber si ésta es similar a alguna distribución paramétrica conocida, la más usada para describir la sobrevivencia en humanos es la **distribución Weibull**. La distribución Weibull, fue sugerida por W. Weibull en 1951, esta distribución juega un papel importante en el análisis de sobrevivencia. La función $S(t)$ relacionada con esta distribución está dada por $S(t) = \exp(-(\lambda t)^p)$, donde $\lambda, p > 0$

Este modelo permite que el riesgo dependa del tiempo. Cuando $p < 1$ el riesgo es monótonamente decreciente, si $p > 1$ el riesgo es monótonamente creciente, y en el caso en el que $p = 1$ la función de sobrevivencia está relacionada con la exponencial. Esto se ve reflejado en la función de riesgo que tiene la siguiente forma: $h(t) = \lambda p (\lambda t)^{p-1}$

Comparación de curvas de sobrevivencia

Se considera la hipótesis nula $H_0: S_1(t) = S_2(t)$, las curvas de sobrevivencia son iguales, frente a las hipótesis alternativas

$H_1: S_1(t) > S_2(t)$, la sobrevivencia del grupo 1 es mayor que la del grupo 2.

$H_1: S_1(t) < S_2(t)$, la sobrevivencia del grupo 1 es menor que la del grupo 2.

$H_1: S_1(t) \neq S_2(t)$, la sobrevivencia del grupo 1 es diferente que la del grupo 2.

En este estudio las variables en cuyas categorías nos interesa comparar la sobrevivencia son: edad, FIGO, tamaño tumoral, y la presencia de HPV, presencia de E7 recombinante, presencia de E7 1-20 (aminoácido 1 a 20), presencia de E7 66-85 (aminoácido 66 a 85), todas ellas se explicarán en el apartado siguiente.

Prueba del Logrank

Sean $t(1) < \dots < t(k)$ los distintos tiempos de las observaciones en los dos grupos, $m(j)$ el número de observaciones no censuradas iguales a $t(j)$, y $r(j)$ el número de pacientes aún expuestos al riesgo en el momento $t(j)$. Consideramos las siguientes funciones del logaritmo de la función de sobrevivencia:

$w_i = 1 - e(t_{(i)})$, para una observación no censurada en $t_{(i)}$ y $w_i = -e(T)$ para una observación censurada en T. En la práctica $w_i = 1 - e(t_{(j)})$ para una observación censurada t_i^+ , donde $t_{(j)}$ es la observación no censurada más grande tal que $t_{(j)} < t_i^+$ y $e(t_{(j)}) = \sum_{j \leq t_{(i)}} \frac{m(j)}{r(j)}$

La prueba de Logrank se basa en la suma S de los w_i de uno de los grupos, y en la varianza $Var(S)$. El estadístico $L = \frac{S}{\sqrt{Var(S)}}$ se distribuye asintóticamente normal estándar. Para la hipótesis alternativa $H_1: S_1 > S_2$, las regiones críticas están dadas por $L < -Z_\alpha$, si S se obtiene del grupo 1, y $L > Z_\alpha$, si S se obtiene del grupo 2. α es el nivel de significancia.

La otra característica de los datos de sobrevivencia que llevó a desarrollar estos métodos estadísticos es que, en la mayoría de las situaciones, no se conoce la distribución de los tiempos de sobrevivencia. Se usan entonces métodos que no asumen una distribución específica. A continuación se describe el modelo de riesgo proporcional de Cox, método de regresión múltiple usado para investigar la asociación entre las posibles variables pronóstico y el tiempo de sobrevivencia.

Modelo de riesgo proporcional de Cox

Este método de regresión múltiple se usa para investigar la asociación entre las posibles variables pronóstico y el tiempo de sobrevivencia.

Para un paciente con factores pronóstico X_1, \dots, X_n , se define la tasa de muerte o tasa de riesgo en el tiempo t como $h(t; X_1, \dots, X_n) = h_0(t) \exp(\beta_1 X_1 + \dots + \beta_n X_n)$, $t > 0$, donde $h_0(t)$ es una función arbitraria del tiempo, y β_1, \dots, β_n son coeficientes de regresión que ponderan las variables pronóstico. Así, un valor negativo de β_i indica un pronóstico desfavorable para pacientes con la característica y un valor positivo indica un pronóstico favorable para pacientes con la característica.

