

## Factores relacionados a las enfermedades cardiovasculares en personas adultas, atendidas en el hospital PNP. Augusto B. Leguía - 2017

*Fausto Francisco Matos Uribe*<sup>1</sup>

*Emma Norma Cambillo Moyano*<sup>2</sup>

**Resumen:** Determinar factores relacionados a las enfermedades cardiovasculares, en personas adultas, atendidas en el Hospital PNP Augusto B. Leguía. Rímac-2017. Material y métodos: El estudio fue observacional, retrospectivo y analítico, el diseño muestral es aleatorio simple, se tomó una muestra al azar de 1284 pacientes e historias clínicas entre los meses de octubre a diciembre 2017, se usó la regresión logística binaria múltiple para determinar las variables (predictores) que se relacionan con las enfermedades cardiovasculares, variable dependiente cualitativa binaria que toman valores 1=tiene la enfermedad cardiovascular, 0=no tiene la enfermedad. Resultados: de un total de 12 predictores (clasificados en factores): Género, edad, raza, índice de masa corporal, alcoholismo, tabaquismo, actividad física, nivel colesterol, nivel de triglicéridos, presión sistólica, presión diastólica, diabetes; resultaron significativo al 5% para predecir la enfermedad cardiovascular los predictores: Edad, alcoholismo, colesterol, triglicéridos, presión arterial sistólica y presión arterial diastólica; la prueba de bondad de ajuste de Hosmer-Lemeshow resulto no significativo ( $p > 0.05$ ), lo cual indica que existe un buen ajuste del modelo logístico a las observaciones; la capacidad predictiva del modelo estimado para clasificar a las observaciones correctamente fue del 85.6%; el área bajo la curva ROC, es 0.92, con Intervalo de confianza 95% (0.905-0.935), lo cual indica que el modelo tiene poder discriminante para clasificar correctamente sanos versus enfermos con una probabilidad del 92%. Conclusión: La regresión logística binario, demostró ser un buen método para determinar factores que se relacionan con las enfermedades cardiovasculares y además realizar predicciones.

**Palabras claves:** Enfermedades cardiovasculares, factores demográficos, estilo de vida, clínicos y regresión logística binaria..

<sup>1</sup>Universidad Nacional Mayor de San Marcos, [fmatosu@unmsm.edu.pe](mailto:fmatosu@unmsm.edu.pe)

<sup>2</sup>Universidad Nacional Mayor de San Marcos, [ecambillom@unmsm.edu.pe](mailto:ecambillom@unmsm.edu.pe)

**Factors related to cardiovascular diseases in adult people served at the PNP.  
Hospital Augusto B. Leguía - 2017**

**Abstract:** Determine factors related to cardiovascular diseases in adults treated at the PNP Augusto B. Leguía Hospital. Rimac-2017. Material and methods: The study was observational, retrospective and analytical, the sample design is simple random, a random sample of 1284 patients and medical records was taken between the months of October to December 2017, multiple binary logistic regression was used to determine the variables (predictors) that are related to cardiovascular diseases, binary qualitative dependent variable that take values 1=has cardiovascular disease, 0=does not have the disease. Results: from a total of 12 predictors (classified into factors): gender, age, race, body mass index, alcoholism, smoking, physical activity, cholesterol level, triglyceride level, systolic pressure, diastolic pressure, diabetes; The predictors were significant at 5% to predict cardiovascular disease: Age, alcoholism, cholesterol, triglycerides, systolic blood pressure and diastolic blood pressure; the Hosmer-Lemeshow goodness-of-fit test was not significant ( $p > 0.05$ ), which indicates that there is a good fit of the logistic model to the observations; the predictive capacity of the estimated model to classify the observations correctly was 85.6%; the area under the ROC curve is 0.92, with a 95% confidence interval (0.905-0.935), which indicates that the model has discriminant power to correctly classify healthy versus sick with a probability of 92%. Conclusion: The binary logistic regression proved to be a good method to determine factors that are related to cardiovascular diseases and also make predictions.

**Keywords:** Cardiovascular diseases, demographic factors, lifestyle, clinical and binary logistic regression.

*Recibido 12/03/2023*

*Aceptado 23/06/2023*

*Publicado online 30/06/2023*

## 1. Introducción

La Organización Mundial de la Salud [28], refiere que: Las enfermedades cardiovasculares (ECV) son la principal causa de muerte en todo el mundo. Desde el año 2000, las muertes se han incrementado en lo que respecta a la cardiopatía isquémica, incrementándose de más de 2 millones de defunciones en 2000 a 8.9 millones en 2019, siendo esta enfermedad, responsable del 16 % del total de muertes en el mundo, y en segundo lugar de causas de muertes es accidente cerebrovascular que representa el 11 % del total de muertes.

La enfermedad cardiovascular es un concepto amplio para problemas con el corazón y los vasos sanguíneos, estos problemas a menudo se deben a la aterosclerosis, ocurre cuando la grasa y el colesterol se acumulan en las paredes de las arterias, esta acumulación grasa se llama placa, con el correr del tiempo estrecha los vasos sanguíneos y causa problemas en todo el cuerpo, si una arteria resulta obstruida, esto puede llevar a que se presente un ataque cardíaco o un accidente cerebrovascular [22].

