



## O-016 - DESARROLLO Y VALIDACIÓN INTERNA DE UN MODELO DE *MACHINE LEARNING* PARA PREDECIR LA OCURRENCIA DE HERNIA INCISIONAL POSTERIOR A UNA LAPAROTOMÍA POR LÍNEA MEDIA

Lozada Hernández, Edgard Efren<sup>1</sup>; Armenta Medina, Dagoberto<sup>2</sup>; Ramírez del Real, Tania Aglae<sup>3</sup>; Salazar Colores, Sebastian<sup>4</sup>

<sup>1</sup>Hospital Regional de Alta Especialidad del Bajío, Guanajuato México; <sup>2</sup>INFOTEC, Guanajuato México;

<sup>3</sup>CentroGEO, Guanajuato México; <sup>4</sup>Centro de Investigaciones en Óptica, Guanajuato México.

### Resumen

**Introducción:** Una de las principales complicaciones posterior a una laparotomía es la hernia incisional (IH), en los estudios de profilaxis de HI no existe un consenso en determinar qué pacientes son de bajo o alto riesgo, por lo que es complicado la comparación de resultados, lo cual a su vez ha derivado en una falta de consenso en la técnica de elección para la prevención de esta complicación. En los últimos diez años se han publicado siete escalas predictoras, estas escalas han usado métodos tradicionales como el de regresión logística múltiple o la regresión de Cox, que, aunque son métodos validados y probados, han resultado al menos en la predicción de esta complicación con un bajo rendimiento pronóstico. No existe un estudio publicado en la actualidad en el cual se aplique la inteligencia artificial para la predicción de HI, así el objetivo de este estudio fue desarrollar un modelo predictivo de HI basado en aprendizaje automático con el uso de la técnica XGBoost, de manera secundaria desarrollar una app-web con los resultados del modelo, que ayude a los cirujanos encargados del cierre de la pared abdominal a tener un respaldo objetivo para determinar a los pacientes de alto riesgo y en ellos proporcionar una asignación más selectiva de las intervenciones de prevención.

**Métodos:** La conducción y reporte de los resultados del estudio se basó en la guía TRIPOD statement. Fue registrado en los comités hospitalarios de investigación y ética del hospital y en la plataforma *Clinical Trials* con número de registro NCT 0571899. Se incluyeron pacientes mayores de 18 años, posoperados de laparotomía exploradora por línea media y que completaran 24 meses de seguimiento posterior a la cirugía inicial. Los modelos predictivos se desarrollaron usando tres algoritmos de aprendizaje automático: árbol de decisión, regresión logística y XGBoost. Todo el procesamiento de datos se realizó mediante escritura de código personalizado en Python dentro del entorno de Jupyter Notebook (anaconda 3). Utilizando el algoritmo de aprendizaje automático XGBoost para la predicción y los valores SHAP para la explicación, se desarrolló una aplicación web basado en la librería Shiny de R.

**Resultados:** 789 pacientes fueron analizados, 161 presentaron HI con una incidencia del 20,4%. El modelo con mejor desempeño fue el clasificador XGBoost con una puntuación de *accuracy* de 0,97, AUC de 0,97, precisión de 1, *recall* 0,89% F1 0,94 para la variable de interés (presencia de HI). En

comparación con regresión logística y árbol de decisiones que obtuvieron un AUC de 0,62 y 0,74 respectivamente. Las seis características más influyentes en el modelo fueron: riesgo de infección, antecedente de cirugía abdominal previa, IMC, Edad, que la cirugía fuera realizada fuera colorrectal y que la cirugía fuera de urgencia.

Comparison of artificial intelligence models and their performance in HI forecasting

Evaluation metric	Logistic regression	Decision tree	XGBoost
<b>AUC</b>	0.62(0.57-0.74)	0.74 (0.62-0.83)	0.97 (0.94-0.98)
<b>Accuracy</b>	0.64(0.58-0.67)	0.73(0.65-0.88)	0.97(0.93-98)
<b>Recall</b>	49%	25%	89%
<b>Specificity</b>	87%	97%	96%
<b>PPV</b>	50%	71%	90%
<b>NPV</b>	87%	83%	94%
<b>F1</b>	0.44	0.67	0.94

AUC= Under curve area, () = 95% confidence interval PPV, NPV positive and negative predictive value

**Conclusiones:** El uso de inteligencia artificial es útil para predecir la ocurrencia de HI en los pacientes posoperados de laparotomía por línea media, el modelo XGBoost obtuvo una sensibilidad de 0,97%, especificidad 0,96 y *accuracy* 0,97, con mejor rendimiento diagnóstico que los modelos tradicionales de regresión logística múltiple y árbol de decisión.