Understanding the Impact of Social Determinants of Health on 90-Day Readmission Rates Following Posterolateral Fusion: A Machine Learning Approach

Samuel Gregory Reyes, Pranav Malhotra Bajaj, Steven Kurapaty, Austin Ryan Chen, Erik Gerlach, Mark A Plantz, Wellington K Hsu[1], Alpesh Ashwin Patel, Srikanth Divi[2]
[1]Dept of Ortho Surg, [2]Northwestern Medicine

INTRODUCTION: Social determinants of health (SDH) are factors outside of clinic that affect the health of a patient. SDH have become increasingly more important across various fields of medicine, yet there remains a dearth of literature in spine surgery. Machine learning (ML) algorithms are well suited to glean new insights from clinical databases with high accuracy, outpacing traditional statistical methods. This study aims to apply ML to understand relationships between SDH and 90-day readmission rates following lumbar posterolateral fusion (PLF).

METHODS: Patients that underwent single or multi-level lumbar PLF at a multi-center academic health system between 2002 –2020 were acquired from the electronic medical record. More than 60 clinical variables including demographics, past medical and surgical history, postoperative complications, 30-day readmission, 90-day readmission, 90-day reoperation, and 1-year reoperation rates were included. The Geocodio platform, which retrieves geographic data from diverse sources including the US Census Bureau, was used to map addresses of patients to census tracts. Patients were assigned SDH characteristics by census tract using the Social Vulnerability Index (SVI). The SVI uses census data to determine social vulnerability by ranking each tract on 15 social factors (i.e. poverty, lack of vehicle access, crowded housing). The primary endpoint was 90-day readmission. ML models were run using custom Python scripts and were validated by Area Under the Curve (AUC). The data was split into training/testing (80/20) sets. Validation was performed on withheld test data following optimization. The best performing model was determined by AUC; SHAP values were calculated for the best model to rank features by impact on output prediction.

RESULTS: A total of 1,454 patients were included in the sample. 30-day and 90-day readmission rates were 13.3% (n=194) and 22.4% (n=325), respectively. Additionally, the 90-day and 1-year reoperation rates were 3.1% (n=45) and 4.9% (n=72), respectively. Of the 5 models, Random Forest performed the best with an AUC of 0.691 (Figure 1). An AUC ROC curve for the best performing model (Random Forest) is shown in Figure 2. SHAP values showed that BMI, acute inpatient rehabilitation discharge status, and oncologic history were globally most impactful predictors (Figure 3). The SHAP plot showed that extremely high and low BMIs, discharge to acute inpatient rehabilitation, and presence of prior oncologic history were strong positive predictors of 90-day readmission. Socioeconomic status, minority status & language, and overall SVI rank were also found to be in the top 10 most predictive variables for readmission following PLF.

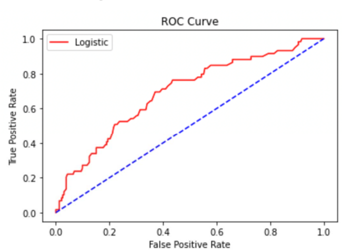
DISCUSSION AND CONCLUSION: This study found that BMI, acute inpatient rehabilitation discharge status, and oncologic diagnosis are most predictive of 90-day readmission following lumbar PLF. Even when including all relevant demographic variables and clinical comorbidities, SDH variables were highly representative in the top 3 most predictive variables of 90-day readmission, as well in the top 10 most predictive factors. This work identified increased risk for 90-day readmission with limited English-language proficiency, higher overall SVI rank, and lower socioeconomic rank. Further studies are needed to clarify the interplay between social determinants of health and patient outcomes.

Figure 1 - Model Validation Metrics

Models	AUC (± CI)
Random Forest	**0.691 (0.610, 0.772)**
Logistic Regression	0.670 (.588, .752)
Gradient Boosted Classifier	0.644 (0.561, 0.723)
XGBoost Classifier	0.621 (0.534, 0.704)
Decision Tree	.566 (.482, .650)

Figure 1: List the AUC for competing models.

Figure 2 – Random Forest ROC-AUC



AUC = 0.691

Figure 3 – Feature Importance SHAP Plot