En América Latina y el Caribe, la hipertensión es una condición muy común que conduce tanto a la enfermedad cardíaca como al accidente cerebrovascular, y según datos de la red de científicos de la salud de todo el mundo que proporciona datos rigurosos y oportunos sobre los factores de riesgo para enfermedades no transmisibles, indicó que 28 % de las mujeres y el 43 % de los varones desconocen su condición de hipertensos [29].

En el caso del Perú según [19], las enfermedades isquémicas del corazón y las enfermedades cerebrovasculares se constituyen como segunda y tercera causa de mortalidad en el adulto mayor respectivamente; además, presentan importantes diferencias por edad, sexo, nivel de educación, quintiles de riqueza entre otras. Entre los principales factores de riesgo que se presentan en este estudio destacan el colesterol alto, triglicérido alto, hipertensión arterial, enfermedades del corazón y diabetes. La cuantificación de estos factores constituye un aspecto esencial para comprender la dimensión de este problema a escala poblacional y para una planificación de los recursos sanitarios.

El objetivo del presente trabajo de investigación es determinar factores relacionados a las enfermedades cardiovasculares en personas adultas, atendidas en el hospital PNP. Augusto B. Leguía-Lima-Perú, por antecedentes de otros estudios relacionados a las enfermedades cardiovasculares [3, 4, 5, 7, 8, 10, 17, 20, 30, 32, 33, 34, 38], se tomó en cuenta 12 variables predictoras entre cualitativas y cuantitativas que se relacionan con la enfermedad cardiovascular, tales como: Género, raza, edad, índice de masa corporal, alcoholismo, tabaquismo, actividad física, colesterol, triglicéridos, presión arterial sistólica, presión arterial diastólica y diabetes, para tal propósito se tomó una muestra elegida al azar de 1284 pacientes adultos con edades de 20 años a más y para determinar los factores que se relacionan con la enfermedad, se utilizó el modelo de regresión logístico binario múltiple, la variable en estudio es la enfermedad cardiovascular que toma valores 1=tiene la enfermedad y 0=no tiene la enfermedad, y con la ayuda de programa estadístico SPSS versión 23, se estimó en primer lugar los coeficientes del modelo; para la significación global de los coeficientes, se usó la prueba estadística de razón de verosimilitud, la cual resultó estadísticamente significativa ( $0.05 > p$ ); para la evaluación del impacto de cada predictor, para explicar la probabilidad de ocurrencia de la enfermedad, se utilizó el estadístico Wald [37], resultando estadísticamente significativa al 5 % los siguientes predictores: La edad, el

alcoholismo, colesterol, triglicéridos, presión arterial sistólica y la presión arterial diastólica, lo cual indica que estas variables en forma conjunta se relacionan con la probabilidad de ocurrencia de la enfermedad cardiovascular; la interpretación de los coeficientes estimados, se presentó al modelo estimado en términos Odds ratio [24], resultando los predictores: Edad, colesterol, triglicéridos, presión arterial sistólica y diastólica factores de riesgos, excepto el alcoholismo que resultó un factor protector para la enfermedad; para la bondad de ajuste del modelo estimado se utilizó la prueba de Hosmer-Lemeshow [18], resultó no significativo la prueba estadística ( $0.206 > 0.05$ ) lo cual indica un buen ajuste del modelo estimado a los datos observados; también se calculó la bondad de ajuste del modelo estimado, a través de los coeficientes de determinación [24, 12, 26] de McFadden=48.8 %, Cox-Snell=47.5 % y Nagelkerke=64.8 %, todos los porcentaje son aceptable para este tipo de estudio; la capacidad predictiva del modelo estimado, el modelo estimado también permitió clasificar a los pacientes en cada una de las categorías determinadas a través de la tabla de confusión, para el trabajo el modelo estimado con punto de corte 0.5, logró clasificar correctamente el 85.6 % del total de la muestra (1284), resultado que es satisfactorio; el área bajo la curva ROC, resultó de 0.92 que corresponde a la probabilidad de que el modelo estimado clasificó muy bien a las observaciones. Con respecto a la adecuación del modelo a través del análisis de los residuos, sólo 3.43 % de los datos observados de la muestra superan desviaciones estándar y ninguna de las observación llegan a desviaciones estándar, y con respecto a la presencia de observaciones influyentes ninguna superó a la distancia de Cook  $> 1$ ; la matriz de correlaciones entre las estimaciones de los parámetros, presentan valores bajos, así como también no hubo la presencia de multicolinealidad entre los predictores, el factor de inflación de varianzas de los predictores del modelo estimado son todos menores a 3. La regresión logística binaria múltiple, es una técnica estadística que identificó variables predictoras que se relacionan o influyen en la enfermedad cardiovascular y por las estadísticas obtenidas el modelo estimado se ajusta bien a los datos observados.

