

**Modelos predictivos de machine learning para estimar la sobrevida del injerto renal,
basados en variables clínicas de la población colombiana**

Yermis Carolina Rocha Arrieta

Asesor

Sixto Enrique Campaña Bastidas

Universidad Nacional Abierta y a Distancia UNAD

Escuela de Ciencias Básicas, Tecnología e Ingeniería ECBTI

Especialización en Ciencia de Datos y Analítica

2025

Resumen

El trasplante renal es la mejor opción terapéutica para pacientes con enfermedad renal terminal. La adecuada selección de la pareja donante-receptor es fundamental para optimizar la sobrevida tanto del injerto como del paciente. En los últimos años, se han propuesto diversos modelos predictivos con el fin de mejorar el aprovechamiento de los órganos disponibles y los desenlaces clínicos postrasplante.

Este estudio tuvo como objetivo desarrollar y evaluar modelos de machine learning para predecir la sobrevida del injerto renal, utilizando un conjunto de datos construido a partir de variables clínicas de una cohorte colombiana, conformada por 2,012 pacientes.

Se entrenaron y compararon dos enfoques: el modelo de riesgos proporcionales de Cox y el modelo de Bosques Aleatorios de Supervivencia. El modelo de Bosques Aleatorios mostró un rendimiento superior (IC = 0,61, IBS = 0,1596) en comparación con el modelo de Cox (IC = 0,58, IBS= 0,1636). Además, permitió la estratificación de pacientes en grupos de riesgo alto y bajo, con diferencias estadísticamente significativas ($p = 0.006$). Estos resultados sugieren que los modelos no paramétricos pueden capturar relaciones complejas y no lineales entre las variables, ofreciendo ventajas frente a los métodos estadísticos tradicionales. Así mismo, este trabajo constituye un paso importante hacia el desarrollo de herramientas predictivas adaptadas al contexto local y al perfil de los pacientes trasplantados en Colombia.

Palabras claves: Machine learning, trasplante, riñón, sobrevida del injerto.

Abstract

Kidney transplantation is the best therapeutic option for patients with end-stage renal disease. Proper matching between donor and recipient is crucial to optimize both graft and patient survival. In recent years, several predictive models have been proposed to improve organ allocation and post-transplant outcomes.

This study aimed to develop and evaluate machine learning models to predict kidney graft survival using a dataset constructed from clinical variables of a Colombian cohort of 2,012 patients.

Two approaches were trained and compared: the Cox proportional hazards model and the Random Survival Forest model. The Random Survival Forest model showed superior performance (C-index = 0.61, IBS = 0.1596) compared to the Cox model (C-index = 0.58, IBS = 0.1636). It also enabled the stratification of patients into high- and low-risk groups, with statistically significant differences ($p = 0.006$). These results suggest that non-parametric models can capture complex and non-linear relationships among variables, offering advantages over traditional statistical methods. This work also represents an important step toward the development of predictive tools adapted to the local context and the clinical profile of transplant patients in Colombia.

Keywords: Machine learning, transplantation, kidney, graft survival.

Tabla de Contenido

Introducción	8
Planteamiento del Problema	10
Justificación	12
Objetivos	14
Objetivo General	14
Objetivos Específicos	14
Marco Teórico.....	15
Criterios y Herramientas en la Evaluación de Riñones para Trasplante	15
Machine Learning y Predicción de la Supervivencia del Injerto Renal.....	18
Asignación de Órganos en Colombia.....	20
Metodología	22
Tipo de Estudio	22
Fuente de Información	22
Consideraciones Éticas.....	22
Método	23
Resultados	25
Identificación de Variables que Afectan la Sobrevida del Injerto Renal	25
Preparación de los Datos.....	26
Codificación y Escalamiento de Variables	28
Características Demográficas de la Población	28
Modelado	28
Evaluación de los Modelos	31

Discusión.....	36
Conclusiones y Recomendaciones	41
Referencias Bibliográficas	42

Lista de Tablas

Tabla 1 <i>Clasificación de Variables y Porcentaje de Datos Nulos</i>	26
Tabla 2 <i>Conjunto Final de Variables Clínicas</i>	30
Tabla 3 <i>Evaluación del Rendimiento de los Modelos Mediante el Índice de Concordancia</i>	32

Lista de Figuras

Figura 1 <i>Curvas de Kaplan-Meier por Grupos de Riesgo según el Modelo de RSF</i>	33
Figura 2 <i>AUC Acumulada Dinámica</i>	34
Figura 3 <i>Importancia de Variables por Permutación</i>	35

Introducción

La enfermedad renal crónica (ERC) es un problema salud pública de gran importancia. Se estima que alrededor del 10 % de la población mundial padece ERC, y en los últimos años se ha observado un incremento en su incidencia, en parte debido al aumento de factores de riesgo como la Diabetes Mellitus y la Hipertensión Arterial (Kovesdy, 2022).

En Colombia, la incidencia de ERC en 2022 fue de 1,5 por cada 100 habitantes, y solo 1,7 de cada 1.000 pacientes accede a un trasplante renal, a pesar de que este procedimiento es la mejor opción terapéutica, ya que mejora tanto la calidad de vida como la supervivencia en comparación con la diálisis (Fondo Colombiano de Enfermedades de Alto Costo, 2023; Lasserre et al., 2012; Nino-Torres et al., 2021). Similar a lo que pasa a nivel mundial, en Colombia la demanda de órganos supera ampliamente la oferta.

Por lo anterior, se han desarrollado métricas como el Kidney Donor Profile Index (KDPI) o el Estimated Post-Transplant Survival (EPTS), implementadas en los Estados Unidos (OPTN, 2020, 2024a). Estas herramientas han permitido reducir el descarte riñones y mejorar el aprovechamiento de los órganos disponibles, incrementando así la supervivencia tanto del injerto como del paciente (OPTN, 2020, 2024a).

En Colombia, no se cuenta con métricas similares construidas a partir de variables clínicas locales. Por esta razón, el objetivo de este estudio fue aplicar modelos de machine learning para predecir la sobrevida del injerto renal, siguiendo la metodología CRISP-DM (Cross Industry Standard Process for Data Mining). Para ello, se recopilaron datos clínicos de pacientes trasplantados en dos instituciones trasplantadoras del país. Tras un proceso riguroso de limpieza, la base de datos fue utilizada para entrenar y evaluar dos modelos de supervivencia: el modelo de

riesgos proporcionales de Cox y el modelo de Bosques Aleatorios de Supervivencia (RSF, por sus siglas en inglés), siendo este último el que mostró un mejor desempeño.

Una de las principales limitaciones del estudio fue el tamaño de la muestra. Aunque inicialmente se contaba con 2.621 pacientes, aproximadamente el 23 % de los registros debieron ser excluidos debido a la ausencia de información en variables clave, lo cual restringe la capacidad de generalización de los modelos y puede introducir sesgos en los resultados.

No obstante, este trabajo representa un esfuerzo por incorporar enfoques analíticos avanzados en el contexto colombiano, con el propósito de desarrollar herramientas que apoyen la toma de decisiones clínicas con base a la realidad epidemiológica. A su vez, ha permitido la formación de una especialista en ciencia de datos con formación previa en el área de la salud, contribuyendo con una visión interdisciplinaria al desarrollo y aplicación de modelos predictivos en la práctica clínica nacional.

Planteamiento del Problema

El trasplante renal presenta varios desafíos uno de ellos es la escasez de órganos, a pesar de los grandes esfuerzos que se realizan se estima que menos del 10 % de las necesidades globales de trasplantes de órganos sólidos son satisfechas anualmente (Weiss et al., 2023).

Factores como el envejecimiento de la población, enfermedades crónicas, a temprana edad, como la diabetes e hipertensión y la mejora de los protocolos de cuidado a pacientes neurocríticos, hacen más compleja la decisión de ofertar un órgano. En Colombia, sólo un 10% de los potenciales donantes, se convierten en donantes reales, disminuyendo la posibilidad de que los pacientes en lista de espera sean trasplantados (Mususú et al., 2024).

En el año 2023 se realizaron alrededor de 1.000 trasplantes de órganos sólidos provenientes de donantes fallecidos; sin embargo al 31 de diciembre del mismo año, 4.002 pacientes seguían en lista de espera de un órgano sólido. Durante este mismo periodo, 106 pacientes murieron esperando un trasplante (Instituto Nacional de Salud, 2023) .

A nivel mundial, el panorama no es muy distinto, en Estados Unidos, mueren en promedio 13 personas cada día esperando un órgano, y aproximadamente 100.000 se encuentran en lista de espera (Health Resources & Services Administration, 2024).

Otro reto, en el área del trasplante renal es la sobrevida del injerto a largo plazo. Si bien los avances científicos y el uso de nuevos medicamentos inmunosupresores han logrado disminuir las tasas de rechazo agudo, el rechazo crónico continúa siendo una preocupación constante para los clínicos (Holt, 2017). Al año del trasplante, la sobrevida del injerto renal es alta (entre el 90 % y el 95 %); sin embargo, esta disminuye progresivamente con el tiempo, alcanzando aproximadamente el 60 % a los 10 años, debido en parte al rechazo crónico y a otras

complicaciones postrasplante (Maggiani-Aguilera et al., 2021; OPTN, 2024a; Poggio et al., 2021).

El rechazo crónico es un proceso irreversible y multifactorial en el cual el tejido funcional del injerto es reemplazado por tejido fibrótico, como resultado de la respuesta inmune, la isquemia inicial, los efectos adversos de los inmunosupresores, entre otros factores, que conduce finalmente a la pérdida del injerto y al reingreso del paciente a la lista de espera, con los desafíos adicionales que implica un retrasplante (Abbas et al., 2021).

