

	Parámetro	Estimación
$\hat{\beta}_0$	Ordenada en el origen	-0,080727
$\hat{\beta}_1$	Edad	0,000497
$\hat{\beta}_2$	Colesterol total	0,000293
$\hat{\beta}_3$	HDL	-0,000672
$\hat{\beta}_4$	Presión Arterial Sistólica	0,000506
$\hat{\beta}_5$	Diabetes	0,023995
$\hat{\beta}_6$	Tabaquismo	0,017748
$\hat{\beta}_7$	Sexo	0,007638

Tabla 4. Resumen de la información dada por la segunda recta de regresión estimada.

regresión $H_0: \beta_1 = \dots = \beta_7 = 0$ es $1,1127 \cdot 10^{-48}$ (el estadístico es 63,82 y usamos la distribución $F_{k;n-k-1}$ para contrastar, con $n = 219$ y $k = 7$), un valor incluso más pequeño que el del análisis anterior, siendo muy razonable aceptar que la realidad se corresponde fielmente con las predicciones que damos con esta recta de regresión.

Por último, al igual que hicimos con la primera recta de regresión, para cada parámetro damos sus intervalos de confianza al 5 % (ver Tabla 5) y estudiamos las hipótesis de contraste $H_0: \beta_j = 0$ (ver Tabla 6).

Intervalos de confianza con $\alpha = 5\%$			
	Parámetro	Extremo inferior	Extremo superior
β_0	Ordenada en el origen	-0,101623	-0,059831
β_1	Edad	0,000354	0,000639
β_2	Colesterol total	0,000223	0,000364
β_3	HDL	-0,000832	-0,000513
β_4	Presión Arterial Sistólica	0,000374	0,000637
β_5	Diabetes	0,016272	0,031717
β_6	Tabaquismo	0,010298	0,025199
β_7	Sexo	0,002399	0,012877

Tabla 5. Con una probabilidad del 95 % cada uno de los parámetros $\beta_j, j = 0, \dots, 7$ mencionados en esta tabla se encuentran en dicho rango de valores.

Los p-valores de las hipótesis de contraste $H_0: \beta_j = 0$ son tan pequeños para todos los parámetros que podemos afirmar que todos ellos son relevantes y aportan información a la recta de regresión. Confirmando nuestras sospechas hechas en el primer análisis que apuntaban a que estos eran los parámetros relevantes. El hecho de que los p-valores sean mucho más pequeños que en el análisis anterior es también un buen síntoma que sugiere que hemos conseguido eliminar duplicidades, es decir, los parámetros que intervienen ahora no presentan tantas relaciones lineales entre ellos.

Con respecto a los intervalos de confianza para un valor $\alpha = 5\%$. Este segundo análisis ha mejorado las estimaciones de los parámetros. Consecuentemente, hemos calibrado el impacto que tienen factores de riesgo y cómo estos afectan de una manera u otra al riesgo de evento

Hipótesis de contraste $H_0: \beta_j = 0$		
	Parámetro	p-valor
β_1	Edad	$6,72 \cdot 10^{-9} \%$
β_2	Colesterol total	$2,75 \cdot 10^{-12} \%$
β_3	HDL	$1,05 \cdot 10^{-12} \%$
β_4	Presión Arterial Sistólica	$1,24 \cdot 10^{-10} \%$
β_5	Diabetes	$4,38 \cdot 10^{-7} \%$
β_6	Tabaquismo	$4,78 \cdot 10^{-4} \%$
β_7	Sexo	0,4466 %

Tabla 6. Realizamos hipótesis de contrastes para todos los parámetros (excepto ordenada en el origen, el hecho de que la ordenada en el origen pueda ser 0 o no, no es relevante ya que simplemente actúa como una corrección de la recta) y damos el p-valor, cuanto más bajo sea este, con mayor contundencia podemos rechazarlo y cuanto más alto, existe menor evidencia estadística para rechazarlo y por tanto lo natural será aceptar la hipótesis de contraste, $\beta_j = 0$.

cardiovascular ya que todos los intervalos de confianza nuevos tienen una longitud inferior a la que tenían en el primer análisis (depurar la recta eligiendo solo los valores relevantes nos ha permitido tener un conocimiento más preciso de la información de la que disponemos), destacando especialmente el parámetro "Colesterol total", la longitud de su intervalo de confianza es del orden de casi 10 veces más estrecho (una excepcional mejora en la estimación de la cuantificación de la influencia del colesterol como factor de riesgo).

