Estado del Arte

[1. Introducción al Problema 1](#_Toc175590925)

[2. Modelos de Predicción de Enfermedades Cardiovasculares 1](#_Toc175590926)

[2.1. Estudios Previos en Modelos Predictivos 2](#_Toc175590927)

[3. Factores de Riesgo Relevantes 2](#_Toc175590928)

[3.1. Identificación de Factores de Riesgo Tradicionales 2](#_Toc175590929)

[3.2. Factores de Riesgo Emergentes 3](#_Toc175590930)

[4. Técnicas de Aprendizaje Automático y Enfoques Actuales 3](#_Toc175590931)

[4.1. Técnicas Clásicas de Machine Learning 3](#_Toc175590932)

[4.2. Enfoques Más Recientes 4](#_Toc175590933)

[5. Limitaciones en los Estudios Previos 4](#_Toc175590934)

[6. Conclusiones del Estado del Arte 5](#_Toc175590935)

[7. Bibliografía 5](#_Toc175590936)

# 1. Introducción al Problema

Las enfermedades cardiovasculares (ECV) representan una de las principales causas de muerte a nivel mundial. Según datos de la Organización Mundial de la Salud (OMS), estas enfermedades son responsables de aproximadamente el 31% de las muertes globales. El impacto de las ECV no solo es devastador a nivel individual, sino también a nivel económico, dado el alto costo de los tratamientos y la pérdida de productividad en pacientes afectados.

Los factores de riesgo asociados con las ECV han sido ampliamente documentados. Entre los más destacados se incluyen la hipertensión, los niveles elevados de colesterol, el tabaquismo, la diabetes, la inactividad física y la obesidad. Estos factores son responsables de un gran porcentaje de los casos, lo que subraya la importancia de implementar estrategias predictivas para prevenir y tratar las ECV de manera más eficiente.

# 2. Modelos de Predicción de Enfermedades Cardiovasculares

## 2.1. Estudios Previos en Modelos Predictivos

A lo largo de los años, se han desarrollado diferentes modelos para la predicción del riesgo cardiovascular. Estos modelos permiten estimar la probabilidad de desarrollar enfermedades cardiovasculares en un periodo determinado, en base a factores de riesgo individuales.

* Framingham Risk Score (FRS): Es uno de los modelos más antiguos y utilizados para predecir ECV. Se basa en datos del Estudio de Framingham, que incluye variables como la edad, el género, la presión arterial, el colesterol total, el tabaquismo y la diabetes. Sin embargo, aunque es ampliamente utilizado, su limitación principal radica en que fue desarrollado sobre una población específica (Framingham, EE.UU.), lo que reduce su capacidad para generalizar a otras poblaciones con distintas características demográficas.

* QRISK Models: Esta serie de modelos avanzados incorpora una gama más amplia de variables, como el índice de masa corporal (IMC), los antecedentes familiares, y factores étnicos. QRISK ha sido validado en varias poblaciones y se ha mostrado superior al Framingham en cuanto a su capacidad para predecir el riesgo en diversas poblaciones. Su mayor personalización lo hace más efectivo para su uso clínico en el Reino Unido y Europa.
* Machine Learning en Predicción de ECV: En años recientes, los modelos de machine learning (ML) han revolucionado el campo de la predicción del riesgo cardiovascular. Estas técnicas permiten el uso de grandes volúmenes de datos y múltiples variables para mejorar la precisión de las predicciones. Algoritmos como Random Forest, Support Vector Machines (SVM) y Redes Neuronales han demostrado su capacidad para ofrecer predicciones más precisas comparadas con los modelos tradicionales, debido a su habilidad para manejar interacciones complejas entre variables.

# 3. Factores de Riesgo Relevantes

## 3.1. Identificación de Factores de Riesgo Tradicionales

Los factores de riesgo tradicionales han sido validados en numerosos estudios. Variables como la hipertensión, diabetes, niveles elevados de colesterol, tabaquismo, obesidad, y la edad son reconocidas como los predictores más comunes de las ECV. Estos factores han sido el pilar de los modelos de predicción, como el Framingham Risk Score y SCORE (Systematic Coronary Risk Evaluation).

Sin embargo, aunque estos factores son efectivos en poblaciones amplias, estudios más recientes sugieren que algunos factores no tradicionales, como la inflamación crónica o ciertos marcadores genéticos, podrían mejorar significativamente las predicciones, ofreciendo una mejor personalización del riesgo.

## 3.2. Factores de Riesgo Emergentes

Con los avances en la investigación médica, han surgido nuevos factores de riesgo que podrían complementar los modelos tradicionales. Algunos de estos factores incluyen:

* Datos genéticos y biomarcadores: Se han desarrollado investigaciones que destacan la relevancia de incluir la genética en los modelos predictivos. Los avances en la secuenciación del genoma y la identificación de biomarcadores específicos podrían aportar un nuevo nivel de precisión en las predicciones.

* Datos de dispositivos wearables: La tecnología de dispositivos portátiles, como los relojes inteligentes y pulseras de actividad física, que monitorizan constantemente la frecuencia cardíaca y otros parámetros de salud, proporciona una fuente continua de datos que puede ser utilizada para refinar aún más los modelos predictivos.