## 2. Materiales y Métodos

### 2.1. Métodos

El presente trabajo de investigación fue observacional, retrospectivo y analítico debido a que se estudia factores relacionados a las enfermedades cardiovasculares en personas adultas con edades de 20 años a más, el modelo utilizado para determinar factores que se relacionan con la enfermedad cardiovascular, fue la regresión logística binaria múltiple.

### 2.2. Materiales

La selección de la muestra fue a través de un muestreo aleatorio simple de pacientes del Hospital

PNP Augusto B. Leguía distrito del Rímac, Lima-Perú, los pacientes atendidos son miembros de la Policía Nacional del Perú, en situación de actividad o retiro y familiares directos. El marco de muestreo estuvo conformado por todos los pacientes del hospital. La Selección del tamaño de la muestra para estimar la prevalencia (proporción) de pacientes con la enfermedad cardiovascular en la población, se realiza con un nivel de confianza del 95 % y un margen de error del 2.641 %, resultando el tamaño de la muestra de 1284 pacientes. Para la muestra se incluyen a pacientes con edades mayor o igual a los 20 años y se excluye a menores de 20 años, el periodo de la recolección de los datos se realizó entre los meses de octubre a diciembre del 2017, los datos se recogieron de sus historias clínicas. Las variables predictoras para el estudio de la relación con la enfermedad cardiovascular son: Género, raza, edad, índice de masa corporal, consumo o no del alcohol, hábito de fumar, actividad física, nivel de colesterol, nivel de triglicéridos, presión arterial sistólica, presión arterial diastólica y diabetes. Para el análisis exhaustivo de los datos, se realizó a través de grupos de variables, factor demográfico (género, raza, edad e índice de masa corporal), factor estilo de vida (actividad física, alcoholismo, tabaquismo) y factor clínico (colesterol, triglicéridos, presión arterial sistólica, presión arterial diastólica y diabetes).

### 3. Resultados y discusión

Los datos recopilados de la variable dependiente enfermedad cardiovascular es cualitativa dicotómica que toma valores  $y=1$  tiene la enfermedad, y  $=0$  no tiene la enfermedad, y un conjunto de predictores que por antecedentes de otras investigaciones se relacionan con la enfermedad, clasificados en factores con el propósito de realizar un análisis más exhaustivo en: Demográficos (género, edad, raza, índice de masa corporal), estilo de vida (alcoholismo, tabaquismo, actividad física) y clínico (colesterol, triglicérido, presión arterial sistólica, presión arterial diastólica y diabetes). Los datos fueron organizados y presentados en tablas de distribución de frecuencias univariado y bivariado, así como medidas de tendencia central y dispersión para las variables predictoras cuantitativas y para el análisis bivariado se estudió asociaciones. Con respecto al género 715 (55.7 %) son varones y 569(44.3 %) son mujeres; con respecto a la edad de los pacientes, la edad mínima fue de 21 años y la máxima fue de 93 años, con una edad promedio de 57.45 años y un coeficiente de variación de 25.98 %, lo cual indica que las edades no presentan mucha variabilidad con respecto a la media; colesterol tiene un nivel mínimo de 114 mg/dl y un máximo de 480 mg/dl, con una media de 189.67 mg/dl y coeficiente de variación del 26.1 %, indica que existe una baja variación; triglicéridos, tiene un nivel mínimo de 48 mg/dl y un máximo de 482 mg/dl con valor promedio de 140.49 mg/dl y coeficiente de variación del 55.4 %, indicando que los niveles de triglicéridos presentan mucha variación; para la presión arterial sistólica, la presión mínima es de 80 mmHg y máxima 190 mmHg, la presión promedio fue 115.65 mmHg y coeficiente de variación de 19.0 % valor que indica una baja variación; presión arterial diastólica, la presión mínima de 40 mmHg y máxima 100 mmHg, la presión promedio 64.99 mmHg y un coeficiente de variación de 10.7 % valor que indica una baja variación; para el índice de masa corporal, el mínimo es de 19.57 Kg/m<sup>2</sup>, la máxima 46.48 kg/m<sup>2</sup>, el índice de masa corporal promedio fue de 26.85 kg/m<sup>2</sup> y un coeficiente de variación de 11.9 %, valor que indica una baja variación. Además las correlaciones entre las variables predictoras fue baja correlación, a excepción la variable presión arterial sistólica y diastólica fue de 0.649 . Se detectó que la variable alcoholismo es confusora entre el tabaquismo y la enfermedad cardiovascular [21]. Con respecto al análisis bivariado las variables: Género, razas, alcoholismo, tabaquismo, actividad física y diabetes, cada una de ellas resultó

asociado a la enfermedad cardiovascular, se usó la prueba estadística de Chi-Cuadrado ( $0.05 > p$ ).