Además, los intervalos de confianza del segundo análisis de los parámetros "Colesterol total", "Presión Arterial Sistólica" y "Diabetes" son contenidos completamente por los intervalos de confianza obtenidos para estos mismos parámetros en el primer análisis. Para los otros tres parámetros, los intervalos de confianza del segundo análisis están contenidos casi completamente en los respectivos intervalos de confianza del primero. Puede ser interesante estudiar si en nuestra muestra existe algún tipo de relación entre los parámetros "Edad", "HDL" o "Tabaquismo" con el parámetro que hemos añadido para este segundo análisis, la variable "Sexo", y que esta sea la razón por la que sus intervalos de confianza no estén contenidos completamente en los del análisis previo. Ya hemos visto que sí que existe esta entre el "Tabaquismo" y el "Sexo", el análisis para los otros dos es más complicado pero sugiere nuevas preguntas interesantes.

Resumimos las conclusiones más interesantes a las que hemos llegado con este segundo análisis depurado que cambia (esencialmente mejora) algunos de los resultados obtenidos antes y aporta nuevas observaciones importantes.

- Con una probabilidad del 95 %, envejecer 10 años conlleva el aumento de entre un 0,35 % y un 0,64 % del riesgo de evento cardiovascular.
- El aumento del colesterol es perjudicial: Un aumento de 10 mg/dL incrementa el riesgo de evento cardiovascular entre un 0,22 % y un 0,36 % con una probabilidad del 95 %, en cambio, el aumento del HDL es beneficioso para prevenir eventos cardiovasculares, un aumento de 10 mg/mL disminuye el riesgo de evento cardiovascular entre un 0,51 % y un 0,83 % con una probabilidad del 95 %.
- Con una probabilidad del 95 %, el incremento en 10 mmHg de la presión arterial sistólica conlleva el aumento de entre un 0,37 % y un 0,64 % del riesgo de evento cardiovascular.

- Los pacientes diabéticos presentan con una probabilidad del 95 % un aumento de entre un 1,63 % y un 3,17 % del riesgo de evento cardiovascular.
- Los fumadores habituales presentan con una probabilidad del 95 % un aumento de entre un 1,03 % y un 2,52 % del riesgo de evento cardiovascular.
- El sexo del paciente influye en el riesgo que este tiene de sufrir un evento cardiovascular, con un 95 % de probabilidad, un hombre tiene un riesgo de evento cardiovascular de entre un 0,24 % y un 1,29 % superior al que tiene una mujer.

4. Estudio de datos influyentes.

En todo estudio de naturaleza bioestadística es habitual encontrar datos influyentes o “anómalos”. La razón de esto es que en la muestra solemos encontrar pacientes con condiciones o patologías especiales que requieren ser tratados de forma personalizada, susceptibles de alejarse demasiado de las estimaciones predichas por el modelo. Por ello, es importante detenerse en los datos anómalos para estudiarlos con mayor cuidado y de manera individualizada. Queremos decidir si estos pacientes requieren un tratamiento excepcional o si el modelo ha fallado para ellos.

Decimos que una observación es influyente si tiene mucho impacto en $\hat{\beta}$ y en \hat{y} , a estas observaciones las llamaremos puntos críticos.

Consideremos \hat{Y} el vector con las estimaciones \hat{y}_i , Y el vector con las observaciones y_i y X la matriz de datos, entonces, se satisface,

$$\hat{Y} = X\hat{\beta} = X(X^T X)^{-1} X^T Y.$$

Definimos la matriz $H := X(X^T X)^{-1} X^T$, y tomamos la coordenada h_{ii} que definimos como leverage del dato i -ésimo. Pasamos ahora a definir la distancia de Cook como,

$$D_i := \frac{1}{k+1} \left(\frac{e_i}{s_R \sqrt{1-h_{ii}}} \right)^2 \frac{h_{ii}}{1-h_{ii}}.$$

Podemos usar el leverage, h_{ii} y la distancia de Cook para identificar puntos críticos [8, 9].

