

## PROPUESTA DE MODELO DE ANÁLISIS DE DATOS PARA CATEGORIZAR EL NIVEL DE RIESGO CARDIOVASCULAR EN EL DEPARTAMENTO DEL ATLÁNTICO

**Luis Rodrigo Consuegra Gonzalez**

Trabajo de Investigación como requisito para optar el título de Maestría en  
Ingeniería de Sistemas y Computación

Tutores

**Paul Adolfo Sanmartín Mendoza, Tutor**  
**Diana Judith Heredia Vizcaíno, Cotutor**

### RESUMEN

La presente propuesta de trabajo de investigación de maestría tiene como objetivo proponer un modelo basado en análisis de datos para categorizar el nivel de riesgo cardiovascular que apoye en la generación de planes para la prevención y control de estas enfermedades en el departamento del Atlántico. La investigación presentó un enfoque mixto, de carácter aplicado, con diseño experimental y una población de 10.270 personas, de las cuales se tomaron 6218 personas en la muestra. En primera medida, se estableció el modelo PROCAM, de acuerdo con investigaciones previas, como el recomendado para el contexto del departamento del Atlántico, en segunda medida, se hallaron como variables intervinientes en el modelo el sexo, edad, tabaquismo, antecedentes familiares (muerte por infarto agudo de miocardio u otra enfermedad cardiovascular), imc (índice masa corporal), diabetes mellitus, presión arterial sistólica y diastólica, colesterol HDL, colesterol LDL y triglicéridos. Como resultado de la investigación se espera que mediante la implementación del modelo se tenga una constante actualización de los datos de las historias clínicas para realizar análisis y filtros de la población estimada con riesgo en la realización de estrategias focalizadas de prevención de los pacientes detectados.

#### **Antecedentes:**

- Las enfermedades cardiovasculares (ECV) constituyen un grave problema de salud pública mundial por ser la primera causa de mortalidad según la Organización Mundial de la Salud.
- Tasa de mortalidad mundial año 2016 fue del 26,7% con 15,2 millones de muertes de un total de 56,9 millones a causa de la cardiopatía isquémica y los accidentes cerebrovasculares pertenecientes al grupo de ECV.
- El desarrollo de ECV es promovido por varios factores de riesgo (FR), algunos de estos son modificables y sujetos a prevención. La OMS indica

lineamientos basados en evidencias, sobre cómo reducir la incidencia, ya sean por primera vez o de forma recurrente, de enfermedades cardíacas coronarias, cerebro vasculares y vasculares periféricas.

- El ministerio de salud de Colombia registra las ECV dentro de las diez principales causas de mortalidad y se proyecta el incremento de estas incidencias.
- En el departamento del Atlántico la tasa de mortalidad 2018 debido a la cardiopatía isquémica y los accidentes cerebrovasculares representan un 25,60%, según cifras del DANE.
- En los países en vía de desarrollo, donde se afectan las poblaciones más pobres, reduciendo los años de vida productiva y aumentando los costos de servicios de salud y obstaculizando el desarrollo social de las comunidades.

### **Objetivos:**

#### **General**

- Proponer un modelo basado en análisis de datos para categorizar el nivel de riesgo cardiovascular que apoye a la generación de planes de prevención y control de estas enfermedades en el departamento del Atlántico.

#### **Específicos**

- Identificar los fundamentos teóricos y epistemológicos asociados al pronóstico de riesgos cardiovasculares.
- Formular el modelo de análisis de datos que permita categorizar el nivel de riesgo cardiovascular del departamento.
- Aplicar el modelo teórico construido utilizando historias clínicas del departamento del Atlántico para validar su funcionamiento y futuras aplicaciones de alerta temprana.

**Materiales y Métodos:** La investigación presentó un enfoque mixto, de carácter aplicado, con diseño experimental y una población de 10.270 personas, de las cuales se tomaron 6218 personas en la muestra.

**Resultados:** Como resultados se destaca la propuesta y aplicación de un modelo de estimación del riesgo cardiovascular acorde con el contexto del departamento del Atlántico.

**Conclusiones:** La validación del modelo resultó efectiva, en la medida que la muestra utilizada para su desarrollo (Departamento del Atlántico) contó con las características previamente establecidas, en la implementación se requiere un alto esfuerzo en el procesado de la fuente de datos clínicos a la medida de cada entidad de salud, con la adopción de la historia clínica interoperable se espera que esta etapa sea homogénea y permita la utilización en cualquier entidad, en cuanto a su futuro funcionamiento, el modelo representó un avance dentro del contexto en el cual se encuentra debido a que no se había realizado previamente esta

modelación dentro del Departamento, permitiendo la inclusión de soluciones tecnológicas en las estrategias de análisis cardiovascular de la población.