Para el análisis multivariado de la variable dependiente enfermedad cardiovascular dicotómica, y las 12 variables predictoras: Género, edad, raza, índice de masa corporal, alcoholismo, tabaquismo, actividad física, colesterol, triglicéridos, presión arterial sistólica, presión arterial diastólica y diabetes, con el objetivo de determinar factores relacionados a las enfermedades cardiovasculares, se utilizó la regresión logística binaria múltiple (1), en primer lugar se estimó los coeficientes de los predictores del modelo logístico, utilizando el método de máxima verosimilitud, que consiste en maximizar la función de verosimilitud (2) derivando parcialmente con respecto a cada coeficiente del modelo e igualando a cero, las ecuaciones de verosimilitud se presenta en (3).

$$E(Y_i) = \pi_i = P(Y_i = 1/x_1, x_2, \dots, x_k) = \frac{e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k}}{1 + e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k}} \quad (1)$$

$$LL(\beta) = \sum_{i=1}^n [y_i(\beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik}) - \ln(1 + e^{\beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik}})] \quad (2)$$

$$\frac{\partial(LL(\beta))}{\partial\beta_j} = \sum_{i=1}^n [y_i x_{ij} - \pi_i x_{ij}] = \sum_{i=1}^n [(y_i - \pi_i) x_{ij}] = 0 \quad i = 1, 2, \dots, n \quad j = 0, 1, 2, \dots, k \quad x_{i0} = 1 \quad \forall i \quad (3)$$

Para calcular la solución  $(\beta_0, \beta_1, \dots, \beta_k)$  de (3) estas ecuaciones (estimación de los parámetros), se utilizó métodos iterativos de Newton Raphson (existen paquetes estadísticos que estiman éstos parámetros), el vector de parámetros estimados  $(\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k)$ , posee la propiedad de eficiencia asintóticas (insesgado, consistencia, normales), después de ajustar el modelo logístico, se realizó el proceso de evaluación de la bondad de ajuste. La prueba de razón de verosimilitud (4) para probar la significación global de los 12 predictores, se basa en la estadística razón de verosimilitud la cual tiene una distribución Chi-cuadrado con grados de libertad igual al número de variables predictoras, bajo la hipótesis nula cierta, la cual permite probar la hipótesis nula de que los coeficientes del modelo logístico son nulos, es decir, las variables predictoras no se relacionan con la probabilidad de ocurrencia de la enfermedad cardiovascular  $P(Y=1)$ .

$$RV = G = -2Ln \left[ \frac{L(\hat{\beta}_0)}{L(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k)} \right] \approx \chi_{(k)gl}^2 / H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0 \quad cierta \quad (4)$$

Para los datos, la prueba resultó significativa al 5% , lo cual indica que por lo menos algún coeficiente no es nulo, para predecir la probabilidad de ocurrencia de la enfermedad cardiovascular; para la significación individual de cada coeficiente se utilizó la prueba estadística de Wald [37], resultando significativas las variables predictoras: Edad, alcoholismo, colesterol, triglicéridos, presión arterial sistólica y presión arterial diastólica; con el OR se interpreta los coeficiente estimados y además las variables predictoras: Edad, colesterol, triglicéridos, presión arterial sistólica

y diastólica, resultó ser de riesgos para la enfermedad; la variable alcoholismo resultó ser una variable protectora para la enfermedad; la prueba de bondad de ajuste de Hosmer-Lemeshow [18], resultó no significativo al 5 % ( $p=0.206>0.05$ ), lo cual indica que las probabilidades estimadas se ajustan bien a los valores de las observaciones; al igual que en la regresión múltiple, también se calculó la bondad de ajuste del modelo estimado, a través de los coeficientes de determinación (5,6,7) de McFadden=48.8 %, Cox-Snell=47.5 % y Nagelkerke=64.8 %, cuyos porcentaje mide la variación de los valores observados en la variable enfermedad cardiovascular explicada por las variables predictoras del modelo estimado, todos estos porcentaje son aceptables para este tipo de estudio Tabla 01.

El  $R^2$  de McFadden [24]

$$R_{McFadden}^2 = 1 - \frac{LL(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k)}{LL(\hat{\beta}_0)} \quad (5)$$

El  $R_{CS}^2$  de Cox & Snell [12]

$$R_{CS}^2 = 1 - \left( \frac{e^{-2LL(\hat{\beta}_0)}}{e^{-2LL(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k)}} \right)^{-\frac{1}{n}} = 1 - e^{-\left(\frac{G}{n}\right)} \quad (6)$$

El  $R_N^2$  de Nagelkerke [26]

$$R_N^2 = \frac{1 - \left( \frac{e^{-2LL(\hat{\beta}_0)}}{e^{-2LL(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k)}} \right)^{-\frac{1}{n}}}{1 - (e^{-2LL(\hat{\beta}_0)})^{-\frac{1}{n}}} = \frac{R_{CS}^2}{1 - e^{-(G+2LL(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k))}} \quad (7)$$