Las distancias de Cook de nuestra muestra son por lo general bastante pequeñas. Esta medida resalta por encima del resto solo a los pacientes 33 (distancia de Cook 0,1742) y 158 (distancia de Cook 0,1613), exceptuando estos dos datos, el resto de elementos de la muestra lleva asociada una distancia de Cook inferior a 0,075.

Para el leverage consideraremos anómalos aquellos datos que sean superiores a $3 \frac{k+1}{n} = 0,17808$. Tenemos 6 datos que satisfacen esto, los pacientes 15 ($h_{ii} = 0,2112$), 29 ($h_{ii} = 0,1869$), 79 ($h_{ii} = 0,2290$), 115 ($h_{ii} = 0,3521$), 159 ($h_{ii} = 0,3813$) y 161 ($h_{ii} = 0,2389$).

Los pacientes 33 y 158 que eran influyentes según la distancia de Cook son también los únicos dos pacientes cuyo riesgo de evento cardiovascular es superior al 15 %, 18 % y 19 % respectivamente. Como su riesgo de evento cardiovascular es tan elevado comparado con el del resto de pacientes es normal que estos dos puntos sean considerados influyentes, la coordenada y se encuentra lejos de la del resto de datos.

Pasamos al estudio personalizado de los datos influyentes según el “leverage” (ninguno coincide con los datos influyentes según la distancia de Cook)

- Paciente 15: Se trata de un paciente pluripatológico en tratamiento con múltiples fármacos, entre ellos *Acecumarol*. Este es un anticoagulante, un tratamiento preventivo para disminuir el riesgo de accidentes tromboembólicos. Por ello, es de esperar que algunos de

sus factores de riesgo hayan sido modificados por el tratamiento farmacológico mientras que otros factores han permanecido inalterados, creando un dato con parámetros poco comunes que puede ser leído de manera influyente por la regresión. Además el paciente era diabético, en tratamiento hipoglucemiante, lo que hace que el valor de las cifras de glucemia estén alteradas artificialmente.

- Paciente 29: Se trata de un paciente con hipertrigliciremia severa con valores de colesterol dentro de la normalidad, en este paciente habría que descartar una hipertrigliceremia familiar, en cuyo caso habría factores genéticos que podrían influir en su riesgo cardiovascular. Este paciente sobrepasa en gran manera los valores habituales de triglicéridos y es probable que por ello sea un dato influyente para la recta de regresión.
- Paciente 79: La distancia de Cook para este paciente es muy pequeña, $1,85 \cdot 10^{-6}$, así que según este criterio, el paciente 79 se encuentra lejos de ser un dato influyente. Su única anomalía es padecer diabetes mellitus tipo I, único paciente con esta disfunción de entre todo los analizados. Esta característica podría hacer que este dato sea influyente. No obstante, no somos capaces de decidir si el modelo ha fallado ya que el dato podría no ser influyente (la predicción y el dato "REGICOR" real fueron muy similares).
- Paciente 115: La tensión diferencial, considerada factor de riesgo por varios autores [18], de esta paciente es muy elevada. Los valores de "LDL" y "Triglicéridos" son demasiado elevados, mientras que los de "HDL" se encuentran por debajo de lo aconsejable. Además, el paciente es diabético y fumador habitual. Al tener tantos parámetros en rangos muy alejados de las cifras habituales es razonable pensar que este paciente deba ser considerado un dato influyente para nuestra recta de regresión. Su riesgo de evento cardiovascular debería ser estudiado de forma especial y personalizada.
- Paciente 159: Paciente de avanzada edad pluripatológico, polimedocado, bien controlado. Los datos de los distintos parámetros son muy satisfactorios para una persona de su edad (posiblemente como consecuencia de su medicación) que lo convertiría en un dato influyente por estar en rangos distintos a lo que suele ser habitual para pacientes de su edad.
- Paciente 161: Paciente diabético fumador de avanzada edad relativamente bien controlado por medicación lo que hace que sus parámetros puedan estar modificados artificialmente.

