



Classification and stratification of patient pain archetypes following total knee arthroplasty: a machine learning approach

Justin Chew, MD, PhD; Junying Wang; Alexandra Sideris, PhD; Victoria Xu; Daniel Maalouf, MD, MPH; Stavros Memtsoudis, MD, PhD, MBA; Jashvant Poeran, MD, PhD; Jiabin Liu, MD, PhD

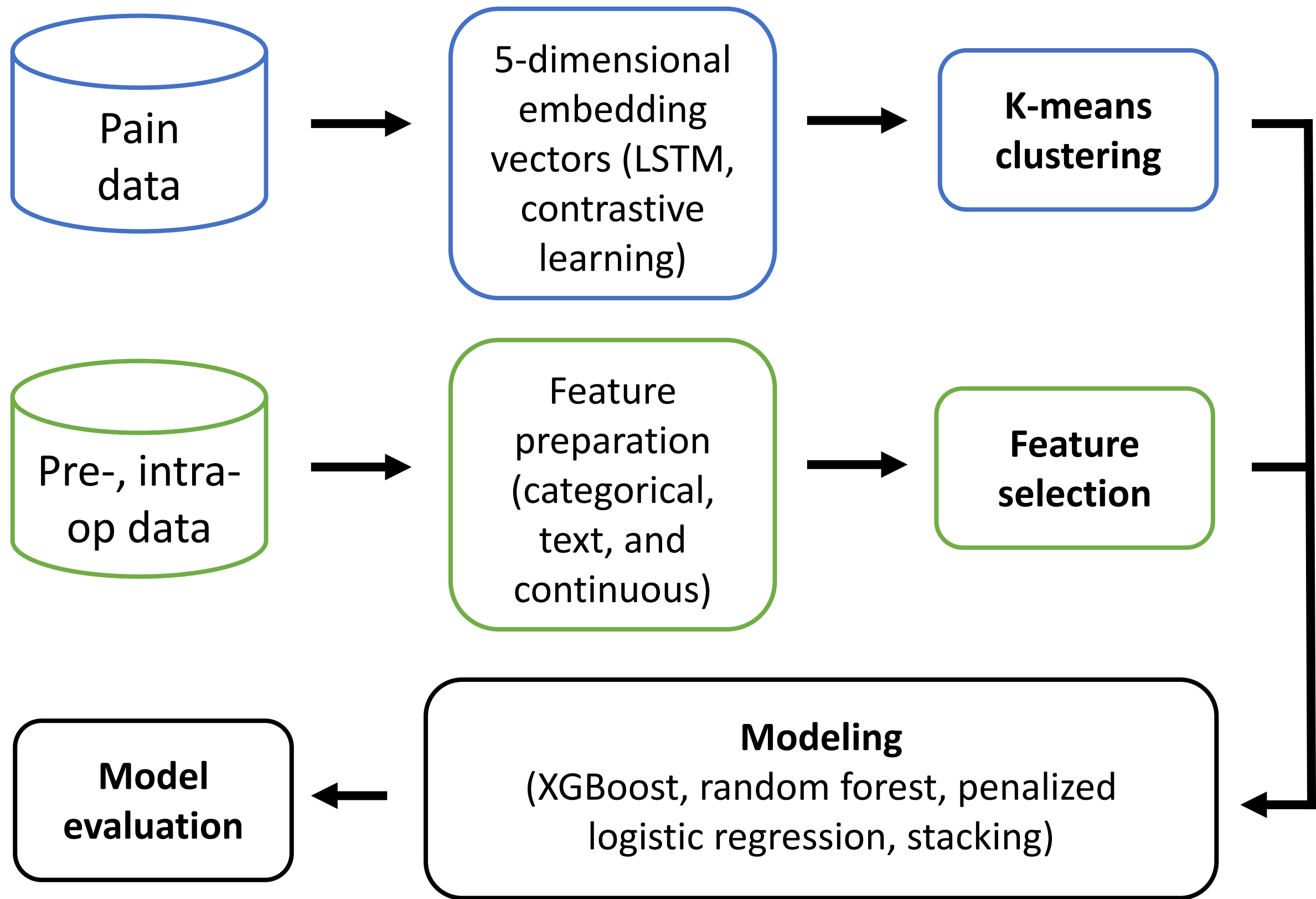
Department of Anesthesiology and Pain Prevention Research Center, Hospital for Special Surgery, New York, NY



