

ENTREGA 1

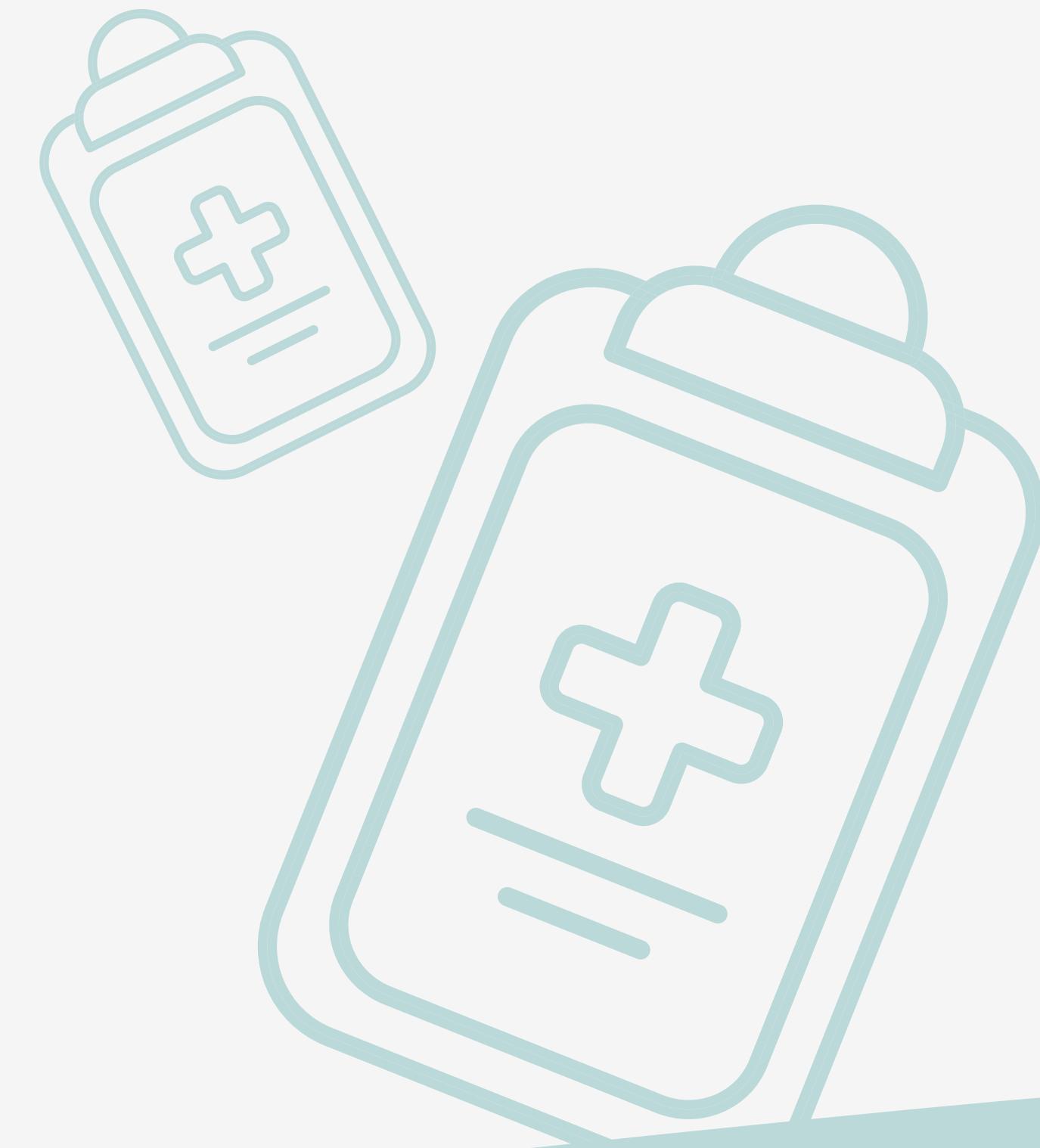
Análisis predictivo de infecciones post-trasplante
hepático mediante técnicas de Machine Learning y
visualización interactiva

