Prediction of Day Case Discharge Failure in Total Hip and Knee Arthroplasty Using Supervised Machine Learning in a Non-Selective Enhanced Recovery System

Amr Selim, Abdelrahman Ibrahim[1], Rebecca Louise Warren, Niall Graham, Daniel Redfern, Geraint Emyr Rhys Thomas
[1]Trauma & Orthopaedics

INTRODUCTION:

Enhanced Recovery After Surgery (ERAS) protocols for hip and knee arthroplasty have been widely adopted to accelerate recovery, improve patient outcomes, reduce costs, shorten hospital stays, and standardise care.

Predicting failure to achieve day-case discharge is critical for optimising perioperative planning, allocating resources, and identifying modifiable risk factors to to improve outcomes. This study aimed to develop and validate supervised machine learning (ML) models to predict discharge failure in a universal, standardised, non-selective ERAS pathway implemented at a high-volume tertiary orthopaedic centre.

METHODS:

All patients undergoing primary total hip arthroplasty (THA), total knee arthroplasty (TKA), or partial knee replacement (PKR) under a universal non-selective ERAS pathway between April 2024 and April 2025 were included. Key components of this pathway include preoperative education through Joint School, peri-operative nutritional support, a standardised anaesthetic and analgesic regimen, early postoperative physiotherapy, and telephone follow-up after discharge.

Prospectively collected data were analysed. Supervised ML models including Logistic Regression, Random Forest, Gradient Boosting, and Decision Tree were developed using Python. An 80/20 train-test split was used, with hyperparameter tuning via GridSearchCV. Five-fold cross-validation was employed to optimise model performance, and SMOTE (Synthetic Minority Over-sampling Technique) was used to addressed class imbalance between discharge successes and failures prior to splitting.

Predictors included in the model were patient demographics, the joint operated on, attendance at Joint School, preoperative nutrition and volume, and anaesthetic variables such as spinal drug type (Bupivacaine or Prilocaine), preparation method (Plain vs Heavy), spinal drug concentration, and spinal volume. The spinal dose was calculated from volume and concentration and stratified into Prilocaine and Bupivacaine subgroups. Additional features included whether general anaesthesia (GA) was given, the intraoperative dexamethasone (DXM) dose, nerve block details, postoperative pain scores, whether an X-ray was performed before return to the ward, requirement for blood transfusion, time to mobilisation, postoperative nutrition and volume, and the designation of rehabilitation staff. SHAP (SHapley Additive exPlanations) values were used to interpret feature importance.

RESULTS:

A total of 3,025 patients were included (median age 70.1 years, IQR 62.3–76.6). Females comprised 49.8% of the cohort. The majority of patients were ASA grade 2 (69.2%), followed by grade 3 (17.7%). THA, TKA, and PKR accounted for 44.4%, 51.5%, and 4.1% of procedures, respectively. The right side was operated on in 52.6% of cases. Median length of stay was 1 day (IQR 0–1), with a median time to discharge of 29 hours (IQR 12.5–33) from admission. Day 0 discharge occurred in 822 patients (26.8%), and 2,481 patients (82.0%) were discharged on Day 0 or 1.

For prediction of failure to discharge at Day 0 or 1, ML models demonstrated strong performance. Across all models, test precision ranged from 0.76 to 0.96, recall from 0.80 to 0.94, F1 scores from 0.78 to 0.95, and ROC AUC from 0.86 to 0.99, with Random Forest and Gradient Boosting showing the highest overall performance.

For Day 0 discharge failure prediction, test precision ranged from 0.78 to 0.82, recall from 0.57 to 0.74, F1 scores from 0.67 to 0.78, and ROC AUC from 0.77 to 0.88. While performance was slightly reduced, predictive accuracy remained robust.

SHAP feature importance analysis identified the top predictors for discharge failure. For Day 0 or 1 failure, the most influential features included the use of higher spinal drug concentrations, particularly Bupivacaine over Prilocaine, higher Bupivacaine doses, absence of intra-operative DXM, older patient age, longer time to mobilisation, and joint type; where TKA and THA were associated with higher failure risk compared to PKR.

For Day 0 discharge prediction (Figure 1), time to mobilisation was by far the most significant predictor, followed by Bupivacaine dose (with lower doses having earlier discharge), patient sex (with males more likely to be discharged earlier), and the designation of rehabilitation staff (with physiotherapists and Enhanced Recovery practitioners associated with greater discharge success). Older age increased the risk of failure, while administrating intra-operative DXM reduced it. Joint type again showed a consistent trend, with PKR patients more likely to be discharged on Day 0 than those undergoing THA or TKA.

DISCUSSION AND CONCLUSION:

Accurately predicting discharge failure in day case hip and knee arthroplasty is essential for optimising peri-operative care and resource allocation. This study demonstrates that ML models can reliably predict both same-day and early discharge failures using standard perioperative variables. Random Forest and Gradient Boosting models achieved the highest

performance. SHAP feature importance analysis identified key modifiable and non-modifiable factors that contribute to discharge failure. Among the non-modifiable factors, older age, female sex, higher ASA grades, and bilateral procedures were associated with a higher likelihood of discharge failure. This could aid organisations in planning and resource allocation before admission. Modifiable factors included use of spinal anaesthesia alone (without GA), choice of spinal drug (Prilocaine associated with better outcomes than Bupivacaine), lower spinal dose where appropriate, administration of DXM, early mobilisation upon return to the ward, rehabilitation led by physiotherapists or Enhanced Recovery-trained staff, and provision of postoperative nutrition.