I. Introduction

- Heterogeneity in pain courses following total knee arthroplasty (TKA) remains poorly characterized
- We utilize machine learning to identify pain archetypes following TKA, determine important features for predicting pain outcomes, and classify patients by risk of severe post-op pain
- This is a retrospective, single center study including all patients undergoing TKA from April 1, 2021 through October 31, 2024 (n = 17,200)

II. Data processing and modeling pipeline



III. Pain clustering results and characterization

Table 1: Baseline demographics

Sex	Female	10475 (60.9%)
	Male	6725 (39.1%)
Race	Asian	689 (4.01%)
	Black	1527 (8.88%)
	Other	1502 (8.73%)
	White	13482 (78.38%)
Ethnicity	Hispanic	1271 (7.39%)
	Not Hispanic	15513 (90.19%)
	Other	416 (2.42%)
Patient Class	Ambulatory	4003 (23.27%)
	Inpatient	13197 (76.73%)
Age		68.25 +/- 12.78
BMI		30.3 +/- 8.1

Figure 1: Hourly average pain scores

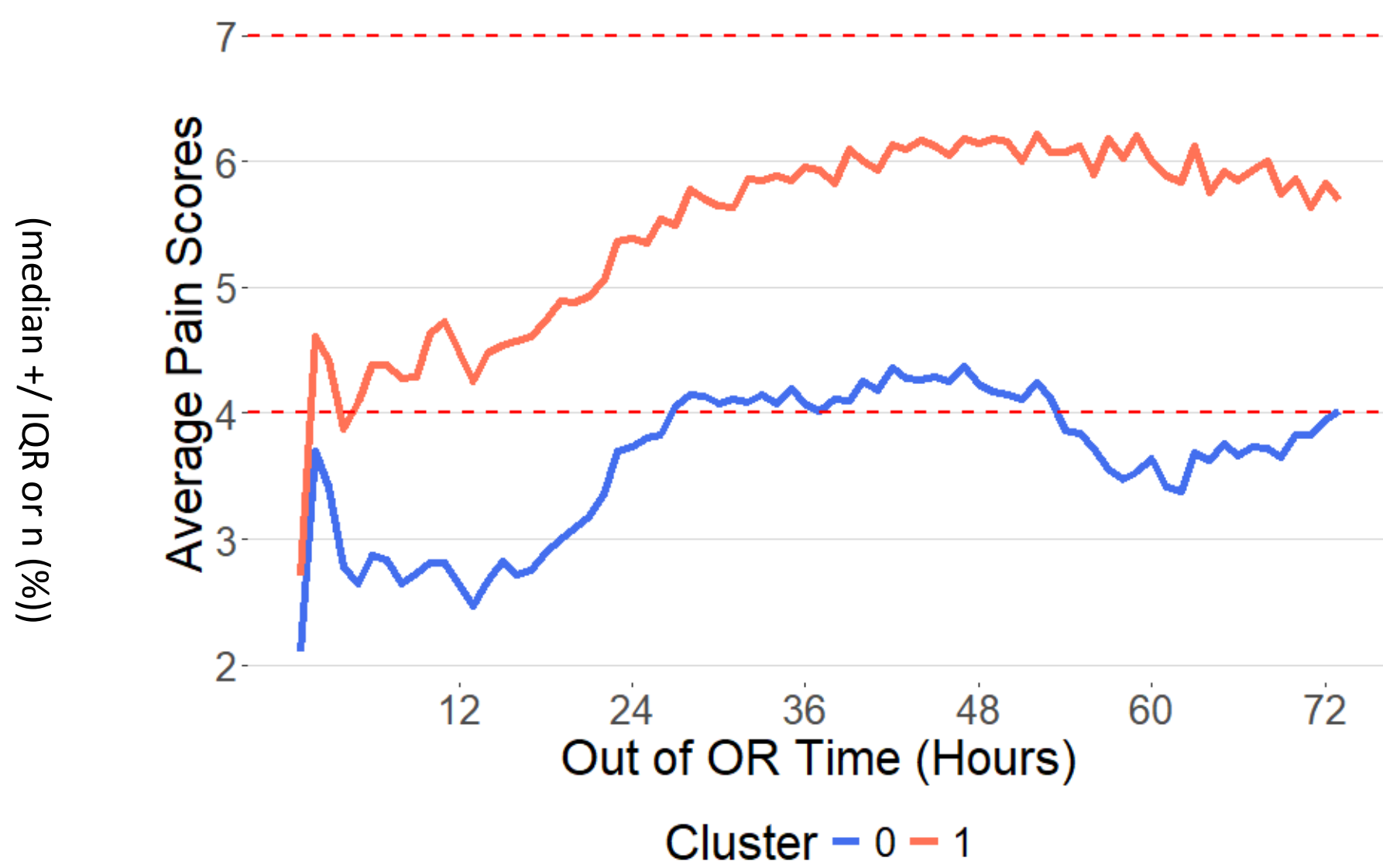


Table 2: Pain metrics/outcomes by cluster (median +/- IQR or n (%))