Método

Muestra

Pacientes con cáncer de cérvix tratadas en el Instituto Nacional de Cancerología, Bogotá que iniciaron el tratamiento entre el 3 de agosto de 1995 y el 31 de julio de 1996, y cumplieron el criterio de inclusión: tener cáncer de cérvix en estadio IIB o estadio IIIB. Fueron Criterios de exclusión el presentar desnutrición o infección y no concluir el tratamiento. El seguimiento se hizo hasta el 31 de agosto de 2001.

Definición de variables

Variable dependiente: Tiempo de sobrevivencia medido en días, desde el comienzo del tratamiento de radioterapia, hasta la muerte o hasta la fecha de la última observación, o la fecha de terminación del estudio, 31 de agosto de 2001.

Variables independientes: Se consideraron 8 variables independientes definidas a continuación:

Edad. Medida en años cumplidos en el momento del comienzo del tratamiento. Según el criterio de estudios anteriores, se categorizó en menor o igual que cuarenta años y mayor que cuarenta años.

Figo. Estadio clínico, definido por la Federación Internacional de Ginecología y Obstetricia. En la muestra se presentaron los estadios IIB y IIIB.

Tamaño tumoral. Tamaño del tumor al comienzo del tratamiento, se categorizó en menor o igual que cinco centímetros y mayor que cinco centímetros.

Presencia HPV. Indica la presencia del virus de papiloma humano antes de la aplicación del tratamiento de radioterapia.

Presencia de E7 recombinante. Indica la presencia de E7 recombinante antes de la aplicación del tratamiento de radioterapia.

Presencia de E7 1-20. Indica la presencia de E7 1-20 (aminoácido 1 a 20) antes de la aplicación del tratamiento de radioterapia.

Presencia de E7 66-85. Indica la presencia de E7 66-85 (aminoácido 66 a 85) antes de la aplicación del tratamiento de radioterapia.

Respuesta al tratamiento. Tiene dos categorías, una que indica la respuesta completa al tratamiento, y otra que indica una respuesta parcial o nula.

Análisis de datos

Para analizar la influencia individual de cada variable en el tiempo de supervivencia, se estimó la función de supervivencia en cada categoría de la variable, utilizando el método de Kaplan-Meier. Luego se compararon estas estimaciones con las prueba de hipótesis del Logrank.

Para identificar los factores de pronóstico, multivariadamente, se usó la distribución Weibull y el modelo de regresión de Cox. Para los análisis mencionados se utilizó el paquete estadístico SAS.

Resultados

Análisis descriptivo

Los datos correspondientes a 58 pacientes, fueron recolectados por personal de la sección de inmunología del Instituto Nacional de Cancerología. Este grupo fue coordinado por la Doctora Maria Mercedes Bravo, directora de dicha sección. La tabla 1 resume las frecuencias y porcentajes de las variables a estudiar.

Tabla 1
Frecuencia y categorías de las variables de estudio.

Variable	Categoría	Frecuencia	Porcentaje
Edad	<=40 años	13	22.4%
	> 40 años	45	77.6%
Recaída	Si	24	41.4%
	No	34	58.6%
Tamaño Tumoral	<5 cms	27	46.6%
	> 5 cms	31	53.4%
FIGO	Estadio IIB	25	43.1%
	Estadio IIIB	33	56.9%

Variable	Categoría	Frecuencia	Porcentaje
Respuesta al tratamiento	Completa	48	82.8%
	Parcial o nula	10	17.2%
Presencia de <i>HPV</i>	Si	47	81%
	No	11	19%
Presencia de E7 recombinante.	Si	32	56.1%
	No	25	43.9%
Presencia de E7 1-20	Si	15	26.3%
	No	42	73.4%
Presencia de E7 66-85	Si	18	31.6%
	No	39	68.4%

Estimación de la función $S(t)$.

A la totalidad de los datos de tiempo de sobrevivencia se le aplicó el método de Kaplan-Meier para estimar las función $S(t)$. Ver figura 1.

