Social Determinants of Health Play a Role in Determining Health Utilization following PLIF or TLIF.

Samuel Gregory Reyes, Steven Kurapaty, PRANAV MALHOTRA BAJAJ, 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 social factors that affect the overall wellbeing of a person, including both their physical and mental states. Even considering their impacts on health, SDH have not been thoroughly investigated within spine surgery. New machine learning (ML) techniques are being employed throughout the medical field to gain novel insights into patient outcomes. This study will apply ML to understand relationships between SDH and postoperative health utilization in patients undergoing posterior lumbar or transforaminal lumbar interbody fusion surgeries (PLIF/TLIF).

METHODS: Patients that underwent single or multi-level PLIF/TLIF at a multi-center academic health system between 2002 –2020 were acquired from the electronic medical record. More than 50 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 different 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 a postoperative health utilization score, which was calculated as the sum of all ED or urgent care visits, invasive procedures (e.g. cardiac catheterization), non-routine testing (EMG), and non-routine imaging (CT, X-Ray, MRI) in the first 90 days after surgery. Several ML models were run using custom Python scripts and were validated by mean absolute error (MAE). 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 MAE; SHAP values were calculated for the best model to rank features by impact on output prediction.

RESULTS: A total of 1,368 patients and 53 variables were included in the final sample. Average utilization score for patients was 5.31 (range 0-49), average BMI was 30.1, and 51.7% of the cohort was male. Of the 5 models ran, Random Forest Regressor performed best with an MAE of 1.94 (Figure 1). SHAP values showed that Socioeconomic Status, Household Composition & Disability, and Minority Status & Language were the most important variables in determining increased postoperative health utilization (Figure 2). In addition, Housing Type & Transportation and overall SVI rank were also included in the top 10 most predictive variables for high health utilization.

DISCUSSION AND CONCLUSION: Several predictors of postoperative health utilization following PLIF/TLIF included SDH characteristics, including Socioeconomic Status, Household Composition & Disability, and Minority Status & Language. The results of this study suggest that even when considering relevant clinical variables, SDH characteristics play an important role in overall health utilization postoperatively.

Figure 1 - Model Validation Metrics

Models	MAE
Random Forest Regressor	**1.94**
Gradient Boosted Regressor	1.95
XGBoost Regressor	2.09
Suport Vector Machine (Regression)	2.66
Logistic Regressor	4.09

Figure 1 illustrates the Mean Absolute Error (MAE) metric for competing models.

Figure 2 - SHAP Plot