

Análisis de Riesgo en Trasplante Hepático Utilizando Inteligencia Computacional

Computational Intelligence Toward Risk Analysis in Liver Transplantation

Presentación: 6 de octubre, 2020

Doctorando:

Parag Chatterjee

Grupo de Investigación y Desarrollo en Bioingeniería (GIBIO)
Facultad Regional Buenos Aires
Universidad Tecnológica Nacional, Argentina
parag@frba.utn.edu.ar

Director/a:

Ricardo Armentano

Co-director/a:

Leandro Cymberknop

Resumen

Los últimos años han visto un cambio fenomenal en los paradigmas de atención médica y el análisis de datos con inteligencia computacional ha sido un jugador clave en este campo. Uno de los principales objetivos de incorporar inteligencia computacional en el análisis de la salud es obtener una mejor comprensión de los pacientes y ofrecer un tratamiento más eficiente. Este estudio se basa en pacientes de trasplante hepático bajo el Programa Nacional de Trasplante Hepático de Uruguay, considerando en detalle los parámetros de salud de los pacientes. El objetivo principal se constituyó en aplicar la inteligencia computacional para separar la población estudiada en grupos, facilitando así analizar eficientemente los grupos de riesgo de los pacientes evaluados bajo el programa de trasplante hepático con respecto a sus parámetros de salud correspondientes, en una perspectiva predictiva previa al trasplante. En este aspecto, se aplicaron algoritmos de agrupamiento en la población considerando sus datos de salud y se analizaron los grupos obtenidos. Los grupos mostraron propiedades distintivas, mostrando que un grupo tiene un riesgo más alto que el otro. Además, esto marca la base de los sistemas de soporte a la decisión clínica (*clinical decision support system*) en el trasplante hepático, que actúan como una herramienta de asistencia para el personal médico para obtener una visión más profunda de la salud del paciente y gracias a la visualización holística del escenario de la atención médica, también ayudan a elegir una estrategia de tratamiento eficiente y personalizada.

Palabras clave: salud, análisis predictivo, sistema de soporte a las decisiones, trasplante hepático, análisis de datos, predicción, riesgo.

Abstract

Recent years have seen a phenomenal change in healthcare paradigms and data analytics along with computational intelligence has been a key player in this field. One of the main objectives of incorporating computational intelligence in healthcare analytics is to obtain better insights about the patients and proffer more efficient treatment. This work is focused to the National Liver Transplant Program of Uruguay, where comprehensive health indicators of the patients were considered and analyzed for intelligent risk prediction and profiling. The principal objective has been the separation of the cohort in risk-groups using computational intelligence, thereby facilitating efficient analysis of the patients assessed under the liver transplantation program in a predictive pre-transplant perspective. In this respect, clustering algorithms were applied on the cohort considering their comprehensive health-data at the entry-point evaluation and the clusters obtained were analyzed. The clusters showed distinctive properties pertaining to risks, providing predictive models and detailed risk-profiling of the patients in the pre-transplant stage. Also, this marks the foundation of Clinical Decision Support Systems in liver transplantation, which act as an assistive tool for the medical personnel in getting deeper insights to patient health in advance and leads to the holistic visualization of the healthcare scenario, helping to choosing a more efficient and personalized treatment strategy.

Keywords: healthcare, predictive analytics, decision support system, liver transplantation, data analysis, prediction, risk.