Ante estas dificultades, predecir la sobrevida del injerto es de gran valor para la asignación de órganos, porque permite seleccionar la mejor pareja donante-receptor y en consecuencia esto contribuye a mejorar los resultados del trasplante, disminuyendo el número de pacientes que deben regresar a lista de espera y favoreciendo un uso más eficiente de los órganos disponibles (Moghadam & Ahmadi, 2022).

En Colombia, no se cuenta con modelos predictivos ajustados a las características clínicas de su población que apoyen a los clínicos en el análisis de múltiples variables del donante y del receptor, aspectos fundamentales al momento de aceptar o rechazar un órgano y seleccionar la mejor pareja donante-receptor.

En consecuencia, surge la necesidad de explorar cómo los modelos de machine learning, contruidos a partir de variables clínicas representativas de la población colombiana, pueden contribuir a predecir la sobrevida del injerto renal. Este acercamiento permitiría avanzar hacia decisiones más personalizadas, objetivas y eficaces en el proceso de asignación y trasplante de órganos en el país.

Justificación

El trasplante renal es la terapia por excelencia para los pacientes con ERC en estadio terminal, ya que mejora la calidad de vida y ofrece una mayor sobrevida en comparación con quienes permanecen en diálisis (Ghelichi-Ghojogh et al., 2022).

A nivel mundial, hace ya varios años que se han propuesto y utilizado modelos predictivos que permiten maximizar la supervivencia del injerto renal y optimizar el uso de los órganos disponibles (OPTN, 2020, 2024a). Sin embargo, en Colombia, la toma de decisiones sobre la aceptación de un órgano o la elección de la mejor pareja donante-receptor se basa en unas pocas variables y en el juicio clínico, sin el apoyo de modelos objetivos basados en datos.

Estudios recientes en la población colombiana sugieren que el uso de estas herramientas contribuiría con una evaluación más exacta de la calidad del donante y la probabilidad de sobrevida del paciente y del injerto, lo cual, a su vez, podría aumentar el número de trasplantes realizados y reducir la tasa de descarte de órganos viables (Nino-Torres et al., 2021; Vanin A et al., 2023).

La importancia de este trabajo radica en que, mediante el uso de técnicas de machine learning, se pueden identificar patrones complejos entre múltiples variables clínicas del donante y del receptor que influyen en la sobrevida del injerto, permitiendo la construcción de modelos predictivos personalizados. Este enfoque no solo es aplicable al trasplante renal, sino que también podría extenderse a otros tipos de trasplante, como el hepático o el cardíaco (Briceño et al., 2022; Naruka et al., 2022). Implementarlo en el contexto colombiano permitiría optimizar el uso de los órganos disponibles y, al mismo tiempo, mejorar los resultados individuales, contribuyendo así a una mayor equidad y eficiencia en el sistema de salud.

Finalmente, este estudio representa una valiosa oportunidad para generar conocimiento basado en datos locales que puede respaldar las decisiones clínicas en trasplante renal. Así mismo, permite la formación de una estudiante especialista en ciencia de datos con formación previa en el área de la salud, aportando una visión práctica e interdisciplinaria al desarrollo de herramientas predictivas aplicables al contexto clínico colombiano.

Objetivos

Objetivo General

Aplicar modelos de machine learning para predecir la sobrevida del injerto renal, utilizando un dataset construido a partir de variables clínicas de la población colombiana, que influyen en el éxito del trasplante.

Objetivos Específicos

Identificar las variables que afectan la sobrevida del injerto renal mediante una búsqueda bibliográfica general, priorizando aquellas relevantes para la población colombiana, con el fin de aplicarlas en modelos predictivos.

Construir un dataset mediante la selección de variables relevantes, validación de características y aplicación de técnicas de reducción de dimensionalidad, con el fin de entrenar modelos de machine learning que predigan la sobrevida del injerto renal en la población colombiana.

Evaluar modelos de machine learning para predecir la sobrevida del injerto renal en la población colombiana, seleccionando el modelo más adecuado con base en su rendimiento y capacidad predictiva.

Marco Teórico

El trasplante renal tiene la capacidad de mejorar la calidad de vida y la supervivencia de los pacientes que tienen enfermedad renal terminal (Axelrod et al., 2018; Bertsimas et al., 2013). Desafíos como la escasez de órganos, el número creciente de pacientes en lista de espera y los costos asociados al proceso de trasplante requieren una meticulosa evaluación de la calidad del riñón y viabilidad del trasplante, con el fin de reducir el descarte de órganos y asignar la mejor pareja donante-receptor que permita el mayor y mejor beneficio para el paciente (Maggiani-Aguilera et al., 2021).

Criterios y Herramientas en la Evaluación de Riñones para Trasplante

Predecir la supervivencia del injerto es un factor clave en el proceso de trasplante y para ello se han propuesto diferentes metodologías (Topuz et al., 2018). Entre ellas, los modelos de Riesgos Proporcionales de Cox y el estimador de Kaplan-Meier, ampliamente utilizadas.

El Índice de Riesgo del Donante Renal (KDRI, Kidney Donor Risk Index) fue desarrollado a partir de un estudio publicado en el 2009, que analizó los datos de 69 440 pacientes adultos sometidos a su primer trasplante de riñón con donante fallecido en Estados Unidos. Utilizando el modelo de riesgos proporcionales de Cox, el estudio evaluó el riesgo de muerte o pérdida del injerto en función de múltiples factores del donante, del receptor y del procedimiento del trasplante (Rao et al., 2009).

El KDRI propuesto incluía 14 variables que estaban independientemente asociadas con la falla del injerto o la muerte, estas son: Edad del donante, raza, altura, peso, antecedentes de diabetes, antecedentes de hipertensión, creatinina sérica, causa de muerte cerebrovascular, estado de donación después de la muerte circulatoria, estatus del virus de la hepatitis C (VHC), número incompatibilidades del Antígeno Leucocitario Humano (HLA, por sus siglas en inglés), HLA-B

y HLA -DR, tiempo de isquemia fría y trasplante doble o en bloque. En el 2014, la Red de Obtención y Trasplante de Órganos (OPTN, Organ Procurement and Transplantation Network) de Estados Unidos, como parte de la implementación del Sistema de asignación renal (KAS, Kidney Allocation System) incluyó el KDPI, pero solo contenía 10 variables (OPTN, 2024a) .

El KDPI se basa en el KDRI, y representa el riesgo asociado al donante en un valor porcentual, que va de 0 a 100, comparando el riesgo de cada órgano donado con los riesgos observados en la población de donantes (OPTN, 2024a). Un puntaje bajo en el KDPI se asocia a una función estimada más prolongada del injerto, y viceversa, un KDPI más alto se asocia a una función estimada más corta.

Hay cuatro categorías del KDPI, estas son: 0-20%, 21-34%, 35-85%, 86-100%. Riñones con un KDPI de 0-20%, se espera que funcionen en promedio 11 años y medio, después del trasplante, comparado a los 12 años de riñones provenientes de donantes vivos(OPTN, 2024a). No obstante, es importante destacar que la supervivencia del injerto también depende de otros factores específicos del paciente como la edad, diagnóstico primario, comorbilidades, incompatibilidades del HLA, adherencia al tratamiento inmunosupresor, entre otros.

En el año 2024, el KDPI actualizó las variables utilizadas pasando de 10 a 8, donde se excluyó del cálculo la raza y el estatus del VHC, esto porque los riñones de donantes fallecidos afroamericanos/negros y positivos para el VHC tienen un KDPI superior al 20%. La raza es un mal indicador de la variación genética humana y, debido a las innovaciones en el tratamiento, los resultados posteriores al trasplante de riñón de donantes fallecidos positivos para el VHC son similares a los de los riñones de donantes negativos para el VHC. Por lo tanto, se consideró que estos factores inflaban el KDPI y no reflejaban la mejor probabilidad de falla del injerto para los riñones provenientes de esta población(OPTN, 2024b).

Las ocho variables con las que cuenta el modelo actual son: Edad, estatura, peso, antecedentes de diabetes, antecedentes de hipertensión, causa de muerte, creatinina sérica, estado de donación después de la muerte circulatoria (OPTN, 2024b).

Otra métrica utilizada en el KAS es el EPTS, también desarrollada bajo el modelo de Cox, utilizada para cuantificar las asociaciones entre cuatro variables del candidato adulto y el tiempo hasta su muerte postrasplante. Las variables que se tienen en cuenta son la Edad, tiempo en diálisis, diagnóstico de diabetes, y si fue o no previamente donante de órganos sólidos(OPTN, 2020).

Cada paciente en lista de espera recibe un puntaje del EPTS que va de 0-100%. El EPTS se usa en conjunto con el KDPI para asegurar un emparejamiento donde la función del injerto sea más longeva, es lo comúnmente conocido como el “longevity matching”. El uso de ambas métricas sólo se utiliza cuando el donante tiene un KDPI $\leq 20\%$ para priorizar a los pacientes con un puntaje de EPTS $\leq 20\%$, con el fin de obtener el mayor beneficio de los riñones de mayor calidad en los pacientes con una expectativa de vida mayor (OPTN, 2020).

Estas métricas no se utilizan por separado sino en conjunto con otras, como el tiempo en lista de espera, PRA calculado (cPRA, Calculated Panel Reactive Antibodies), compatibilidad del HLA-DR, entre otros, para priorizar los trasplantes de acuerdo con la compatibilidad y la expectativa de supervivencia del injerto, mejorando así la asignación de riñones en función de los factores de riesgo específicos del donante y las características del receptor.