Es interesante observar que la gran mayoría de los datos influyentes son diabéticos (pacientes que no constituyen un porcentaje tan representativo de la población). Por ello, para realizar un análisis más depurado y fiable, sería aconsejable estudiar la población diabética y no diabética de forma separada.

Para nuestro estudio, al contar solo con 6 datos influyentes sobre un total de 219, retirarlos del análisis no cambiaría de forma significativa los resultados obtenidos.

5. Regla de clasificación

Queremos predecir si un paciente con unas determinadas características es susceptible de tener un riesgo de evento cardiovascular bajo o moderado (inferior al 10%), clasificado en la población 0, o de padecer un riesgo alto o muy alto (igual o superior al 10%), población 1.

En nuestra muestra contamos con 206 pacientes en la población 0 (riesgo bajo o moderado) y con 13 en la población 1 (riesgo alto o muy alto).

La regla de clasificación que hemos aplicado es la regresión logística, usando el comando "Solver" del programa Excel para resolver la combinación de $\hat{\beta}_j$ que maximiza el estimador log VEROSIMILITUD.

Con esto, la regla obtenida consiste en calcular el valor

$$\begin{aligned}
 f(x_i) = & \hat{\beta}_0 + \hat{\beta}_1 \cdot \text{Edad} + \hat{\beta}_2 \cdot \text{Colesterol Total} + \hat{\beta}_3 \cdot \text{HDL} + \hat{\beta}_4 \cdot \text{LDL} \\
 & + \hat{\beta}_5 \cdot \text{Triglicéridos} + \hat{\beta}_6 \cdot \text{PAS} + \hat{\beta}_7 \cdot \text{PAD} + \hat{\beta}_8 \cdot \text{Glúcidos} \\
 & + \hat{\beta}_9 \cdot \text{Filtrado Glomerular} + \hat{\beta}_{10} \cdot \text{Albuminuria} \\
 & (+\hat{\beta}_{11} \text{ si el paciente es diabetico}) \\
 & (+\hat{\beta}_{12} \text{ si el paciente es fumador}) \\
 & (+\hat{\beta}_{13} \text{ si el paciente es hombre}),
 \end{aligned}$$

donde los valores de los coeficientes $\hat{\beta}_j$ vienen dados en la Tabla 7.

Regresión Logística		
	Parámetro	Estimación para la regresión logística
$\hat{\beta}_0$	Ordenada en el origen	-800
$\hat{\beta}_1$	Edad	1,0139631
$\hat{\beta}_2$	Colesterol total	0,1004485
$\hat{\beta}_3$	HDL	1,0092167
$\hat{\beta}_4$	LDL	1,03056338
$\hat{\beta}_5$	Triglicéridos	1,0258
$\hat{\beta}_6$	Presión Arterial Sistólica	1,030683
$\hat{\beta}_7$	Presión Arterial Diastólica	1,0169
$\hat{\beta}_8$	Glucosa	1,0222
$\hat{\beta}_9$	Filtrado Glomerular	1,000156
$\hat{\beta}_{10}$	Albuminuria	1,0002
$\hat{\beta}_{11}$	Diabetes	1,0000833
$\hat{\beta}_{12}$	Tabaquismo	1,000083
$\hat{\beta}_{13}$	Sexo	1,00015

Tabla 7. Valores propuestos por la función "Solver" de Excel para la regla de clasificación de regresión logística.

De esta forma, si el valor de $f(x_i)$ es menor que 0 clasificamos el dato x_i en la población 0, y en caso contrario en la población 1. Es cierto que la función "Solver" a veces puede ser imprecisa. No obstante, la regla de clasificación que hemos dado ha sido acertada para un 93,15 % de los datos (204 del total de 219), aunque al haber usado los mismos datos para comprobar la fiabilidad y crear la regla de clasificación este valor tan elevado de aciertos puede haber sido adulterado, deberíamos comprobar esta regla con datos nuevos. En caso de no poder conseguir datos nuevos, para evaluar la regla, para cada dato de los que disponemos podemos crear una regla de clasificación con esta metodología (regresión logística), hacer la predicción de ese dato y evaluar si la predicción ha acertado.