Introducción

El análisis de datos en salud ha generado interesantes aplicaciones en virtud del advenimiento de paradigmas tales como Inteligencia Artificial, e-Salud e informática médica. La atención médica siempre ha sido uno de los principales sectores motivados para adoptar nuevas tecnologías, dado que el objetivo principal es proporcionar un mejor tratamiento a los pacientes. En este sentido, la minimización del tiempo de respuesta en el diagnóstico y tratamiento constituye un componente crucial en los servicios de salud eficientes, donde el análisis de datos manifiesta su potencial a través del uso de métodos inteligentes. En la actualidad, el trasplante hepático constituye la última opción terapéutica en pacientes con enfermedad hepática. Es un proceso de elevado costo, el cual que requiere la experiencia de un equipo interdisciplinario especializado y, por sobre todas las cosas, un seguimiento cercano de los pacientes durante el transcurso de su enfermedad. Esta actividad trae aparejada la generación de un gran volumen de información, donde son incorporados datos de naturaleza múltiple. Es por ello para lograr un manejo clínico adecuado de los pacientes (que obviamente tendrá influencia directa en su pronóstico) deben tomarse decisiones acertadas basadas en la interacción de múltiples variables simultáneas. El objetivo principal de este plan de trabajo fue el utilizar herramientas de Inteligencia Computacional de modo de efectuar una predicción de los principales determinantes de la morbilidad y mortalidad en pacientes registrados en el marco del Programa Nacional de Trasplante Hepático en Uruguay. Esencialmente, se trabajó en el riesgo inherente de progresión de la enfermedad hepática, en términos de la morbimortalidad del paciente, a lo largo de su estancia en el programa.

Desarrollo

A. Antecedentes y Descripción de la Población Estudiada

Este estudio se llevó a cabo sobre una base de datos de pacientes registrados bajo el Programa Nacional de Trasplante Hepático en Montevideo, Uruguay. Los pacientes considerados en este estudio se inscribieron en el programa entre los años 2014 y 2017 y fueron evaluados al momento de ingreso en la lista (pre-trasplante) [5]. Los pacientes fueron aceptados (o rechazados) para entrar al programa en función de la gravedad de sus enfermedades y comorbilidades. Para la organización (y priorización) de los pacientes en la lista de espera, se utiliza la escala numérica proporcionada por el Modelo para Enfermedades Hepáticas en Etapas Terminales (*MELD, por sus siglas en inglés*), aplicado en mayoría de los centros en América Latina (Salvalaggio et al., 2014). Otro parámetro de referencia lo constituye el Riesgo Cardiovascular (RCV), el cual proporciona la probabilidad de padecer un accidente cardiovascular a 10 años y determina en base a la edad, sexo, presión arterial, colesterol, tabaquismo y presencia de diabetes (D'Agostino et al., 2008). En virtud de poder efectuar el análisis propuesto, se contó con datos relacionados a la sobrevivencia de los pacientes y aquellos que fueron

derivados a trasplante o se encontraban en lista de espera a lo largo del período correspondiente a un intervalo temporal de 3 años. La población evaluada para el presente estudio se encontró constituida por 104 pacientes, con edad media de 47 años (rango entre 14-70 años) y una cantidad casi similar de hombres y mujeres. La indicación más frecuente para el trasplante de hígado fue la cirrosis, seguida del carcinoma hepatocelular y la insuficiencia hepática aguda.

B. Análisis de Datos y Separación de la Cohorte en Función de los Indicadores de Salud Evaluados

El enfoque original de este estudio fue el de analizar la totalidad de los parámetros constitutivos de la población y de modo de obtener un modelo predictivo de evolución de la enfermedad basado en los datos iniciales del reclutamiento. En este aspecto, las técnicas de agrupamiento (*clustering*) suelen ser de gran utilidad para separar grupos de individuos y luego analizar en detalle las características distintivas entre los mismos. Cuando las relaciones y el impacto de los diferentes parámetros que forman el conjunto de datos no se conocen bien, se apela a este tipo de técnicas en virtud de definir los factores influyentes. Particularmente en este estudio se apeló a la utilización de la técnica de agrupamiento *k-medias* de aprendizaje no supervisado, la cual permite, a partir de múltiples variables, definir un conjunto de *k* agrupaciones dentro del conjunto general de datos. Cada elemento del conjunto será asignado a uno de los *k*-centros (generados recursivamente en forma aleatoria), de modo de minimizar la suma de las distancias entre e intra-grupos (Healthcare.ai, 2018). Debido a que la técnica requiere la definición a priori del número de agrupamientos, se aplicó el criterio de silueta al conjunto total de datos (Fig. 1), obteniéndose un número óptimo de dos grupos.