Este cambio en el sistema KAS ha sido fundamental para optimizar el uso de órganos disponibles, ya que ayuda a asegurar que los riñones de menor riesgo se asignen a receptores con mayor expectativa de vida del injerto, mientras que riñones con mayor riesgo se pueden destinar

a receptores con menos tiempo en lista de espera o mayor urgencia clínica. Así mismo, ha mejorado la supervivencia del injerto renal (OPTN, 2020, 2024a; Poggio et al., 2021).

Por otro lado, también se ha utilizado el modelo de Kaplan-Meier para predecir la supervivencia del injerto utilizando variables del paciente (edad, sexo, raza, compatibilidad HLA) y del proceso del trasplante (tiempo de isquemia fría, rechazo agudo, estado funcional del injerto postrasplante). No obstante, tanto el modelo de Kaplan-Meier como el de riesgos proporcionales de Cox presentan limitaciones en el contexto del trasplante renal, donde la sobrevida del injerto está determinada por relaciones complejas, interacciones y posibles no linealidades entre múltiples variables clínicas, del procedimiento y del postrasplante (Naqvi et al., 2021; Ravindhran et al., 2023; Senanayake et al., 2019; Topuz et al., 2018).

Machine Learning y Predicción de la Supervivencia del Injerto Renal

La inteligencia artificial (IA) ha venido a revolucionar la forma en la que trabajamos, incluyendo el área de la medicina, donde la IA puede analizar una gran cantidad de datos y transformarlos en recomendaciones clínicas.

El Machine learning es una rama de la IA, donde un algoritmo computacional aprende de una serie de datos o eventos pasados para luego generar predicciones reproducibles o clasificación de datos (Balch et al., 2021). El ML permite analizar una gran cantidad de datos y variables identificando asociaciones ocultas entre estas, que difícilmente podrían ser captadas por los métodos estadísticos tradicionales (Ravindhran et al., 2023; Senanayake et al., 2019).

Los cuatro métodos comúnmente utilizados en ML son el aprendizaje supervisado, el no supervisado, el semisupervisado y el aprendizaje por refuerzo. Los métodos de aprendizaje supervisados se utilizan para predecir resultados o eventos futuros, tal como la supervivencia del injerto renal (Balch et al., 2021).

En un metaanálisis realizado por Ravindhran y colaboradores evaluaron si los modelos de ML pueden mejorar la predicción de la supervivencia de un riñón trasplantado. En esta revisión se incluyeron 31 artículos a los cuales se les evaluó la calidad metodológica y el riesgo de sesgo, utilizando la herramienta de Quiao y el PROBAST, respectivamente. En los artículos incluidos se evaluó el área bajo la curva ROC (AUROC), sensibilidad y especificidad. También construyeron una curva ROC resumida jerárquica (HSROC, hierarchical summary ROC) para comparar el rendimiento de los modelos basados en ML versus los de regresión, para predecir la supervivencia del injerto a corto (menos de 1 año) y largo plazo (más de 3 años) (Ravindhran et al., 2023).

Ellos encontraron que la mayoría de los estudios incluyeron variables preoperatorias del donante y receptor, y se usaron 29 modelos diferentes de ML, entre ellos, Redes neuronales artificiales (ANN, Artificial Neural Networks), Árboles de decisión (DT, Decision Trees), Bosques aleatorios (RF, Random Forest), Máquinas de Vectores de Soporte (SVM, support vector machines) o modelos híbridos; y aunque los autores refieren que es un tema ampliamente debatido y controversial acerca del cual sería el mejor modelo, los resultados de la revisión sugieren que los modelos híbridos, variaciones del SVM y RF tuvieron el mejor desempeño en la predicción de la supervivencia del injerto (Ravindhran et al., 2023).

En términos generales, los modelos basados en ML tuvieron una curva HSROC significativamente más alta (0,82 para todos los estudios y 0,85 basado en el análisis de sensibilidad que incluyó solamente estudios de buena calidad metodológica, bajo riesgo de sesgo y validación de los datos en un dataset separado), una razón de probabilidades de diagnóstico más alta (18,24) y una curva AUROC de 0,82, rendimientos que hasta ahora no se han logrado con muchos modelos estadísticos tradicionales, incluyendo el KDPI cuyo índice de concordancia

está alrededor de 0,60, es decir, un poder predictivo moderado(OPTN, 2020, 2024a; Ravindhran et al., 2023) .

En contraste, un estudio publicado en el 2020 comparó el rendimiento de ML y el modelo de regresión convencional utilizando las mismas variables para predecir diferentes resultados del trasplante renal. Los autores concluyeron que ambos métodos tuvieron un rendimiento casi idéntico y que en general el ML no demostró una clara ventaja para predecir los resultados del trasplante renal. No obstante, no descartan que exista otro algoritmo de ML que supere a los algoritmos estudiados y que los resultados no se pueden generalizar a estudios que incluyan otro tipo de variables, especialmente información clínica no tabular, donde los modelos de ML han demostrado ser más ventajosos (Bae et al., 2020).

Asignación de Órganos en Colombia

En Colombia, la asignación de órganos está a cargo del Instituto Nacional de Salud, por medio de la Red Nacional de Donación y Trasplante de Órganos y Tejidos, y se realiza teniendo en cuenta criterios geográficos y técnico-científicos (Instituto Nacional de Salud, 2018).

Los criterios geográficos hacen referencia al orden de asignación de los órganos, el cual se realiza de la siguiente manera: en primer lugar, de forma local, es decir, se asigna a un receptor en la misma Institución Prestadora de Servicios de Salud (IPS) donde fue rescatado el órgano, siempre que cuente con un programa de trasplantes habilitado y se cumplan los criterios técnico-científicos establecidos. Si no es posible asignarlo localmente, se informará a la coordinación regional para buscar un receptor en otras instituciones habilitadas para trasplante dentro de la misma región. Finalmente, si tampoco hay un receptor compatible en la región, la coordinación nacional se encargará de asignarlo en otra región del país, conforme a los criterios técnicos de asignación establecidos (Ministerio de la Protección Social, 2004).

Los criterios técnico-científicos que se tienen en cuenta para la asignación de los órganos incluyen: compatibilidad de grupo sanguíneo, compatibilidad HLA, edad del receptor, estado compasivo, antecedente de donante vivo o manifestación positiva hacia la donación, y tiempo en lista de espera. A cada uno de estos criterios se les asigna un puntaje determinado, y los pacientes con mayor puntaje serán priorizados para recibir el trasplante (Instituto Nacional de Salud, 2018).

Si bien esta herramienta ha sido útil por muchos años, excluye otras variables que se ha demostrado impactan la sobrevida del injerto renal (Hickey et al., 2017; OPTN, 2020, 2024a; Sharif, 2022).

La supervivencia del injerto renal depende de múltiples factores, del donante, del receptor, del procedimiento del trasplante y también del programa de trasplante. El objetivo de este trabajo fue aplicar modelos de machine learning para predecir la sobrevida del injerto renal, utilizando un dataset construido a partir de variables clínicas de la población colombiana, que influyen en el éxito del trasplante.

Metodología

Tipo de Estudio

Estudio observacional retrospectivo, que analiza datos históricos de pacientes sometidos a trasplante renal, evaluando la sobrevida del injerto a través de modelos de machine learning.

Los criterios de inclusión fueron: receptores adultos (mayores de 18 años) que recibieron su primer trasplante renal. Los criterios de exclusión fueron: receptores menores de edad, pacientes con retrasplante y casos con información clínica incompleta en variables críticas para el modelo.

Fuente de Información

Los datos utilizados en este estudio fueron provistos por dos IPS especializadas en trasplantes de órganos, previa aprobación por parte de los comités de ética institucionales y la firma de acuerdos de confidencialidad y tratamiento de datos, conforme a lo establecido en las Leyes 23 de 1981, 1581 de 2012 y 1266 de 2008.

La información fue entregada de forma anonimizada, impidiendo la identificación de los pacientes.

Se recibieron 2 621 registros correspondientes a pacientes mayores de edad trasplantados con donantes vivos o fallecidos durante el periodo Agosto del 2007 - Julio 2023.

Consideraciones Éticas

Este proyecto fue aprobado por comité de ética y cumple con las recomendaciones de la Declaración de Helsinki, el informe de Belmont y la Resolución 8430 de 1993 del Ministerio de Salud “Normas Científicas, Técnicas y Administrativas para la Investigación en Salud”.

Método

Si bien existen diversas metodologías ampliamente reconocidas para el desarrollo de procesos de minería de datos como SEMMA (Sample, Explore, Modify, Model, Assess), KDD (Knowledge Discovery in Databases) y CRISP-DM (Cross-Industry Standard Process for Data Mining), en el presente estudio se optó por la aplicación del enfoque CRISP-DM. Esta elección se fundamenta en su versatilidad, su independencia frente al dominio industrial y su neutralidad respecto a las tecnologías empleadas, lo que la convierte en una metodología adecuada para el área médica (García, 2018; Niaksu, 2015).