Nuestra regla ha clasificado de forma altamente precisa los datos de la población 0, solo 7 fallos de entre 206 pacientes de esta población, pero clasifica con poca precisión a los pacientes de la población 1, 8 fallos de entre 13 pacientes.

6. Conclusiones

En este estudio hemos realizado dos rectas de regresión lineal para mejorar los resultados proporcionados por esta herramienta. De esta manera se ha conseguido cuantificar de forma muy precisa el impacto de 7 de los 13 factores de riesgo de los cuales sospechábamos una potencial relación con el riesgo cardiovascular. De los 6 factores de riesgo iniciales restantes, uno de ellos, "LDL" está relacionado implícitamente de forma lineal con otros dos de los factores incluidos en el segundo análisis, "Colesterol total" y "HDL". Otro de los potenciales factores de riesgo descartados para la segunda recta de regresión, el "Filtrado glomerular" ha sido estudiado cualitativamente de forma independiente en el Apéndice A. Para futuros estudios se sugiere estudiar la posible independencia de los 4 potenciales factores de riesgo restantes ("Triglicéridos", "Presión Arterial Diastólica", "Glucosa", "Albuminaria"). También se sugiere corroborar los resultados aquí obtenidos con otros datos.

En este trabajo también hemos propuesto una regla de clasificación por medio de una regresión logística. Para un próximo estudio proponemos la creación de reglas óptimas de clasificación para estas dos poblaciones. Con este fin, como funciones para cada una de las poblaciones se propone: usar funciones tales que para los parámetros $\beta_1 - \beta_{10}$ sean una normal de dimensión 10 con media y varianza muestral (intentando obtener el dato de la albuminaria y no solo la clasificación en A1, A2 o A3) y para los parámetros $\beta_{11}, \beta_{12}, \beta_{13}$ Bernoullis tomando como p el siguiente valor

$$p = \frac{\text{casos valor 1 en la población } i}{\text{casos totales en la población } i}, \quad i = 1, 2$$

independientes cada una de ella del resto de parámetros que intervienen en la función de densidad. Tras esto se aconseja incluir costes de mala clasificación si se desea evitar clasificar pacientes con riesgo alto o muy alto como pacientes de riesgo bajo o moderado, mejorando su función como una herramienta de prevención de riesgo. También sería interesante usar algún test para cerciorarnos de que estas funciones de densidad reflejan la realidad.

A. El filtrado glomerular es un factor de riesgo cardiovascular no lineal.

La Enfermedad Renal Crónica (ERC) tiene en nuestro medio una prevalencia aproximada del 15%, siendo mayor en varones, personas con edad avanzada y pacientes con enfermedad cardiovascular. Fue incluida como factor de riesgo independiente de Enfermedad Cardiovascular por el *Joint National Committee on Prevention Detection and Treatment of High Blood pressure* en 2003 [7]. Partiendo de estas premisas, parece interesante utilizar los marcadores analíticos de función renal para el diagnóstico precoz de patología cardiovascular ya que la marcada acumulación de factores de riesgo cardiovascular en estos nefrópatas podría hacer considerar a la enfermedad renal crónica como un trastorno cardiovascular (e.g. [12, 14]). En esta Sección discutiremos que existe una dependencia entre estos dos factores pero que esta no es lineal.

La metodología empleada es un contraste χ^2 de independencia (supondremos los valores de las probabilidades desconocidos y los estimaremos por cantidad de cada clase dividido por el tamaño de la muestra). Vamos a contrastar la siguiente hipótesis de rechazo.

H_0 : el filtrado glomerular y el riesgo cardiovascular se distribuyen según dos variables aleatorias independientes.

Por último, antes de la recopilación de datos, conviene explicar el significado de las distintas categorías en relación al filtrado glomerular, así como los criterios de clasificación.

