prosthesis

MDPI

Review

Machine Learning-Based Prediction of Postoperative Complications Using Laboratory and Clinical Parameters in Total Hip and Total Knee Arthroplasty: A Scoping Review

Platon S. Papageorgiou [1,†], Dimitra P. Papagelopoulos [1,†], Aris Sopilidis [1], Andreas Tsoupras [2], Rafail C. Christodoulou [3], Ioannis E. Kougioumtzis [4], Vasileios Karampikas [1], Andreas F. Mavrogenis [1], Panayiotis J. Papagelopoulos [1] and Panayiotis D. Megaloikonomos [1,*]

[1] First Department of Orthopaedic Surgery, Attikon University General Hospital, National and Kapodistrian University of Athens, 12462 Athens, Greece
[2] Paediatric Orthopedics Unit, Geneva University Hospitals, University of Geneva, 1205 Geneva, Switzerland
[3] Department of Radiology, School of Medicine, Stanford University, Stanford, CA 94305, USA
[4] Sports Medicine and Orthopaedic Traumatology, General Hospital of Nikaia, 18454 Piraeus, Greece
* Correspondence: pdmeg@med.uoa.gr
† These authors contributed equally to this work.

Abstract

Background: Total joint arthroplasties (TJAs) are well-established orthopedic procedures that provide effective pain relief and restore functional mobility. Although these procedures are generally highly successful, postoperative complications may still occur. This review examined the literature on how machine learning (ML) can predict postoperative complications based on laboratory and clinical parameters for total knee (TKA) and total hip arthroplasty (THA). **Methods:** A scoping review was conducted and reported in accordance with the PRISMA criteria for scoping reviews (PRISMA-ScR). We searched PubMed/MEDLINE, Scopus, and EMBASE for studies evaluating the use of ML to predict adverse events following TKA and/or THA, using routinely available laboratory values combined with demographics and clinical parameters. **Results:** The initial search yielded 144 studies, of which 18 were finally deemed eligible for inclusion. Key predictors of complications in the examined studies included hemoglobin, hematocrit, white blood cell count, platelet count, and renal markers. Need for transfusion was the most frequently studied outcome, followed by severe adverse events and postoperative delirium. The ML models, especially gradient boosting machines (GBMs), random forests (RFs), and neural networks (NNs), demonstrated promising discriminative performance in predicting adverse events. **Conclusions:** Combining laboratory data with demographics and clinical information may improve the accuracy of complication prediction. This review emphasizes the potential of ML analysis of laboratory parameters to create risk assessment tools, which could detect complication risk and potentially improve outcomes for patients undergoing TKA or THA.

Keywords: total hip arthroplasty; total knee arthroplasty; complications; laboratory parameters; machine learning; risk prediction; personalized perioperative care

Check for updates

Academic Editors: George Mihai Avram and Marco Cicciu

Received: 9 August 2026
Revised: 13 October 2026
Accepted: 28 October 2026
Published: 2 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Total joint arthroplasty (TJA) gained popularity during the 20th century, with hip and knee replacements representing the most common and effective elective procedures in orthopaedic surgery [1]. Traditionally, these procedures have been indicated for end-stage osteoarthritis, inflammatory arthritis, osteonecrosis, or post-traumatic conditions [2].

It is estimated that by 2030, 3.5 million primary total knee arthroplasties (TKAs) and 500,000 primary total hip arthroplasties (THAs) will be performed annually [2]. These procedures are highly effective in restoring mobility and relieving pain. Although not very common, some patients may experience postoperative adverse events. These include prolonged length of stay (LOS), postoperative delirium (POD), anemia, and the need for blood transfusion. More serious complications include acute kidney injury (AKI), deep vein thrombosis (DVT), pulmonary embolism (PE), surgical site infection (SSI), periprosthetic joint infection (PJI), and sepsis [3–7].

Machine learning (ML) recently emerged as a transformative tool in healthcare, offering innovative solutions for predicting and mitigating postoperative complications in patients undergoing primary TJA [8]. Recent advances in ML algorithms, ranging from random forests (RFs) and gradient boosting machines (GBMs) to neural networks (NNs), have demonstrated superior accuracy in predicting complications compared with traditional statistical methods [9,10]. The development of interpretable ML frameworks and web-based risk calculators promise personalized perioperative decision-making feedback, enabling clinicians to identify high-risk patients and tailor their interventions accordingly [10,11]. This review aimed to explore the current literature on the use of ML for predicting postoperative complications following primary TKA and THA using laboratory and clinical inputs.

2. Materials and Methods

2.1. Study Design

This scoping review was conducted to systematically map, synthesize, and analyze the existing literature regarding the use of ML algorithms to predict postoperative complications following TKA and THA, using preoperative laboratory and clinical parameters. The protocol was not prospectively registered, as registration is not a mandatory element for scoping reviews. Nevertheless, the review was conducted according to prespecified methodology and eligibility criteria, while search strategy and data synthesis are transparently presented in detail. Reporting was performed according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) guidelines (Supplementary File, A) [12].

2.2. Eligibility Criteria

Eligibility criteria were established using the Population–Concept–Context (PCC) framework, a widely adopted approach for scoping reviews. The population comprised adult patients undergoing primary, elective TKA and/or THA. The concept of interest was the development, evaluation, or re-evaluation of ML, deep learning (DL), or artificial intelligence (AI) models incorporating routinely available laboratory parameters, combined with demographic and/or clinical data, as input features. Laboratory parameters included, but were not limited to, complete blood count (CBC) indices, biochemical measures of renal, nutritional and metabolic function, and coagulation parameters. The context comprised the prediction of postoperative adverse events, including postoperative complications of variable severity, prolonged LOS and unplanned readmission.

We defined ML models as multivariable predictive models using clinical, laboratory, or patient-level variables to predict postoperative outcomes. Hence, logistic regression (LR), nomograms, clustering, and other predictive algorithms were included when specifically developed for prediction. Simple statistical analyses and linear regression models without a predictive modeling framework were excluded.

Studies were excluded if they evaluated patients undergoing arthroplasty for fractures (i.e., hip fractures). Reports that did not incorporate laboratory parameters as model input

or did not examine postoperative adverse events were also excluded. Only studies published in English were considered eligible.