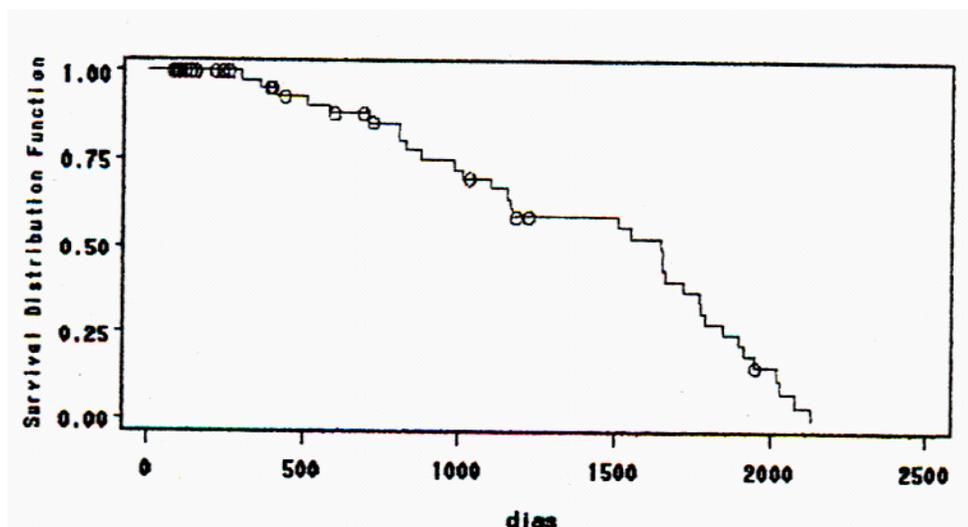


Figura 1. Gráfica de $S(t)$ para todo el grupo

El 41.38% corresponde a observaciones censuradas, la interpolación lineal permite concluir que la mediana de $S(t)$ es 1646 días.

La observación de la figura 2 permite concluir que la función de sobrevivencia $S(t)$ corresponde a una distribución exponencial para $t \leq 1000$, pero como 1000 es muy inferior a la mediana, no se considera que el tiempo de sobrevivencia se distribuya exponencialmente.

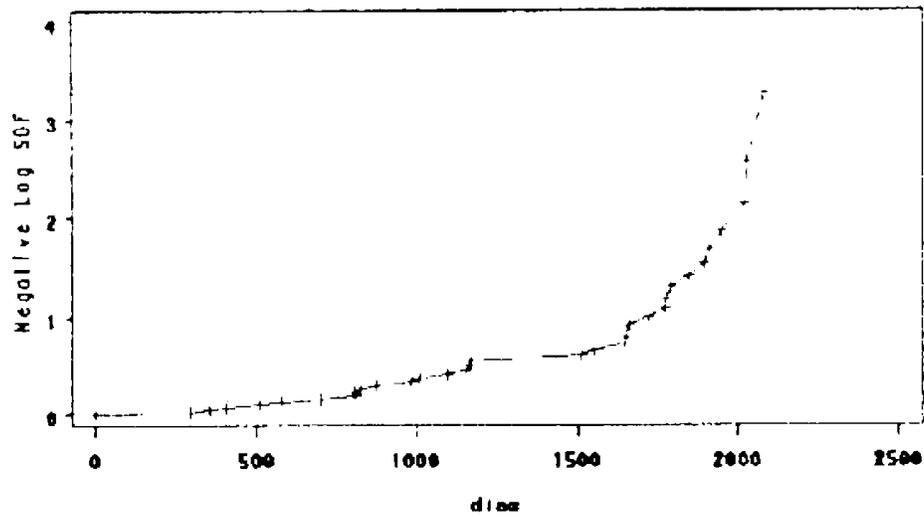


Figura 2: Gráfica de t vs $-\ln S(t)$

La gráfica de la figura 3 muestra que ésta se puede aproximar por una línea recta, entonces, se puede concluir que la función de sobrevivencia $S(t)$ es similar a la distribución paramétrica Weibull, cuyos parámetros se determinan teniendo en cuenta que la pendiente de la recta proporciona una estimación del parámetro p , y el intercepto nos permite estimar a $p \ln(\lambda)$.

