

# Ronda clínica y epidemiológica. Introducción al análisis multivariable (parte I)

Daniela Garcés<sup>1</sup>, Fabián Jaimes Barragán<sup>2</sup>

## INTRODUCCIÓN

El nuestro es un mundo “multivariable”. La mayoría de los eventos de cualquier índole que suceden en nuestra vida tienen múltiples causas, aunque veamos solo un resultado. Por esto se diseñan los estudios analíticos epidemiológicos para evaluar la asociación entre una o varias exposiciones y el riesgo de obtener un desenlace específico en un sujeto determinado (1). Aunque los estudios epidemiológicos se centren en una exposición de interés, por lo general se deben considerar otros factores en el análisis ya que pueden distorsionar o variar la asociación encontrada entre la exposición y el desenlace.

Buscando minimizar este problema, los estudios epidemiológicos utilizan herramientas analíticas como el análisis estratificado y el análisis multivariable para encontrar y controlar las variables que puedan confundir los resultados (variables de confusión), las variables que no dejan ver la asociación entre el factor de exposición y el desenlace (variables supresoras), y aquellas variables que puedan cambiar el impacto de la exposición en el desenlace (interacción) (1). De dichas herramientas, el análisis multivariable ofrece una mayor ventaja analítica ya que permite estudiar de manera simultánea diferentes variables independientes para evaluar el grado de asociación que tienen estas con el desenlace (2).

El presente artículo busca brindar al lector una idea clara y concisa de la definición de análisis multivariable, sus características y sus usos, además de proporcionar una orientación en la elección de la técnica estadística requerida de acuerdo con la naturaleza de la variable de desenlace. En una segunda entrega profundizaremos en aspectos relacionados con el análisis multivariable y la interpretación de sus resultados.

## GENERALIDADES

En los diseños experimentales típicos se busca una posible relación causal entre la variable independiente y el desenlace por medio de la asignación aleatoria de una intervención (3). La asignación aleatoria le permite al investigador organizar la población de estudio en grupos que

---

<sup>1</sup> Estudiante de Medicina, Universidad de Antioquia, Medellín, Colombia.

<sup>2</sup> Profesor Titular, Grupo Académico de Epidemiología Clínica (GRAEPIC), Departamento de Medicina Interna, Facultad de Medicina, Universidad de Antioquia, e Investigador, Unidad de Investigaciones, Hospital Pablo Tobón Uribe, Medellín, Colombia.

Correspondencia: fabian.jaimes@udea.edu.co

Recibido: abril 28 de 2014

Aceptado: abril 29 de 2014

compartirán de manera similar variables independientes, que pueden o no afectar a la variable dependiente del estudio. En los estudios observacionales, o no experimentales, no es posible hacer una asignación aleatoria de la muestra, por lo que se hace necesario contar con un mecanismo adicional que permita el control de las variables de confusión o covariables con respecto a sus efectos en la variable dependiente. Es en este punto donde el análisis multivariable juega un papel fundamental actuando como mecanismo de control mediante el análisis simultáneo de las variables independientes; de este modo se logra una mayor economía en el análisis de los datos y se tiene el potencial para hacer predicciones y obtener una mayor consistencia en la inferencia estadística (4).

Para comprender mejor cómo funciona el análisis multivariable utilizaremos el siguiente ejemplo. En 1995, Hasdai y colaboradores hicieron un estudio para conocer los efectos producidos por el hábito de fumar en pacientes sometidos a revascularización coronaria percutánea. Para ello, siguieron una cohorte de 5.437 pacientes que fueron agrupados en no fumadores, exfumadores (quienes dejaron de fumar al menos 6 meses antes de la intervención), exfumadores recientes (quienes dejaron de fumar inmediatamente luego de la intervención) y fumadores persistentes (5). El desenlace principal para el que se buscaba medir la asociación era el riesgo de muerte por cualquier causa. En la tabla 1 se muestran los resultados obtenidos en el análisis univariado de la investigación.