2.3. Search Strategy and Study Selection

We conducted a literature search in PubMed/MEDLINE, Scopus, and EMBASE between 20 January and 20 February 2026. The search strategy included controlled vocabulary and free-text keywords related to the prediction of complications and adverse events using ML models based on laboratory parameters in patients undergoing primary TKA and/or THA. For each database, the search strategy was adapted to its specific indexing system and search syntax. In PubMed/MEDLINE, Medical Subject Headings (MeSH) terms were combined with relevant free-text terms. In Scopus, the corresponding terms were searched using the title, abstract, and keyword fields. In EMBASE, Emtree terms were combined with free-text keywords. Boolean operators ("AND", "OR") were used to combine and identify relevant studies (Supplementary File, B).

The screening process was conducted in two sequential phases. First, the titles and abstracts of all retrieved records were assessed for potential eligibility. Articles that appeared to meet the inclusion criteria, as well as those for which eligibility could not be determined based on the abstract alone, were subsequently assessed through full-text review. The initial screening was conducted by one reviewer (PSP), while a second reviewer (DPP) independently reviewed the eligibility assessments and verified the extracted study records. Any disagreements were resolved through discussion until consensus was reached. The numbers of records identified, screened, excluded, and finally included, along with the specific reasons for exclusion at the full-text assessment stage, are presented in the PRISMA-ScR flow diagram (Figure 1).

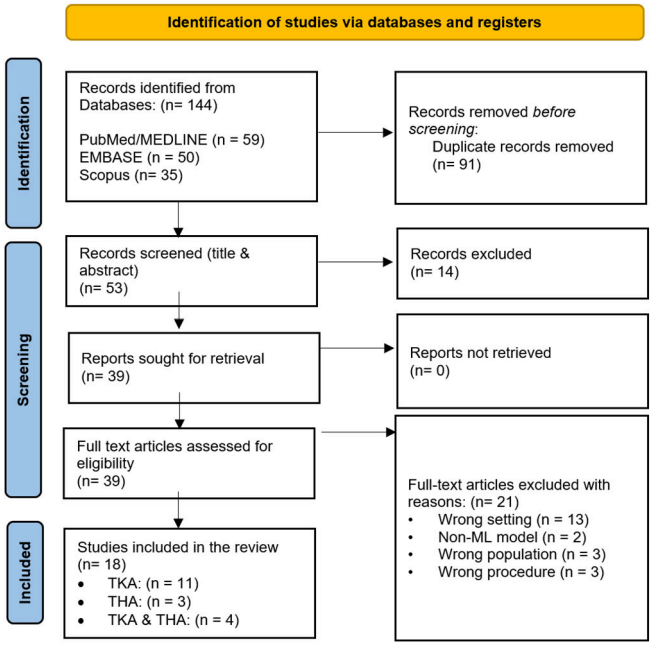


Figure 1. PRISMA-ScR flow diagram illustrating the study identification, screening, eligibility assessment, and inclusion process. *TKA*: Total Knee Arthroplasty; *THA*: Total Hip Arthroplasty; *ML*: Machine Learning.

2.4. Data Synthesis

Results were summarized using both structured and narrative approaches and organized according to the adverse events assessed, including prolonged LOS and unplanned postoperative readmissions. Findings were charted and synthesized descriptively, consis-

tent with the recommended scoping review methodology. Data extracted from each study included the primary author and publication year, study design, arthroplasty type (TKA and/or THA), and examined patient cohort size. Laboratory feature inputs were recorded, including blood biomarkers and metabolic variables incorporated into the predictive models, alongside relevant clinical and demographic information. The ML architecture was characterized according to the specific predictive models employed, while the targeted complications were documented based on the postoperative adverse events predicted by the respective algorithmic pipelines. Model performance was captured using discrimination metrics (Supplementary File, C).

A critical appraisal of the included prediction-model studies was performed using the Prediction model Risk of Bias Assessment Tool (PROBAST) (Supplementary File, D) [13]. Two reviewers independently assessed risk of bias (ROB) across the domains of participants, predictors, outcomes, and analysis, as well as the applicability of each study with respect to participants, predictors, and outcomes. Disagreements were resolved by discussion. The PROBAST findings were summarized descriptively and were not used as an exclusion criterion.

3. Results

3.1. Study Selection

The database searches identified 144 entries, including 59 records from PubMed/MEDLINE, 50 from EMBASE, and 35 records from Scopus. After removal of 91 duplicate records, 53 unique articles were screened by title and abstract. Fourteen records were excluded during this initial screening, and 39 reports proceeded to full-text assessment. All 39 reports were successfully retrieved. Following full-text assessment, 21 reports were excluded; 13 due to non-appropriate setting, two did not report the use of ML, three examined the wrong population and three assessed other procedures (Figure 1).

Eighteen studies met the inclusion criteria and were finally included in the scoping review. Of these, 11 studies (61.1%) focused on TKA, three studies (16.7%) on THA, and four studies (22.2%) focused on both procedures [9,10,14–29]. Two studies assessed the same patient cohort [16,25]. Additionally, three studies used data from the same database, with potentially overlapping patient populations [9,10,29]. Therefore, the cumulative sample size across studies cannot be reliably calculated, as the extent of overlap is unclear. Thus, sample sizes are reported separately for each study in the tables.

The application of ML algorithms to laboratory and clinical data demonstrated potential in enhancing predictive accuracy for complications, supporting the development of personalized risk assessment tools to improve perioperative care. Notably, transfusion requirement was the most frequently examined outcome (Figure 2).

3.2. Risk of Bias and Applicability Assessment

The PROBAST ROB and applicability assessment showed considerable variation in the methodological quality of the included studies [13]. Of the 18 studies assessed, three were classified as having a low overall ROB, seven as having unclear ROB, and eight as having a high overall ROB. Applicability was judged to be high in six studies, unclear in seven, and low in five studies, indicating that most models addressed populations, predictors, and outcomes relevant to TJA, despite limitations in their potential transferability to other clinical settings (Supplementary File, D).