Este proceso iterativo y dinámico incluye las siguientes fases:

1. **Comprensión del negocio:** Es el entendimiento del propósito del análisis de los datos. Se debe identificar cuáles son las necesidades, en este caso del área del trasplante renal en Colombia, y qué decisiones se espera apoyar con el modelo. En otras palabras, es el entendimiento del problema, con el fin de diseñar los objetivos y un plan para el desarrollo del proyecto (García, 2018; Niaksu, 2015).
2. **Comprensión de los datos:** Esta fase incluye la recolección inicial y la exploración de los datos disponibles. Se identifican las variables clínicas, se revisa integridad, consistencia y calidad de los datos (García, 2018; Niaksu, 2015).
3. **Preparación de los datos:** Esta etapa comprende un trabajo técnico de preprocesamiento para transformar los datos crudos en un conjunto de datos adecuado para el modelado. Incluye tratamiento de datos nulos o inconsistentes, detección y corrección de valores atípicos o clínicamente improbables. También se agrupan o reclasifican variables, codificación de variables categóricas, normalización o estandarización de variables y creación de nuevas variables (García, 2018; Niaksu, 2015).

4. Modelamiento: En esta fase se realiza la selección de las técnicas de modelado más apropiadas para el problema (clasificación, regresión, supervivencia, clustering) y se revisa nuevamente que los datos estén en el formato requerido por el modelo (García, 2018; Niaksu, 2015).

5. Evaluación: El objetivo es determinar si el modelo cumple con los criterios clínicos y técnicos definidos en la fase de comprensión del negocio.

Esta evaluación se realiza por medio de métricas como la sensibilidad, especificidad, exactitud, puntaje F1, índice de concordancia (IC), curvas ROC y AUC, entre otros.

También se debe evaluar la usabilidad y confiabilidad del modelo desde el punto de vista del usuario final (personal clínico) (Niaksu, 2015).

6. Implementación: En esta fase, el conocimiento generado se transforma en acciones concretas (García, 2018). Esto incluye presentar el modelo desarrollado de forma clara y comprensible al usuario final, con el objetivo de facilitar la toma de decisiones basadas en datos y apoyar su incorporación en la práctica clínica. Las decisiones derivadas con base en el resultado del modelo deben ser transparentes, éticas y complementarias al juicio clínico (Niaksu, 2015).

Resultados

Identificación de Variables que Afectan la Sobrevida del Injerto Renal

Se realizó una búsqueda en bases de datos científicas como PubMed, Embase y Scopus, con el fin de identificar las variables comúnmente utilizadas para predecir la sobrevida del injerto renal. Para la estrategia de búsqueda se emplearon las siguientes palabras claves: Kidney, Transplant, Machine Learning, Graft Survival.

Las variables identificadas fueron revisadas por profesionales clínicos con experiencia en trasplante renal. Por medio de una encuesta estructurada y aplicada mediante Google Forms, los expertos seleccionaron las variables que, según su criterio clínico, tienen mayor relevancia como predictores de la sobrevida del injerto. Se incluyeron las variables que tuvieron una votación igual o superior al 80%, tales como:

Donante: Edad, sexo, índice de masa corporal (IMC), creatinina sérica, presencia de diabetes, hipertensión arterial, causa de muerte, criterios expandidos y tipo de donante.

Receptor: Edad, sexo, IMC, tipo de diálisis, tiempo en diálisis, presencia de diabetes, hipertensión arterial, diagnóstico primario, número de incompatibilidades HLA (ABDR), y panel reactivo de anticuerpos (PRA).

Como variable intraoperatoria se incluyó el tiempo de isquemia fría.

La variable o evento de interés fue la sobrevida del injerto censurada por muerte, que es el tiempo entre el trasplante y el diagnóstico de falla del injerto. Los datos fueron censurados en tres situaciones: (i) en el último control registrado para pacientes con injerto funcional, (ii) en la fecha del último seguimiento en caso de pérdida de contacto, y (iii) en la fecha de fallecimiento para aquellos que murieron antes de presentarse la falla del injerto.

Preparación de los Datos

Se recibió un total de 2 621 registros correspondientes a donantes y receptores, con 22 variables clínicas y demográficas asociadas al proceso de trasplante renal. A continuación, se realiza una descripción de las variables junto con el porcentaje de datos faltantes identificados en estas:

Tabla 1

Clasificación de Variables y Porcentaje de Datos Nulos

Variable	Clasificación	Porcentaje de datos nulos, %
Receptor		
Edad (años)	Continua	0%
Sexo	Nominal	0%
IMC	Continua	20%
Tipo de diálisis	Nominal	22.59%
Tiempo en diálisis (Meses)	Continua	20.14%
Diabetes	Nominal	17.17%
Diagnóstico primario	Nominal	0.08%
Compatibilidades HLA (ABDR)	Ordinal	22.59%
PRA Clase I (PRA_CI)	Nominal	20.03%
PRA Clase II (PRA_CII)	Nominal	20.03%
Donante		
Edad (años)	Continua	0%
Sexo	Nominal	0%

Variable	Clasificación	Porcentaje de datos nulos, %
IMC	Continua	22.59%
Diabetes	Nominal	2.3%
Hipertensión arterial	Nominal	22.59%
Creatinina sérica (mg/dL)	Continua	22.59%
Causa de muerte	Nominal	18.85%
Criterios expandidos	Nominal	0%
Tipo de donante	Nominal	0%
Tiempo de Isquemia Fría(horas)	Continua	2.29%
Perdida de Injerto	Nominal	0%
Tiempo de Seguimiento (días)	Continua	0.9%

Durante el proceso de limpieza de datos, se eliminaron los registros de pacientes sin información de seguimiento clínico o sin registro del tiempo en diálisis, al considerarse variables clave para el análisis de supervivencia. Del mismo modo, y debido a su relevancia clínica y carácter no redundante, los valores faltantes de creatinina sérica no fueron imputados; por tanto, se excluyeron los registros correspondientes. Los datos con valores erróneos o atípicos fueron tratados mediante eliminación directa (por ejemplo, valores de índice de masa corporal superiores a 100) o mediante winsorización al percentil 1 y 99.

Dado que el objetivo del estudio fue evaluar la capacidad predictiva de las variables pretrasplante en la sobrevida del injerto, no se incluyeron variables intraoperatorias como el

tiempo de isquemia fría. Luego del proceso de limpieza, se conservaron 2 012 registros válidos para el análisis posterior.

Codificación y Escalamiento de Variables

Las variables binarias fueron codificadas mediante Label Encoder, mientras que aquellas con más de dos categorías se transformaron utilizando One-Hot Encoding, asegurando así su adecuada representación para los modelos de machine learning.

El escalamiento de las variables se realizó mediante StandardScaler con el objetivo de normalizar los rangos de las variables numéricas y evitar que alguna de ellas dominara el entrenamiento del modelo.

Características Demográficas de la Población

Los receptores del sexo femenino representaron el 41 % de la muestra, con una mediana de edad de 41 años (RIC: 31–52), mientras que los del sexo masculino correspondieron al 59 %, con una mediana de edad de 46 años (RIC: 35–55).

En cuanto a los donantes, la distribución fue similar: el 45 % correspondió al sexo femenino, con una mediana de edad de 46 años (RIC: 34–55), y el 55 % al sexo masculino, con una mediana de 38 años (RIC: 27–51).

Modelado

Para evaluar la sobrevida del injerto renal se aplicaron dos enfoques de modelado: el modelo de Riesgos Proporcionales de Cox, ampliamente utilizado en estudios clínicos, y el algoritmo de Bosques Aleatorios de Supervivencia, que permite capturar relaciones no lineales y de alta complejidad.

Riesgos Proporcionales de Cox

Se evaluó inicialmente la colinealidad entre variables categóricas mediante el coeficiente Cramér's V, identificándose correlaciones altas ($V > 0.8$) entre las variables tipo de donante y causa de muerte, así como entre PRA clase I (PRA CI) y clase II (PRA CII). También se observó una correlación moderadamente alta ($V > 0.7$) entre diabetes en el receptor y el diagnóstico primario. Debido a estas redundancias, se decidió eliminar las variables causa de muerte y diagnóstico primario para evitar multicolinealidad en los modelos. En el caso del PRA, se consolidó en una sola variable binaria, donde se consideró positivo si al menos una de las clases (CI o CII) lo era, y negativo únicamente cuando ambas eran negativas.

Bosques Aleatorios de Supervivencia

Para la construcción del modelo se consideraron dos enfoques: ajuste directo del modelo con hiperparámetros predeterminados, y selección de hiperparámetros mediante búsqueda en rejilla utilizando validación cruzada de cinco pliegues para evaluar el rendimiento de este.

La búsqueda en rejilla identificó como combinación óptima los hiperparámetros: $n_estimators=100$, $min_samples_split=10$, $min_samples_leaf=50$, $max_features='sqrt'$. Para el subconjunto de datos que eran pacientes trasplantados con donantes fallecidos fueron: $n_estimators=200$, $min_samples_split=10$, $min_samples_leaf=10$, $max_features='sqrt'$.

Para el modelo base (sin hiperparámetros), se emplearon los parámetros: $n_estimators=100$, $min_samples_split=10$, $min_samples_leaf=15$, $max_features='sqrt'$.

Selección de Variables

Se exploraron dos métodos para la reducción de la dimensionalidad: la selección de características basada en la importancia por permutación y el análisis de componentes principales (PCA). Sin embargo, la aplicación de estos métodos resultó en una disminución del desempeño

del modelo, posiblemente debido a la pérdida de variables con valor predictivo clínico o a la pérdida de interpretabilidad.