Además, es importante tener en cuenta la heterogeneidad de las poblaciones. Factores de riesgo que son críticos en una población pueden no ser igualmente relevantes en otras, lo que subraya la necesidad de adaptar los modelos a contextos locales.

# 4. Técnicas de Aprendizaje Automático y Enfoques Actuales

## 4.1. Técnicas Clásicas de Machine Learning

Las técnicas clásicas de machine learning han sido ampliamente aplicadas para la predicción de ECV. Entre las más comunes se encuentran:

* Regresión logística
* k-Nearest Neighbors (k-NN)
* Random Forest
* Support Vector Machines (SVM)
* Árboles de decisión

Estos modelos se han evaluado a través de métricas como la precisión, la sensibilidad y el área bajo la curva ROC (AUC), lo que ha permitido determinar su efectividad en predicciones a corto y largo plazo.

## 4.2. Enfoques Más Recientes

* Deep Learning y Redes Neuronales: Estas técnicas, que permiten procesar grandes volúmenes de datos de manera más eficiente, han demostrado un gran potencial. No obstante, una de sus principales limitaciones es la interpretabilidad, lo que ha dificultado su adopción en entornos clínicos donde se requiere una explicación clara de los resultados.

* AutoML (Automated Machine Learning): Herramientas de AutoML están ganando terreno ya que permiten automatizar el proceso de selección de modelos y parámetros óptimos para conjuntos de datos específicos, facilitando la creación de modelos predictivos sin requerir un conocimiento profundo en ciencia de datos.
* Explainable AI (XAI): Dado que los modelos de machine learning tienden a ser “cajas negras”, los esfuerzos en torno a inteligencia artificial explicable han cobrado importancia. Se están desarrollando enfoques que buscan hacer que los modelos sean más interpretables para los clínicos, quienes necesitan comprender cómo un modelo llegó a una predicción para tomar decisiones más informadas.

# 5. Limitaciones en los Estudios Previos

Aunque los avances en la predicción de enfermedades cardiovasculares son prometedores, los estudios previos presentan ciertas limitaciones:

* Calidad de los Datos: La exactitud de los modelos predictivos depende de la calidad de los datos utilizados. Los datos incompletos, sesgados o provenientes de poblaciones limitadas pueden reducir la precisión de los modelos y su capacidad para generalizar a otros grupos de personas.

* Transferencia de Resultados: Los modelos desarrollados en una población pueden no ser aplicables o transferibles a otra, especialmente si existen diferencias significativas en cuanto a genética, estilo de vida o factores socioeconómicos.

* Interpretabilidad vs. Precisión: Existe una tensión inherente entre la precisión de los modelos avanzados (como las redes neuronales profundas) y la facilidad de interpretar los resultados. En el contexto clínico, es crucial que los médicos puedan entender cómo y por qué se generó una predicción, lo que ha impulsado el desarrollo de métodos explicables.

# 6. Conclusiones del Estado del Arte

La revisión realizada muestra que, aunque los modelos tradicionales de predicción de enfermedades cardiovasculares, como el Framingham Risk Score, han sido ampliamente utilizados y validados, las técnicas de machine learning ofrecen una oportunidad única para mejorar la precisión y la personalización de las predicciones.

Nuestro estudio se posiciona dentro de esta tendencia, buscando mejorar la predicción de ECV mediante el uso de nuevas técnicas de aprendizaje automático. Al explorar nuevos factores de riesgo o mejorar la precisión en poblaciones específicas, buscamos avanzar en la capacidad de predicción y, en última instancia, en la prevención de enfermedades cardiovasculares en un contexto clínico.

# 7. Bibliografía

* **Organización Mundial de la Salud (OMS)**, Observatorio de Salud Global, enfermedades cardiovasculares.
* **Hippisley-Cox, J., & Coupland, C.** (2017). Desarrollo y validación de los algoritmos de predicción de riesgo QRISK3 para estimar el riesgo futuro de enfermedad cardiovascular: estudio de cohorte prospectivo. *BMJ*, 357.
* **D'Ascenzo, F., et al.** (2019). Predicción basada en aprendizaje automático de eventos adversos tras un síndrome coronario agudo (PRAISE): un estudio de modelado de conjuntos de datos agrupados. *The Lancet*, 394(10206), 705-711.
* **Khera, A. V., et al.** (2018). Predicción poligénica de trayectorias de peso y obesidad desde el nacimiento hasta la adultez. *Cell*, 173(7), 1695-1710.
* **Rawshani, A., et al.** (2017). Factores de riesgo, mortalidad y resultados cardiovasculares en pacientes con diabetes tipo 2. *New England Journal of Medicine*, 376(15), 1407-1418.
* **Rajkomar, A., et al.** (2018). Aprendizaje profundo escalable y preciso con registros electrónicos de salud. *NPJ Digital Medicine*, 1(1), 1-10.
* **Lundberg, S. M., & Lee, S. I.** (2017). Un enfoque unificado para interpretar las predicciones de modelos. *Advances in Neural Information Processing Systems*, 4765-4774.