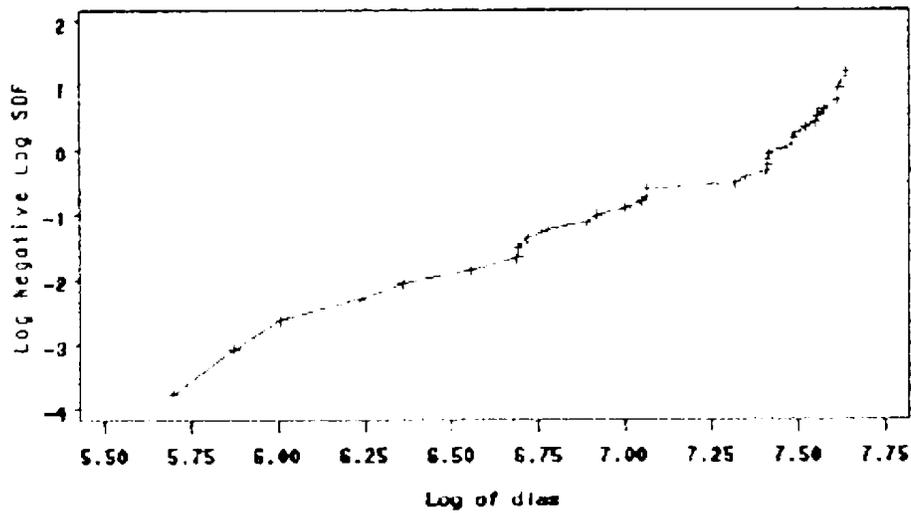


Figura 3: Gráfica de $\ln t$ vs $\ln (-\ln S(t))$.

Comparación de curvas de supervivencia.

Para estudiar la influencia de cada variable individualmente en el tiempo de supervivencia, se compararon las curvas $S(t)$ entre categorías de la variable.

E7 recombinante.

Se dividió el grupo de estudio dependiendo de la presencia de E7 recombinante. En la figura 4 se nota una leve superioridad de la curva de supervivencia de aquellas mujeres que presentaron E7 recombinante. Sin embargo, al aplicar la prueba de Logrank para la variable categorizada se obtuvo un valor Chi-cuadrado de 2.0672, $P > 0.05$, lo que lleva a que no se rechaza la hipótesis de igualdad entre las curvas de supervivencia a un nivel $\alpha = 0.05$.

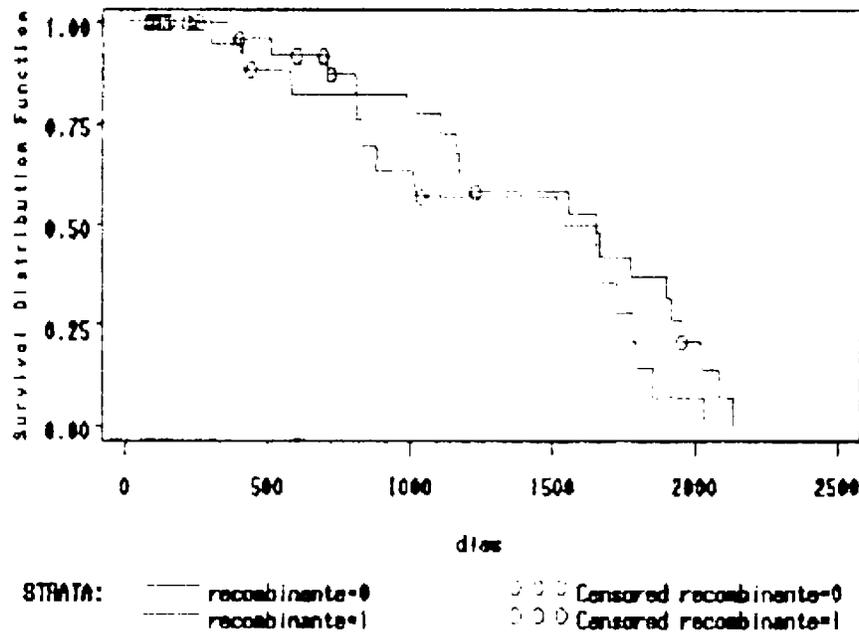


Figura 4: Estimación de Kaplan- Meier para la función de Supervivencia en cada categoría de la variable E7 recombinante.