La capacidad predictiva del modelo estimado clasificó correctamente con punto de corte 0.5 al 85.6 % de los pacientes con o sin la enfermedad cardiovascular (Tabla 04); el área bajo la curva ROC, es 0.92 (Figura 01), indica la probabilidad de que el modelo estimado clasificó correctamente a dos pacientes, el sano en sano y al enfermo en enfermo, con un intervalo de confianza (0.905; 0.935) que no contiene al 0.5. Con respecto a la evaluación de la adecuación del modelo estimado, se realizó a través del análisis de los residuos resultó 44 observaciones atípicas (discordantes mayor a desviaciones estándar) que sólo representa el 3.43 % del total de observaciones resultan no significativos para eliminarlos; con respecto a las observaciones influyentes a los parámetros estimados o las predicciones, ninguna observación superó a la distancia de . Además la matriz de correlaciones entre las estimaciones de los parámetros del modelo presentan valores bajos (Tabla 02), así como también no se detectó la presencia de multicolinealidad entre los predictores, el factor de inflación de varianzas de los predictores del modelo estimado son todos menores a 3 (Tabla 03). Conclusión: El modelo logístico binario estimado demostró ser un buen método para determinar factores relacionados con las enfermedades cardiovasculares y por las estadísticas obtenidas, el modelo estimado se ajusta bien a los datos de la muestra. También se hace presente que no se encontró interacciones significativas entre las variables predictoras.

Tabla 01

*Variables relacionados a las enfermedades cardiovasculares, en pacientes adultos atendidos en el hospital PNP Augusto B Leguía-Rímac, modelo regresión logística binario (n=1284)*

Variables	Coefficientes estimados	significación estadística	$OR = e^{\hat{\beta}_j}$	IC(95%)
Edad	0.025	0.000(**)	1.025	1.012-1.039
Alcoholismo	-0.441	0.018(*)	0.644	0.447-0.928
Colesterol	0.012	0.000(**)	1.012	1.009-1.016
Triglicéridos	0.008	0.000(**)	1.008	1.005-1.010
PAS	0.081	0.000(**)	1.085	1.070-1.099
PAD	0.050	0.002(**)	1.051	1.018-1.085
Intercepto	-18.042	0.000(**)		
Bondad de ajuste:				
Estadística G	827.131	0.000(**)		
$-2LL(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3, \hat{\beta}_4, \hat{\beta}_5, \hat{\beta}_6)$	867.099			
Hosmer-Lemeshow	10.930	0.206(ns)		
R <sup>2</sup> McFadden	48.8%			
R <sup>2</sup> Cox y Snell	47.5%			
R <sup>2</sup> Nagelkerke	64.8%			
% global correcto	85.6%			
Curva ROC	0.920	0.000(**)		0.905-0.935

(\*) Significación 5%a

(ns) Prueba no significativa

(\*\*) Significación 1%

Nota: Elaboración propia.



Tabla 02

*Matriz de correlaciones entre las estimaciones de los parámetros.*

<b>Predictores</b>	<b>Const.</b>	<b>Edad</b>	<b>Alcohol(1)</b>	<b>Colesterol</b>	<b>Triglicéridos</b>	<b>PAS</b>	<b>PAD</b>
<b>Const.</b>	1.000	-.342	-.140	-.323	-.076	-.302	-.640
<b>Edad</b>	-.342	1.000	.344	.005	.070	-.089	.033
<b>Alcohol</b>	-.140	.344	1.000	-.083	-.002	-.079	.019
<b>Colesterol</b>	-.323	.005	-.083	1.000	-.272	.081	.001
<b>Triglicéridos</b>	-.076	.070	-.002	-.272	1.000	-.040	.025
<b>PAS</b>	-.302	-.089	-.079	.081	-.040	1.000	-.385
<b>PAD</b>	-.640	.033	.019	.001	.025	-.385	1.000

Nota: Elaboración propia

Tabla 03

*Factor de inflación de varianzas (FIV) de los predictores del modelo estimado(n=1284)*

<b>Predictores</b>	<b>Coefficiente estimado</b>	<b>Error estándar del coeficiente estimado.</b>	<b>FIV</b>
<b>Constante</b>	-18.04	1.22	
<b>Edad</b>	0.02506	0.00669	1.15
<b>Colesterol</b>	0.01234	0.00195	1.10
<b>Triglicéridos</b>	0.00756	0.00120	1.09
<b>Presión sistólica</b>	0.08131	0.00677	1.19
<b>Presión diastólica</b>	0.05010	0.0163	1.18
<b>Alcoholismo</b>	-0.441	0.187	1.15

Nota: Elaboración propia (Obtenido por Minitab versión 14)

Tabla 04

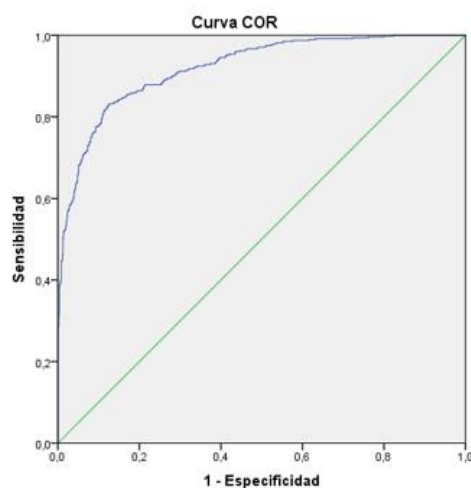
*Clasificación de pacientes, mediante el modelo logístico estimado, con punto de corte  $c=0.50$  (n=1284)*

<b>Observados</b>		<b>Pronosticado</b>		<b>Porcentaje correcto</b>
		<b>No=0</b>	<b>Si=1</b>	
<b>Enfermedad cardiovascular</b>	<b>No=0</b>	735	72	91.1%
	<b>Si=1</b>	113	364	76.3%
<b>Porcentaje global</b>				85.6%

Nota: Elaboración propia.