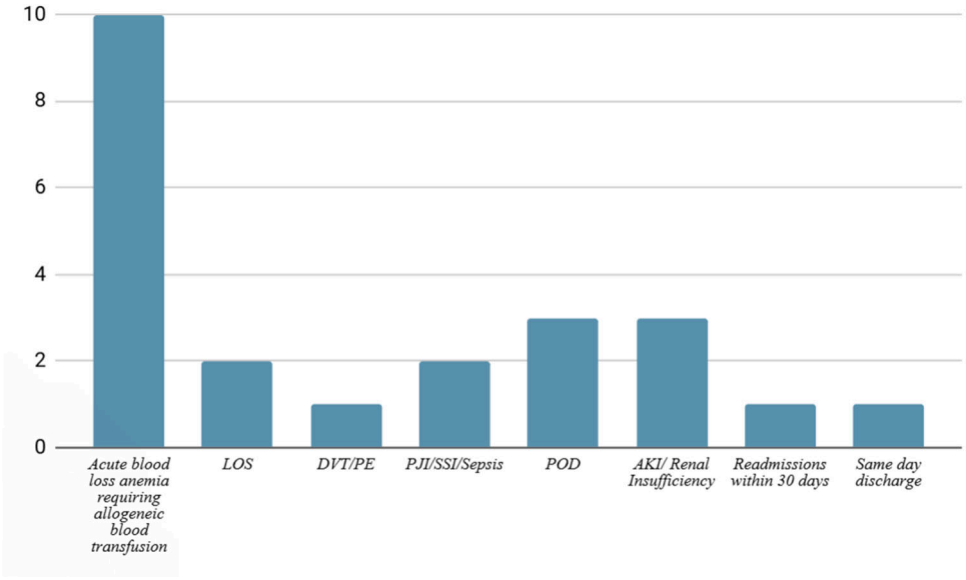


Figure 2. The graph shows the assessed complications across the included studies. *LOS*: prolonged length of stay; *DVT*: deep vein thrombosis; *PE*: pulmonary embolism; *PJI*: prosthetic joint infection; *SSI*: surgical site infection; *POD*: postoperative delirium; *AKI*: acute kidney injury.

3.3. Blood Loss and Transfusion Need

Ten studies studied blood loss and transfusion needs. Eight of them identified preoperative hemoglobin (HGB) as the most significant predictor, with levels below 13 g/dL strongly correlating with increased transfusion needs (Table 1). Cavazos et al. achieved an area under the curve (AUC) of 0.89 by incorporating HGB, creatinine, and albumin into a multilayer perceptron neural network (MPNN), demonstrating the interplay between anemia, nutritional status, metabolic reserve, and hemodynamic stability in TKA patients [14]. Similarly, Cohen et al. applied multiple algorithms to predict transfusion rates in THA, with AUCs ranging from 0.80 to 0.83, using bleeding disorders and preoperative HCT as predictors [15]. A GBM model in another study achieved an AUC of 0.97 by incorporating HGB concentration and platelet (PLT) count, emphasizing the contribution of thrombocytopenia to perioperative bleeding risk [16]. Recent research has also investigated HGB thresholds for predicting transfusion risk in fast-track THA and TKA [17]. Additionally, a GBM-based model demonstrated promising predictive performance for transfusion after TKA, achieving an AUC of 0.84 using HGB and PLT as predictors [18].

Kolin et al. applied LR to predict both postoperative anemia and blood transfusion following TKA with AUCs of 0.88 and 0.90, respectively [19]. Another study used K-means clustering to identify vulnerable patient subgroups at increased risk for various complications, including blood transfusion among Medicare patients undergoing TKA or THA [10]. Further research demonstrated the highest predictive performance with an MPNN, which achieved an AUC of 0.99 for predicting transfusion risk after TKA, using HGB, preoperative anticoagulant use, and antiplatelet use as predictors [20]. Wu et al. used univariate analysis and LR to identify risk factors for substantial hemorrhage following TKA, highlighting PLT as a predictive variable [21]. Huang et al. compared LR with six machine-learning algorithms to predict postoperative transfusion in 12,642 patients undergoing primary THA/TKA, identifying RF and LSTM as the best-performing models (AUC 0.84 and 0.83, respectively), with preoperative HGB, HCT, and ASA class among the key predictors [22].

Table 1. Studies predicting blood loss and transfusion need after total joint arthroplasty (TJA).

Study (Year)	Procedures	Patients	Predictors	Models	AUC
Cavazos et al. (2023) [14]	TKA	2093	HGB, Crea, ALB	MPNN	0.89
Cohen et al. (2023) [15]	THA	7265	Bleeding disorders, HCT	ANN, SGB, SVM, ENP	0.80–0.83
Faure et al. (2025) [16]	TKA	774	HGB, PLT	GBM	0.97
Jørgensen et al. (2024) [17]	THA, TKA	3913	HGB	N/R	N/R
Jo et al. (2020) [18]	TKA	1686	HGB, PLT	GBM	0.84
Kolin et al. (2023) [19]	TKA	14,188	HGB	LR	0.88 (anemia) 0.90 (transfusion)
Ranti et al. (2020) [10]	THA, TKA	134,252	Na, BUN, CREA, ALB, bilirubin, WBC, HCT, PLT, INR	K-means clustering	N/R
Sayeed et al. (2022) [20]	TKA	1328	HGB, anticoagulants, antiplatelets	MPNN	0.99
Wu et al. (2022) [21]	TKA	3350	PLT, HGB	Univariate, LR	N/R
Huang et al. (2021) [22]	THA, TKA	12,642	HGB, HCT	LSTM, RF, LR, DT, KNN, SVM, Naïve Bayes	0.65–0.86 (multiple algorithms)

TKA: Total Knee Arthroplasty; *THA*: Total Hip Arthroplasty; *HGB*: Hemoglobin; *HCT*: Hematocrit; *PLT*: Platelet Count; *Crea*: Creatinine; *ALB*: Albumin; *INR*: International Normalized Ratio; DT: Decision Tree; KNN: k-nearest neighbors; *MPNN*: Multilayer Perceptron Neural Network; *ANN*: Artificial Neural Network; LSTM: Long Short-Term Memory; RF: Random Forest; LR: Logistic Regression; *SGB*: Stochastic Gradient Boosting; *SVM*: Support Vector Machine; *ENP*: Elastic Net Penalization; *GBM*: Gradient Boosting Machine; *LR*: Logistic Regression; *AUC*: Area Under the Curve; *N/R*: Not Reported; *Na*: Sodium; *WBC*: White Blood Cells.