E7 1-20

La prueba de Logrank aplicada a las curvas de supervivencia de la variable E7 1 - 20 categorizada, no permite, a un nivel $\alpha = 0.05$ rechazar la hipótesis de igualdad entre éstas, como se muestra en la figura 5.

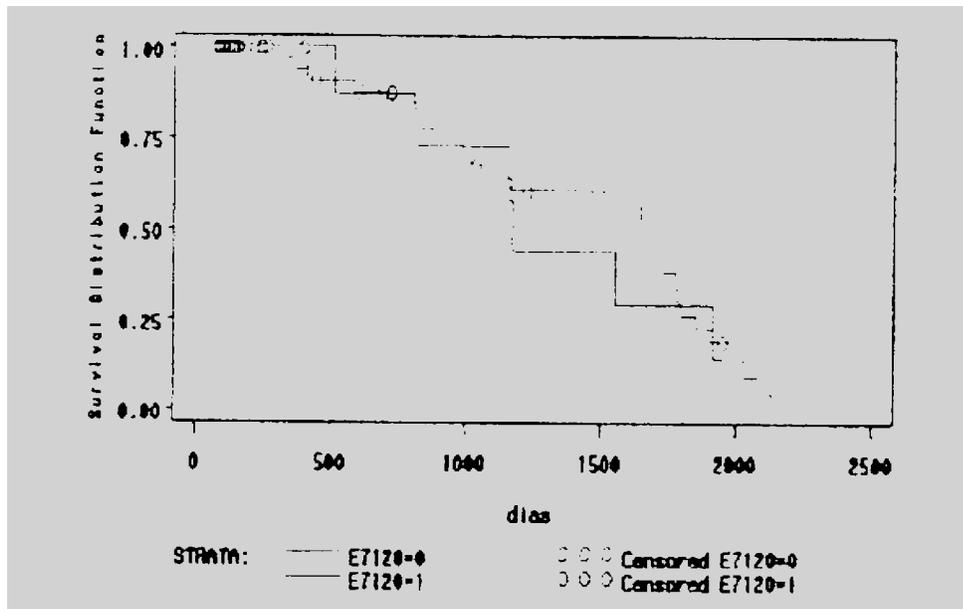


Figura 5: Estimación de Kaplan-Meier para la función de Supervivencia en cada categoría de la variable $E7\ 1 - 20$.

E1 66-85

Al aplicar las prueba de Logrank se obtiene el valor 0.3453, aunque no se rechaza la hipótesis de igualdad entre las curvas para la variable $E7\ 66 - 85$, se destaca en la figura 6 la superioridad de la curva $S(t)$ para quienes no presentaron $E7\ 66 - 85$.

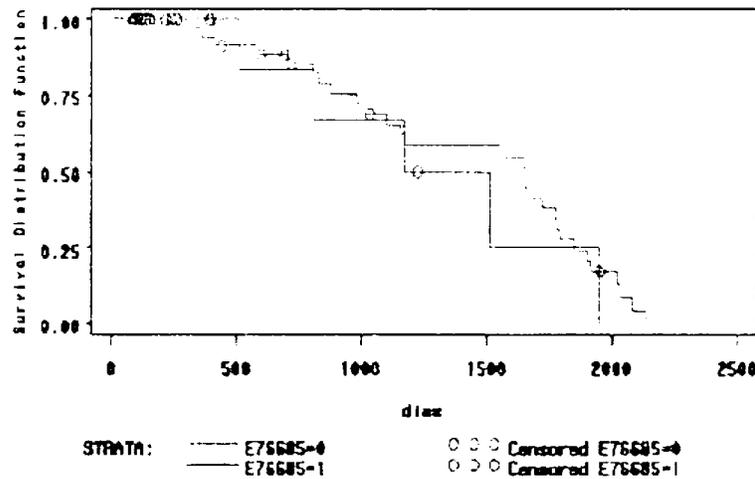


Figura 6. Estimación de Kaplan-Meier para la función de Supervivencia en cada categoría de la variable $E7\ 66-85$.

FIGO

La variable *FIGO* contiene las categorías *IIB* y *II/B*, la comparación gráfica de las curvas de sobrevivencia muestra una leve superioridad para las mujeres en estadio clínico *IIB* (figura 7). La prueba de Logrank proporcionó un valor 0.1264, valor con el que no se rechaza la hipótesis de igualdad de las curvas.