Estadios del filtrado glomerular estimado ($mL/min/1,73 m^2$) en función de creatinina con fórmula CKD-EP¹

- G1 $\geq 90\%$: FG normal o elevado.
- G2 60 – 89%: FG ligeramente disminuido.
- G3a 45 – 59%: FG ligera-moderadamente disminuido.
- G3b 30 – 44%: FG moderada-gravemente disminuido.
- G4 15 – 29%: FG gravemente disminuido.
- G5² < 15%: Fallo renal.

Tabla de observados				
Fil. Glomerular/RC (REGICOR)	Bajo	Moderado	Alto	Muy Alto
G1	70	17	4	0
G2	53	39	6	2
G3a	10	7	1	0
G3b	3	5	0	0
G4	2	0	0	0

Tabla 8. Tabla de observados relación filtrado glomerular con riesgo cardiovascular (según test REGICOR).

Tabla de esperados				
Fil. Glomerular/RC (REGICOR)	Bajo	Moderado	Alto	Muy Alto
G1	57.34	28.26	4.57	0.83
G2	63.01	31.05	5.02	0.91
G3a	11.34	5.59	0.90	0.16
G3b	5.04	2.48	0.40	0.07
G4	1.26	0.62	0.10	0.02

Tabla 9. Tabla de esperados relación filtrado glomerular con riesgo cardiovascular (según test REGICOR).

El estadístico de Pearson toma el valor $b = 19,00$, lo cual implica que el p -valor es 8,84% (contrastamos con χ^2_{12}). Esto permite rechazar la hipótesis de contraste H_0 con un nivel de confianza superior al 90%.

Realicemos un análisis cualitativo sobre la relación entre la filtración glomerular y el riesgo cardiovascular. Las diferencias más abultadas que encontramos entre observados y esperados se encuentran en las clases G1 y G2. Para G1 el riesgo observado parece más bajo que el esperado, mientras que para G2 los resultados parecen indicar lo contrario. En cambio, para las clases G3a, G3b y G4, los resultados observados concuerdan con los esperados, lo cuál sugiere que una vez alcanzado el estado G3a (o superior) la calidad del filtrado glomerular es independiente del riesgo cardiovascular.

En base a estos resultados podemos concluir que los pacientes en estadio G1 (filtrado glomerular normal o elevado) presentan en la práctica un RCV menor que el que cabría esperar si ambos fenómenos fueran independientes. El adecuado filtrado glomerular sería capaz de compensar, al menos parcialmente, el efecto nocivo de estos factores de riesgo.

¹Individuos de raza negra: multiplicar por 1,159

²En nuestra muestra, ningún paciente ha alcanzado la categoría de fallo renal y en consecuencia no la tendremos en cuenta para nuestro estudio.

En el estadio G2 ya el filtrado glomerular se encuentra discretamente disminuido (60 – 89 %), la uremia aumenta levemente y sería responsable de un estado de inflamación crónica a nivel de toda la anatomía con el consiguiente aumento de RCV como se aprecia en nuestro estudio.

Cuando la insuficiencia renal ya está consolidada, es decir filtrado glomerular inferior al 60 % (estadios G3a a G5), la conjunción de los distintos factores de riesgo ya han lesionado de forma importante e irreversible tanto el sistema cardiovascular como la función renal y en consecuencia el deterioro renal no se ve reflejado en un aumento del RCV que ya es de por sí elevado.

Del análisis de estos resultados se desprende que la detección precoz de los pacientes con disminución ligera del filtrado glomerular (G2), previa a la insuficiencia renal crónica, permitiría implementar las actuaciones sobre los Factores de Riesgo Cardiovascular modificables y con ello disminuir de forma costo-efectiva la morbimortalidad cardiovascular así como la evolución hacia la enfermedad renal crónica.