**Tabla 1. Análisis univariado para riesgo relativo (RR) de muerte de acuerdo con el hábito de fumar**

Evento	No fumadores	Exfumadores	Exfumadores recientes	Fumadores persistentes
Muerte por cualquier causa (RR; intervalo del confianza del 95%)	1,0 (referencia)	1,08 IC 95% = 0,92-1,96	0,56 IC 95% = 0,40-0,77	0,74 IC 95% = 0,59-0,94

Adaptación de Hasdai D. Effect of smoking status on the long-term outcome after successful percutaneous coronary revascularization. N Engl J Med 1997; 336: 755-61.

Los resultados del análisis univariado indican que el riesgo relativo de morir por cualquier causa después de haberse sometido a una revascularización coronaria percutánea es menor en los fumadores persistentes y en los exfumadores recientes que en los no fumadores. ¿Cómo puede ser esto posible luego de que muchos estudios han demostrado la asociación entre el hábito de fumar y una alta mortalidad? La respuesta a este interrogante se puede encontrar en la tabla 2.

Observando detalladamente la tabla 2, encontramos varias diferencias significativas en las características clínicas y angiográficas de los cuatro grupos al empezar el estudio. En comparación con los fumadores persistentes y los exfumadores recientes, los no fumadores eran mayores, habían tenido anginas por períodos más largos, la EAC (enfermedad arterial coronaria) era de mayor extensión, tenían una mayor frecuencia de cirugía de revascularización previa y menor frecuencia de revascularización completa. Además, las principales condiciones coexistentes tales como la

hipertensión arterial y la diabetes mellitus fueron casi dos veces más comunes en los no fumadores que en los exfumadores recientes y los persistentes. Teniendo todas estas características en cuenta, no es raro encontrar que los dos últimos grupos tengan un menor riesgo relativo de morir por cualquier causa que los no fumadores y los exfumadores.

Al ver que las características basales de los cuatro grupos eran diferentes entre sí y que estas diferencias podían tener no solo la evidente relación con el factor de riesgo sino también afectar al desenlace de interés, los investigadores utilizaron el análisis multivariable para ajustar por todas aquellas variables independientes que actuaban como variables de confusión en el estudio: edad, sexo, cirugía de revascularización previa, infarto agudo de miocardio, insuficiencia cardíaca congestiva, historia de diabetes, historia de hipertensión, enfermedad de múltiples vasos, historia familiar de enfermedad coronaria y números de vasos dilatados. El resultado se observa en la tabla 3.

**Tabla 2. Características de base de la población en estudio**

Característica	No fumadores (n = 2.009)	Exfumadores (n = 2.259)	Exfumadores recientes (n = 435)	Fumadores persistentes (n = 734)
Edad promedio en años	67	65	56	55
Duración de la angina en meses	4 ± 15	4 ± 17	2 ± 8	3 ± 7
DM - N° de pacientes (%)	424 (21)	402 (18)	36 (8)	77 (10)
HTA - N° de pacientes (%)	1.082 (54)	1.095 (48)	165 (38)	286 (39)
Hipercolesterolemia - N° de pacientes (%)	810 (40)	877 (39)	142 (33)	282 (38)
Cirugía de revascularización previa - N° de pacientes (%)	351 (17)	514 (23)	45 (10)	89 (12)
Extensión de la EAC				
1 vaso	1.008 (50)	1.143 (51)	246 (57)	406 (55)
2 vasos	725 (36)	815 (36)	147 (34)	261 (36)
3 vasos	276 (14)	301 (13)	42 (10)	67 (9)
Revascularización completa - N° de pacientes (%)	1.154 (57)	1.299 (58)	280 (64)	441 (60)

Adaptación de Hasdai D. Effect of smoking status on the long-term outcome after successful percutaneous coronary revascularization. N Engl J Med 1997; 336: 755-61. DM: diabetes mellitus. HTA: hipertensión arterial. EAC: enfermedad arterial coronaria.