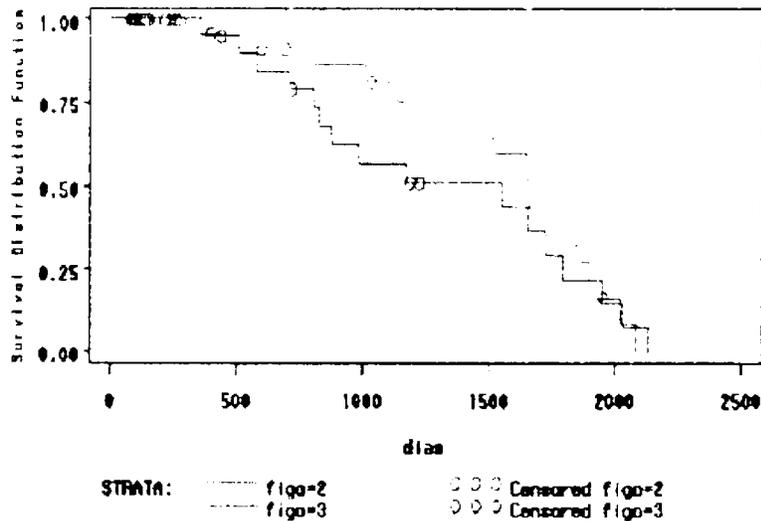


Figura 7: Estimación de Kaplan-Meier para la función de Sobrevivencia en cada categoría de la variable FIGO.

Edad

Las mujeres se dividieron en dos grupos dependiendo de su edad al inicio del estudio, se obtuvieron las curvas de sobrevivencia para aquellas menores de 40 años (inclusive con 40 años), y para las mayores. Ver figura 8. El resultado de la prueba de Logrank, 0.4282, indica que no se debe rechazar la hipótesis de igualdad entre las curvas.

Tamaño tumoral

La variable tamaño tumoral se categorizó dependiendo de si el tumor generado por la enfermedad era superior o inferior a 5 centímetros. La observación de las gráficas de la figura 9 destaca la superioridad de $S(t)$ para quienes presentaron tumores menores a 5 centímetros. La prueba de comparación de las curvas de sobrevivencia no permite rechazar la hipótesis de igualdad entre éstas. El valor obtenido fue 2.7179 para Logrank.

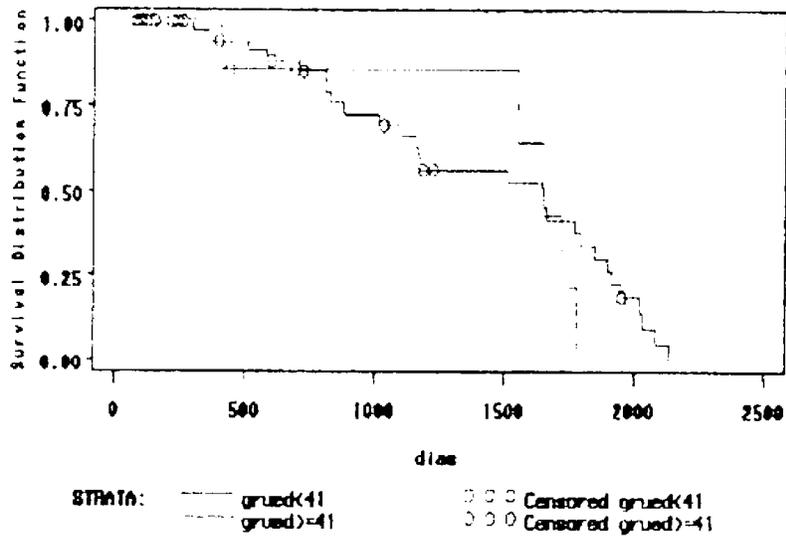


Figura 8. Estimación de Kaplan-Meier para la función de Supervivencia en cada categoría de la variable *Edad*

