Research in Health & Medical Sciences *Review Article*

2025 October; 4(4)

Applications of machine learning in predicting complications and revision rates following total hip arthroplasty: A narrative literature review

Ali Ghanbarloo[*1]

[1] Department of Operating Room, School of Allied Medical Sciences, Iran University of Medical Sciences, Tehran, Iran

***Corresponding Author Email: ghanbarlooali85@gmail.com**

Received: 2026/3; Revised: 2026/7; Accepted: 2026/7

Abstract

Total hip arthroplasty (THA) is one of the most successful orthopaedic operations. However, the increased volume of these treatments is accompanied by an increased number of post-operative problems and revision rates. Traditional statistical models for risk stratification sometimes fail to account for complicated non-linear interactions among patient-specific variables. Thus, machine learning (ML), a subset of artificial intelligence, has been proposed as a viable method in the orthopedic literature to optimize predictive accuracy. The purpose of this narrative review was to critically appraise the current research on the application of machine learning approaches to predict short-term problems, long-term implant survival, and revision rates after THA. A literature search of relevant publications on the development and clinical validation of ML algorithms was performed in PubMed and Scopus. A review of the literature available shows that ML models, and specifically ensemble methods such as Random Forest and Gradient Boosting, have outperformed traditional logistic regression in predicting surgical site infections, dislocations, and readmissions, and have produced robust survival analyses for long-term implant failure. However, currently, these benefits are constrained by the reliance on retrospective data, algorithmic bias and the 'black-box' character of sophisticated models that limit the broad clinical integration of ML. ML has the transformative potential to tailor preoperative counseling, but additional prospective, multicenter studies using comprehensible artificial intelligence (XAI) are required to ensure safe clinical adoption.

Keywords: Machine learning, Total hip arthroplasty, Complication prediction, Revision arthroplasty, Artificial intelligence.

Introduction

The frequency of primary and revision total hip arthroplasty (THA) is predicted to increase exponentially in the next decades due to aging populations, broadening of surgical reasons, and rising prevalence of osteoarthritis (1, 2). THA is generally considered an affordable intervention; however, post-operative complications, including surgical site infections, periprosthetic joint infections (PJI), venous thromboembolism, instability, and aseptic loosening, impose a severe physiological burden on patients and a significant financial burden on healthcare systems (3-6).

Thus, preoperative risk stratification is of utmost importance. Surgeons have traditionally employed linear scoring methods such as the American Society of Anesthesiologists (ASA) physical status classification, Charlson Comorbidity Index (CCI), and classic multivariable logistic regression models to predict adverse outcomes (7, 8). However, classic statistical methods face frequent challenges when dealing with high-dimensional data, missing variables, and the non-linear interaction between a patient's comorbidities and intra-operative factors.

In this regard, machine learning (ML) has emerged as a disruptive tool in orthopedic surgery (9, 10). ML employs algorithms that can learn iteratively from enormous volumes of data without explicit programming and is able to identify subtle clinical trends that are not detectable with standard analysis (11-14). Although the literature on this topic is rapidly increasing, the techniques and clinical applications of these models reported in different studies are significantly varied. This narrative review critically synthesizes the available literature on the application of machine learning techniques in the prediction of short-term complications, long-term outcomes, and revision rates following primary and revision THA, with emphasis on the comparison of the algorithmic performance against traditional models.

Literature Search Strategy

A narrative review of the current literature on machine learning applications for the prediction of THA outcomes. A systematic search of the literature was conducted in the PubMed and Scopus databases from inception to June 2026. Keywords and Medical Subject Headings (MeSH) were combinations of "Machine Learning", "Artificial Intelligence", "Deep Learning", "Total Hip Arthroplasty", "Complications", and "Revision". Articles were selected based on relevance to predictive modelling of postoperative adverse events, implant survivorship, and comparison of algorithm performance. Studies were included if they reported the development, internal/external validation, or clinical application of ML models in the context of THA. Studies lacking specific predictive measures (e.g., Area Under the Curve [AUC]), studies of non-orthopaedic use, expert opinion without original data, and studies in a language other than English were excluded.

Review of the Literature

To understand the current state of machine learning in THA, it is necessary to review the basic concepts, the most frequently referenced algorithms in the literature, and their specific clinical applications for predicting short- and long-term postoperative events.

1. Machine learning concepts and algorithms reported in THA literature

A paradigm change has been observed in the analysis of patients' data, moving from classical biostatistics to artificial intelligence, as demonstrated in a review of recent studies. The traditional models measure the isolated influence of one risk factor. ML algorithms, however, are effective at processing thousands of interconnected data pieces simultaneously.