**Tabla 3. Resultados del análisis multivariable ajustado para riesgo relativo (RR) de muerte de acuerdo con el hábito de fumar**

Evento	No fumadores	Exfumadores	Exfumadores recientes	Fumadores persistentes
Muerte por cualquier causa (RR; intervalo del confianza del 95%)	1,0 (referencia)	1,34 IC 95% = 1,14 - 1,57	1,21 IC 95% = 0,87 - 1,70	1,76 IC 95% = 1,37 - 2,26

Adaptación de Hasdai D. Effect of smoking status on the long-term outcome after successful percutaneous coronary revascularization. N Engl J Med 1997; 336: 755-61.

Luego del ajuste, la asociación entre el hábito de fumar y el riesgo relativo de muerte por cualquier causa cambió rotundamente. El análisis multivariable permitió hacer un ajuste simultáneo, sin necesidad de estratificar la población de estudio, de todas las características que actuaban como variables de confusión.

## USOS COMUNES DEL ANÁLISIS MULTIVARIABLE

La investigación clínica utiliza los modelos de análisis multivariable con más frecuencia en los estudios observacionales acerca de la etiología y el pronóstico de una enfermedad, y ocasionalmente en los estudios de diagnóstico o de intervención. Dependiendo del tipo de diseño seleccionado y de la pregunta de

investigación que se busca responder, el análisis multivariable presenta ventajas y limitaciones con respecto al análisis de los datos.

En los estudios etiológicos no experimentales, el análisis multivariable permite fortalecer el argumento de la causalidad entre un factor de riesgo y un desenlace específico mediante la exclusión de las variables de confusión. Sin embargo, presenta una limitación para este ajuste ya que no pueden considerarse las variables no medidas o las variables desconocidas (2). En el contexto de una entidad clínica ya establecida, algunas de las preguntas más difíciles que enfrentan los médicos diariamente son las relacionadas con la gravedad de la enfermedad y el riesgo de un desenlace

desfavorable en el tiempo. A pesar de la dificultad para predecir la respuesta exacta para una sola persona, el análisis multivariable puede proporcionar información confiable acerca del pronóstico en grupos de pacientes con un conjunto particular de factores conocidos. En este tipo de estudios, el análisis multivariable ayuda a demostrar que la supervivencia o el curso clínico de un grupo de individuos con una misma enfermedad pueden variar dependiendo de las características iniciales de la población de estudio y ajustando por cada una de esas características que influyen en el desenlace. En los estudios de diagnóstico, por otra parte, el análisis multivariable busca definir la mejor combinación entre signos, síntomas u otras características del paciente y el diagnóstico de interés, sin importar la relación causal que pueda existir entre todos estos factores (2).

Cuando se habla de estudios de intervención se piensa principalmente en la asignación aleatoria, y es esta la que permite hacer el ajuste de las variables no medidas que el análisis multivariable no puede ajustar. Sin embargo, en algunas preguntas de intervención no es posible hacer la asignación aleatoria por limitaciones logísticas, problemas éticos o por diversas razones. En estos casos el análisis multivariable puede realizar, aunque con ciertas restricciones, el ajuste de las características basales de los grupos para lograr una aproximación estadística razonable a la estimación del efecto de la intervención.

## MÉTODOS DE ANÁLISIS MULTIVARIABLE

La clase de variable dependiente o de desenlace es lo que dicta, con el cumplimiento de ciertos requerimientos o supuestos estadísticos, el tipo de análisis multivariable indicado en cada pregunta o problema de investigación. Los principales métodos de análisis multivariable utilizados en la investigación clínica son la regresión lineal múltiple, la regresión logística con sus adaptaciones y la regresión proporcional de Cox.

### Regresión lineal múltiple

El concepto de regresión hace referencia a la ley experimental o fórmula matemática que traduce la relación entre variables correlacionadas (6). Cuando se quiere poner una variable en función de otra o de otras se acude al bien conocido recurso de la regresión

lineal. Esta, como su nombre lo indica, determina la mejor línea para la predicción de la variable de resultado basada en los valores de una o más variables independientes (2). El primer caso, solo una variable independiente, corresponde a la regresión lineal simple; mientras que la regresión lineal múltiple expresa el efecto de más de una variable independiente. En este modelo la variable de resultado es continua o de intervalo, en la que todos los valores representan magnitudes y la distancia entre los números de su escala es igual. La presión arterial, el peso y la temperatura son ejemplos de este tipo de variable, ya que un cambio en una unidad en cualquier punto de la escala es igual a un milímetro de mercurio, un kilogramo o un grado centígrado, respectivamente.