Por otra parte, durante el análisis exploratorio, se evaluó la asociación entre las variables clínicas y la sobrevida del injerto mediante curvas de Kaplan-Meier y pruebas log-rank. En este análisis, la variable tipo de diálisis (hemodiálisis, diálisis peritoneal y prediálisis) no mostró diferencias estadísticamente significativas en la supervivencia del injerto, y dado que se contaba con la variable meses en diálisis, que proporciona una medida cuantitativa precisa del tratamiento sustitutivo renal previo al trasplante, se optó por excluir tipo de diálisis como una estrategia de parsimonia para los modelos.

De igual forma, se identificó que la variable diabetes en el donante presentaba una baja variabilidad, con menos del 2 % de los casos, por lo cual se descartó su inclusión al no aportar valor predictivo sustancial. La lista final de variables consideradas para el entrenamiento de los modelos se presenta en la Tabla 2.

Tabla 2

Conjunto Final de Variables Clínicas

Variables Clínicas	
Donante	Receptor
Edad (años)	Edad (años)
Sexo	Sexo
IMC	IMC
Hipertensión arterial	Tiempo en diálisis (Meses)
Creatinina sérica (mg/dL)	Diabetes
Criterios expandidos	Compatibilidades HLA (ABDR)
Tipo de donante	PRA

Evaluación de los Modelos

Como métrica de evaluación general del modelo se utilizó el índice de concordancia (IC), que mide la capacidad discriminativa del modelo, es decir, qué tan bien logra ordenar a los pacientes según su riesgo estimado. Un par de individuos se considera concordante si aquel con mayor riesgo predicho presenta un evento (pérdida del injerto) antes que el otro. El IC se calcula como la proporción de pares concordantes sobre el total de pares comparables (Paquette et al., 2022).

Para evaluar la precisión del modelo se utilizó el Puntaje Integrado de Brier (IBS, por sus siglas en inglés) en el periodo comprendido entre 1 y 5 años postrasplante. El puntaje de Brier en un punto específico del tiempo cuantifica el error cuadrático medio entre las probabilidades de supervivencia predichas por el modelo y los eventos observados, combinando tanto el componente de calibración como de discriminación (Pölsterl, 2020). El IBS, por su parte, representa el promedio del puntaje de Brier calculado a lo largo de todos los puntos del intervalo de tiempo evaluado, proporcionando una estimación global del error de predicción del modelo en dicho horizonte. En otras palabras, el IBS resume cuán precisas son las predicciones del modelo en múltiples momentos del seguimiento, y un menor valor de IBS indica mejor desempeño general del modelo de supervivencia (Pölsterl, 2020).

El modelo de Cox obtuvo un IC inicial de 0,56. Con el objetivo de mejorar el desempeño del modelo se exploraron técnicas de regularización, incluyendo Lasso y Elastic Net, sin embargo, no se observaron mejoras en el rendimiento. Por su parte, el modelo de RSF alcanzó un IC de 0,57 sin ajuste de hiperparámetros y de 0,59 tras la optimización.

Dado el bajo rendimiento de ambos modelos y con el fin de mejorar la capacidad predictiva, se decidió restringir el análisis a pacientes trasplantados con donantes fallecidos, dado que la sobrevida del injerto en este grupo difiere de la observada en receptores de donantes vivos (OPTN, 2020; Patel et al., 2024). Con esta decisión también se buscó reducir la heterogeneidad de los factores pronósticos, y capturar con mayor precisión los determinantes de la falla del injerto en esta población específica.

En este subgrupo, el modelo de Cox mejoró levemente su desempeño, alcanzando un IC de 0,58; mientras que el modelo de RSF logró un IC de 0,60 sin optimización y de 0,61 con ajuste de hiperparámetros.

Tabla 3

Evaluación del Rendimiento de los Modelos Mediante el Índice de Concordancia

Modelo	Cox	RSF	
		Base	Optimizado
Trasplantes Totales	0,56	0,57	0,59
Trasplantes con Donantes Fallecidos	0,58	0,60	0,61

Dado que el mejor desempeño de los modelos se observó en la cohorte de receptores de donantes fallecidos, los análisis posteriores se centraron en esta subpoblación. En este grupo se calculó el IBS en un horizonte clínicamente relevante como cinco años postrasplante, obteniéndose un valor de 0,1636 y 0,1596 para el modelo Cox y RSF, respectivamente.

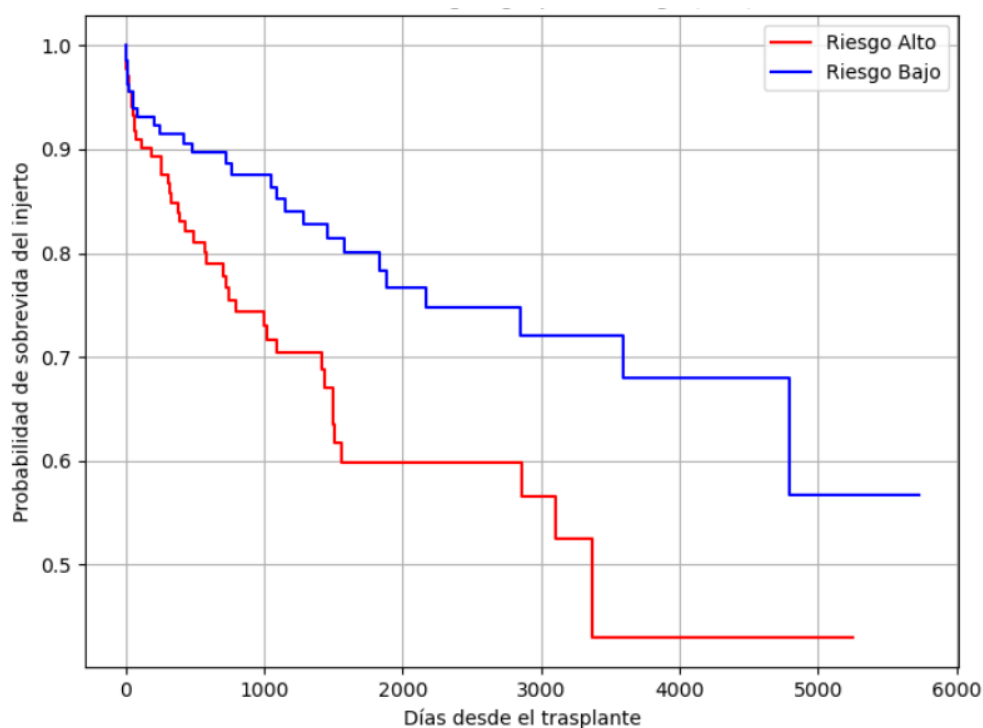
Un aspecto clave en los modelos predictivos aplicados a trasplante renal es su capacidad para generar un puntaje de riesgo individual basado en una población de referencia. Con el fin de evaluar esta capacidad discriminativa en los modelos, se utilizaron las probabilidades individuales de sobrevida del injerto a cinco años. Los pacientes fueron clasificados en grupos de

alto y bajo riesgo usando la mediana de dichas probabilidades como punto de corte. La sobrevida observada en ambos grupos se comparó mediante curvas de Kaplan-Meier.

En el modelo de Cox no se encontraron diferencias estadísticamente significativas entre los grupos (prueba log-rank, $p = 0,1186$), mientras que el modelo RSF sí logró una separación significativa ($p = 0,006$), lo que sugiere una mayor capacidad de estratificación del riesgo clínicamente relevante por parte del modelo no paramétrico (Figura 1).

Figura 1

Curvas de Kaplan-Meier por Grupos de Riesgo según el Modelo de RSF

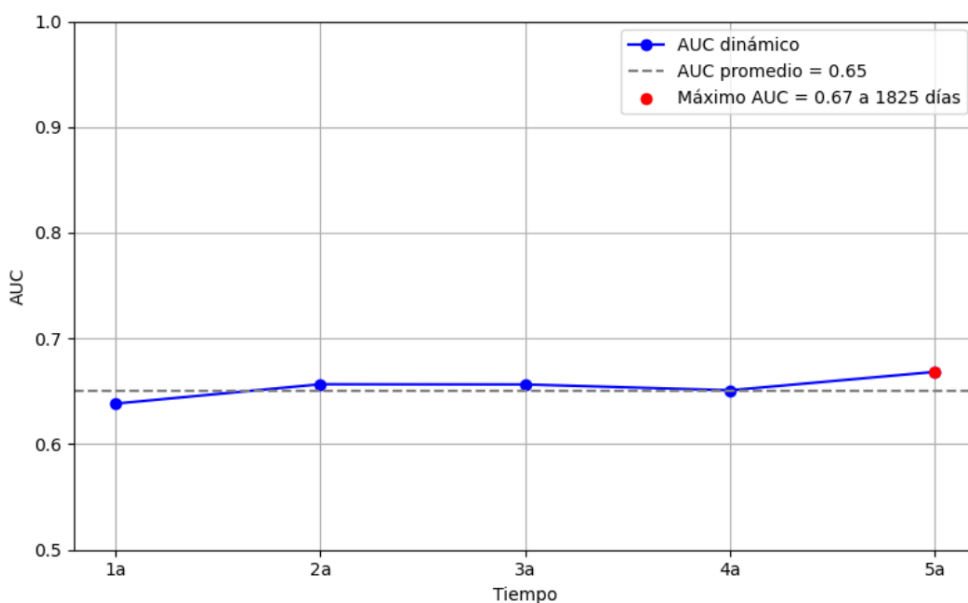


Dado que el modelo de RSF presentó el mejor rendimiento global (mayor índice de concordancia, menor IBS y capacidad para discriminar grupos de riesgo), se utilizó para evaluar la discriminación a lo largo del tiempo mediante el área bajo la curva acumulada dinámica (cumulative dynamic AUC). Esta métrica estima la capacidad del modelo para distinguir entre

individuos que perderán el injerto antes o después de un punto temporal específico (Pölsterl, 2020). Se calcularon los valores anuales desde el primer hasta el quinto año postrasplante, alcanzando su máximo a los cinco años ($AUC = 0,67$), coincidiendo con el horizonte clínico habitual. El AUC en el primer año fue de 0,64 y el AUC promedio en el período fue de 0,65 (Figura 2).