3.4. Severe Adverse Events

Severe adverse events were defined as DVT, PE, SSI, PJI, AKI, renal insufficiency, and metabolic disturbances. Four studies investigating these complications were identified and are summarized in Table 2.

Ranti et al. identified 3153 patients who developed complications following TJA and found that those with substantially elevated preoperative levels of blood urea nitrogen (BUN), creatinine, and international normalized ratio (INR) experienced nearly twice the rate of severe adverse events [10].

Another study demonstrated that preoperative PLT levels < 249,890/mm^3, sodium levels < 139.5 mEq/L, ASA score > 4, high body mass index (BMI), and prolonged LOS were the better predictors of PJI, with a histogram-based gradient boosting (Hist-GBM) model achieving an AUC of 0.88 [23].

Ko et al. achieved an AUC of 0.89 by using creatinine levels > 1.2 mg/dL and eGFR < 60 mL/min/1.73 m^2 in a web-based algorithm, highlighting the predictive value of impaired renal function markers for postoperative outcomes [24]. In contrast, Tran et al. assessed electrolyte disturbances and achieved an AUC of 0.98 using a GBM model with preoperative GFR as key predictor [25].

Table 2. Studies assessing predictors of severe adverse events after total joint arthroplasty (TJA).

Study (Year)	Procedures	Patients	Predictors	Adverse Events	Models	AUC
Ranti et al. (2020) [10]	THA, TKA	134,252	BUN, Crea, INR	PJI, DVT/PE, RI, AKI, sepsis	K-means clustering	N/R
Salimy et al. (2025) [23]	THA	51,053	PLT, Na	PJI	Hist-GBM, ANN, KNN, RF, SVM	0.88 (GBM)
Ko et al. (2022) [24]	TKA	5757	Crea, eGFR, HGB, PLT, BUN	AKI	GBM	0.89
Tran et al. (2025) [25]	TKA	774	GFR, Crea	AKI, K disturbances	GBM	0.98

TKA: Total Knee Arthroplasty; *THA*: Total Hip Arthroplasty; *BUN*: Blood Urea Nitrogen; *Crea*: Creatinine; *INR*: International Normalized Ratio; *PLT*: Platelet Count; *Na*: Sodium; *K*: Potassium; *CCI*: Charlson Comorbidity Index; *eGFR*: Estimated Glomerular Filtration Rate; *GFR*: Glomerular Filtration Rate; *HGB*: Hemoglobin; *HCT*: Hematocrit; *GBM*: Gradient Boosting Machine; *ANN*: Artificial Neural Network; *KNN*: K-Nearest Neighbors; *RF*: Random Forest; *SVM*: Support Vector Machine; *SGB*: Stochastic Gradient Boosting; *ENP*: Elastic Net Penalization; *PJI*: Periprosthetic Joint Infection; *DVT*: Deep Vein Thrombosis; *PE*: Pulmonary Embolism; *RI*: Renal Insufficiency; *AKI*: Acute Kidney Injury; *AUC*: Area Under the Curve; *N/R*: Not Reported.

3.5. Postoperative Delirium

Three studies targeted POD prediction using laboratory markers indicative of nutritional and metabolic status before surgery (Table 3). In a cohort of 3570 patients, with a POD rate of 4.9%, the authors used extreme gradient boosting (XGBoost) algorithms to develop predictive models based on ASA score, preoperative HGB levels, albumin, renal function markers such as the BUN/creatinine ratio, and non-laboratory variables including age, depression, and medication use. The model achieved an AUC of 0.82 [26]. A prospective study of 364 patients undergoing TKA or THA achieved an AUC of 0.96 using LR based on serum albumin < 3.5 g/dL and prealbumin < 15 mg/dL. These findings highlighted the role of malnutrition in postoperative neurocognitive decline and its potential contribution to POD, possibly mediated by changes in total tau (T-tau) and phosphorylated tau (P-tau) levels [27]. In another retrospective study, 681 elderly patients undergoing THA were evaluated for predictors of POD, including the Barthel Index, preoperative HGB levels, diabetes history, and cerebrovascular disease history. A combined least absolute shrinkage and selection operator (LASSO) regression and LR model achieved an AUC of 0.84 [28].

3.6. Prolonged Hospital Stay

Two studies explored the prediction of prolonged hospital stay following TJA using preoperative laboratory parameters and ML models (Table 4). Chen et al. analyzed 267,966 patients undergoing TKA and evaluated prolonged LOS, defined as LOS exceeding the 75th percentile of the cohort. They employed artificial neural network (ANN), RF, and GBM models to predict prolonged LOS, achieving an AUC of 0.71. Their models identified

preoperative transfusion history, HCT < 39% and an American Society of Anesthesiologists (ASA) score ≥ 3 as critical predictors of prolonged LOS [29]. Jørgensen et al. conducted an analysis of an existing ML algorithm in 3913 patients undergoing THA or TKA to identify patients at increased risk of LOS > 4 days or readmission due to medical complications, with preoperative HGB emerging as an important predictor [17].

Table 3. Studies predicting postoperative delirium (POD) after total joint arthroplasty (TJA).

Study (Year)	Procedures	Patients	Predictors	Models	AUC
Jung et al. (2022) [26]	TKA	3570	HGB, BUN/Crea, ALB, WBC, HCT, CRP	XGBoost	0.82
Wang et al. (2024) [27]	TKA, THA	364	ALB, Prealbumin	LR	0.96
Zhang et al. (2024) [28]	THA	681	HGB	LASSO, LR	0.84

TKA: Total Knee Arthroplasty; *THA*: Total Hip Arthroplasty; *HGB*: Hemoglobin; *BUN*: Blood Urea Nitrogen; *Crea*: Creatinine; *ALB*: Albumin; *WBC*: White Blood Cell Count; *HCT*: Hematocrit; *XGBoost*: Extreme Gradient Boosting; *LR*: Logistic Regression; *LASSO*: Least Absolute Shrinkage and Selection Operator; *AUC*: Area Under the Curve.