En la regresión lineal asumimos que existe una relación en forma de línea recta entre la variable dependiente, que se encuentra convencionalmente en el eje de la Y, y la variable independiente representada en el eje de la X (7). Cuando la variable independiente aumenta o disminuye, la media o el valor esperado de la variable de desenlace aumenta o disminuye, respectivamente, de una manera lineal. Analicemos el siguiente ejemplo:

En 2013 Akande y colaboradores publicaron una investigación en la que buscaban encontrar la asociación entre la resistencia a la insulina y la hipertensión esencial (8). Tomaron una muestra de 70 pacientes de los cuales 35 eran normotensos y 35 presentaban hipertensión esencial. Ningún paciente tenía historia de diabetes, intolerancia a la glucosa o resistencia a la insulina. A todos se les midió la presión arterial y se les tomó una muestra de sangre para medir insulina y glucosa plasmáticas en ayunas, con el fin de evaluar la resistencia a la insulina mediante el método HOMA-IR (*homeostasis model assessment-insulin resistance*). Además, se tomaron medidas antropométricas como peso, estatura, índice de masa corporal (IMC) y circunferencia de la cadera. Los resultados se muestran en la tabla 4.

Se observa que el HOMA-IR fue mayor en los pacientes hipertensos que en los normotensos. Sin embargo, aunque algunos parámetros muy relacionados con la hipertensión y la resistencia a la insulina como el IMC y la circunferencia de cadera varían poco entre ambos grupos, la edad muestra diferencias estadísticamente significativas en esa misma comparación. Por lo tanto,

es necesario establecer con mayor certeza la forma en la que estos parámetros pueden modificar el resultado del HOMA-IR. Para esto, los investigadores decidieron utilizar el método de regresión lineal múltiple con los

valores de HOMA-IR como variable dependiente o de resultado, y la edad, el IMC, la circunferencia de cadera y la presencia de HTA como variables independientes. Los resultados del ajuste se encuentran en la tabla 5.

**Tabla 4. Características clínicas y metabólicas de la población de estudio**

Parámetro	Hipertensos (n = 35) Media	Controles (n = 35) Media	P
Edad	55,8 ± 10,3	50,5 ± 9,1	0,026
PAS	142,5 ± 16,2	115,8 ± 11,5	0,001
PAD	94,5 ± 10,7	76,4 ± 7,3	0,001
IMC	27,4 ± 5,8	26,8 ± 5,3	0,69
Circunferencia de cadera	96,3 ± 13,7	90,0 ± 11,8	0,26
GPA	93,1 ± 19,9	89,9 ± 13,3	0,43
IPA	15,5 ± 5,3	12,9 ± 3,4	0,02
HOMA-IR	3,6 ± 1,9	2,9 ± 0,9	0,04

Adaptación de Akande TO, Adeleye JO, Kadiri S. Insulin resistance in Nigerians with essential hypertension. Afr Health Sci 2013; 13(3): 655-660. PAS: presión arterial sistólica. PAD: presión arterial diastólica. IMC: índice de masa corporal. GPA: glucosa plasmática en ayunas. IPA: insulina plasmática en ayunas.

**Tabla 5. Regresión lineal múltiple para HOMA-IR ajustando por presencia de hipertensión y otros factores potencialmente asociados con la resistencia a la insulina**

	Coefficiente de regresión	IC 95%	P
Edad	0,007	-0,034; 0,047	0,74
HTA	0,736	0,038; 1,434	0,04
IMC	0,115	0,052; 0,177	0,001
Circunferencia de cadera	0,023	-0,015; 0,062	0,23

Adaptación de Akande TO, Adeleye JO, Kadiri S. Insulin resistance in Nigerians with essential hypertension. Afr Health Sci 2013; 13(3): 655-660. HTA: hipertensión arterial. IMC: índice de masa corporal.