Figura 2

AUC Acumulada Dinámica

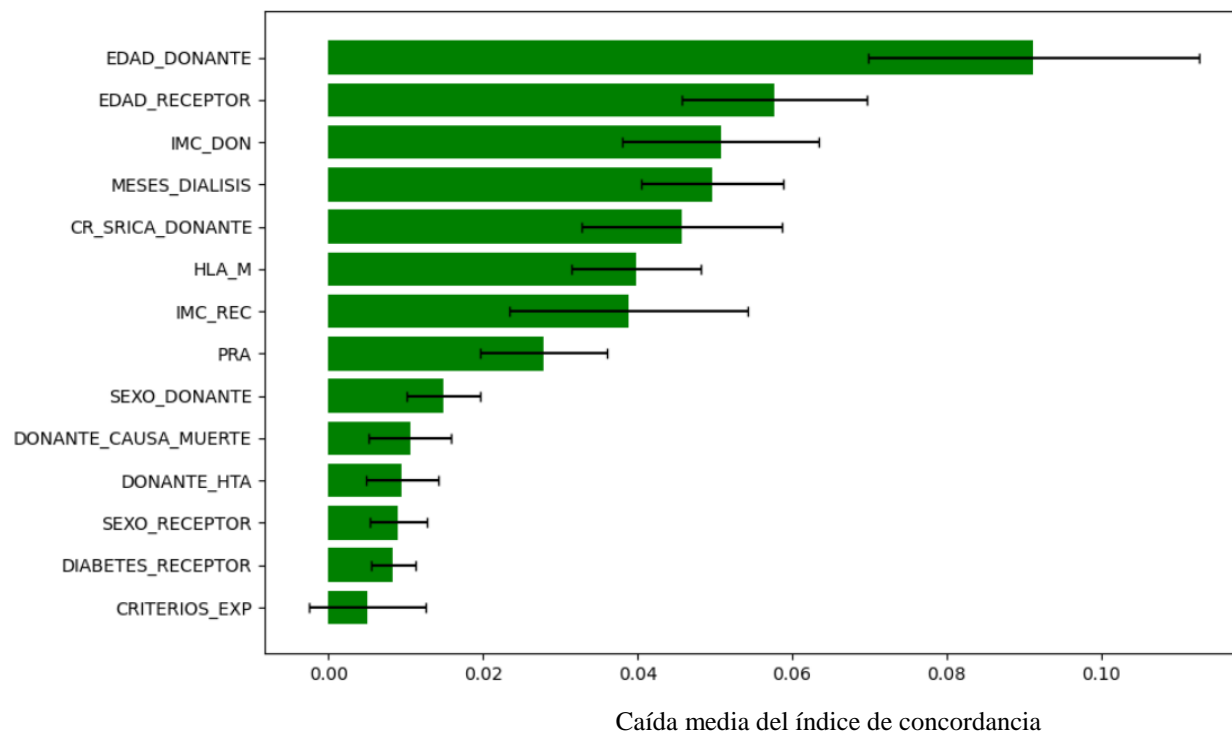


Finalmente, se evaluó la importancia de las variables en el modelo de RSF mediante el método de permutación con validación cruzada, utilizando como métrica la caída media del índice de concordancia y su desviación estándar (DE). Las cinco variables con mayor impacto en la predicción de la supervivencia del injerto fueron: edad del donante ($IC=0,091$, $DE= 0,02$), edad del receptor ($IC=0,058$, $DE= 0,01$), IMC del donante ($IC=0,051$, $DE=0,01$), meses en diálisis ($IC=0,050$, $DE=0,009$) y creatinina sérica del donante ($IC=0,046$, $DE=0,01$). Vale la pena destacar que, si bien estas variables fueron las más relevantes, el análisis mostró que la exclusión

individual de cualquiera variable disminuye el desempeño del modelo, lo que sugiere que el conjunto completo de estas aporta información valiosa y complementaria para la predicción (Figura 3).

Figura 3

Importancia de Variables por Permutación



Discusión

Este estudio evaluó la capacidad predictiva de variables clínicas pretrasplante para estimar el riesgo de falla del injerto renal en una cohorte de pacientes trasplantados en Colombia. Si bien la literatura ha propuesto diversos algoritmos de machine learning para abordar este tipo de problemas, en el presente trabajo se priorizó el uso de modelos que consideran la censura en los datos de supervivencia (Mark et al., 2019; Moghadam & Ahmadi, 2022; Naqvi et al., 2021; Ravindhran et al., 2023; Tolstyak et al., 2021; Topuz et al., 2018). Esta decisión metodológica se fundamentó en la necesidad de aprovechar de manera más eficiente la información disponible, especialmente en contextos donde el seguimiento es variable y no todos los eventos de interés ocurren durante el período de observación (Paquette et al., 2022; Senanayake et al., 2021).

El modelo de riesgos proporcionales de Cox, ampliamente utilizado en la literatura médica para evaluar la sobrevida en pacientes trasplantados, permite estimar el efecto de múltiples covariables sobre el riesgo del evento de interés a lo largo del tiempo (Garcia-Lopez et al., 2025). Sin embargo, este modelo presenta limitaciones relevantes, como la suposición de proporcionalidad de riesgos, la relación lineal entre covariables y el logaritmo de la tasa de riesgo (*log-hazard*), su baja tolerancia a grandes volúmenes de datos o alta dimensionalidad, y la inestabilidad frente a la colinealidad entre variables (Garcia-Lopez et al., 2025; Ishwaran & Lu, 2019).

Por su parte, el modelo de Bosques Aleatorios de Supervivencia se basa en los principios fundamentales del algoritmo Random Forest, que incluyen: (i) la generación de árboles mediante muestreo con reemplazo (bootstrap); (ii) la selección aleatoria de un subconjunto de variables para dividir cada nodo; (iii) el crecimiento profundo de los árboles; y (iv) la predicción del riesgo

individual a partir del promedio de las estimaciones de todos los árboles, cada uno entrenado con una muestra diferente (Ishwaran & Lu, 2019).

Entre las ventajas del modelo RSF destaca su carácter no paramétrico, lo que elimina la necesidad de asumir una forma funcional específica para la relación entre covariables y el riesgo. Esto le permite capturar interacciones complejas y no lineales entre variables clínicas. Además, el RSF es robusto frente a valores atípicos, no presenta problemas de convergencia y es aplicable en contextos de alta dimensionalidad, como los que suelen encontrarse en estudios clínicos de trasplante(Ishwaran & Lu, 2019) .

Al comparar el desempeño de los modelos mediante el IC, se observó que el modelo de RSF supera consistentemente al modelo de Cox, tanto en su versión básica como optimizada. Esta tendencia se mantuvo en el subconjunto de los trasplantes con donantes fallecidos, donde además el RSF obtuvo el menor IBS, lo que sugiere que su enfoque no paramétrico permite una mejor discriminación del riesgo en contextos clínicos caracterizados por múltiples covariables y relaciones potencialmente no lineales (Ravindhran et al., 2023).

En el estudio de Paquette y colaboradores, desarrollado con una cohorte de 180 141 pacientes trasplantados con donantes fallecidos y 35 variables, el modelo RSF alcanzó un IC de 0,64 para predecir la sobrevida del injerto. Por su parte, Senanayake, utilizando una muestra de 7 365 pacientes y 20 variables, reportaron un IC de 0,67 con RSF (Paquette et al., 2022; Senanayake et al., 2021). Estos hallazgos refuerzan la noción de que el tamaño de la muestra es un factor determinante en la calidad predictiva del modelo(Paquette et al., 2022). En este sentido, ampliar la cohorte utilizada en este análisis podría contribuir a mejorar la capacidad de generalización del modelo.

Otro factor clave es la selección de variables predictoras significativas. Aunque en la literatura se reportan estudios que utilizan desde unas pocas hasta cientos de variables, debe existir un equilibrio para evitar el sobreajuste del modelo sin excluir información relevante (Bae et al., 2020; Moghadam & Ahmadi, 2022; Naqvi et al., 2021; Paquette et al., 2022). En este sentido, Senanayake y colaboradores resaltan que, a pesar de que los modelos de RSF y Cox obtuvieron un mismo IC (0,67), el modelo de Cox logró este desempeño con solo siete variables, lo que lo convierte en un modelo más parsimonioso y fácilmente adaptable al entorno clínico (Senanayake et al., 2021).

Las variables incluidas en el presente estudio fueron seleccionadas con base en la experiencia clínica de profesionales expertos en trasplante renal, la evidencia disponible en la literatura y la completitud de los datos registrados. Se identificó que la edad del donante representa un predictor crítico de la supervivencia del injerto, lo cual es coherente con múltiples estudios previos. Esta relación puede explicarse por qué a mayor edad disminuye la reserva funcional del riñón y se incrementa la probabilidad de comorbilidades vasculares y estructurales que comprometen la viabilidad a largo plazo del órgano trasplantado (Dayoub et al., 2018; Pinto-Ramirez Id et al., 2022; Senanayake et al., 2021; Sharif, 2022).