Variable		Cluster 0 (N=8686)	Cluster 1 (N=8514)	P-value
Average NRS score		2.90 ± 2.08	4.39 ± 1.77	<.0001
Median NRS score		3.00 ± 3.50	4.50 ± 2.50	<.0001
Maximum NRS score		7.00 ± 1.00	8.00 ± 2.00	<.0001
Normalized area under pain vs. time curve		2.68 ± 2.11	4.49 ± 1.75	<.0001
Number of pain scores > 4 (first 24 hours)		5.00 ± 6.00	8.00 ± 7.00	<.0001
Number of pain scores > 7 (first 24 hours)		2.00 ± 4.00	4.00 ± 4.00	<.0001
Total MMEs consumed over hospitalization		40.00 ± 40.00	62.50 ± 49.00	<.0001
Variable	Level	Cluster 0 (N=8686)	Cluster 1 (N=8514)	P-value
Required chronic pain consult?	N	8427 (97.02%)	8033 (94.35%)	<.0001
	Y	259 (2.98%)	481 (5.65%)	
Required acute pain consult?	N	7755 (89.28%)	7871 (92.45%)	<.0001
	Y	931 (10.72%)	643 (7.55%)	

*: For categorical variables, p-values were based on the Chi-squared test;
For continuous variables, p-values were based on the Wilcoxon rank sum tests.

IV. Identification of features significant for prediction

Figure 2: Bar plot of ranked feature importance (XGBoost)