Daniel Esteban Aguilera Figueroa - María Alejandra Pérez Petro

Diego Felipe Carvajal Lombo - Jesús Manuel Ospino Bernal

Índice

01	Definición de problemática
02	Objetivos y metodología
03	Productos e indicadores
04	Procesos actuales y desafíos
05	Enfoque analítico
06	Recolección de datos y entendimiento
07	Conclusiones iniciales



Definición de problemática

La Fundación Santa Fe de Bogotá es una institución de salud.

Atención a pacientes con enfermedades hepáticas - **trasplantes de hígado.**

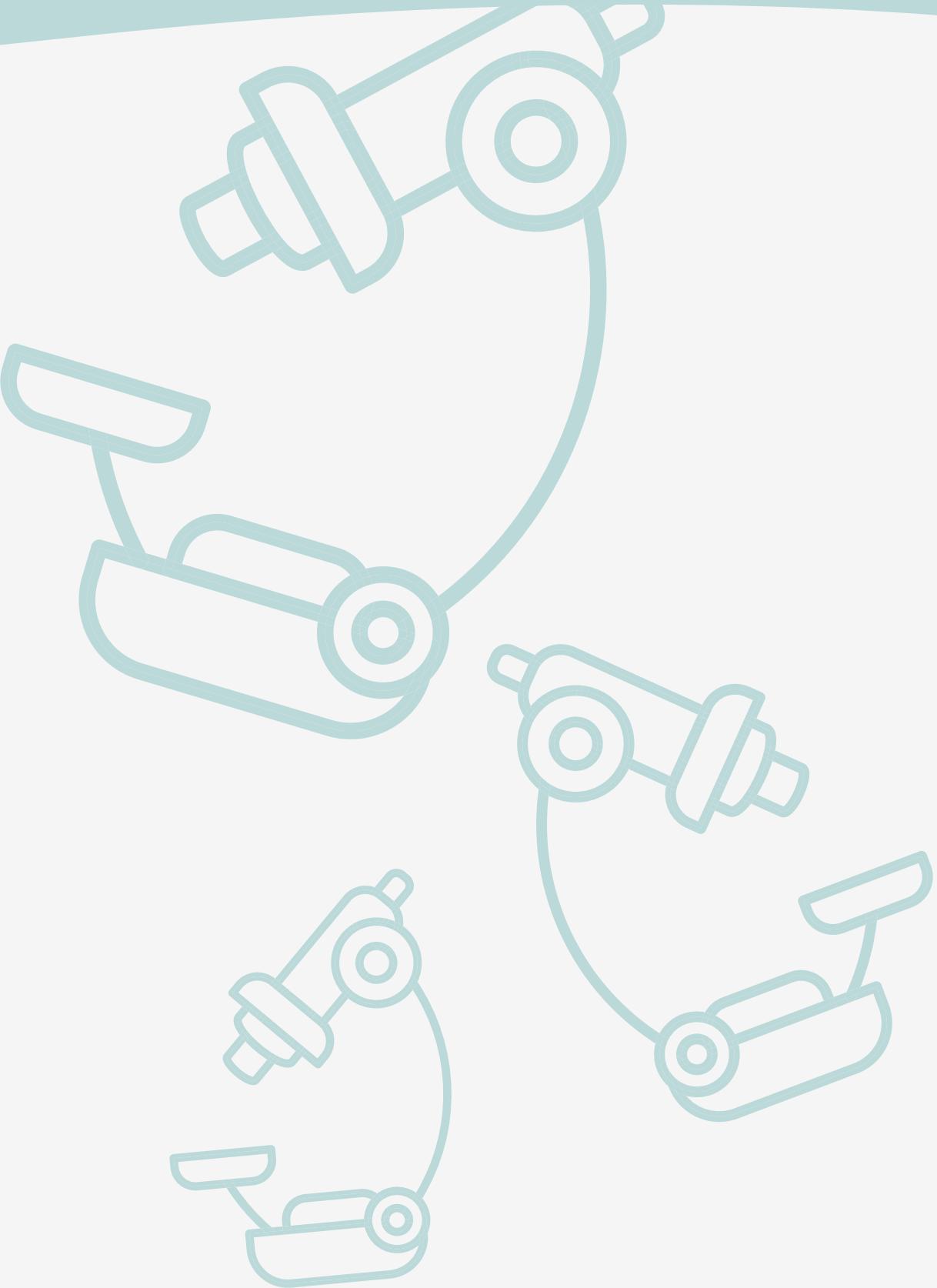
Actualmente, la organización busca comprender la incidencia y los factores de riesgo asociados a las infecciones postrasplante hepático,