**Palabras clave:** Enfermedad cardiovascular, factores de riesgos cardiovasculares, riesgo cardiovascular, modelos de predicción cardiovascular

## ABSTRACT

The present proposal of master's research work aims to propose a model based on data analysis to categorize the level of cardiovascular risk that supports the generation of plans for the prevention and control of these diseases in the department of Atlántico. The research presented a mixed approach, of an applied nature, with an experimental design and a population of 10,270 people, of which 6,218 people were taken in the sample. In the first measure, the PROCAM model was established, in accordance with previous research, such as the one recommended for the context of the department of Atlántico, in the second measure, sex, age, smoking, family history (death due to acute myocardial infarction or other cardiovascular disease), BMI (body mass index), diabetes mellitus, systolic and diastolic blood pressure, HDL cholesterol, LDL cholesterol and triglycerides. As a result of the research, it is expected that through the implementation of the model there will be a constant update of the data of the medical records to perform analyzes and filters of the estimated population at risk in the implementation of focused prevention strategies of the detected patients.

### Background:

- Cardiovascular diseases (CVD) constitute a serious global public health problem because they are the leading cause of mortality according to the World Health Organization.
- Global mortality rate in 2016 was 26.7% with 15.2 million deaths out of a total of 56.9 million due to ischemic heart disease and strokes belonging to the CVD group.
- The development of CVD is promoted by several risk factors (RF), some of these are modifiable and subject to prevention. The WHO indicates evidence-based guidelines on how to reduce the incidence, whether for the first time or recurrently, of coronary heart disease, cerebrovascular disease and peripheral vascular disease.
- The Colombian Ministry of Health registers CVD among the ten main causes of mortality and an increase in these incidences is projected.
- In the department of Atlántico, the 2018 mortality rate due to ischemic heart disease and cerebrovascular accidents represents 25.60%, according to DANE figures.
- In developing countries, where the poorest populations are affected, reducing the years of productive life and increasing the costs of health services and hindering the social development of the communities.

**Objective:****General**

- Propose a model based on data analysis to categorize the level of cardiovascular risk that supports the generation of prevention and control plans for these diseases in the department of Atlántico.

**Specific**

- Identify the theoretical and epistemological foundations associated with the prognosis of cardiovascular risks.
- Formulate the data analysis model that allows categorizing the level of cardiovascular risk in the department.
- Apply the theoretical model built using medical records from the Atlántico department to validate its operation and future early warning applications.

**Materials and Methods:** The research presented a mixed approach, of an applied nature, with an experimental design and a population of 10,270 people, of which 6,218 people were taken in the sample.

**Results:** As results, the proposal and application of a cardiovascular risk estimation model in accordance with the context of the department of Atlántico stands out.

**Conclusions:** The validation of the model was effective, to the extent that the sample used for its development (Atlántico Department) had the previously established characteristics, in the implementation a high effort is required in the processing of the source of clinical data tailored to Each health entity, with the adoption of the interoperable clinical history, it is expected that this stage will be homogeneous and allow its use in any entity, in terms of its future operation, the model represented an advance within the context in which it is found due to that this modeling had not been previously carried out within the Department, allowing the inclusion of technological solutions in the cardiovascular analysis strategies of the population.

**KeyWords:** Cardiovascular disease, cardiovascular risk factors, cardiovascular risk, cardiovascular prediction models

**REFERENCIAS**

1. Alagona, P., & Ahmad, T. A. (2015). Cardiovascular Disease Risk Assessment and Prevention. Current Guidelines and Limitations. Medical Clinics of North America, 99(4), 711–731. <https://doi.org/10.1016/j.mcna.2015.02.003>
2. Alonso Lanza, J. L. (2005). La historia clínica electrónica: ideas, experiencias y reflexiones. Acimed, 13(5), 1-1.