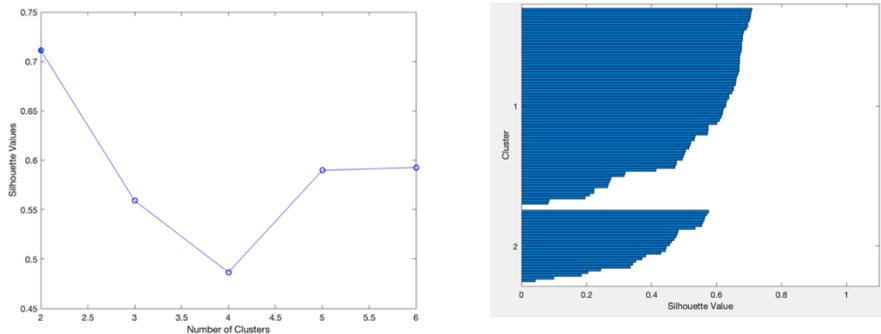


Fig. 1. Separation of Cohort into Two Clusters Including All Health Indicators

El primer análisis de agrupamiento se basó en la evaluación completa de la población, teniendo en cuenta todas sus características, pero de modo de analizar los grupos desde un punto de vista independiente y menos sesgado, se excluyó la información vinculada a RCV, MELD-Na, sobrevida y derivación a trasplante en el proceso de agrupación (Tabla IV). En el segundo análisis, se aplicó un enfoque similar al punto anterior, pero para llevar a cabo el agrupamiento por separado en los pacientes en lista de espera y los pacientes trasplantados.

C. Resultados

En el primer análisis (agrupamiento efectuado incluyendo la totalidad de la población, pero sin considerar el RCV, MELD-Na, sobrevida y derivación a trasplante), se obtuvieron dos grupos de 61 y 36 pacientes cada uno. El grupo más pequeño mostró valores más altos de FR (7), MELD (19), y un mayor porcentaje de muerte (25%) y pacientes trasplantados (69%) que el grupo mayor cantidad de individuos (FR: 6, MELD: 16, muerte: 16%, trasplante: 54%). A valores similares de edad media, IMC y presión arterial, el grupo más reducido mostró un valor significativamente menor de HDL y valores más altos de LDL, triglicéridos, colesterol total / HDL, plaquetas, linfocitos, neutrófilos, monocitos, eosinófilos, basófilos y glucemia. Una tendencia similar se observó en los grupos obtenidos de los pacientes pertenecientes a la lista de espera y aquellos derivados a trasplante.

Tabla I. Características de los Clusters
(Agrupamiento basado en la evaluación de la población completa sin los parámetros de evaluación de la predicción)

Parameters	Cluster 1 (n=61)	Cluster 2 (n=36)
Age at the moment of evaluation (years)	46 ± 16	48 ± 14
Gender (%)	Male: 49% Female: 51%	Male: 55% Female: 45%

BMI (kg/m ²)	27 ± 5	27 ± 5
Systolic blood pressure (mm Hg)	117 ± 12	116 ± 12
Diastolic blood pressure (mm Hg)	67 ± 8	68 ± 8
Total Cholesterol (mmol/L)	148 ± 58	175 ± 136
Triglycerides	91 ± 40	134 ± 102
HDL (mmol/L)	42 ± 24	31 ± 24
LDL (mmol/L)	88 ± 45	109 ± 79
Total Cholesterol/HDL	9 ± 18	20 ± 40
Platelets (x1000)	102 ± 59	165 ± 86
Lymphocytes	1100 ± 679	1667 ± 961
Neutrophils	2616 ± 1237	5808 ± 1735
Monocytes	475 ± 274	825 ± 383
Eosinophils	187 ± 202	233 ± 323
Basophils	11 ± 32	36 ± 59
Glycemia	100 ± 40	97 ± 42

Values are expressed as Mean ± Standard Deviation / Percentage

Conclusiones

En el presente trabajo, se utilizaron técnicas de Inteligencia Computacional en virtud de efectuar la clasificación de pacientes que ingresan al programa nacional de trasplante de hepático, cuya implementación ha podido diferenciar a la población analizada en dos grupos de riesgo. Los resultados arrojados por el análisis ponen en evidencia que los parámetros de evaluación de la predicción (RCV, MELD-Na, sobrevida y derivación a trasplante) no resultan de peso para la llevar a cabo el agrupamiento y que por ende puede utilizarse únicamente la información obtenida al ingreso del paciente al programa.