Complicación que afecta hasta el 71% de los pacientes en los seis meses posteriores al procedimiento (Tezcan et al., 2023).

El objetivo: identificar patrones, diseñar estrategias preventivas y proponer intervenciones clínicas diferenciadas según los perfiles de riesgo identificados.

Objetivo del proyecto

Aprovechar las técnicas de ciencia de datos para analizar, estructurar y optimizar la información clínica de los pacientes sometidos a trasplante hepático en la Fundación Santa Fe de Bogotá, con el fin de identificar factores de riesgo asociados a infecciones post-trasplante y fortalecer la toma de decisiones clínicas y administrativas a partir de evidencia basada en datos.



Indicadores KPI

Reducción de la tasa
de infecciones
postrasplante (6
Meses - **14% → 10%**)

Disminución de
reingreso a UCI.
Pasar de **62%** a **55%**.

Incremento de la
supervivencia
postrasplante (**67% → 70%**)



Optimización de
estancia en UCI
post-operatorio.

Reducción del
impacto de
comorbilidades en el
riesgo infeccioso.

Productos



Proceso de recolección de datos utilizando herramientas de PowerApp

Aplicación digital para la captura estandarizada de información de pacientes trasplantados.

Personal médico y administrativo encargado de registro

Formularios estandarizados, integración de bases, guía técnica



Tablero Power BI

Herramienta interactiva de visualización que integra indicadores clave (KPIs) permitiendo monitorear la evolución clínica y apoyar decisiones estratégicas del hospital.

Jefe de departamento, investigadores y médicos tratantes

Conexión tiempo real, filtros dinámicos, generación reportes

Productos



Modelo de Machine Learning

Algoritmo predictivo diseñado para identificar grupos de pacientes según características clínicas y demográficas, y clasificar el riesgo de infección postrasplante hepático

Equipo médico e investigadores del área de hepatología

Entrenamiento modelo
Clasificación nuevos pacientes
Documentación

Procesos actuales y desafíos



- ✓ Fragmentación de datos
- ✓ Falta de estandarización
- ✓ Ausencia de trazabilidad
- ✓ Baja calidad de datos

ENFOQUE ANALÍTICO

Hipótesis

- Mayor tiempo quirúrgico y tiempos de isquemia prolongados se asocian a mayor riesgo de infección.
- La infección activa al momento del trasplante y el uso terapéutico de antibióticos pre-transplante predicen eventos infecciosos tempranos
- características del receptor (edad, IMC, comorbilidades) junto con el tipo de reconstrucción biliar modulan el riesgo de complicaciones e infecciones en los primeros 6 meses.

Técnicas estadísticas

- Análisis exploratorio con distribuciones, frecuencia y porcentajes. Además de medidas de tendencia central.
- Nivel de significancia **p <0.05**
- Análisis bivariado (prueba Chi2 de Pearson o test de Fisher).
- Pruebas t-student o U-Mann de Fisher.
- Pruebas de normalidad Shapiro-Wilk o Kolmogorov-Smirnov.