3. Alonso A, Krijthe BP, Aspelund T, et al. (2013). Simple risk model predicts incidence of atrial fibrillation in a racially and geographically diverse population: the CHARGE-AF consortium. *J Am Heart Assoc.* 2013;2 (2): e000102.
4. Amaya, R. (2007). *La investigación en la práctica educativa: Guía metodológica de investigación para el diagnóstico y evaluación en los centros docentes.* Madrid: Secretaría General Técnica.
5. Analysts, B., Howson, C., Richardson, J., Sallam, R., & Kronz, A. (2020). *Magic Quadrant for Analytics and Business Intelligence Platforms.* Gartner, Inc.: Stamford, CT, USA, 1-58.
6. Arias, F. (2006). *Introducción a la metodología científica.* Caracas-Venezuela: Episteme.
7. Assman G, Cullen P, Schulte H (2002). Simple scoring scheme for calculating the risk of acute coronary events based on the 10-years follow-up of the Prospective Cardiovascular Münster Study (PROCAM). *Circulation.* 2002; 105: 310-5
8. Balakumar, P., Maung-U, K., & Jagadeesh, G. (2016). Prevalence and prevention of cardiovascular disease and diabetes mellitus. *Pharmacological Research,* 113, 600–609. <https://doi.org/10.1016/j.phrs.2016.09.040>
9. Berry, M., & Linoff, G. (1997). *Data mining techniques: For marketing, sales, and marketing support.*
10. Best, J. W. (1982). *Cómo investigar en educación.* España: Ediciones Morata
11. Blakley, B., McDermott, E., & Geer, D. (2001, September). Information security is information risk management. In *Proceedings of the 2001 workshop on New security paradigms* (pp. 97-104).
12. Brotons, C., Moral, I., Fernández, D., Puig, M., Bonacho, E. C., Muñoz, P. M., ... & Aguado, L. J. Q. (2019). Estimation of lifetime risk of cardiovascular disease (IBERLIFERISK): a new tool for cardiovascular disease prevention in primary care. *Revista Española de Cardiología (English Edition),* 72(7), 562-568.
13. Campos de Aldana, M. S., Moya Plata, D., Mendoza Matajira, J. D., & Duran Niño, E. Y. (2014). Chronic Non-Communicable Diseases and Use of Information and Communication Technology: Systematic Review. *Revista Cuidarte,* 5(1), 661–669.
14. Cazallo, A., Meñaca, I., Lechuga, J., Medina, H., Uribe, C., & Barragán, C. (2019). Modeling foreign investment received in the Colombian oil sector during the period 1996-2016. *Revista Espacios,* 40(20), 13-23.
15. Celermajer, D. S., Chow, C. K., Marijon, E., Anstey, N. M., & Woo, K. S. (2012). Cardiovascular disease in the developing world: Prevalences, patterns, and the potential of early disease detection. *Journal of the American College of Cardiology,* 60(14), 1207–1216. <https://doi.org/10.1016/j.jacc.2012.03.074>

16. Chi, M., Plaza, A., Benediktsson, J. A., Sun, Z., Shen, J., & Zhu, Y. (2016). Big data for remote sensing: Challenges and opportunities. *Proceedings of the IEEE*, 104(11), 2207-2219.
17. Conpes, D. N. P. 113 de 2008. Política de seguridad alimentaria y nutricional.
18. Córdova-Villalobos, J. Á., Barriguete-Meléndez, J. A., Lara-Esqueda, A., Barquera, S., Rosas-Peralta, M., & Hernández-Ávila, M. (2008). Chronic noncommunicable diseases in Mexico: epidemiological synopsis and comprehensive prevention [Las enfermedades crónicas no transmisibles en México: sinopsis epidemiológica y prevención integral]. *Salud Publica de Mexico*, 50(5), 419–427. <https://doi.org/10.1590/S0036-36342008000500015>
19. Cullen Paul, Helmut Schulte, and Gerd Assmann. 1997. "The Münster Heart Study (PROCAM)." 96(7):2128–36
20. DANE. (2019). Boletín Técnico Estadísticas Vitales (EEVV) Tercer trimestre 2019 pr. Retrieved from [https://www.dane.gov.co/files/investigaciones/poblacion/bt\\_estadisticasvitales\\_IIItrim\\_2019pr-20-diciembre-2019.pdf](https://www.dane.gov.co/files/investigaciones/poblacion/bt_estadisticasvitales_IIItrim_2019pr-20-diciembre-2019.pdf)
21. De Barrera, J. H., & Morales, M. F. B. (2000). Metodología de la investigación holística. Instituto Universitario de Tecnología Caripito.
22. DE PRINCIPIOS, D. E. C. L. A. R. A. C. I. Ó. N. LEY 23 DE 1981.
23. De Salud Pública, P. D. (2013). Pública 2012-2021. Bogotá: Ministerio de Salud y Protección Social.
24. Díaz-Realpe, J. E., Muñoz-Martínez, J., & Sierra-Torres, C. H. (2007). Factores de riesgo para enfermedad cardiovascular en trabajadores de una institución prestadora de servicios de salud, Colombia. *Revista de Salud pública*, 9, 64-75.
25. Eysenbach, G. (2001). What is e-health? *Journal of medical Internet research*, 3(2), e20.
26. Francula-Zaninovic, S., & Nola, I. A. (2018). Management of Measurable Variable Cardiovascular Disease' Risk Factors. *Current Cardiology Reviews*, 14(3), 153–163. <https://doi.org/10.2174/1573403x14666180222102312>
27. García, J., Cazallo, A., Barragan, C. E., Mercado, M., Olarte, L., Meza, V. (2019). Indicadores de Eficacia y Eficiencia en la gestión de procura de materiales en empresas del sector construcción del Departamento del Atlántico, Colombia. *Revista ESPACIOS*, 40(22).
28. Gawryszewski, V. P., & de Souza, M. de F. M. (2014). Mortality due to cardiovascular diseases in the Americas by region, 2000-2009. *Sao Paulo Medical Journal*, 132(2), 105–110. <https://doi.org/10.1590/1516-3180.2014.1322604>
29. Goff DC Jr, Lloyd-Jones DM, Bennett G, et al; American College of Cardiology/American Heart Association Task Force on Practice Guidelines. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Circulation*. 2014; 129(25) (suppl 2): S49-S73.