Referencias

- [1] M. Al de Francisco, L. Aguilera y V. Fuster. Enfermedad cardiovascular, enfermedad renal y otras enfermedades crónicas. Es necesaria una intervención más temprana en la enfermedad renal crónica. *Nefrología* 29.1 (2009), págs. 6-9.
- [2] R. Alcanzar, L. Orte y A. Otero. Enfermedad renal crónica. *Nefrología* (2008), págs. 3-6.
- [3] R. Alcázar Arroyo, L. Orte, E. González Parra, J. Górriz, J. Navarro, A. Martín de Francisco, M. Egocheaga y F. Álvarez Guisasaola. Documento de consenso SEN-semFYC sobre la enfermedad renal crónica. *Nefrología* 28.3 (2008), págs. 273-282.
- [4] R. W. Alexander. Hypertension and the pathogenesis of atherosclerosis: oxidative stress and the mediation of arterial inflammatory response: a new perspective. *Hypertension* 25.2 (1995), págs. 155-161.
- [5] A. Arias Morales, R. García Hernández y M. Oliva Pérez. Riesgo cardiovascular global en pacientes ancianos hipertensos. *Revista Cubana de Medicina* 53.2 (2014), págs. 178-188.
- [6] A. Asesio Sánchez, A. Fernández y A. Morcillo. *Influencia de la Dieta Mediterránea en la percepción de la salud: Estudio en población adolescente del IES Antonio Machado de Soria y población adulta de la provincia de Soria*. Editorial Académica Española, 2018.
- [7] A. V. Chobanian, G. L. Bakris, H. R. Black, W. C. Cushman, L. A. Green, J. L. Izzo Jr, D. W. Jones, B. J. Materson, S. Oparil, J. T. Wright Jr et al. Seventh report of the joint national committee on prevention, detection, evaluation, and treatment of high blood pressure. *hypertension* 42.6 (2003), págs. 1206-1252.
- [8] R. D. Cook. Detection of influential observation in linear regression. *Technometrics* 19.1 (1977), págs. 15-18.
- [9] R. D. Cook. Influential observations in linear regression. *Journal of the American Statistical Association* 74.365 (1979), págs. 169-174.
- [10] J. M. B. Díez, J. L. del Val García, J. T. Pelegrina, J. L. M. Martínez, R. M. Peñacoba, I. G. Tejón, E. M. R. Quintana, M. P. Sajkiewicz, A. A. Boronat, B. Á. Pérez et al. Epidemiología de las enfermedades cardiovasculares y factores de riesgo en atención primaria. *Revista española de cardiología* 58.4 (2005), págs. 367-373.
- [11] S. G. García, R. M. Barmudez, J. B. Sanjuan, A. C. Amenós, R. D. Piquet y A. L. M. de Fran. Recomendaciones sobre la utilización de ecuaciones para la estimación del filtrado glomerular en adultos. *Nefrología* 26.6 (2006), págs. 658-665.