Table 4. Studies assessing predictors of prolonged length of hospital stay (LOS) after total joint arthroplasty (TJA).

Study (Year)	Procedures	Patients	Predictors	Models	AUC
Chen et al. (2023) [29]	TKA	267,966	HCT, PLT, WBC, preoperative transfusion	ANN, RF, GBM, KNN	0.71
Jørgensen et al. (2024) [17]	THA, TKA	3913	HGB	N/R	N/R

TKA: Total Knee Arthroplasty; *THA*: Total Hip Arthroplasty; *HCT*: Hematocrit; *PLT*: Platelet Count; *WBC*: White Blood Cells; *HGB*: Hemoglobin; *ANN*: Artificial Neural Network; *RF*: Random Forest; *GBM*: Gradient Boosting Machine; *KNN*: K-Nearest Neighbors; *AUC*: Area Under the Curve; *N/R*: Not Reported.

3.7. Readmission and Discharge Planning

Two studies examined discharge planning and readmission prediction following TJA, using laboratory and non-laboratory factors to assess patient readiness for discharge or the risk of complications requiring rehospitalization within 30 days after surgery (Table 5). Wei et al. identified preoperative sodium, preoperative INR, BMI, age, anesthesia type, operative time, dyspnea status, functional status, race, anemia status, and chronic obstructive pulmonary disease (COPD) as key predictors of successful same-day discharge following TKA. Their ANN and LR models achieved AUCs of 0.80 and 0.79, respectively [9]. Ranti et al. further demonstrated the potential clinical utility of BUN, creatinine, and INR for identifying patients at high risk for unplanned readmission, based on an analysis of 5068 readmission cases [10].

Table 5. Studies assessing discharge planning and readmissions after total joint arthroplasty (TJA).

Study (Year)	Procedures	Patients	Predictors	Models	AUC
Wei et al. (2021) [9]	TKA	28,742	Na, INR, HGB	ANN, LR	0.80, 0.79
Ranti et al. (2020) [10]	THA, TKA	134,252	BUN, crea, INR	K-means clustering	N/R

TKA: Total Knee Arthroplasty; *THA*: Total Hip Arthroplasty; *Na*: Sodium; *INR*: International Normalized Ratio; *HGB*: Hemoglobin; *BUN*: Blood Urea Nitrogen; *Crea*: Creatinine; *ANN*: Artificial Neural Network; *LR*: Logistic Regression; *AUC*: Area Under the Curve; *N/R*: Not Reported.

4. Discussion

The ML is a subset of AI that enables computer systems to automatically learn and improve from experience without being explicitly programmed for each specific task [30]. The fundamental principle underlying ML involves analyzing data to identify patterns and make predictions, using algorithms capable of efficiently processing complex datasets [30]. The integration of AI in joint reconstructive surgery has expanded its role beyond surgical precision to predicting clinical outcomes [31]. Recent research indicates that AI-powered algorithms can demonstrate promising discriminative performance in predicting postoperative complications. These models have demonstrated predictive utility across a broad spectrum of outcomes, ranging from relatively mild adverse events, such as transfusion requirements and prolonged LOS, to severe postoperative complications [32]. Moreover, the emergence of personalized data-driven practice (PDDP), powered by AI, enables the integration of various data sources, including imaging, device data, and real-time laboratory metrics, into patient-specific treatment plans, shifting orthopaedic care from reactive to proactive management [33,34].

This review summarizes the emerging evidence on ML models that use routinely available laboratory parameters to predict complications after primary TJA. Across the included studies (Tables 1–5), HGB, HCT, PLT, and markers of renal function were the laboratory variables most consistently associated with postoperative complications, particularly when combined with demographic and clinical characteristics such as age, BMI, ASA score and comorbidity burden [9,10,14–29]. Transfusion requirement was the most frequently examined outcome, followed by severe adverse events, POD, LOS and 30-day readmission rates, highlighting the clinical relevance of these outcomes in the perioperative management of TJA patients and the need for reliable predictive tools to identify high-risk patients and support personalized perioperative care [14–16,18–20,23,25–29].

The reviewed ML models demonstrated high discriminative performance in predicting postoperative adverse events, with the highest predictive performance reported for transfusion risk, AKI, PJI, and POD [14–20,23–29]. Tree-based ensemble methods, MPNN, and RFs demonstrated favorable performance across several studies; however, heterogeneity in datasets, outcomes, and validation approaches precludes firm conclusions regarding the superiority of a specific algorithm [14–16,18,20,22–27]. Thus, the highest reported AUC values should be interpreted cautiously, particularly in studies with limited external validation, as apparent gains in discrimination may reflect dataset-specific signals, outcome prevalence, or overfitting, rather than results that can be reliably reproduced in other clinical settings [22,23,25–29]. Cluster-based approaches nevertheless provide complementary value by identifying clinically recognizable high-risk phenotypes, such as patients with concomitant elevations in BUN, creatinine, and INR who experienced higher rates of severe adverse events and unplanned readmission [10]. Because these variables are routinely available before surgery, they may support preoperative risk stratification for transfusion

and renal complications. Similarly, models predicting LOS, same-day discharge, and readmission may support more individualized preoperative discharge planning. However, further prospective studies are needed to determine whether these promising predictive findings translate into model-guided interventions that reduce adverse events and improve quality of care and clinical outcomes.

From our perspective, the clinical value of preoperative laboratory parameters lies not only in their ability to predict complications but also in their potential to enhance preoperative optimization. Several predictors identified in this review relate to potentially modifiable clinical features. Haemoglobin and haematocrit can inform the assessment and management of preoperative anaemia and patient blood management, while albumin and prealbumin may help identify patients who require further nutritional assessment. Renal function and coagulation markers can likewise inform the evaluation and perioperative management of underlying physiological dysfunction [10,14,19,24,27]. However, existing models primarily use these variables as predictors of risk rather than linking their outputs to defined optimization strategies, and no common framework currently connects ML-based risk estimates with specific, evidence-based interventions. Avram et al. recently highlighted the need to move beyond reliance on BMI thresholds and standardized enhanced recovery protocols towards more individualized metabolic, inflammatory, and nutritional assessment [35]. The growing use of GLP-1 receptor agonists in patients with obesity and type 2 diabetes further illustrates the evolving complexity of perioperative metabolic management, although their role in arthroplasty optimization requires further evaluation. In this context, ML models based on routinely available laboratory data could provide a quantitative basis for identifying patients who may benefit from more detailed assessment and targeted optimization. Nevertheless, laboratory-based risk prediction alone does not establish which interventions are appropriate or whether they improve outcomes. Future research should evaluate the integration of laboratory markers with broader metabolic and inflammatory profiling and determine whether ML-guided, individualized optimization improves clinical outcomes after TJA.