Figura 01

*Curva ROC del modelo, tasa de falsos positivos (1-especificidad) y tasa de verdaderos positivos (sensibilidad)*



También, con la curva ROC, se estimó un nuevo punto de corte de  $c=0.426$ , valor que permite obtener una alta sensibilidad (0.805) y alta especificidad (0.89), los resultados son: del total de pacientes (807) que no tienen la enfermedad el 89.0% de los pacientes fueron correctamente clasificados, mientras que el 11.0% no fueron correctamente clasificados; del total de pacientes (477) que tienen la enfermedad, el 80.5% fue clasificado correctamente, mientras que el 19.5% no fue clasificado correctamente. La tabla de clasificación presentó además que con este modelo estimado, se clasificó correctamente el 85.8% y 14.2%, se clasificó incorrectamente (tasa de error) Tabla 05.

Tabla 05

*Clasificación de pacientes mediante el modelo logístico estimado  
con punto de corte  $c=0.426$  ( $n=1284$ )*

Observados		Pronosticado		Porcentaje correcto
		Enfermedad		
		No=0	Si=1	
Enfermedad cardiovascular	No=0	718	89	89.0%
	Si=1	93	384	80.5%
Porcentaje global				85.8%

Nota: Elaboración propia

Además, se clasificó a las observaciones, a través de la validación cruzada, consiste en eliminar la *ésima* observación y con las restantes observaciones, se estima el modelo logístico y luego se estima y así sucesivamente [6]. Con punto de corte de  $c=0.5$ , los resultados son: Del total de pacientes (807) que no tienen la enfermedad el 90.6% de los pacientes fueron correctamente clasificados, mientras que el 9.4% no fueron correctamente clasificados; del total de pacientes (477) que tienen la enfermedad, el 76.1% fueron clasificados correctamente, mientras que el 23.9% no fueron clasificados correctamente. La tabla de clasificación presenta además que con este modelo estimado, fueron clasificados correctamente el 85.2% y 14.8%, no fue clasificado correctamente (tasa de error por validación cruzada) Tabla 06.

Tabla 06.

*Clasificación de pacientes, a través de la validación cruzada con punto de corte  $c=0.5$ , ( $n=1284$ ),*

Observados		Pronosticado		Porcentaje correcto
		Enfermedad		
		No=0	Si=1	
Enfermedad cardiovascular	No=0	731	76	90.6%
	Si=1	114	363	76.1%
Porcentaje global				85.2%

Nota: Elaboración propia

Los resultados obtenidos con respecto a los predictores encontrados (edad, alcoholismo, colesterol, triglicéridos, presión arterial sistólica y presión arterial diastólica) que se relacionan a las enfermedades cardiovasculares, se corroboran con otras investigaciones tales como:

La edad como factor de riesgo: Luis (2021), Álvarez y otros(2020) y Zuni-Chávez y otros (2019); alcoholismo como factor protector: Cedeño-Zambrano y otros (2016) y Sacanella y otros (2019); colesterol factor de riesgo: Luis (2021) y Zuni-Chávez y otros(2019); triglicéridos factor de riesgo: Carranza-Madrigal (2017) y Patiño-Villada y otros (2011); presión arterial sistólica factor de riesgo: Hernández-Martínez y otros(2019), Luis (2021); presión arterial diastólica factor de riesgo: Hernández-Martínez y otros (2019).

Además se comparó a dos pacientes (observación 235 y observación 314), el primero con valores por debajo de los límites críticos para las predictores de riesgo a excepción el alcoholismo (edad, alcoholismo, colesterol, triglicéridos, presión sistólica y presión diastólica), la probabilidad de tener la enfermedad cardiovascular es muy baja; todo lo contrario ocurre para valores por encima de los límites críticos.

#### 4. Conclusión

El objetivo fue determinar factores relacionados a las enfermedades cardiovasculares en personas adultas, en una muestra al azar de 1284 pacientes adultos de 20 años a más, atendidos por consultas externas en el Hospital PNP Augusto B. Leguía-Rímac, Lima-Perú, entre los meses de Octubre a Diciembre 2017.