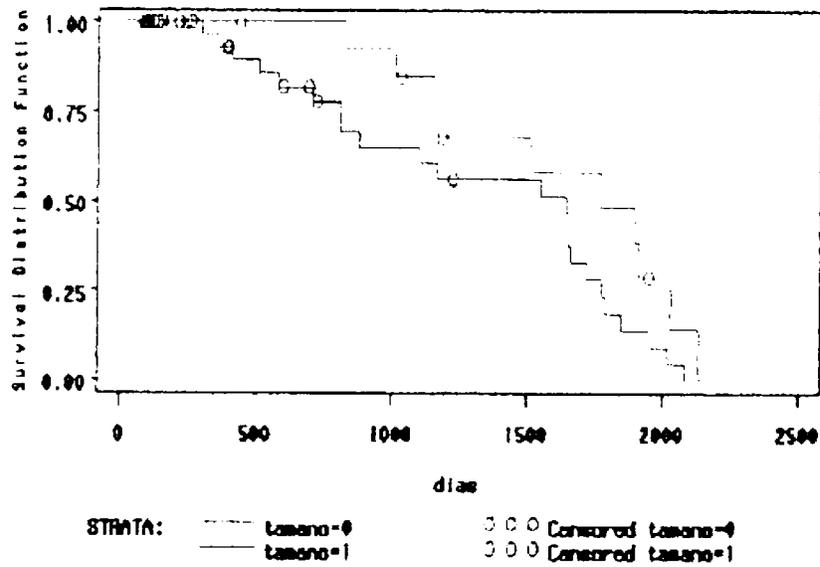


Figura 9: Estimación de Kaplan-Meier para la función de Supervivencia en cada categoría de la variable *tamaño tumoral*

Influencia de las variables en el tiempo de sobrevivencia.

La observación de la figura 3 permite concluir que la función de sobrevivencia $S(t)$ obedece a una distribución Weibull. La estimación de los parámetros para $S(t)$ se realizó utilizando el procedimiento *lifereg* del programa SAS. Los resultados obtenidos con este procedimiento se presentan en la tabla 2

Tabla 2.
Resumen de los parámetros estimados con lifereg

Variable	DF	Estimación	Chi e	p-valor
Intercepto	1	7.49767	167.1557	< .0001
E7 recombinante	1	0.10910	0.5281	0.4674
E7 1 - 20	1	-0.11290	0.3037	0.5816
E7 66 - 85	1	-0.0080021	0.0013	0.9716
FIGO	1	-0.04155	0.0911	0.7628
Grupo de edad	1	-0.0011455	0.0246	0.8755
Tamaño tumoral	1	-0.08251	0.3335	0.5636
Escala	1	0.36696		

Los p valores observados para cada una de las variables de interés son bastante grandes, lo que quiere decir que no podemos rechazar la hipótesis nula que afirma, para cada variable, que ésta no tiene influencia en el tiempo de sobrevivencia.

A partir de los resultados presentados en la tabla 2, se estimaron los valores para los parámetros de la distribución Weibull usando la función $S(t)$ con la forma:

$$S(t) = \exp(-\exp((t-\lambda)/p))$$

Que nos permite hallar las estimaciones de los parámetros:

$$P = 2.7251$$

$$\lambda = 1/1803.8346$$

Entonces la función de riesgo toma la forma: $H(t) = (.7251/1803.835)(t/1803.8346)^{2.7251-1}$

Y la función de sobrevivencia está dada por: $S(t) = \exp(-(t/1803.8346)^{2.7251})$

La gráfica de $S(t)$ aparece en la figura 10.

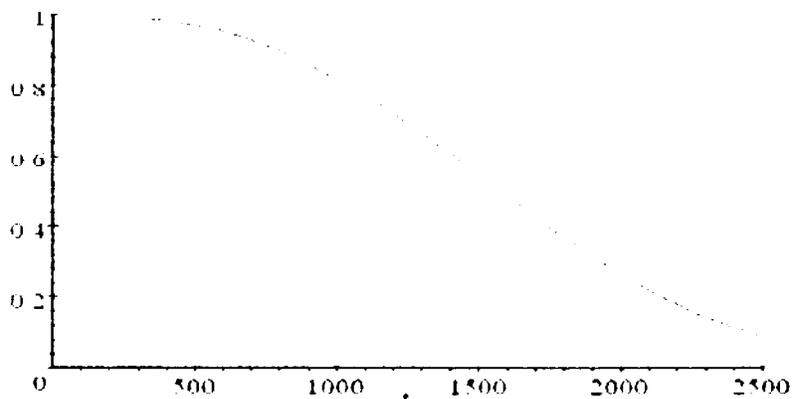


Figura 10. Estimación de $S(t)$ mediante el modelo Weibull

Para contrastar los resultados obtenidos mediante el análisis paramétrico se realizó el análisis multivariado del modelo de Cox, éste se llevo a cabo mediante el procedimiento *phreg* del programa SAS. Los resultados se presentan en tabla 3.