Beyond the development of optimization strategies, successful clinical implementation remains an important challenge. The practical value of ML models depends on their interpretability, calibration, integration into clinical workflows, and accessibility to clinicians at the point of care. Few studies have evaluated integration into electronic health records, clinician use of model outputs, or cost-effectiveness. Future research should therefore focus on studies assessing whether ML-guided perioperative care can reduce complications, LOS, readmissions, and healthcare resource use, and whether these models can be effectively integrated into clinical workflows.

Despite these challenges, ML has considerable potential to enhance decision-making, reduce postoperative complications, and improve patient outcomes through more personalized approaches to care. Addressing data standardization, algorithm transparency, and clinical validation across healthcare settings will be essential for translating this potential into routine clinical practice in joint reconstructive surgery [34,36–38].

5. Limitations

While this review provides valuable insights into the predictive potential of routinely available laboratory parameters and ML models for postoperative complications following primary TKA and THA, several limitations should be acknowledged. The included studies were heterogeneous in terms of data sources, patient populations, outcome definitions, laboratory thresholds, and ML methodologies. This heterogeneity limited direct comparisons across studies and precluded formal quantitative synthesis. Because the included studies assessed different outcomes and populations, their AUC values should be compared

cautiously. Many models were developed using retrospective data, which may introduce selection bias and limit the assessment of unmeasured confounders. Some studies used the same patient cohort or examined the same database; hence, the cumulative sample size of our review cannot be reliably calculated. In addition, only a minority of studies underwent external or multicenter validation, raising concerns about model calibration, overfitting, and reproducibility to different clinical settings. Differences in perioperative protocols, transfusion thresholds, case mix, coding practices, and laboratory reference ranges may further affect model performance across different institutions. The lack of standardized reporting in some studies also limited the assessment of model performance and interpretability. Future research should prioritize prospective, multicenter validation using standardized outcome definitions and reporting frameworks to establish the reliability of these models and their use in clinical practice across diverse TJA populations.

6. Conclusions

The ML models using routinely available laboratory and clinical parameters show promising predicting potential regarding postoperative complications following primary TJA. Across the available evidence, laboratory measures such as HGB, HCT, PLT, and renal function markers were frequently used and associated with predictive performance across outcomes including transfusion, severe adverse events, POD, LOS, and readmission. These findings support the potential role of ML in risk stratification and more personalized perioperative care. However, methodological heterogeneity, retrospective study designs, and limited validation remain important limitations. Future prospective, multicenter studies are needed to establish the clinical utility of these models in joint reconstructive surgery.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/prosthesis8110116/s1, Supplementary File: A. Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) Checklist. B. Search queries for PubMed/MEDLINE, Scopus and EMBASE databases. C. Detailed Data-Charting Characteristics of Included Studies. D. Prediction model risk of bias assessment tool (PROBAST) assessing the risk of bias and applicability of included studies.

Author Contributions: Conceptualization and methodology: P.S.P., D.P.P. and P.D.M.; software, investigation, resources, visualization, and project administration: P.S.P. and D.P.P.; formal analysis and data curation: P.S.P., D.P.P., A.S., A.T., R.C.C., I.E.K. and V.K.; validation: P.S.P., D.P.P., A.F.M., P.J.P. and P.D.M.; writing—original draft preparation: P.S.P., D.P.P. and P.D.M.; writing—review and editing: P.S.P., D.P.P., A.F.M., P.J.P. and P.D.M.; supervision: P.D.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data was created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
AKI	Acute Kidney Injury
ALB	Albumin
ANN	Artificial Neural Network

ASA	American Society of Anesthesiologists Score
AUC	Area Under the Curve
BMI	Body Mass Index
BUN	Blood Urea Nitrogen
CBC	Complete Blood Count
CCI	Charlson Comorbidity Index
COPD	Chronic Obstructive Pulmonary Disease
Crea	Creatinine
CRP	C-reactive Protein
DVT	Deep Vein Thrombosis
eGFR	Estimated Glomerular Filtration Rate
ENP	Elastic Net Penalization
GBM	Gradient Boosting Machine
GFR	Glomerular Filtration Rate
HCT	Hematocrit
HGB	Hemoglobin
Hist-GBM	Histogram-Based Gradient Boosting
INR	International Normalized Ratio
KNN	K-Nearest Neighbors
LOS	Length of Stay
LR	Logistic Regression
MeSH	Medical Subject Heading
ML	Machine Learning
MPNN	Multilayer Perceptron Neural Network
N/R	Not Reported
PCC	Population, Concept, and Context
PDDP	Personalized Data-Driven Practice
PE	Pulmonary Embolism
PJI	Periprosthetic Joint Infection
PLT	Platelet Count
POD	Postoperative Delirium
PRISMA-ScR	Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews
PROBAST	Prediction Model Risk of Bias Assessment Tool
RF	Random Forest
ROB	Risk of Bias
SGB	Stochastic Gradient Boosting
SVM	Support Vector Machine
SSI	Surgical Site Infection
TJA	Total Joint Arthroplasty
TKA	Total Knee Arthroplasty
THA	Total Hip Arthroplasty
WBC	White Blood Cell Count
XGBoost	Extreme Gradient Boosting