Asimismo, se reconoce que la inclusión de otras variables clínicas e incluso sociodemográficas disponibles en la etapa pretrasplante, tales como comorbilidades del receptor, nivel educativo o acceso al sistema de salud, podría aportar valor predictivo adicional. La incorporación de estos factores podría no solo mejorar el rendimiento del modelo, sino también permitir una estratificación más precisa del riesgo individual, favoreciendo toma de decisiones personalizadas.

Por otra parte, el modelo de RSF también demostró una mayor capacidad para estratificar a los pacientes en grupos de riesgo alto y bajo, donde las curvas de Kaplan-Meier construidas a partir de este modelo fueron las únicas que mostraron diferencias estadísticamente significativas en la supervivencia del injerto. Esto sugiere que el modelo no solo mejora la discriminación individual del riesgo, sino que también es útil para segmentaciones clínicas con relevancia práctica, pudiendo seguir un enfoque similar al utilizado por las métricas del KDPI y EPTS (OPTN, 2020, 2024a), donde el puntaje de riesgo se hace con base a una población de referencia, permitiendo priorizar intervenciones, uso eficiente de órganos disponibles y comunicar de forma más comprensible el pronóstico estimado, adaptado al contexto clínico local.

A su vez, la evaluación del AUC acumulado dinámico entre el primer y el quinto año postrasplante mostró una capacidad discriminativa aceptable pero no óptima. El AUC observado en el primer año fue ligeramente menor (0,63), posiblemente debido a la menor variabilidad clínica en esta etapa inicial, donde cerca del 90% de los injertos se mantienen funcionales (Poggio et al., 2021). Al quinto año, el AUC aumentó (0,67) lo que sugiere que el modelo discrimina mejor a largo plazo. Sin embargo, el rendimiento observado aún se mantiene por debajo del umbral de 0,70. A pesar de que el modelo incluyó variables clínicas pretrasplante con un alto grado de completitud, este mostró sensibilidad a su eliminación, lo que evidencia la relevancia de éstas en la predicción del desenlace. Esto refuerza que el reto no radica únicamente en la ausencia de información, sino también en aspectos claves como la calidad y cantidad de datos, variabilidad clínica y frecuencia del evento de interés, factores que podrían estar restringiendo el poder discriminativo del modelo. A esto se suman las limitaciones inherentes a los estudios retrospectivos como posibles sesgos de selección y problemas de calidad en los registros, los cuales también pueden afectar el desempeño del modelo predictivo.

Finalmente, estos hallazgos refuerzan la necesidad de fomentar una mayor colaboración entre instituciones trasplantadoras del país, con el fin de construir bases de datos multicéntricas, heterogéneas y representativas. Este tipo de esfuerzos permitiría el desarrollo de modelos predictivos más robustos, generalizables y clínicamente útiles, que contribuyan a optimizar la asignación de órganos y mejorar los desenlaces postrasplantes.

Conclusiones y Recomendaciones

Este estudio representa un acercamiento al uso de modelos de supervivencia basados en machine learning para predecir la falla del injerto renal en una cohorte colombiana utilizando variables clínicas pretrasplante.

Los resultados mostraron que, aunque ambos modelos tuvieron un rendimiento discriminativo limitado, el modelo RSF superó al modelo de Cox en todas las configuraciones analizadas. Razón por la cual, estos modelos no paramétricos pueden ofrecer ventajas en contextos clínicos complejos, como lo es el trasplante renal.

En vista de estos hallazgos, se recomienda:

- Fomentar la colaboración entre centros trasplantadores del país para aumentar el tamaño muestral y la heterogeneidad de la cohorte, favoreciendo el entrenamiento de modelos predictivos con mayor capacidad de generalización.
- Explorar otros algoritmos de machine learning como redes neuronales o métodos basados en gradiente, así como técnicas de selección de características y calibración, que podrían contribuir a mejorar el rendimiento predictivo.
- Fortalecer los sistemas de información clínica para mejorar la calidad, completitud y estandarización de los datos registrados.

Finalmente, aunque el rendimiento alcanzado no permite su aplicación inmediata en la práctica clínica, este trabajo constituye un paso importante hacia el desarrollo de herramientas predictivas adaptadas al contexto local y al perfil de los pacientes trasplantados en Colombia.

Referencias Bibliográficas

- Abbas, A. K., Lichtman, A. H., Pillai, S., & Baker, D. (2021). *Cellular and Molecular Immunology* (10th Edition). Elsevier. <https://shop.elsevier.com/books/cellular-and-molecular-immunology/abbas/978-0-323-75748-5>
- Axelrod, D. A., Schnitzler, M. A., Xiao, H., Irish, W., Tuttle-Newhall, E., Chang, S. H., Kasiske, B. L., Alhamad, T., & Lentine, K. L. (2018). An economic assessment of contemporary kidney transplant practice. *American Journal of Transplantation*, *18*(5), 1168–1176. <https://doi.org/10.1111/AJT.14702>
- Bae, S., Massie, A. B., Caffo, B. S., Jackson, K. R., & Segev, D. L. (2020). Machine learning to predict transplant outcomes: helpful or hype? A national cohort study. *Transplant International*, *33*(11), 1472–1480. <https://doi.org/10.1111/TRI.13695>
- Balch, J. A., Delitto, D., Tighe, P. J., Zarrinpar, A., Efron, P. A., Rashidi, P., Upchurch, G. R., Bihorac, A., & Loftus, T. J. (2021). Machine Learning Applications in Solid Organ Transplantation and Related Complications. In *Frontiers in Immunology* (Vol. 12). Frontiers Media S.A. <https://doi.org/10.3389/fimmu.2021.739728>
- Bertsimas, D., Farias, V. F., & Trichakis, N. (2013). Fairness, Efficiency, and Flexibility in Organ Allocation for Kidney Transplantation. *Https://Doi.Org/10.1287/Opre.1120.1138*, *61*(1), 73–87. <https://doi.org/10.1287/OPRE.1120.1138>
- Briceño, J., Calleja, R., & Hervás, C. (2022). Artificial intelligence and liver transplantation: Looking for the best donor-recipient pairing. *Hepatobiliary and Pancreatic Diseases International*, *21*(4), 347–353. <https://doi.org/10.1016/J.HBPD.2022.03.001>

- Dayoub, J. C., Cortese, F., Anžič, A., Grum, T., & de Magalhães, J. P. (2018). The effects of donor age on organ transplants: A review and implications for aging research. *Experimental Gerontology, 110*, 230. <https://doi.org/10.1016/J.EXGER.2018.06.019>
- Fondo Colombiano de Enfermedades de Alto Costo, C. de alto costo (CAC). (2023). Situación de la enfermedad renal crónica, la hipertensión arterial y la diabetes mellitus en Colombia 2022. *Cuenta de Alto Costo* .
- García, G. (2018). *Aplicación de la metodología CRISP-DM a la recolección y análisis de datos georreferenciados desde Twitter* [Universidad Militar Nueva Granada]. Chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/<https://repository.umng.edu.co/server/api/core/bitstreams/955a9b47-3c07-4359-a203-56fa2ee0d768/content>
- Garcia-Lopez, A., Jiménez-Gómez, M., Gomez-Montero, A., Gonzalez-Sierra, J. C., Cabas, S., & Giron-Luque, F. (2025). Survival analysis using machine learning in transplantation: a practical introduction. *BMC Medical Informatics and Decision Making, 25*(1), 1–15. <https://doi.org/10.1186/S12911-025-02951-7/FIGURES/7>
- Ghelichi-Ghojogh, M., Mohammadizadeh, F., Jafari, F., Vali, M., Jahanian, S., Mohammadi, M., Jafari, A., Khezri, R., Nikbakht, H. A., Daliri, M., & Rajabi, A. (2022). The global survival rate of graft and patient in kidney transplantation of children: a systematic review and meta-analysis. *BMC Pediatrics, 22*(1), 1–15. <https://doi.org/10.1186/S12887-022-03545-2/FIGURES/9>
- Health Resources & Services Administration. (2024). *Organ Donation Statistics* / [organdonor.gov](https://www.organdonor.gov). OPTN. <https://www.organdonor.gov/learn/organ-donation-statistics>
- Hickey, M. J., Zheng, Y., Valenzuela, N., Zhang, Q., Krystal, C., Lum, E., Tsai, E., Lipshutz, G., Gritsch, A., Danovitch, G., Veale, J., Gjertson, D., Cecka, M., & Reed, E. (2017).