Se identificó factores relacionados a las enfermedades cardiovasculares, en personas adultas, mediante la regresión logística binaria múltiple estos fueron: factor demográfico: Edad (factor de riesgo); factor estilo de vida: Alcoholismo (factor protector) y factor clínicos: Colesterol (factor de riesgo), triglicéridos (factor de riesgo), presión arterial sistólica (factor de riesgo) y presión arterial diastólica (factor de riesgo).

La regresión logística binaria múltiple, es una técnica estadística que identificó variables predictoras cualitativas o cuantitativas que se relacionan o influyen en la probabilidad de ocurrencia

de la enfermedad cardiovascular, y además el modelo estimado clasificó con un porcentaje aceptable del total de pacientes de la muestra (enfermos o sanos) y por último realizar predicciones.

## Referencias bibliográficas

- [1] Agresti, A. (2013). *Categorical Data Analysis*. Tercera edición.
- [2] Akaike, Hirotugu (1974). *A new look at the statistical model identification*. *IEEE Transactions on Automatic Control* 19 (6): 716-723. doi:10.1109/TAC.1974.1100705.
- [3] Álvarez, L., Frías, J., Fernández de Bobadilla, J., Díaz, M. (2020). *Prevalencia de factores de riesgo cardiovascular en trabajadores de un hospital terciario de Madrid*. *Rev. Asoc. Esp. Med Trab. Diciembre 2020 • vol. 29 • Núm. 4 • 257-392*. <https://scielo.isciii.es/pdf/medtra/v29n4/1132-6255-medtra-29-04-274.pdf>
- [4] Barrera, D. (2015). *Factores de riesgo para enfermedades cardiovasculares según los determinantes de la salud presentes en los choferes de transporte público 2014*. [Tesis de licenciatura, Universidad Nacional Mayor de San Marcos].
- [5] Berdú, J., Chacón, T., Fonseca, A. & Pérez, R. (2020). *Factores de riesgo de las enfermedades cardiovasculares en pacientes de 50 años y menos*. *Revista Médica. Granma RPNS-1853 773*. <http://scielo.sld.cu/pdf/mmed/v24n4/1028-4818-mmed-24-04-772.pdf>.
- [6] Berrendero, J., Cárcamo, J. (2019). *Las componentes lineales de una regla de clasificación cuadrática*. <https://caminosaleatorios.wordpress.com/2018/09/04/las-componentes-lineales-de-una-regla-de-clasificacion-cuadratica>.
- [7] Chambergó-Michilot, D., Velit-Ríos, B., Cueva-Parra, A. (2020). *Prevalencia de enfermedades cardiovasculares en el Hospital Nacional Dos de Mayo de Perú*. *Revista Mexicana Angiología. 48(3):84-89*, [https://www.rmangiologia.com/frame\\_esp.php?id=33](https://www.rmangiologia.com/frame_esp.php?id=33)
- [8] Cedeño-Zambrano, J., Vásquez-Jaramillo, P. y Roca-Lino, V. (2016). *Riesgo cardiovascular relacionado con el consumo de alcohol*. *REVISTA CIENTÍFICA DOMINIO DE LA CIENCIA, 2(4), 17-27*. file:///C:/Users/Lenovo/Downloads/Dialnet Riesgo Cardiovascular Relacionado Con El Consumo De Alcoh-5761632.pdf
- [9] Chapra, S., Canale, R. (2006). *Métodos Numérico para ingenieros, 5ta. Ed.*, Mc Graw Hill.
- [10] Cisneros, L., Carrazanas, E. (2013). *Factores de riesgo de la cardiopatía Isquémica*. *Revista Cubana de Medicina General Integral, 29(4),369-378*.
- [11] Cook, R. & Weisberg, S. (1982). *Residuos e influencia en la regresión*. Chapman & Hall.
- [12] Cox, D. & Snell, E. (1989). *Analysis of Binary Data*. Chapman & Hall.
- [13] De la Garza, J., Morales, B., Gonzales, B. (2013). *Análisis estadístico Multivariante*. Editorial: Mc Graw Hill. (pp.60-107).
- [14] Evans, M., Rosenthal, J. (2013). *Probabilidad y estadística*. Editorial Reverte, S.A. Barcelona.