References

1. Blom, A.W.; Donovan, R.L.; Beswick, A.D.; Whitehouse, M.R.; Kunutsor, S.K. Common elective orthopaedic procedures and their clinical effectiveness: Umbrella review of level 1 evidence. *BMJ* **2021**, *374*, n1511. [CrossRef] [PubMed]
2. Kurtz, S.; Ong, K.; Lau, E.; Mowat, F.; Halpern, M. Projections of primary and revision hip and knee arthroplasty in the United States from 2005 to 2030. *J. Bone Jt. Surg. Am.* **2007**, *89*, 780–785. [CrossRef] [PubMed]
3. Heo, S.M.; Harris, I.; Naylor, J.; Lewin, A.M. Complications to 6 months following total hip or knee arthroplasty: Observations from an Australian clinical outcomes registry. *BMC Musculoskelet. Disord.* **2020**, *21*, 602. [CrossRef] [PubMed]
4. Ferguson, K.; Winter, A.; Russo, L.; Khan, A.; Hair, M.; MacGregor, M.; Holt, G. Acute kidney injury following primary hip and knee arthroplasty surgery. *Ann. R. Coll. Surg. Engl.* **2017**, *99*, 307–312. [CrossRef] [PubMed]
5. Rodríguez-Merchán, E.C. The stiff total knee arthroplasty: Causes, treatment modalities and results. *EFORT Open Rev.* **2019**, *4*, 602–610. [CrossRef] [PubMed]
6. Filippone, E.J.; Yadav, A. Acute kidney injury after hip or knee replacement: Can we lower the risk? *Cleve Clin. J. Med.* **2019**, *86*, 263–276. [CrossRef] [PubMed]
7. Gómez-Barrena, E.; García-Rey, E. Complications in total joint arthroplasties. *J. Clin. Med.* **2019**, *8*, 1891. [CrossRef] [PubMed]

8. Liu, Y.; Ai, J.; Teng, X.; Huang, Z.; Wu, H.; Zhang, Z.; Wang, W.; Liu, C.; Zhang, H. Risk factor analysis and establishment of a nomogram model to predict blood loss during total knee arthroplasty. *BMC Musculoskelet. Disord.* **2024**, *25*, 7570. [CrossRef] [PubMed]
9. Wei, C.; Quan, T.; Wang, K.Y.; Gu, A.; Fassihi, S.C.; Kahlenberg, C.A.; Malahias, M.-A.; Liu, J.; Thakkar, S.; Della Valle, A.G.; et al. Artificial neural network prediction of same-day discharge following primary total knee arthroplasty based on preoperative and intraoperative variables. *Bone Jt. J.* **2021**, *103-B*, 1358–1366. [CrossRef] [PubMed]
10. Ranti, D.; Warburton, A.J.; Hanss, K.; Katz, D.; Poeran, J.; Moucha, C. K-means clustering to elucidate vulnerable subpopulations among Medicare patients undergoing total joint arthroplasty. *J. Arthroplast.* **2020**, *35*, 3488–3497. [CrossRef] [PubMed]
11. Groot, O.Q.; Ogink, P.T.; Lans, A.T.; Twining, P.K.; Kapoor, N.D.; DiGiovanni, W.; Bindels, B.J.; Bongers, M.E.; Oosterhoff, J.H.; Karhade, A.V.; et al. Machine learning prediction models in orthopedic surgery: A systematic review in transparent reporting. *J. Orthop. Res.* **2022**, *40*, 2627–2639. [CrossRef] [PubMed]
12. Tricco, A.C.; Lillie, E.; Zarin, W.; O'Brien, K.K.; Colquhoun, H.; Levac, D.; Moher, D.; Peters, M.D.J.; Horsley, T.; Weeks, L.; et al. PRISMA Extension for Scoping Reviews (PRISMAScR): Checklist and Explanation. *Ann. Intern. Med.* **2018**, *169*, 467–473. [CrossRef] [PubMed]
13. Wolff, R.F.; Moons, K.G.; Riley, R.D.; Whiting, P.F.; Westwood, M.; Collins, G.S.; Reitsma, J.B.; Kleijnen, J.; Mallett, S. PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. *Ann. Intern. Med.* **2019**, *170*, 51–58. [CrossRef] [PubMed]
14. Cavazos, D.R.; Sayeed, Z.; Court, T.; Chen, C.; Little, B.E.; Darwiche, H.F. Predicting factors for blood transfusion in primary total knee arthroplasty using a machine learning method. *J. Am. Acad. Orthop. Surg.* **2023**, *31*, e845–e858. [CrossRef] [PubMed]
15. Cohen-Levy, W.B.; Klemt, C.; Tirumala, V.; Burns, J.C.; Barghi, A.; Habibi, Y.; Kwon, Y.-M. Artificial neural networks for the prediction of transfusion rates in primary total hip arthroplasty. *Arch. Orthop. Trauma Surg.* **2023**, *143*, 1643–1650. [CrossRef] [PubMed]
16. Faure, N.; Knecht, S.; Tran, P.; Tamine, L.; Orban, J.-C.; Bronsard, N.; Gonzalez, J.-F.; Micicoi, G. Prediction of transfusion risk after total knee arthroplasty: Use of a machine learning algorithm. *Orthop. Traumatol. Surg. Res.* **2025**, *111*, 103985. [CrossRef] [PubMed]
17. Jørgensen, C.C.; Michelsen, C.; Petersen, T.; Kehlet, H. Preoperative hemoglobin thresholds for increased risk of "medical" complications in fast-track total hip and knee arthroplasty, a secondary analysis of a machine-learning algorithm. *Transfusion* **2024**, *64*, 438–442. [CrossRef] [PubMed]
18. Jo, C.; Ko, S.; Shin, W.C.; Han, H.-S.; Lee, M.C.; Ko, T.; Ro, D.H. Transfusion after total knee arthroplasty can be predicted using the machine learning algorithm. *Knee Surg. Sports Traumatol. Arthrosc.* **2020**, *28*, 1757–1764. [CrossRef] [PubMed]
19. Kolin, D.A.; Lyman, S.; Della Valle, A.G.; Ast, M.P.; Landy, D.C.; Chalmers, B.P. Predicting postoperative anemia and blood transfusion following total knee arthroplasty. *J. Arthroplast.* **2023**, *38*, 1262–1266.e2. [CrossRef] [PubMed]
20. Sayeed, Z.; Cavazos, D.; Court, T.; Chen, C.; Little, B.; Darwiche, H. Machine learning for prediction of blood transfusion rates in primary total knee arthroplasty. In Proceedings of the 2022 11th International Conference on Computing and Pattern Recognition, Beijing, China, 17–19 November 2022; pp. 590–594.
21. Wu, E.B.; Hung, K.C.; Juang, S.E.; Chin, J.C.; Lu, H.F.; Ko, J.Y. Are risk factors for postoperative significant hemorrhage following total knee arthroplasty potentially modifiable? A retrospective cohort study. *J. Pers. Med.* **2022**, *12*, 434. [CrossRef] [PubMed]
22. Huang, Z.; Martin, J.; Huang, Q.; Ma, J.; Pei, F.; Huang, C. Predicting postoperative transfusion in elective total hip and knee arthroplasty: Comparison of different machine learning models of a case-control study. *Int. J. Surg.* **2021**, *96*, 106183. [CrossRef] [PubMed]
23. Salimy, M.S.; Buddhiraju, A.; Chen, T.L.W.; Mittal, A.; Xiao, P.; Kwon, Y.-M. Machine learning to predict periprosthetic joint infections following primary total hip arthroplasty using a national database. *Arch. Orthop. Trauma Surg.* **2025**, *145*, 05757. [CrossRef] [PubMed]
24. Ko, S.; Jo, C.; Chang, C.B.; Lee, Y.S.; Moon, Y.; Youm, J.W.; Han, H.; Lee, M.C.; Lee, H.; Ro, D.H. A web-based machine-learning algorithm predicting postoperative acute kidney injury after total knee arthroplasty. *Knee Surg. Sports Traumatol. Arthrosc.* **2022**, *30*, 545–554. [CrossRef] [PubMed]
25. Tran, P.; Knecht, S.; Tamine, L.; Faure, N.; Orban, J.-C.; Bronsard, N.; Gonzalez, J.-F.; Micicoi, G. Risk prediction of kalaemia disturbance and acute kidney injury after total knee arthroplasty: Use of a machine learning algorithm. *Orthop. Traumatol. Surg. Res.* **2025**, *111*, 103958. [CrossRef] [PubMed]
26. Jung, J.W.; Hwang, S.; Ko, S.; Jo, C.; Park, H.Y.; Han, H.-S.; Lee, M.C.; Park, J.E.; Ro, D.H. A machine-learning model to predict postoperative delirium following knee arthroplasty using electronic health records. *BMC Psychiatry* **2022**, *22*, 4067. [CrossRef] [PubMed]
27. Wang, B.; Xin, Y.; Tang, X.; Wang, F.; Hua, S.; Yang, Y.; Xu, S.; Gong, H.; Dong, R.; Lin, Y.; et al. Potential value of serum prealbumin and serum albumin in the identification of postoperative delirium in patients undergoing knee/hip replacement: An observational study and internal validation study. *Front Neurol.* **2024**, *15*, 1375383. [CrossRef] [PubMed]