- New priorities: Analysis of the New Kidney Allocation System on UCLA patients transplanted from the deceased donor waitlist. *Human Immunology*, 78(1), 41–48.
<https://doi.org/https://doi.org/10.1016/j.humimm.2016.10.020>
- Holt, C. D. (2017). Overview of Immunosuppressive Therapy in Solid Organ Transplantation. *Anesthesiology Clinics*, 35(3), 365–380. <https://doi.org/10.1016/J.ANCLIN.2017.04.001>
- Instituto Nacional de Salud. (2018). Criterios de Asignación para Trasplante Renal en Colombia. *Instituto Nacional de Salud*.
- Instituto Nacional de Salud. (2023). Informe Ejecutivo Red Nacional de Donación y Trasplantes, 2023. *Instituto Nacional de Salud*. <https://www.ins.gov.co/BibliotecaDigital/informe-ejecutivo-colombia-2023-anual.pdf>
- Ishwaran, H., & Lu, M. (2019). Random Survival Forests. *Wiley StatsRef: Statistics Reference Online*, 1–13. <https://doi.org/10.1002/9781118445112.STAT08188>
- Kovesdy, C. P. (2022). Epidemiology of chronic kidney disease: an update 2022. *Kidney International Supplements*, 12(1), 7. <https://doi.org/10.1016/J.KISU.2021.11.003>
- Lasserre, J., Arnold, S., Vingron, M., Reinke, P., & Hinrichs, C. (2012). Predicting the outcome of renal transplantation. *J Am Med Inform Assoc*, 2, 255–262.
<https://doi.org/10.1136/amiajnl-2010-000004>
- Maggiani-Aguilera, P., Hernández-Estrada, S., Cano-Cervantes, J. H., Chávez-Iñiguez, J. S., Pérez-Flores, C., Díaz-Avedaño, O. del C., Oseguera-Vizcaíno, M. C., Matías-Carmona, M., Ovando-Morga, D. F., Ramírez-Ramírez, M. G., Navarro-Blackaller, G., Covarrubias-Velasco, M. A., Torres-Mayorga, J. A., Gómez-Navarro, B., García-García, G. (2021). The Sum of Kidney Donor Profile Index and Estimated Post-Transplant Survival Scale and Their Correlation with eGFR Decline in Deceased Donor Kidney

Recipients. *Revista de Investigación Clínica*, 73(4), 216–221.

<https://doi.org/10.24875/RIC.20000618>

Mark, E., Goldsman, D., Gurbaxani, B., Keskinocak, P., & Sokol, J. (2019). Using machine learning and an ensemble of methods to predict kidney transplant survival. *PloS One*, 14(1). <https://doi.org/10.1371/JOURNAL.PONE.0209068>

Ministerio de la Protección Social. (2004). *Decreto 2493 de 2004*.

<https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=14525>

Moghadam, P., & Ahmadi, A. (2022). A machine learning framework to predict kidney graft failure with class imbalance using Red Deer algorithm. *Expert Systems with Applications*, 210, 118515. <https://doi.org/10.1016/J.ESWA.2022.118515>

Mususú, W. C., García-Lopez, A., Lozano-Suarez, N., Gómez-Montero, A., Orellano-Salas, M., Vargas-Pérez, L., Escobar-Chaves, X., & Girón-Luque, F. (2024). Understanding the dynamics of deceased organ donation and utilization in Colombia. *Rev Panam Salud Publica*;48, Mar. 2024, 48, 2024. <https://doi.org/10.26633/RPSP.2024.24>

Naqvi, S. A. A., Tennankore, K., Vinson, A., Roy, P. C., & Abidi, S. S. R. (2021). Predicting Kidney Graft Survival Using Machine Learning Methods: Prediction Model Development and Feature Significance Analysis Study. *Journal of Medical Internet Research*, 23(8). <https://doi.org/10.2196/26843>

Naruka, V., Arjomandi Rad, A., Subbiah Ponniah, H., Francis, J., Vardanyan, R., Tasoudis, P., Magouliotis, D. E., Lazopoulos, G. L., Salmasi, M. Y., & Athanasiou, T. (2022). Machine learning and artificial intelligence in cardiac transplantation: A systematic review. *Artificial Organs*, 46(9), 1741–1753. <https://doi.org/10.1111/AOR.14334>,

- Niaksu, O. (2015). CRISP Data Mining Methodology Extension for Medical Domain . . *Baltic Journal of Modern Computing*, 3, 92–109.
- Nino-Torres, L., Garcia-Lopez, A., Giron-Luque, F., & Nino-Murcia, A. (2021). Retrospective Analysis of the Kidney Donor Profile Index to Predict Patient and Graft Survival at 5 Years Posttransplantation in a Colombian Cohort. *Transplantation Proceedings*, 53(7), 2197–2203. <https://doi.org/10.1016/J.TRANSPROCEED.2021.07.041>
- OPTN. (2020). A Guide to Calculating and Interpreting the Estimated Post-Transplant Survival (EPTS) Score Used in the Kidney Allocation System (KAS) What is the EPTS score? *The Organ Procurement and Transplantation Network (OPTN)*.
- OPTN. (2024a). A Guide to Calculating and Interpreting the Kidney Donor Profile Index (KDPI). *The Organ Procurement and Transplantation Network (OPTN)*.
- OPTN. (2024b). *Refit Kidney Donor Profile Index without Race and Hepatitis C Virus*. OPTN. <https://hrsa.unos.org/policies-bylaws/public-comment/refit-kidney-donor-profile-index-without-race-and-hepatitis-c-virus/>
- Paquette, F. X., Ghassemi, A., Bukhtiyarova, O., Cisse, M., Gagnon, N., Vecchia, A. Della, Rabearivelo, H. A., & Loudiyi, Y. (2022). Machine Learning Support for Decision-Making in Kidney Transplantation: Step-by-step Development of a Technological Solution. *JMIR Medical Informatics*, 10(6). <https://doi.org/10.2196/34554>
- Patel, K., Brotherton, A., Chaudhry, D., Evison, F., Nieto, T., Dabare, D., & Sharif, A. (2024). Survival Advantage Comparing Older Living Donor Versus Standard Criteria Donor Kidney Transplants. *Transplant International*, 37, 12559. <https://doi.org/10.3389/TI.2024.12559/BIBTEX>

- Pinto-Ramirez Id, J., Garcia-Lopez Id, A., Salcedo-Herrera, S., Patino-Jaramillo, N., Garcia-Lopez, J., Barbosa-Salinas, J., Riveros-Enriquez, S., Hernandez-Herrera, G., & Giron-Luque, F. (2022). *Risk factors for graft loss and death among kidney transplant recipients: A competing risk analysis*. <https://doi.org/10.1371/journal.pone.0269990>
- Poggio, E. D., Augustine, J. J., Arrigain, S., Brennan, D. C., & Schold, J. D. (2021). Long-term kidney transplant graft survival—Making progress when most needed. *American Journal of Transplantation*, *21*(8), 2824–2832. <https://doi.org/10.1111/ajt.16463>
- Pölsterl, S. (2020). scikit-survival: A Library for Time-to-Event Analysis Built on Top of scikit-learn. *Journal of Machine Learning Research*, *21*(212), 1–6. https://scikit-survival.readthedocs.io/en/stable/api/generated/sksurv.metrics.cumulative_dynamic_auc.html
- Rao, P. S., Schaubel, D. E., Guidinger, M. K., Andreoni, K. A., Wolfe, R. A., Merion, R. M., Port, F. K., & Sung, R. S. (2009). A comprehensive risk quantification score for deceased donor kidneys: The kidney donor risk index. *Transplantation*, *88*(2), 231–236. <https://doi.org/10.1097/TP.0B013E3181AC620B>
- Ravindhran, B., Chandak, P., Schafer, N., Kundalia, K., Hwang, W., Antoniadis, S., Haroon, U., & Zakri, R. H. (2023). Machine learning models in predicting graft survival in kidney transplantation: meta-analysis. *BJS Open*, *7*(2). <https://doi.org/10.1093/BJSOPEN/ZRAD011>
- Senanayake, S., Kularatna, S., Healy, H., Graves, N., Baboolal, K., Sypek, M. P., & Barnett, A. (2021). Development and validation of a risk index to predict kidney graft survival: the kidney transplant risk index. *BMC Medical Research Methodology*, *21*(1), 1–11. <https://doi.org/10.1186/S12874-021-01319-5/FIGURES/4>

- Senanayake, S., White, N., Graves, N., Healy, H., Baboolal, K., & Kularatna, S. (2019). Machine learning in predicting graft failure following kidney transplantation: A systematic review of published predictive models. *International Journal of Medical Informatics*, *130*. <https://doi.org/10.1016/j.ijmedinf.2019.103957>
- Sharif, A. (2022). Deceased Donor Characteristics and Kidney Transplant Outcomes. *Transplant International*, *35*, 10482. <https://doi.org/10.3389/TI.2022.10482/BIBTEX>
- Tolstyak, Y., Zhuk, R., Yakovlev, I., Shakhovska, N., Gregus MI, M., Chopyak, V., & Melnykova, N. (2021). The Ensembles of Machine Learning Methods for Survival Predicting after Kidney Transplantation. *Applied Sciences* 2021, Vol. 11, Page 10380, *11*(21), 10380. <https://doi.org/10.3390/APP112110380>
- Topuz, K., Zengul, F. D., Dag, A., Almeahmi, A., & Yildirim, M. B. (2018). Predicting graft survival among kidney transplant recipients: A Bayesian decision support model. *Decision Support Systems*, *106*, 97–109. <https://doi.org/10.1016/J.DSS.2017.12.004>
- Vanin A, A., Valderrama Cometa, L. A., Acuña Roldan, C. F., Alhadj, N. A., & Devia Santacruz, C. J. (2023). Graft and Patient Survival in Kidney Transplant with Deceased Donor Using KDRI (Kidney Donor Risk Index), KDPI (Kidney Donor Profile Index), and EPTS (Estimated Post-Transplant Survival) in Colombia. *Annals of Transplantation*, *28*, e940522-1. <https://doi.org/10.12659/AOT.940522>
- Weiss, M. J., Cantarovich, M., Chaudhury, P., Dieudé, M., Hartell, D. P., Martel, A. C., Patriquin, C., Shemie, S. D., Simard, M. J., Woolfsmith, J., Delmonico, F., & Domínguez-Gil, B. (2023). International Donation and Transplantation Legislative and Policy Forum: Methods and Purpose. *Transplantation Direct*, *9*(5), e1351. <https://doi.org/10.1097/TXD.0000000000001351>