Machine Learning

- Modelo de clasificación binaria: Regresión logística - Random Forest.
- PCA y agrupamiento no supervisado (k-means o clustering jerárquico)

Análisis y Limpieza de Datos

Procesamiento inicial

Se dividieron las variables en cuatro grupos según su significado clínico.

Se abordaron los valores atípicos y las inconsistencias con el diccionario.

Las celdas con valores nulos o no validos representaban un alto porcentaje de los datos, estos se mapearon a -1.

Exploración y visualización

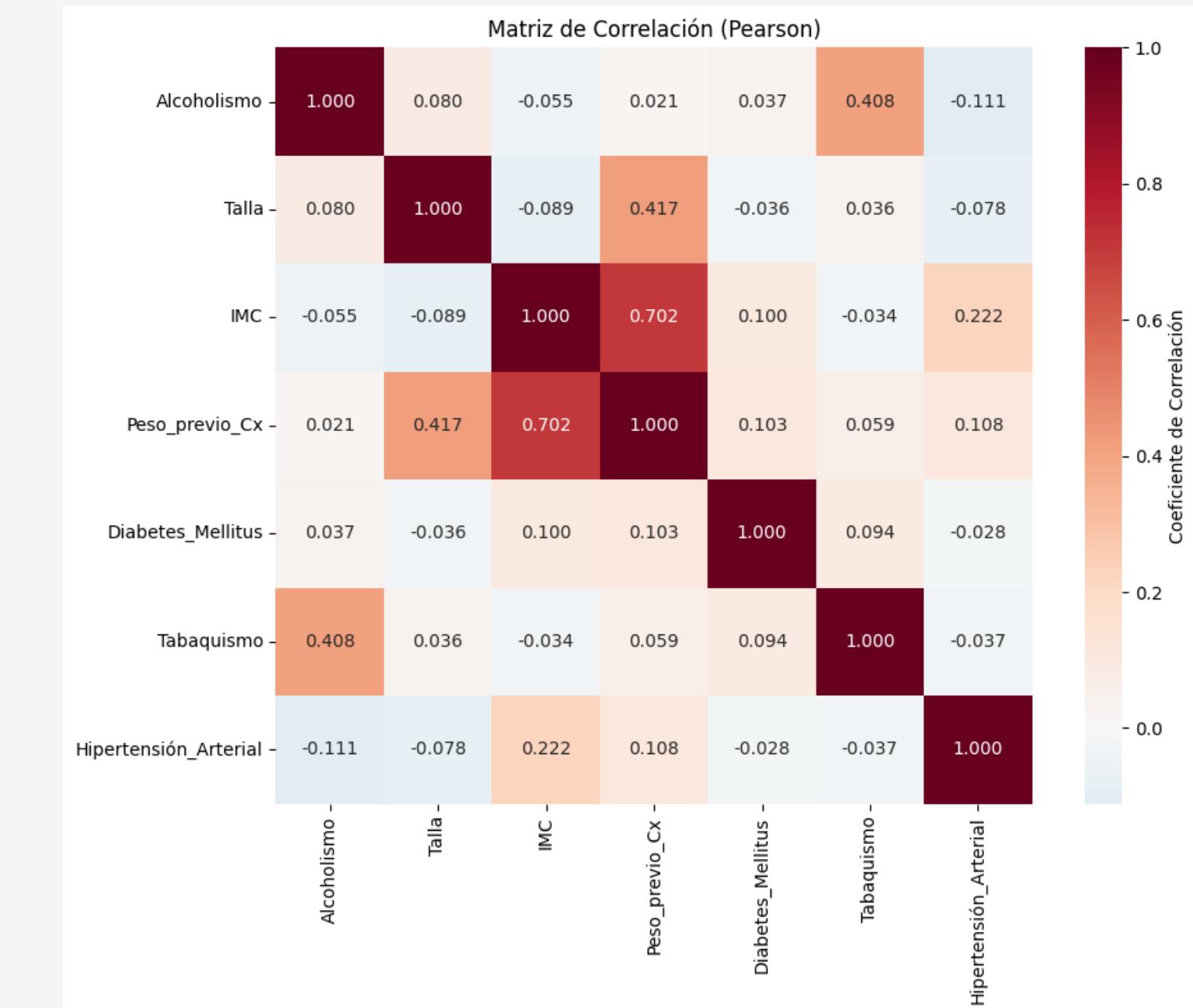
Las categóricas se analizaron mediante diagramas de barras

Las numéricas a través de boxplots e histogramas.