Luego del ajuste, los factores asociados con la resistencia a la insulina medida por HOMA-IR son la hipertensión arterial y el IMC. En pacientes con diagnóstico de hipertensión esencial se espera, en promedio y luego de ajustar por edad, IMC y circunferencia de cadera, un valor mayor en el HOMA-IR de 0,736 (IC 95% = 0,038; 1,434) con respecto a los sujetos normotensos. Nótese que esta diferencia es muy similar al valor de 0,7 de la diferencia de medias de la tabla 4. Adicionalmente, luego de ajustar por edad, diagnóstico de hipertensión y circunferencia de cadera, se observa

que por cada unidad que aumente el IMC aumenta en 0,15 (IC 95% = 0,052; 0,177) el HOMA-IR.

### Regresión logística

Cuando la variable de resultado es dicotómica, es decir, que presenta únicamente dos valores discretos en un punto determinado en el tiempo (como estar vivo o muerto a los 28 días de hospitalización, o tener o no un infarto en el primer año posterior a una revascularización), se debe usar el modelo de regresión

logística (2), mediante el cual se pretende encontrar la probabilidad de que ocurra el desenlace en estudio con la influencia o no de otras variables. Por lo tanto, la regresión logística consiste en obtener una función matemática, con base en las variables independientes, que permita clasificar a los individuos en una de las dos subpoblaciones o grupos establecidos por los dos valores de la variable de resultado o dependiente (6).

La función logística es aquella que determina, para cada individuo y según los valores de una serie de variables, la probabilidad de que presente el resultado estudiado. Para realizar la función logística es necesario que el desenlace dicotómico tome un valor numérico como 0 y 1 para representar su "ausencia" o "presencia", respectivamente, y así los datos puedan ser analizados en paquetes estadísticos (6).

De manera similar a lo que se pretende en la regresión lineal, en la función logística también se busca determinar "la mejor línea" para la predicción de la variable de resultado basada en los valores de una o más variables independientes. Sin embargo, dado que el desenlace en este caso solo tiene las opciones de 0 y 1, la ecuación de partida de los modelos de regresión logística presenta una distribución logística o sigmoidea. Por lo tanto, para poder conocer la ecuación que mejor representa los datos observados y poder interpretar sus resultados, la regresión logística necesita hacer dos transformaciones sucesivas:

1. La probabilidad de ocurrencia del evento ( $p = 1/n$ ) se transforma en un *odds* o razón ( $o = p/1-p$ ), y
2. El *odds* de ocurrencia del evento se convierte en el logaritmo de base natural de ese *odds* ( $\text{logit} = \log[\text{odds}]$ ). Esta transformación permite que el desenlace de la ecuación sea una variable continua y por lo tanto hace que los resultados pasen de tener una distribución sigmoidea a una distribución lineal. Similar a la regresión lineal, en donde cada coeficiente de regresión nos dice el cambio que se espera en la variable de resultado de acuerdo con cada unidad de cambio en la variable independiente, en la regresión logística cada coeficiente se presenta como un *odds ratio* (OR, riesgo relativo indirecto) que nos dice el riesgo de ocurrencia, o de no ocurrencia cuando el OR es menor de 1, de ese desenlace dicotómico de acuerdo con el respectivo cambio en la variable independiente.

La regresión logística es el modelo de análisis multivariable más utilizado en la literatura médica y lo analizaremos con el siguiente ejemplo: en 2013 De la Rosa y colaboradores publicaron los resultados de un estudio hecho en Medellín para determinar la asociación entre los niveles de insulina de pacientes admitidos a la unidad de cuidados intensivos (UCI) y la mortalidad hospitalaria. Se tomó una cohorte de 160 pacientes mayores de 15 años, que no habían recibido insulina o glucosa antes de la admisión a la UCI. Dado que el desenlace es la variable dicotómica mortalidad hospitalaria (vivo o muerto al alta del hospital), la regresión logística es el modelo de análisis indicado para establecer la asociación potencial entre la insulinemia de cada paciente y el riesgo de muerte, ajustando por las potenciales variables de confusión.