Tabla 3.
Resumen de los parámetros estimados con *phreg*

<i>Variable</i>	<i>DF</i>	<i>Estimación</i>	<i>Chí e</i>	<i>p-valor</i>	<i>HR</i>
<i>E7 Recombinante</i>	1	-0.74647	2.7621	0.0965	0.474
<i>E7 1 - 20</i>	1	0.66871	1.1892	0.2755	1.952
<i>E766-85</i>	1	0.10461	0.0259	0.8720	1.110
FIGO	1	-0.09025	0.0530	0.8179	0.914
Grupo de edad	1	-0.00575	0.0811	0.7759	0.994
Tamaño tumoral	1	0.52345	1.6414	0.2001	1.688

Este análisis corrobora los resultados del análisis paramétrico en cuanto a que no se puede rechazar la hipótesis nula que indica que las variables no tienen influencia en el tiempo de sobrevivencia. Por lo tanto concluimos que ningunas variable de estudio tiene una influencia significativa en el tiempo de sobrevivencia de las pacientes.

Conclusiones

El análisis de la información permitió establecer las siguientes conclusiones.

1. La respuesta al tratamiento de radioterapia, al que fueron sometidas las pacientes, muestra ser mejor para aquellas que no presentaron *E7 1-20*, este análisis también permite

concluir que para aquellas pacientes que no presentaron $E7$ 66 - 85 la respuesta es significativamente mejor. La variable $E7$ recombinante no muestra influencia en la respuesta al tratamiento.

2. Al realizar la comparación de las curvas de sobrevivencia entre categorías de cada una de las variables, se llegó a la conclusión que en ningún caso existía diferencia significativa, aunque gráficamente las curvas de sobrevivencia de las pacientes en estadio clínico IIB , y de las pacientes que presentaron tumores de menos de 5 centímetros, son un poco superiores.

3. Durante el desarrollo de este trabajo se estableció que los datos se podían describir por el modelo paramétrico de Weibull. Sin embargo se llevó a cabo el análisis con estos dos modelos, llegando a concluir que ninguna de las variables de estudio influye significativamente en el tiempo de sobrevivencia de las pacientes, conclusión que corrobora los resultados obtenidos al realizar las comparaciones de las funciones $S(t)$ para las variables categorizadas.

Bibliografía

- Bravo, M. et al. (1997). *Significado pronóstico de los anticuerpos antiproteína E7 del virus del papiloma humano tipo 16 en la supervivencia de los pacientes con carcinoma invasivo de cérvix*.
- Collett, D. (1994). *Modelling Survival Data in Medical Research*. Londres: Chapman & Hall.
- Dawson - Saunders B. (1994). *Bioestadística Médica*. México: Manual Moderno.
- Lee E. (1992). *Statistical Methods for survival analysis data*. New York: John Wiley & Sons.
- London, D. (1988). *Survival Models and Their Estimation*. (Segunda edición), Winsted, ACTEX Publication.
- Marubini E. et al. (1994). *Analysing Survival Data from Clinical trials and Observational Studies*. New York: John Wiley & Sons.
- Viladiu P., et al. (1997). Human Papillomavirus DNA and Antibodies to Human Papilloma viruses 16 E2, L2, and E7 peptides as Predictors of Survival in Patients With Squamous Cervical Cancer. *Journal of Clinical Oncology*, 15, 610-619.
- SAS/ STAT. (1990). *User's guide*. Version 6, fourth edition, Volume 2. USA: SAS Inc.

Autoras

Jenny Carolina Serrato Rincón. Matemática. Universidad Nacional de Colombia. Bogotá, Colombia.

Constanza Quintero Guzmán. Profesora Asociada. Departamento de Matemáticas. Universidad Nacional de Colombia. Bogotá, Colombia. E-Mail: cquinterog@unal.edu.co