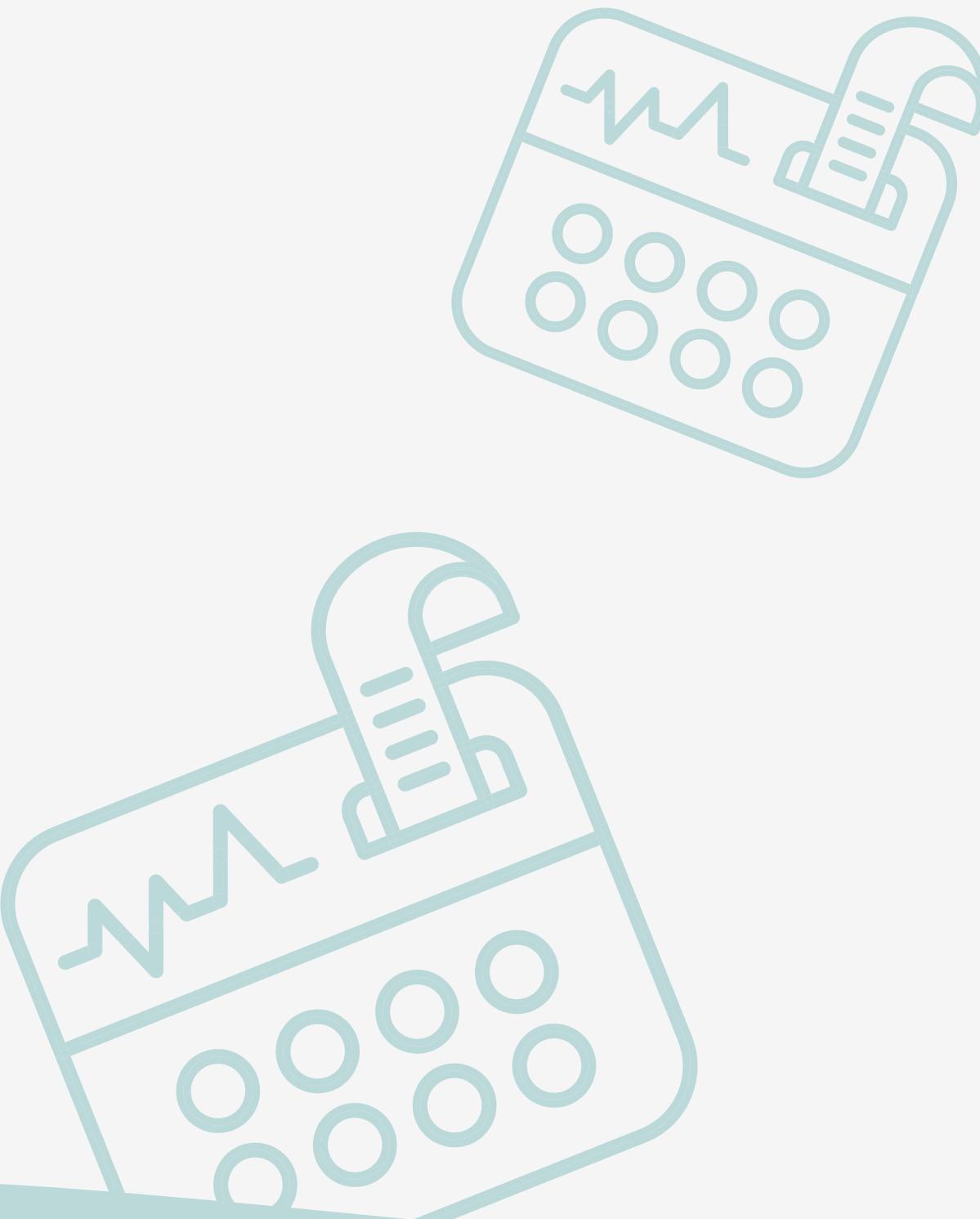
Estas representaciones facilitaron la comprensión del comportamiento de los datos y la detección de patrones relevantes.