Los pacientes que murieron eran de mayor edad y presentaban una puntuación APACHE II más alta y mayor frecuencia de diagnóstico de sepsis. Asimismo, se presentó una mayor mortalidad en los pacientes con valores de insulina por debajo de 5 U/mL (37%) y por encima de 15 U/mL (51%), comparados con los que tenían valores de 5 a 15 U/mL (21%) (9). Los resultados de la regresión logística simple o análisis univariado, es decir, teniendo en cuenta solo una de las variables independientes cada vez, y los de la regresión logística múltiple o análisis multivariable se encuentran en la tabla 6.

La regresión logística simple nos indica que los pacientes con valores bajos ( $< 5$  U/mL) y altos ( $> 15$  U/mL) de insulina, tienen un mayor riesgo de muerte comparados con los que presentan dichos valores en el rango de 5 a 15 U/mL (*odds ratio* [OR] = 2,14; IC 95% = 0,93-4,95 y OR = 3,89; IC 95% = 1,6-9,44, respectivamente). Nótese, sin embargo, que los valores bajos de insulina no alcanzan significado estadístico en dicha asociación porque el límite inferior del intervalo de confianza (0,93) de ese OR está por debajo del punto de no efecto (1). En la regresión logística múltiple se confirma, luego de ajustar por el tipo de tratamiento para control glucémico, la edad, el antecedente de diabetes, los valores de SOFA y APACHE II al ingresar a la UCI y el diagnóstico de sepsis, que los pacientes admitidos con valores de insulinemia mayores de 15 U/mL tienen un riesgo 3,58 veces mayor de morir en el hospital que aquellos con valores de 5 a 15 U/mL. Además, se observa que los pacientes admitidos con

diagnóstico de sepsis tienen un riesgo 2,42 veces mayor de morir que los que ingresan con otros diagnósticos, con cualquier valor que tengan de insulinemia,

SOFA, APACHE, edad e independientemente del tipo de control glicémico que reciban y del antecedente de diabetes.

**Tabla 6. Resultados de la regresión logística simple y múltiple para la mortalidad hospitalaria en pacientes admitidos a la unidad de cuidado intensivo**

Variable	Regresión logística simple OR (95% CI)	Regresión logística múltiple OR (95% CI)	P
Niveles de insulina (U/mL)			
5-15	1 (referencia)	1 (referencia)	
< 5	2,14 (0,93-4,95)	1,68 (0,62-4,54)	0,233
>15	3,89 (1,60-9,44)	3,58 (1,18-10,84)	0,016
Tratamiento intensivo	0,82 (0,43-1,56)	1,00 (0,99-1,02)	0,492
Edad	1,01 (1,00-1,03)	1,00 (0,99-1,02)	0,492
Antecedente de diabetes	0,76 (0,19-3,07)	0,23 (0,04-1,28)	0,317
APACHE II (cada punto)	1,07 (1,02-1,13)	1,05 (0,99-1,11)	0,651
SOFA ( <i>sequential organ failure assessment</i> ) (cada punto)	1,01 (0,99-1,02)	1,01 (0,98-1,02)	0,773
Diagnóstico de sepsis	2,61 (1,26-5,37)	2,42 (1,03-5,65)	0,020

Adaptación de De La Rosa et al. The potencial impact of admission insuline levels on patient outcome in the intensive care unit. J Trauma Acute Care Surg 2013; 74 (1): 270-5.