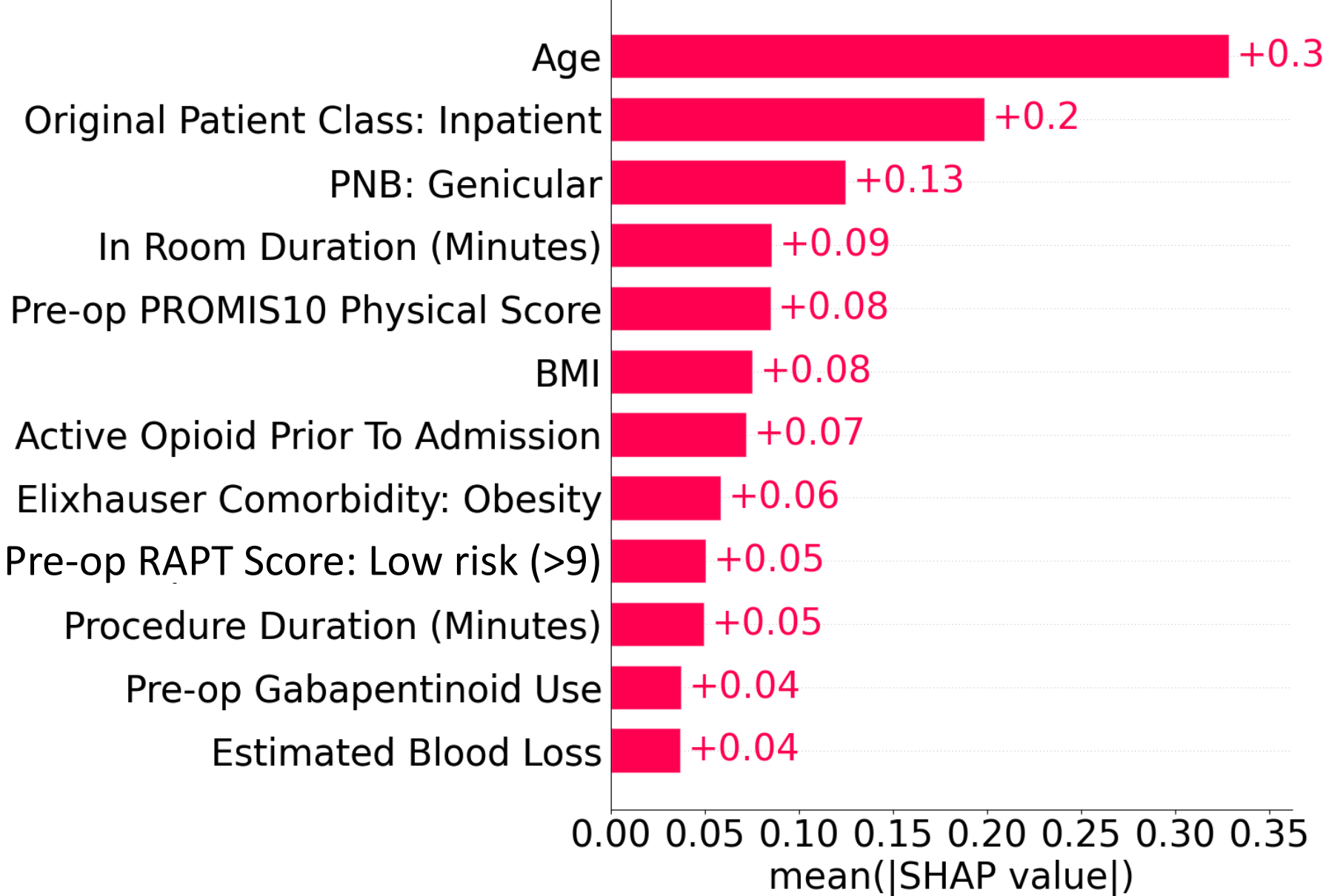
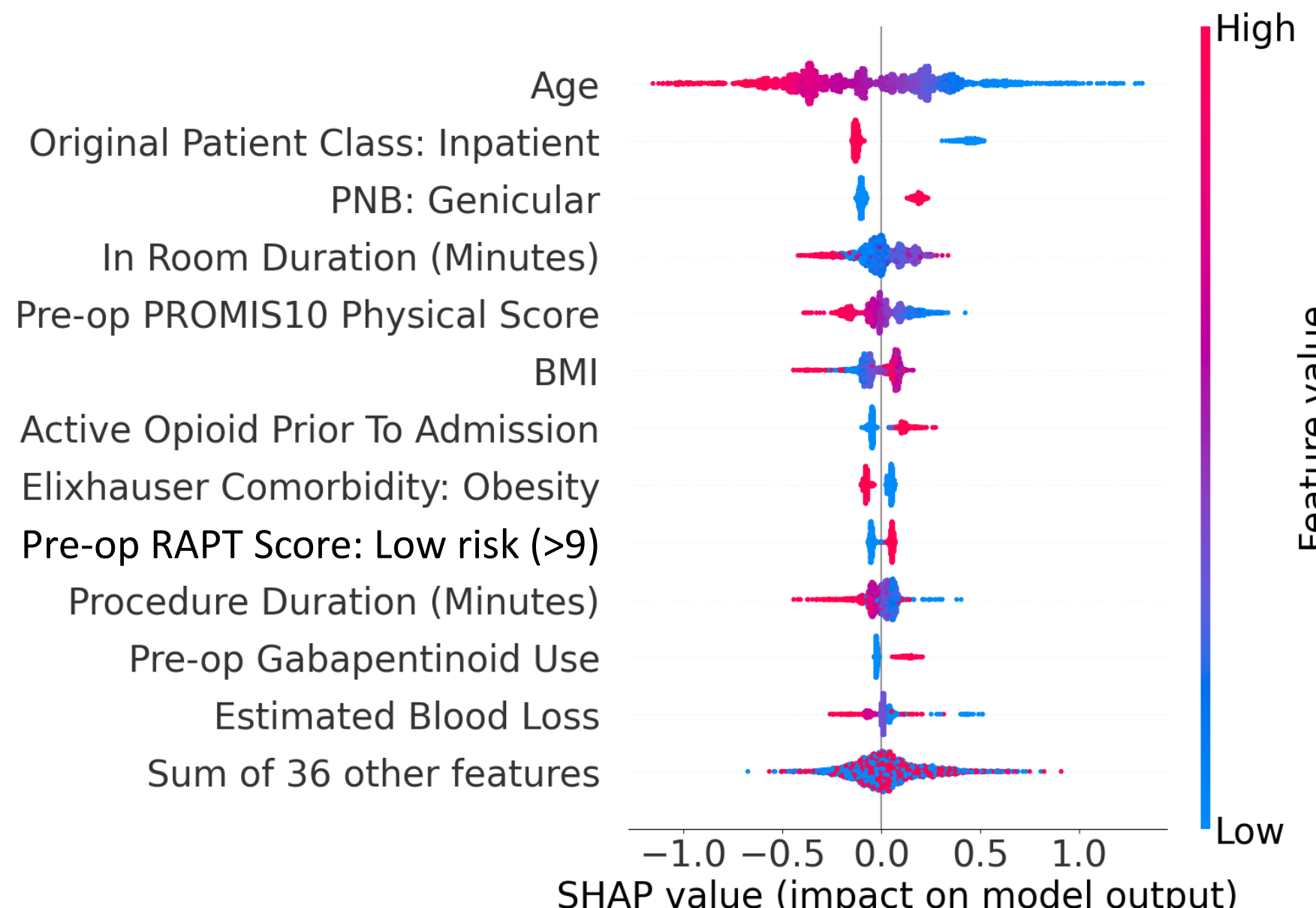


Figure 3: Beeswarm plot for features effect on prediction (XGBoost)



V. Model prediction results using XGBoost

Data	TP	FP	TN	FN	Accuracy	Sensitivity	Precision	F1-score	AUC
Training	4106	2266	4717	2671	0.6412	0.6059	0.6444	0.6245	0.6962
Testing	968	572	1131	769	0.6102	0.5573	0.6286	0.5908	0.6473

VI. Conclusion

- Clustered patients into two groups with markedly different pain trajectories after TKA (by pain scores, MMEs consumed, and pain consult rate), stratifying into patients with either well controlled or poorly controlled post-op pain
- Identified the most important predictive factors as age, inpatient status, genicular block, OR duration, BMI, prior opioid or gabapentin use, and EBL
- Overall, achieved a 61% accuracy rate of classification on testing data subset