Análisis y Limpieza de Datos

- El 51% de los pacientes son hombres mientras que el 49% son mujeres.
- El 35% reporta alcoholismo, 30,8% tabaquismo, el 29,4% diabetes, 19,3% hipertensión arterial.
- La distribución de la edad se caracteriza por tener una media en 55 años, pero una oblicuidad con tendencia a valores mayores, por lo que los datos se agrupan en edades más altas, lo que es consistente con lo que se esperaba de estos.
- La media de la estatura fue de 160 cm y del peso previo al trasplante de 68 kg.
- El 17,2% de los pacientes presentó infección al momento del trasplante, mientras que el 82,8% no.
- Hay registro de 27 enfermedades (etiología) distintas que conllevaron a la enfermedad hepática siendo la más prevalente NASH con un 25,7% de los pacientes.



Conclusiones



Se evidenció que el método actual de registro de datos en la Fundación Santa Fe, basado en un único archivo Excel compartido, es ineficiente y propenso a errores por la falta de validación y estandarización. Esto genera inconsistencias que dificultan el análisis y limitan la confiabilidad de los resultados.

Existen vacíos y discrepancias en variables clave

- Tipos de infección
- Estados postoperatorios

Estos afectan la trazabilidad de los pacientes. Es necesario implementar un sistema estructurado y validado que mejore la calidad, consistencia y aprovechamiento de los datos clínicos.

Bibliografia

- Cleveland Clinic. (2025, agosto 14). MELD score: Calculating & interpreting results. <https://my.clevelandclinic.org/health/diagnostics/meld-score>
- Facultad de Medicina, Universidad Francisco Marroquín. (s. f.). Escala de Child-Pugh. <https://medicina.ufm.edu/eponimo/escala-de-child-pugh/>
- Christian van Delden, Susanne Stampf, Hans H Hirsch, Oriol Manuel, Pascal Meylan, Alexia Cusini, Cédric Hirzel, Nina Khanna, Maja Weisser, Christian Garzoni, Katja Boggian, Christoph Berger, David Nadal, Michael Koller, Ramon Saccilotto, Nicolas J Mueller, Swiss Transplant Cohort Study , Burden and Timeline of Infectious Diseases in the First Year After Solid Organ Transplantation in the Swiss Transplant Cohort Study, Clinical Infectious Diseases, Volume 71, Issue 7, 1 October 2020, Pages e159-e169, <https://doi.org/10.1093/cid/ciz1113>
- Fundación Santa Fe de Bogotá. (s. f.). Sobre nosotros. Recuperado el 17 de octubre de 2025, de <https://fundacionsantafedebogota.org/intelecto/nosotros/quienes-somos>
- Tezcan H, Altunsoy A, Turan Gökçe D, Gökcan H, Arı D, Aydın O, Bostancı EB, Akdoğan Kayhan M. Multidrug-Resistant Infections After Liver Transplantation, Etiology and Risk Factors: A Single-Center Experience. Exp Clin Transplant. 2023 Dec;21(12):952-960. doi: 10.6002/ect.2023.0081. PMID: 38263782.
- República de Colombia. (1991). Constitución Política de la República de Colombia. Recuperada de <https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=4125>
- Superintendencia de Industria y Comercio. (2024, 21 de agosto). Circular Externa No. 002: Lineamientos sobre el tratamiento de datos personales en sistemas de inteligencia artificial.

¡Gracias!