30. Goldstein, B. A., Navar, A. M., & Pencina, M. J. (2016). Risk prediction with electronic health records: the importance of model validation and clinical context. *JAMA cardiology*, 1(9), 976-977.
31. González, P.P., Desarrollo de técnicas de minería de datos en procesos industriales: Modelización en líneas de producción de acero. Julio de 2006: Universidad de la Rioja
32. González, M., & Salazar, F. A. (2009). Aspectos Básicos del Estudio de Muestra y Población para la Elaboración de los Proyectos de Investigación (Doctoral dissertation).
33. González, Gabriela. (15 de febrero de 2021). Estudio prospectivo: características, ejemplos y limitaciones. Lifeder. Recuperado de <https://www.lifeder.com/estudio-prospectivo/>.
34. Hernández Sampieri, R., Fernández Collado, C., & Baptista Lucio, P. (2010). Metodología de la investigación.
35. Jiménez Mesino, J. E. (2019). Autopercepción del estado de salud en individuos con riesgo cardiovascular en Colombia (Masters thesis, Universidad del Norte).
36. Kazi, D. S., Penko, J. M., & Bibbins-Domingo, K. (2017). Statins for Primary Prevention of Cardiovascular Disease: Review of Evidence and Recommendations for Clinical Practice. *Medical Clinics of North America*, 101(4), 689–699. <https://doi.org/10.1016/j.mcna.2017.03.001>
37. Keylee, B., Florez, A., Lugo, V., Rubiano, A., Maestre, V., & Portillo, Y. (2019). Factores de riesgo cardiovascular en trabajadores de tres ciudades de la región caribe colombiana.
38. Kimball, R., & Caserta, J. (2004). The data warehouse ETL toolkit. John Wiley & Sons.
39. Kunstmann, S., de Grazia, R., & Gainza, D. (2012). Atherosclerosis en la mujer: factores de riesgo y prevención TT - Atherosclerosis in women: risk factors and its prevention. *Revista Chilena de Cardiología*, 31(2), 142–147. <https://doi.org/10.4067/S071885602012000200009>
40. LIBRES, C. D. A. T. (2005). Estudio sociológico y del conocimiento de los factores de riesgo de las enfermedades cardiovasculares en la Costa Caribe Colombiana (Estudio Caribe). *Revista Colombiana de Cardiología*, 12(3), 0120-5633.
41. Llanusa Ruiz, S. B., Rojo Pérez, N., Caraballosa Hernández, M., Capote Mir, R., & Pérez Piñero, J. (2005). Las tecnologías de información y comunicación y la gestión del conocimiento en el sector salud. *Revista Cubana de Salud Pública*, 31(3), 0-0.
42. López-Jaramillo, P., Sánchez, R. A., Díaz, M., Cobos, L., Bryce, A., Carrillo, J. Z. P., ... Zanchetti, A. (2013). Latin American consensus on hypertension in patients with diabetes type 2 and metabolic syndrome. *Journal of Hypertension*, 31(2), 223–238. <https://doi.org/10.1097/HJH.0b013e32835c5444>
43. Luna, D., Soriano, E., & González Bernaldo de Quirós, F. (2007). Historia clínica electrónica. *Rev. Hospital B Aires*, 27(2), 77-85.

