

Modelo de machine learning basado en parámetros clínicos y hematológicos permite predecir ingreso a Unidad de Cuidados Intensivos en pacientes con COVID-19

Neftalí Guzmán-Oyarzo¹, Alfonso Hernández -Monsalves¹, Pablo Letelier-Arias¹, Camilo Morales², Eduardo Rojas-Maturana¹, Mauricio Saez-Venegas¹, Nicolas Coña¹, Javiera Díaz¹, Andrés San Martín³, Paola Garcés⁴, Jesús Espinal-Enriquez⁵

1. Laboratorio de Investigación en Salud de Precisión, Departamento de Procesos Diagnósticos y Evaluación, Facultad de Ciencias de la Salud, Universidad Católica de Temuco, Chile, 2. Departamento de Procesos Terapéuticos, Facultad de Ciencias de la Salud, Universidad Católica de Temuco, Chile, 3. Laboratorio Clínico, Hospital Dr. Hernán Henríquez Aravena, Temuco, Chile, 4. Centro Médico Alergoinmuno Araucanía, Temuco, Chile, 5. Computational Genomics Department. National Institute of Genomic Medicine. Mexico City, Mexico.

Introducción A nivel global, el número de casos y fallecimientos asociados a COVID-19 han mostrado una tendencia a disminuir. Sin embargo, persisten pacientes que progresan a cuadros severos y que pueden presentar secuelas y complicaciones asociadas, especialmente personas mayores y con comorbilidades. El uso de herramientas de inteligencia artificial (IA) contribuye a mejorar el manejo clínico de pacientes que ingresan a Unidades de Cuidados Intensivos (UCI), sin embargo, no existen modelos previos validados en población chilena. Así, el objetivo de este estudio fue validar un modelo de machine learning para predecir ingreso a UCI en pacientes hospitalizados por COVID-19. **Metodología** Un total de 201 pacientes hospitalizados por COVID-19 en un hospital público fueron incluidos en el estudio, previamente aprobado por Comité de Ética Científica. Pacientes fueron diagnosticados por criterios establecidos, confirmado por métodos moleculares (RT-PCR) para SARS-CoV-2 y agrupados de acuerdo a severidad. Datos clínicos, epidemiológicos y demográficos fueron obtenidos desde el historial clínico de cada paciente. Resultados de laboratorio fueron obtenidos desde el sistema de informático de laboratorio (LIS) al día de la admisión hospitalaria (día 1). Dataset fue estratificado utilizando 70 % para entrenamiento y 30 % para prueba. Para determinar el mejor modelo predictivo, regresión LASSO (least absolute shrinkage and selection operator) fue utilizada. Tres modelos de machine learning fueron generados, entrenados y validados internamente: regresión logística (LR), random forest (RF), y extreme gradient boosting (XGBoost). Se evaluaron métricas de sensibilidad (Sn), especificidad (Sp), área bajo la curva (AUC), precisión (P) y SHapley Additive exPlanation (SHAP) values. Finalmente, la utilidad clínica de los modelos predictivos fue evaluada mediante análisis de curva de decisión (DCA). **Resultados** Las 6 variables más significativas correspondieron a: diabetes mellitus tipo 2 (DM2), obesidad, recuento absoluto de neutrófilos y basófilos, razón neutrófilos-linfocitos (NLR) y dímero D al día de la admisión hospitalaria. Regresión logística mostró Sn de 0,67; Sp 0,65; AUC 0,74, y P 0,66. Random Forest presentó Sn de 0,87; Sp 0,83; AUC 0,96 y P 0,85. Finalmente, XGBoost presentó Sn de 0,87; Sp 0,85; AUC 0,95 y P 0,86. **Conclusión** Los resultados muestran que el modelo XGBoost, basado en variables DM2, obesidad, recuento absoluto de neutrófilos y basófilos, NLR y dímero D, muestra un rendimiento predictivo robusto y un favorable beneficio clínico, lo que confirma su capacidad para predecir el ingreso a UCI de individuos chilenos con COVID-19, lo que permite mejorar la toma de decisiones clínicas. La validación externa del modelo en otros centros clínicos reforzaría su eficacia y fiabilidad para su uso clínico.