En relación con el primer análisis (agrupamiento con la totalidad de la población e incluyendo la totalidad de los parámetros) el grupo de menor cantidad de individuos manifestó valores más altos de LDL, triglicéridos, colesterol total / HDL, plaquetas, linfocitos, neutrófilos, monocitos, eosinófilos, basófilos y glucemia, sugiriendo que dicho grupo se encontraría en mayor riesgo que su contraparte más grande. De este modo, la herramienta utilizada fue capaz de diferenciar la población, en grupos de distinto riesgo, a partir de la utilización de datos puros obtenidos del ingreso del paciente al programa de trasplante. Partiendo de los resultados alcanzados, y considerando un punto de vista médico, un mayor recuento de linfocitos indica posibilidades de linfocitosis, frecuentemente asociadas con infecciones crónicas, inflamaciones y enfermedades autoinmunes. Además, un mayor recuento de monocitos indica un riesgo potencial de infección y señales de recuento de neutrófilos y plaquetas para el estado de inflamación. Bajo dicha premisa, los resultados de agrupamiento resultan coherentes, al arrojar valores más elevados de dichos parámetros para los grupos de mayor riesgo. Aún sin conocimiento de los eventos futuros (incremento de la severidad o derivación a trasplante), se pudo efectuar la separación en grupos diferenciados, considerando sólo los parámetros obtenidos durante la evaluación inicial. En este sentido, teniendo en consideración que el tiempo medio para el trasplante desde el momento de la entrada al programa es de 3 meses (incluso hasta 3 años), este modelo de agrupamiento podría clasificar a un paciente dentro de un grupo con mayor o menor riesgo, en virtud de focalizar en los aspectos esenciales de su tratamiento. Los sistemas de apoyo a la toma de decisiones conforman en la actualidad componentes interesantes de los sistemas de atención médica (Chatterjee et al., 2017). El enfoque propuesto apunta al diseño de un sistema predictivo de apoyo a la decisión clínica, con el objetivo de clasificar automáticamente la población de pacientes con enfermedad hepática al momento de su alistamiento, facilitando así los diversos aspectos involucrados en su atención (Chatterjee et al., 2019).

Referencias

Chatterjee, P., Armentano, R. L., & Cymberknop, L. J. (2017). Internet of Things and Decision Support System for eHealth - Applied to Cardiometabolic Diseases. *2017 International Conference on Machine Learning and Data Science (MLDS)*. <https://doi.org/10.1109/mls.2017.22>

Chatterjee, P., Noceti, O., Menendez, J., Gerona, S., Harguindeguy, N., Toribio, M., ... Armentano, R. L. (2019). Predictive Risk Analysis for Liver Transplant Patients - eHealth Model Under National Liver Transplant Program, Uruguay. *2019 IEEE 9th International Conference on Advanced Computing (IACC)*. <https://doi.org/10.1109/iacc48062.2019.8971514>

D'Agostino, R. B., Vasan, R. S., Pencina, M. J., Wolf, P. A., Cobain, M., Massaro, J. M., & Kannel, W. B. (2008). General Cardiovascular Risk Profile for Use in Primary Care. *Circulation*, 117(6), 743–753. <https://doi.org/10.1161/circulationaha.107.699579>

Data Science Blog. (2017, July 21). *Step by Step to K-Means Clustering*. healthcare.ai. <https://healthcare.ai/step-step-k-means-clustering/>.

Salvalaggio, P. R., Caicedo, J. C., Albuquerque, L. C., Contreras, A., Garcia, V. D., Felga, G. E., ... Vilatoba, M. (2014). Liver Transplantation in Latin America. *Transplantation*, 98(3), 241-246. doi:10.1097/tp.000000000000198