## Regresión de Cox (Modelo de riesgos proporcionales)

Los llamados estudios de supervivencia, a diferencia de otros diseños, no solo buscan estimar el riesgo de tener un desenlace específico, sino que también incorporan el concepto del tiempo. En estos estudios la variable de resultado es el tiempo transcurrido hasta un determinado evento. El método más utilizado para saber el efecto de una variable independiente sobre la supervivencia, o sobre la probabilidad de estar libre de la ocurrencia de un evento, en una población después de un tiempo determinado es el de Kaplan-Meier. Este método estima la función de supervivencia de uno o varios grupos y la representa gráficamente. También es posible, mediante una comparación por medio de la prueba de Log-Rank, establecer si existen diferencias estadísticamente significativas entre distribuciones de supervivencia (10).

Sin embargo, el método de Kaplan-Meier no permite valorar simultáneamente más de una variable

independiente y tampoco permite estimar la magnitud del efecto que pueda tener una variable en el riesgo de ocurrencia de un evento en el tiempo. Es aquí donde la regresión de Cox juega un papel fundamental como modelo de análisis multivariable en los estudios de supervivencia. La regresión de Cox valora el efecto de distintas variables independientes sobre la función de supervivencia a lo largo de todo el período de observación de los pacientes, sea cual sea el punto temporal que se elija para la comparación (2). Además, la principal ventaja de este tipo de análisis es que permite incorporar al estudio individuos con distintos tiempos de seguimiento, porque es posible obtener información relevante de aquellos sujetos que durante el tiempo de seguimiento definido no presentaron el desenlace de interés. Estos individuos, que dejaron de observarse a partir de un tiempo establecido previamente, se denominan "censurados" porque no es posible conocer su tiempo real de supervivencia. Esta censura asume que si los sujetos pudiesen haber sido seguidos más allá de la fecha de la última observación, todos tendrían en algún

momento posterior el mismo resultado de aquellos con el desenlace observado (10). Al igual que en la regresión logística, en la regresión de Cox también es necesaria una función matemática que transforme el desenlace y permita entender la relación entre las variables por medio de un modelo similar al de la línea recta. En este caso, el tiempo hasta el evento es transformado en una función denominada el *Hazard*, que puede simplificarse en términos prácticos como una medida del riesgo instantáneo de ocurrencia de dicho evento. En la regresión de Cox, por lo tanto, el efecto de las variables independientes se presenta como un *Hazard Relativo* (HR), y expresa la magnitud en la que una variable aumenta o disminuye el riesgo de ocurrencia de un desenlace en el tiempo. En resumen, la regresión de Cox permite establecer cuál o cuáles factores se asocian con un aumento o disminución

de la “velocidad” con que se presenta el desenlace de interés (10).

Recientemente, Uchikawa y colaboradores publicaron un estudio que exploraba el papel de los niveles de albúmina sérica en la predicción de la mortalidad y la incidencia de eventos cardiovasculares en pacientes con falla renal crónica (FRC) sometidos a terapia de resincronización cardíaca (TRC) (11). La media de seguimiento para estos pacientes fue de  $2,6 \pm 2,06$  años y la muestra fue dividida en dos grupos: pacientes con niveles de albúmina normales y pacientes con hipoalbuminemia. Con el método de Kaplan-Meier se observó que los pacientes con albúmina normal tenían menor mortalidad y menor frecuencia de eventos cardiovasculares que aquellos con hipoalbuminemia. El análisis uni y multivariable para mortalidad con la regresión de Cox se muestra en la tabla 7.

**Tabla 7. Factores relacionados con la mortalidad en pacientes con falla renal crónica sometidos a resincronización cardíaca**

	Análisis univariado		Análisis multivariado	
	HR (IC 95%)	P	HR (IC 95%)	P
Albúmina normal	0,33 (0,12-0,87)	<b>0,024</b>	0,32 (0,11-0,98)	<b>0,048</b>
Edad (por cada año)	0,99 (0,97-1,02)	0,701		
ECl	1,21 (0,59-2,45)	0,607		
TFG (por cada mL/min)	0,98 (0,96-1,01)	0,139		
PCR (por cada mg/mL)	1,1 (0,98-1,25)	0,111		
Log BNP (por cada unidad)	2,12 (0,84-5,38)	0,114		
Log PNA (por cada unidad)	4,11 (1,42-11,96)	0,009	2,6 (0,83-8,13)	0,100
FEVI (por cada punto porcentual)	0,99 (0,94-1,05)	0,310		

