Utilizing Machine Learning to Predict Inpatient Mortality Following Revision Total Hip Arthroplasty in Patients with Periprosthetic Joint Infections

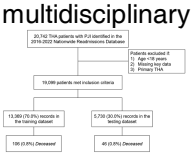
Nora A Galoustian[1], Jeffrey Balian, Christopher D Hamad, Armin Alipour[2], Peyman Benharash, Alexandra Stavrakis
[1]Department of Orthopaedic Surgery, [2]Department of General Surgery

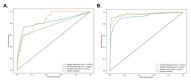
INTRODUCTION: Periprosthetic joint infection (PJI) following total hip arthroplasty (THA) is associated with high morbidity and a 25% 5 year mortality. This is the first study to utilize machine learning (ML) robust oversampling techniques to identify risk factors and predict postoperative mortality occurring during the same admission after revision THA for PJI. These techniques potentially enable improved mortality risk prediction critical to identifying high-risk patients and implementing preventative strategies.

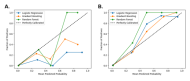
METHODS: This retrospective cohort study utilized the Nationwide Readmissions Database (NRD) to identify adult patients undergoing revision THA for PJI. ML classification models were developed to predict in-hospital mortality and compared to logistic regression. Synthetic Minority Oversampling Technique (SMOTE) was applied to address data imbalance by duplicating observations in the deceased cohort to improve classification accuracy. Model performance was assessed using AUROC, Brier score, and F1 score, with SHapley Additive exPlanation (SHAP) analysis identifying key predictors.

RESULTS: Among 19,099 patients undergoing revision THA for PJI, there was a 0.8% mortality rate (Figure 1). The deceased cohort was older and had higher rates of congestive heart failure, coagulopathy, and renal failure (Table 1). Logistic regression (AUROC 0.855) and gradient boosting (AUROC 0.862) outperformed random forest (AUROC 0.765), with gradient boosting demonstrating the highest F1 score (0.257) (Table 2). Following the application of SMOTE, all models demonstrated improved AUROC, model calibration, and enhanced F1 (Figures 2 and 3). SHAP analysis identified fluid and electrolyte disorders, advancing age, and cardiac arrhythmia as key predictors of mortality (Figure 4). Conversely, elective admission had the greatest protective effect against postoperative inpatient death.

DISCUSSION AND CONCLUSION: This study introduces a novel, streamlined risk score utilizing gradient boosting ML and SMOTE to identify high-risk revision THA patients with PJI, enabling targeted perioperative interventions to mitigate mortality risk. While SMOTE has improved ML accuracy in other surgical specialties, this is its first known application investigating patient outcomes in orthopaedics. Gradient boosting demonstrated strong predictive performance for in-hospital mortality in revision THA PJI patients. SMOTE improved model calibration and precision, highlighting its utility in imbalanced orthopaedic datasets. Fluid and electrolyte abnormalities were the strongest comorbid predictor of mortality, highlighting its critical role in PJI outcomes. Given the established link between high-dose antibiotic-loaded bone cement spacers and calcium sulfate antibiotic impregnated beads with AKI and subsequent electrolyte disturbances, optimizing electrolyte levels perioperatively may be a potential mortality-reducing intervention. Institutional protocols are warranted to provide adequate post-operative maintenance fluids, strict monitoring and repletion of electrolytes, monitoring of systemic antibiotic levels, and tracking systemic calcium levels in patients with greater calcium sulfate bead loads >10cc. Cardiac arrhythmias were identified as the second greatest medical risk factor for mortality, reinforcing the importance of perioperative management strategies such as interdisciplinary care. Notably, elective admission was the strongest protective factor against mortality, likely due to its role in facilitating preoperative optimization and coordinated multidisciplinary care.







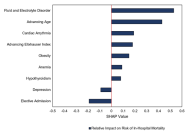


Table 1: Demographic and clinical characteristics of patients with PJI undergoing THA were stratified by whether the patient survived to discharge.
Includes surgery, centrained, and other
*Abbreviations: IQR, interquartile range, SD, standard deviation,

	Alive (n = 19,047)	Deceased (n = 52)	P value
Age (years, median, IQR)	67 [57-77]	73 [64-81]	0.001
Female, %	61.6	51.9	0.28
Elixhauser score (median, IQR)	3 [2-4]	6 [4-7]	<0.001
Elective admission	42.8	25.0	<0.001
Income Quartile, %			0.27
1%, 100	21.2	24.1	
2%-75	26.7	22.2	
76-90	26.3	13.2	
91-100	25.8	32.5	
Payer, %			0.01
Private	27.8	11.1	
Medicare	64.6	77.8	
Medicaid	7.6	0.0	
Uninsured	1.2	1.0	
Other	3.2	2.6	
Comorbidities, %			
Diabetes mellitus	22.0	29.0	0.22
Renal failure	32.8	53.1	<0.001
Fluid and electrolyte disorder	25.0	63.2	<0.001
Unintentional weight loss	1.4	10.4	<0.001
Anemia	9.4	9.8	0.98
Cancer	2.1	6.7	0.08
Liver disorder	2.1	1.0	0.68
Blood transfusion	20.1	37.7	0.28
Rheumatoid arthritis	5.6	5.6	0.94
Collagen vascular disorder	2.5	5.4	0.01
Smoker	11.9	5.0	0.02
Alcohol Abuse	2.9	3.0	0.98
Opioid Abuse	2.9	3.1	0.98
Malnutrition	2.9	18.4	<0.001
Psychosis	1.3	3.0	0.21
Paralysis	2.5	3.0	0.001
Peripheral vascular disease	5.2	0.0	0.11
Pulmonary edema	1.2	10.4	<0.001
Coronary artery disease	10.4	26.7	0.01
Hospital volume, %			0.35
Low Volume	2.9	2.2	
Medium Volume	35.3	19.4	
High Volume	61.8	78.4	
Hospital type, %			0.60
Metropolitan teaching	4.8	2.0	
Metropolitan non-teaching	76.0	79.6	
Non-metropolitan	19.2	18.3	

Table 2: Performance metrics between logistic regression, random forest and gradient boosting models before oversampling.
Performance metrics reported as mean with 95% confidence intervals and obtained via 10-fold cross validation.
LR: Logistic Regression, RF: Random Forest, GB: Gradient Boost, CI, Confidence Interval, AUROC, Area Under the Receiver Operating Characteristic.

Metric	LR (95% CI)	RF (95% CI)	GB (95% CI)
AUROC	0.855 (0.840 – 0.871)	0.765 (0.752 – 0.779)	0.862 (0.846 – 0.876)
Brier Score	0.058 (0.050 – 0.062)	0.063 (0.056 – 0.069)	0.059 (0.051 – 0.067)
F1 Score	0.242 (0.231 – 0.247)	0.149 (0.142 – 0.156)	0.257 (0.251 – 0.263)