

# Risk Stratification of 92,593 Patients with Class 3 Obesity Receiving Total Knee Arthroplasty Using Machine Learning Clustering Analysis

Alejandra Chavarria, Anthony J Kaywood, Alishah Ahmadi, Oserekpamen Favour Omobhude, Adam C Kiss, Mateusz Falryn, Jason Shih Hoellwarth

## INTRODUCTION:

Class 3 obesity, defined as a BMI of 40 or higher, is a well-established risk factor for postoperative complications following total knee arthroplasty (TKA). Patients in this category often present with multiple comorbidities that can further increase surgical risk and impair postoperative recovery. These challenges highlight the need for risk stratification methods tailored specifically to this high-risk population. This study uses machine learning based clustering algorithms to identify unique patient profiles that can provide insight into how certain comorbidities affect TKA outcomes in patients with class 3 obesity.

## METHODS:

The 2015-2021 NIS was queried using ICD-10 CM/PCS coding to identify patients with class 3 obesity who had a TKA. Machine learning clustering analysis evaluated the population based on 49 comorbidities, complications and clinical covariates. The optimal number of clusters was determined using the Davies-Bouldin Index (DBI) and Calinski-Harabasz Index (CHI). Between-cluster multivariate logistic regression analysis was performed to assess risk of non-routine discharge (NRD). Kruskal-Wallis H-Testing was performed to assess variance in length-of-stay between clusters. Statistical analysis was performed using Python, noting  $p < .05$  as significant.

## RESULTS:

92,593 patients were included. The optimal number of clusters was determined to be 6 based on composite DBI and CHI scoring; however, cluster 6 was comprised of less than 10 patients so it was excluded from this study. Most clusters demonstrated a high prevalence of osteoarthritis, hyperlipidemia, and hypertension. Cluster 1 had the lowest rate of NRD (65.12%) and the lowest prevalence of comorbidities. Cluster 2 had a NRD of 68.51% and a high prevalence of anxiety and depression. Cluster 3 had a NRD of 75.07% and a high prevalence of arrhythmia, diabetes, heart failure, and sleep apnea. Cluster 4 had a NRD of 85.09% and a high prevalence of acute kidney failure and periprosthetic joint infection. Cluster 5 had the highest rate of NRD (85.46%) and the highest prevalence of anemia and chronic kidney disease. The adjusted odds ratio of NRD for Cluster 5 was 2.51 (95% CI: 1.98-3.18,  $p < .05$ ) compared to Cluster 1. Kruskal-Wallis H-testing and post-hoc pairwise testing of length-of-stay distributions showed significant differences between all clusters.

## DISCUSSION AND CONCLUSION:

This analysis demonstrates the utility of machine learning based clustering in identifying 5 clinically meaningful subgroups among class 3 obese patients undergoing TKA. Cluster 1 was characterized by the lowest burden of comorbidities and had the lowest NRD rate, representing a low risk group. In contrast, Cluster 5 exhibited the highest NRD rate and had the highest rates of anemia and chronic kidney disease, identifying a significantly high risk population. Notably, mental health conditions predominated in Cluster 2, cardiometabolic disease was more prevalent in Cluster 3, and postoperative infection defined Cluster 4. Understanding the complex relationship between obesity and TKA provides insights that can be used to develop targeted clinical protocols that address the specific needs of each patient subgroup, potentially improving implant success.