Adaptación de Uchikawa T. Serum albumin levels predict clinical outcomes in chronic kidney disease (CKD) patients undergoing cardiac resynchronization therapy. Intern Med 53: 555-561, 2014 DOI: 10.2169/internalmedicine.53.1209. ECl: enfermedad cardíaca isquémica. TFG: tasa de filtración glomerular. PCR: proteína C reactiva. BNP: péptido natriurético cerebral. PNA péptido natriurético atrial. FEVI: fracción de eyección del ventrículo izquierdo.

En el análisis univariable únicamente los valores de albúmina y de péptido natriurético atrial (PNA, en transformación logarítmica) se asocian con el riesgo de muerte. Luego del análisis multivariable con estos dos factores, se observa que la albuminemia es la única variable asociada con mortalidad. En pacientes con FRC sometidos a TRC tener valores normales de albúmina en suero, en comparación con la presencia de hipoalbuminemia, se comporta como un factor

protector al disminuir el riesgo instantáneo de muerte en un 68%. Esta disminución está presente al ajustar por, o independientemente de, los valores de PNA.

## CONCLUSIONES

El análisis multivariable resulta ser la herramienta más eficiente y de mayor confiabilidad para analizar simultáneamente el comportamiento de diferentes



variables independientes sobre un desenlace. Conocer sus usos y limitaciones aumenta las destrezas del médico para analizar e interpretar de forma apropiada los resultados de estudios que día a día sugieren nuevas estrategias en la práctica clínica.

## FINANCIACIÓN

Trabajo apoyado parcialmente por la Estrategia de Sostenibilidad 2013-2104 de la Universidad de Antioquia.

## REFERENCIAS BIBLIOGRÁFICAS

1. Szklo M, Nieto J. *Epidemiology: beyond the basics*. 2nd ed. Sudbury (Massachusetts): Jones and Bartlett Publishers; 2007.
2. Katz MH. *Multivariable analysis: a practical guide for clinicians and public health researchers*. 3rd ed. Cambridge: Cambridge University Press; 2011.
3. Hernández-Avila M, Garrido-Latorre F, López-Moreno S. Diseño de estudios epidemiológicos. *Salud Publica Mex*. 2000;42(2):144–54.
4. Lozares Colina C, López-Roldán P. El análisis multivariado: definición, criterios y clasificación. *Revista de Sociología*. 1991;(37):9–30.
5. Hasdai D, Garratt KN, Grill DE, Lerman A, Holmes DR. Effect of smoking status on the long-term outcome after successful percutaneous coronary revascularization. *N Engl J Med*. 1997 Mar 13;336(11):755–61.
6. Fiuza Pérez MD, Rodríguez Pérez JC. La regresión logística: una herramienta versátil. *Nefrología*. 2000;20(6):495–500.
7. Daniel WW, Croos CL. *Biostatistics: a foundation for analysis in the health sciences*. 9th ed. Hoboken (Nueva Jersey): John Wiley & Sons; 2008.
8. Akande TO, Adeleye JO, Kadiri S. Insulin resistance in Nigerians with essential hypertension. *Afr Health Sci*. 2013 Sep;13(3):655–60.
9. De La Rosa G, Vasquez EM, Quintero AM, Donado JH, Bedoya M, Restrepo AH, et al. The potential impact of admission insulin levels on patient outcome in the intensive care unit. *J Trauma Acute Care Surg*. 2013 Jan;74(1):270–5.
10. Martínez M, De Irala J. Análisis de supervivencia y análisis multivariado. In: López Jiménez F, editor. *Manual de medicina basada en la evidencia*. 2nd ed. Ciudad de México: JGH Editores; 2001. p. 213–229.
11. Uchikawa T, Shimano M, Inden Y, Murohara T. Serum albumin levels predict clinical outcomes in chronic kidney disease (CKD) patients undergoing cardiac resynchronization therapy. *Intern Med*. 2014 Jan;53(6):555–61.

