Using Applied Machine Learning to Predict Healthcare Utilization Based on Socioeconomic Determinants of Care

Soy Chen, MS; Danielle Bergman, BSN, RN; Kelly Miller, DNP, APRN, FNP-BC; Allison Kavanagh, MS; John Frownfelter, MD, MSIS; and John Showalter, MD

Objective
To determine if it is possible to risk-stratify avoidable utilization without clinical data and with limited patient-level data.

Study Design
The aim of this study was to demonstrate the influences of socioeconomic determinants of health (SDH) with regard to avoidable patient-level healthcare utilization. The study investigated the ability of machine learning models to predict risk using only publicly available and purchasable SDH data. A total of 138,115 patients were analyzed from a de-identified database representing 3 health systems in the United States.

Methods
A hold-out methodology was used to ensure that the model’s performance could be tested on a completely independent set of subjects. A proprietary decision tree methodology was used to make the predictions. Only the socioeconomic features—age group, gender, and race—were used in the prediction of a patient’s risk of admission.

Results
The decision tree–based machine learning approach analyzed in this study was able to predict inpatient and emergency department utilization with a high degree of discrimination using only purchasable and publicly available data on SDH.

Conclusion
This study indicates that it is possible to risk-stratify patients’ risk of utilization without interacting with the patient or collecting information beyond the patient’s age, gender, race, and address. The implications of this application are wide and have the potential to positively affect health systems by facilitating targeted patient outreach with specific, individualized interventions to tackle detrimental SDH at not only the individual level but also the neighborhood level.
Key Takeaways

This investigation demonstrates that it is possible to predict individual hospital and emergency department utilization using publicly available data on socioeconomic determinants of care and purchased behavioral data, without requiring clinical risk factors.

It is possible to risk-stratify patients’ risk of utilization without interacting with the patient or collecting information beyond the patient’s age, gender, race, and address.

The implications of this application are wide and have the potential to positively affect health systems by facilitating targeted patient outreach with specific, individualized interventions to tackle detrimental social determinants of health at not only the individual level but also the neighborhood level.

About Jvion

Clinical AI that prevents avoidable patient harm

Jvion™ helps healthcare organizations prevent avoidable patient harm, lower costs, improve quality and patient outcomes with clinical AI-enabled prescriptive analytics solutions. Jvion goes beyond simple predictive analytics and machine learning to identify patients on a trajectory to becoming high risk and for whom interventions will likely be successful. Jvion determines the interventions that will more effectively reduce risk and enable clinical action. It accelerates time to value by leveraging established patient-level intelligence to drive engagement across hospitals, populations, and patients.