28. Zhang, Y.; Xie, L.J.; Wu, R.J.; Zhang, C.L.; Zhuang, Q.; Dai, W.T.; Zhou, M.X.; Li, X.H. Predicting the risk of postoperative delirium in elderly patients undergoing hip arthroplasty: Development and assessment of a novel nomogram. *J. Investig. Surg.* **2024**, *37*, 2381733. [CrossRef] [PubMed]
29. Chen, T.L.W.; Buddhiraju, A.; Seo, H.H.; Shimizu, M.R.; Bacevich, B.M.; Kwon, Y.-M. Can machine learning models predict prolonged length of hospital stay following primary total knee arthroplasty based on a national patient cohort data? *Arch. Orthop. Trauma Surg.* **2023**, *143*, 7185–7193. [CrossRef] [PubMed]
30. Sarker, I.H. Machine learning: Algorithms, real-world applications and research directions. *SN Comput. Sci.* **2021**, *2*, 160. [CrossRef] [PubMed]
31. Devana, S.K.; Shah, A.A.; Lee, C.; Roney, A.R.; van der Schaar, M.; SooHoo, N.F. A novel, potentially universal machine learning algorithm to predict complications in total knee arthroplasty. *Arthroplast. Today* **2021**, *10*, 135–143. [CrossRef] [PubMed]
32. Oettl, F.C.; Oeding, J.F.; Samuelsson, K. Explainable artificial intelligence in orthopedic surgery. *J. Exp. Orthop.* **2024**, *11*, 12103. [CrossRef] [PubMed]
33. Choi, J.H.; Choi, Y.; Lee, K.S.; Ahn, K.H.; Jang, W.Y. Explainable model using Shapley additive explanations approach on wound infection after wide soft tissue sarcoma resection: "Big Data" analysis based on Health Insurance Review and Assessment Service Hub. *Medicina* **2024**, *60*, 327. [CrossRef] [PubMed]
34. Kader, D.F.; Coppola, A.; Vijay, A.; Fontalis, A.; Haddad, F.S. The future of precision orthopaedics: Personalized data-driven practice. *Bone Jt. Open* **2025**, *6*, 836–840. [CrossRef] [PubMed]
35. Avram, G.M.; Indelli, P.F.; Violante, B.; Graichen, H.; Lustig, S.; Boettner, F.; von Eisenhart-Rothe, R.; Tandogan, R.; Hirschmann, M.T.; Kayaalp, M.E. GLP-1 receptor agonists in hip, knee and shoulder arthroplasty: Implications for obesity, osteoarthritis, metabolic optimisation and complications. *Knee Surg. Sports Traumatol. Arthrosc.* **2026**. *ahead of print*. [CrossRef] [PubMed]
36. Han, F.; Huang, X.; Wang, X.; Chen, Y.; Lu, C.; Li, S.; Lu, L.; Zhang, D. Artificial intelligence in orthopedic surgery: Current applications, challenges, and future directions. *MedComm* **2025**, *6*, e70260. [CrossRef] [PubMed]
37. Vaishya, R.; Sibal, A.; Kar, S.; Reddy, S. Integrating artificial intelligence into orthopedics: Opportunities, challenges, and future directions. *J. Hand Microsurg.* **2025**, *17*, 100257. [CrossRef] [PubMed]
38. Jeyaraman, M.; Ratna, H.V.K.; Jeyaraman, N.; Venkatesan, A.; Ramasubramanian, S.; Yadav, S. Leveraging artificial intelligence and machine learning in regenerative orthopedics: A paradigm shift in patient care. *Cureus* **2023**, *15*, e49756. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.