44. Mannila, H. (2000). Theoretical frameworks for data mining. *ACM SIGKDD Explorations Newsletter*, 1(2), 30-32.
45. Mendivelso, F., & Rodríguez, M. RIESGO RELATIVO. <https://doi.org/10.26852/01234250.39>
46. Ministerio de salud. 2010. "Principales causas de mortalidad en Colombia." Retrieved from <https://www.minsalud.gov.co/salud/Paginas/Enfermedades-cardiovasculares.aspx>
47. Ministerio Salud y Protección Social. (2012). Plan Decenal de Salud Pública PDSP, 2012-2021. Retrieved from <https://www.minsalud.gov.co/plandecenal/Paginas/home2013.aspx>
48. Muñoz, O. M., Rodríguez, N. I., Ruiz, Á., & Rondón, M. (2014). Validación de los modelos de predicción de Framingham y PROCAM como estimadores del riesgo cardiovascular en una población colombiana. *Revista Colombiana de cardiología*, 21(4), 202-212.
49. Murillo, J. (2008). Sobre la metodología de investigación en estudios sobre el discurso de la cortesía: a propósito del empleo de cuestionarios de hábitos sociales. *Cortesía y conversación: De lo escrito a lo oral*, 53-71.
50. OPS/OMS. (2013). Plan de acción mundial para la prevención y el control de las enfermedades no transmisibles 2013-2020.
51. Padrón-Martínez MM, Perea-Martínez A, López Navarrete GE. (2016) Relación cintura/estatura, una herramienta útil para detectar riesgos cardiovascular y metabólico en niños. *Acta Pediatría Mex.*;37(5):297-301.
52. Piatetsky-Shapiro, G. y W. J Frawley. *Knowledge Discovery in Data bases*". AAI/MIT Press, 1991.
53. Quintero, V., Heredia Vizcaíno, D., & Sanmartín Mendoza, P. (2018). Datos, información, tendencias, tres miradas sobre un contexto cambiante. In Ediciones Universidad Simón Bolívar. <https://hdl.handle.net/20.500.12442/3948>.
54. República de Colombia. Ley 1335, de Estilos de Vida Saludables. Bogotá: Congreso de la Republica; 2009.
55. República de Colombia. Ley 1355, de Estilos de Vida Saludables. Bogotá: Congreso de la Republica; 2009.
56. República de Colombia. Ley 2015, Historia Clínica Electrónica Interoperable. Bogotá: Congreso de la Republica; 2020.
57. Rodríguez, D., de Voil, P., Rufino, M. C., Odendo, M., & van Wijk, M. T. (2017). To mulch or to munch? Big modelling of big data. *Agricultural Systems*, 153, 32-42.
58. Ruiz A. (2010) Control del riesgo cardiovascular en el anciano. Servicio de Medicina Interna. Hospital de la Serranía. Ronda. Málaga. [Internet]. Retrieved from <https://docplayer.es/22744334-Control-del-riesgo-cardiovascular-en-el-anciano-alberto-ruiz-cantero-servicio-de-medicina-interna-hospital-de-la-serrania-ronda-malaga.html>
59. Ruiz JS. (2012) Modelos europeos: PROCAM y SCORE. En: Control Global del Riesgo Cardio metabólico La disfunción endotelial como diana preferencial. Madrid: Díaz de Santos

60. Salazar, P., Manterola, C., Quiroz, G., García, N., Otzen, T., Mora, M., & Duque, G. (2019). Estudios de cohortes. 1ª parte. Descripción, metodología y aplicaciones. *Revista de cirugía*, 71(5), 482-493.
61. Sans Menéndez, S. (2000). Las enfermedades cardiovasculares: *Revista Panamericana de Salud P*, 7(1), 49–50. <https://doi.org/10.1590/S1020-49892000000100012>
62. Tamayo, T. (1997). Población y muestra. Tamayo y Tamayo.
63. Van Dalen, D. B., & Meyer, W. J. (2006). Estrategia de la investigación descriptiva. *Manual de técnica de la investigación educacional*.
64. Villarroel, R. H., Gómez, Y. M., & Krause, C. B. (2013). Incorporación de seguridad en el modelado conceptual de procesos extracción-transformación-carga. *Información tecnológica*, 24(6), 101-110.
65. Wolfert, S., Ge, L., Verdouw, C., & Bogaardt, M. J. (2017). Big data in smart farming—a review. *Agricultural Systems*, 153, 69-80.
66. World Health Organization. (2014). *Global Status Report on Noncommunicable Diseases 2014*. 298. Retrieved from <https://www.who.int/nmh/publications/ncd-status-report-2014/en/>.
67. World Health Organization. 2018. "The top 10 causes of death." Retrieved from <http://www.who.int/en/news-room/fact-sheets/detail/the-top-10-causes-of-death>
68. World Health Organization. 2018. "Enfermedades no transmisibles" Retrieved from <https://www.who.int/es/news-room/fact-sheets/detail/noncommunicable-diseases>