- [15] FUNDACIÓN ESPAÑOLA DEL CORAZÓN, (2022). *La raza como factor de riesgo cardiovascular* [Archivo PDF]. <https://fundaciondelcorazon.com/prevencion/marcadores-de-riesgo/raza-etnia-linaje.html>
- [16] Gujarati, N., Garmendia, D. y Arango, G. (2005). *Econometría* (pp. 167-181) (4° ed.) México: McGraw-Hill.
- [17] Hernández-Martínez, J., Varona-Urbe, M., Hernández, G (2019). Prevalencia de factores asociados a la enfermedad cardiovascular y su relación con el ausentismo laboral de los trabajadores de una entidad oficial, *Revista Colombiana de Cardiología 2020*; 27(2):109-116. <https://www.rccardiologia.com/previos/RCC>
- [18] Hosmer, D., Lemeshow, S (2000). *Applied logistic regression*. John Wiley & Sons.
- [19] INEI. (2012). *Condiciones de Riesgo Cardiovasculares. Perú: Situación de Salud de la Población Adulta Mayor*. [Archivo PDF]. [https://www.inei.gob.pe/media/MenuRecursivo/publicaciones\\_digitales/Est/Lib1146/cap02.pdf](https://www.inei.gob.pe/media/MenuRecursivo/publicaciones_digitales/Est/Lib1146/cap02.pdf)
- [20] Luis, A. (2021). *Riesgo cardiovascular en adultos que acuden a la consulta médica en un policlínico privado en Carabayllo*. *Revista Científica Cuidado y Salud Pública*. 2021; 1(1): 18-23. <https://www.cuidadoysaludpublica.org.pe/index.php/cuidadoysaludpublica/article/view/11/135>.
- [21] Mantel y Haenszel (1959). Aspectos estadísticos del análisis de datos de estudios retrospectivos de enfermedades. *Revista del Instituto Nacional del Cáncer*, 22 (4), 719–748.
- [22] MedlinePlus en español [Internet]. Bethesda (MD): Biblioteca Nacional de Medicina (EE. UU.) [Actualizado 22 nov. 2021]. Enfermedad Cardiovascular; [actualizado 22 nov. 2021; consulta 30 ago. 2022]; [aprox. 2 p.]. Disponible en: <https://medlineplus.gov/spanish/ency/patientinstructions/000759.htm>
- [23] McCullagh, P. & Nelder, J. (1983). *Generalized Linear Models*. Chapman and Hall.
- [24] McFadden, D. (1974). Análisis logit condicional del comportamiento de elección cualitativa, en P. Zarembka, editor, *Frontiers in Econometrics* Nueva York, 1974.
- [25] Montgomery, D., Peck, E. & Vining, G. (2004). *Introducción al análisis de regresión Lineal* (pp. 399-414) (3° Ed.). Compañía Editorial Continental.
- [26] Nagelkerke, N.J.D. (1991). A note on a general definition of the coefficient of determination. *Biometrika*, 78, 691-692.
- [27] Nelder, J. and Wedderburn. R (1972). *Generalized Linear Models*. Chapman and Hall. London.
- [28] OMS. (2020). Temas de salud. Enfermedades cardiovasculares. Obtenido de [https://www.who.int/es/health-topics/cardiovascular-diseases#tab=tab\\_1](https://www.who.int/es/health-topics/cardiovascular-diseases#tab=tab_1)
- [29] OPS. (2021). Las enfermedades del corazón siguen siendo la principal causa de muerte en las Américas. <https://www.paho.org/es/noticias/29-9-2021-enfermedades-corazon-siguen-siendo-principal-causa-muerte-americas>
- [30] Patiño-Villada, F., Arango-Vélez, E., Quintero-Velásquez, M. & Arenas-Sosa, M. (2011). Factores de riesgo cardiovascular en una población urbana de Colombia. *Revista de Salud pública*, 13(3), 433-445.

- [31] Pérez, C. (2014). Técnicas Estadísticas Predictivas Con IBM SPSS Modelos (pp.126-138). Editorial Garceta.
- [32] Rivas, E. (2016). Enfermedades cardiovasculares y actividad física: Recomendaciones para la Atención Primaria de Salud en Cuba. *CorSalud*, 8(3),139-143.
- [33] Rojas, S., Querales, M. Villarino, A. (2016). Evaluación de los factores de riesgos que predisponen a la hipertensión arterial a través de un modelo de regresión logística. *Salus.*, 20(2),18-23.
- [34] Sacanella, I., Casas, R., Viñas, E., Castro, S. y Sacanella, E. (2019). Prevención de la enfermedad cardiovascular y bebidas alcohólicas fermentadas. ¿Realidad o ficción?. Servicio de Medicina Interna. Hospital Clínic. Barcelona. Institut d'Investigació Biomèdica August Pi i Sunyer (IDIBAPS). Universidad de Barcelona. Barcelona. CIBER de Obesidad y Nutrición. Instituto de Salud Carlos III. Madrid.<https://dialnet.unirioja.es/servlet/articulo?codigo=7048944>.
- [35] TEXAS HEART INSTITUTE, (2022) Factores de riesgo cardiovascular, consultado Julio 2022, disponible en: <https://www.texasheart.org/heart-health/heart-information-center/topics/factores-de-riesgo-cardiovascular>.
- [36] Uriel, E. & Aldas, J. (2005). Análisis Multivariante Aplicado (5ª ed.). Thomson
- [37] Wald, A (1943) Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Trans.Amer. Math. Soc.* 54 426-482.
- [38] Zuni-Chávez, K., More-Sandoval, B., Fernández-Vargas, C., García-Fuentes, B., Ruiz-Olano, J. & Pérez-Rodríguez, V. (2019). Prevalencia de factores de riesgo cardiovascular en pacientes hospitalizados en un hospital de Lima. *Revista de la Facultad de Medicina Humana, Lima-Perú*, 19(4), 68-73. <http://revistas.urp.edu.pe/index.php/RFMH/article/view/2343>.