- [12] M. Gorostidi, M. Sánchez-Martínez, L. M. Ruilope, A. Graciani, J. Juan, R. Santamaría, M. D. del Pino, P. Guallar-Castillón, F. de Álvaro, F. Rodríguez-Artalejo et al. Prevalencia de enfermedad renal crónica en España: impacto de la acumulación de factores de riesgo cardiovascular. *Nefrología* 38.6 (2018), págs. 606-615.
- [13] R. R. Gorrita Pérez, A. Gilvonio Cárdenas e Y. Hernández Martínez. Caracterización del hábito de fumar en un grupo de escolares adolescentes. *Revista Cubana de Pediatría* 84.3 (2012), págs. 256-264.
- [14] M. I. Gutiérrez Pérez y J. C. Rodríguez Sánchez. *Problemas nefrourológicos y genito-urinarios*.
- [15] S. M. Haffner, S. Lehto, T. Rönnemaa, K. Pyörälä y M. Laakso. Mortality from coronary heart disease in subjects with type 2 diabetes and in nondiabetic subjects with and without prior myocardial infarction. *New England journal of medicine* 339.4 (1998), págs. 229-234.
- [16] W. Insull Jr. The pathology of atherosclerosis: plaque development and plaque responses to medical treatment. *The American journal of medicine* 122.1 (2009), S3-S14.
- [17] J. P. S. de Lafuente, Y. Sáez, M. Vacas, M. Santos, J. D. Sagastagoitia, E. Molinero y J. A. Iriarte. Diferencias de sexo en los factores de riesgo cardiovascular en pacientes con enfermedad coronaria comprobada angiográficamente. *Clínica e investigación en arteriosclerosis* 21.4 (2009), págs. 173-178.
- [18] J. A. Linares, J. B. Simó, F. S. Sancho, S. R. Martínez, M. Y. Marco y M. M. Pérez. La presión diferencial como factor independiente de riesgo cardiovascular. *Atención Primaria* 36.1 (2005), págs. 19-24.
- [19] E. M. T. Lira. Impacto de la hipertensión arterial como factor de riesgo cardiovascular. *Revista Médica Clínica Las Condes* 26.2 (2015), págs. 156-163.
- [20] Á. C. Matía Cubillo, J. M. Soto Esteban y M. E. Martín Pascual. *Riesgo cardiovascular*.
- [21] J. Millán, A. Hernández-Mijares, J. F. Ascaso, M. Blasco, A. Brea, Á. Díaz, P. González-Santos, T. Mantilla, J. Pedro-Botet, X. Pintó et al. La auténtica dimensión del colesterol-no-HDL: colesterol aterogénico. *Clínica e Investigación en Arteriosclerosis* 28.6 (2016), págs. 265-270.
- [22] V. Moreno, A. García Raso, M. García Bueno, C. Sánchez-Sánchez, E. Meseguer, R. Mata y P. Llamas. Factores de riesgo vascular en pacientes con ictus isquémico. Distribución según edad, sexo y subtipo de ictus. *Rev Neurol* 46.10 (2008), págs. 593-8.
- [23] M. Morillas Bueno. ¿Hasta dónde bajar la glucemia para la prevención cardiovascular en la diabetes? *Cardiología hoy 2015* (2015), págs. 593-598.
- [24] M. Naghavi, H. Wang, R. Lozano, A. Davis, X. Liang, M. Zhou et al. GBD 2013 Mortality and Causes of Death Collaborators. Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet* 385.9963 (2015), págs. 117-171.
- [25] Ó. Pérez-Méndez. Lipoproteínas de alta densidad (HDL). ¿Un objetivo terapéutico en la prevención de la aterosclerosis? *Archivos de cardiología de México* 74.1 (2004), págs. 53-67.
- [26] M. Ramirez Iñiguez de la Torre et al. "Determinación del Riesgo Cardiovascular en una población laboral aparentemente sana. Relación con Variables Sociodemográficas y Laborales". Tesis doct. Universitat de les Illes Balears, 2016.
- [27] G. A. Roth, M. H. Forouzanfar, A. E. Moran, R. Barber, G. Nguyen, V. L. Feigin, M. Naghavi, G. A. Mensah y C. J. Murray. Demographic and epidemiologic drivers of global cardiovascular mortality. *New England Journal of Medicine* 372.14 (2015), págs. 1333-1341.
- [28] D. Sarre-Álvarez, R. Cabrera-Jardines, F. Rodríguez-Weber y E. Díaz-Greene. Enfermedad cardiovascular aterosclerótica. Revisión de las escalas de riesgo y edad cardiovascular. *Medicina interna de México* 34.6 (2018), págs. 910-923.

- [29] Wikipedia. *Provincia de Soria* — *Wikipedia, The Free Encyclopedia*. <http://es.wikipedia.org/w/index.php?title=Provincia%20de%20Soria&oldid=145180302>. [Online; accessed 11-August-2022]. 2022.
- [30] S. Yusuf, S. Hawken, S. Ôunpuu, T. Dans, A. Avezum, F. Lanas, M. McQueen, A. Budaj, P. Pais, J. Varigos et al. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. *The lancet* 364.9438 (2004), págs. 937-952.

Sobre el/los autor/es:

Nombre: Alejandro Fernández-Jiménez

Correo electrónico: alejandro.fernandezjimenez@maths.ox.ac.uk

Institución: University of Oxford

Orcid: 0000-0002-4548-4374.

Nombre: Álvaro Fernández Jiménez

Correo electrónico: alvfj1404@gmail.com

Institución: Universidad Católica de Valencia

Orcid: 0000-0002-1580-9764.

Nombre: Abilio José Fernández Vicente

Correo electrónico: abiliofernandez@yahoo.es

Institución: Sanidad de Castilla y León (SaCyL)

Orcid: 0000-0002-9507-7937.