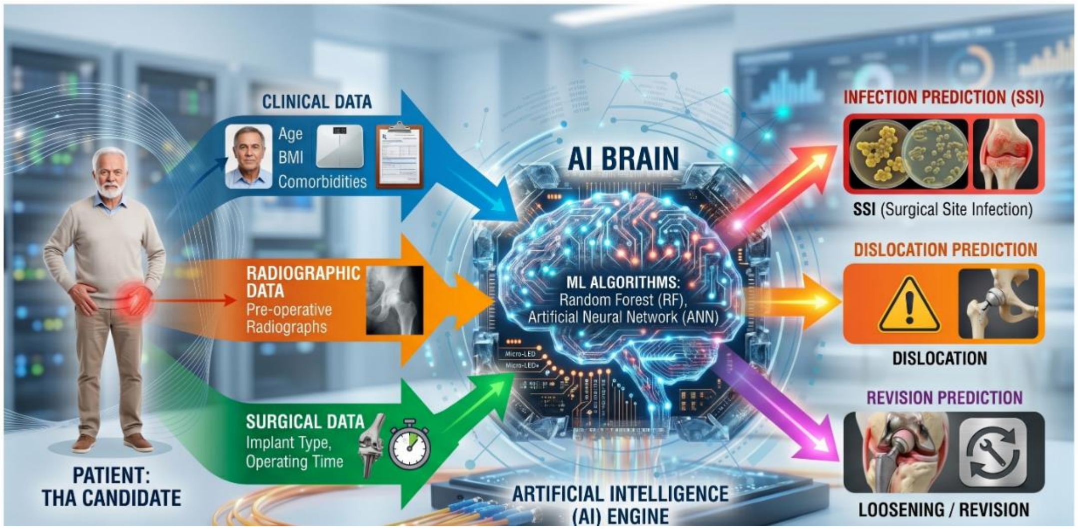

Figure 1. Schematic representation of a machine learning workflow in total hip arthroplasty. The model includes multiple preoperative, radiographic, and intraoperative variables to estimate individual risks of short- and long-term complications.

In the literature, a few well-known ML architectures are often employed in arthroplasty research. Ensemble learning approaches such as Random Forest (RF) and Extreme Gradient Boosting (XGBoost) are extremely robust to overfitting since they are able to generate several decision trees and merge their predictions (15, 16). Studies dealing with complicated inputs, e.g., preoperative radiographs for templating, commonly use Artificial Neural Networks (ANNs) and Deep Learning (DL) models (17). There is also growing research on the use of Natural Language Processing (NLP) to directly extract undocumented comorbidities from unstructured clinical notes in Electronic Health Records (EHR) (18-21).

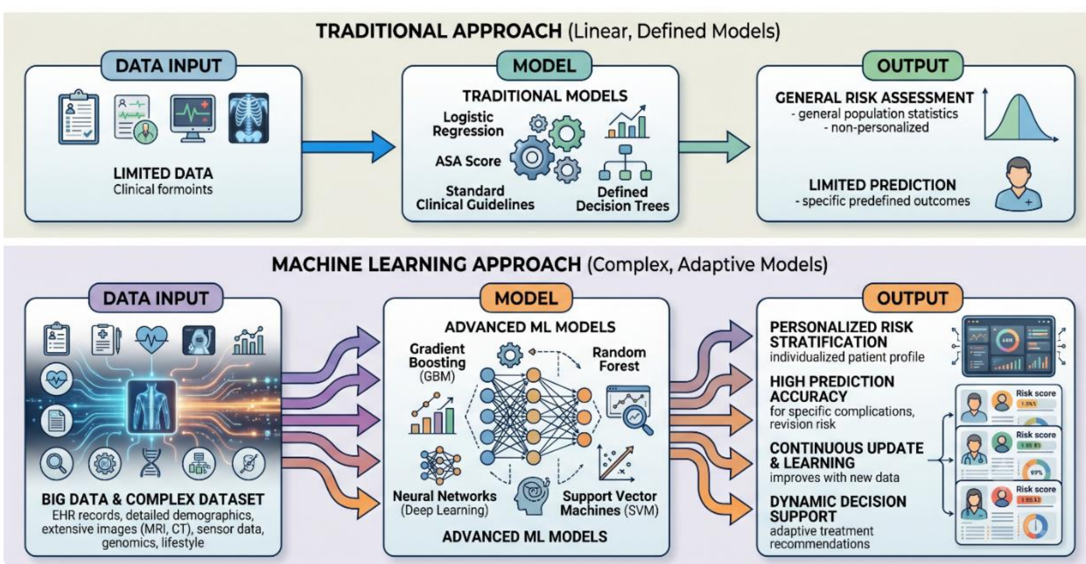

Figure 2. Machine learning versus traditional statistical modelling for risk stratification in orthopaedic surgery: from generalised scoring to personalised prediction.

2. Literature on predicting short-term complications

The literature on predictive modeling primarily focuses on short-term adverse events (30–90 days postoperative) because they are the main drivers of extended hospital length of stay (LOS) and early readmission.

2.1. Surgical site infections and venous thromboembolism

ML algorithms that included preoperative serological markers, BMI, diabetes control metrics, and operative time have shown significantly higher sensitivity in predicting SSI than traditional logistic regression in several comparative studies (18). In a review of registry-based studies, XGBoost models are useful in weighting the non-linear risk of different levels of preoperative anemia and malnutrition. Additionally, the literature describes successful application of ML models to predict venous thromboembolism (VTE) and pulmonary embolism (PE). These models analyze national databases on a large scale to identify hypercoagulable phenotypes, as surgeons can adapt pharmacological prophylaxis protocols based on individual risk profiles (22-24).

2.2. Dislocation and readmission rates

The literature abounds with reports of postoperative dislocation as the most common cause of early revision surgery. Recent studies presenting ML frameworks have included preoperative radiographic parameters (e.g., spinopelvic mobility) with patient demographics to successfully stratify patients into high- and low-risk groups for instability (25, 26). In addition, prediction of 90-day readmission is a major theme of the recent literature on healthcare quality. Some authors have described that deep learning algorithms based on full EHR data, including socioeconomic determinants of health, have a high discriminative ability to identify patients who will need intensive outpatient support (27).

3. Literature on predicting long-term outcomes and revision rates

Mechanical stability and biological response are important factors for long-term success after THA. The literature stresses the importance of advanced capabilities of survival analysis to predict late-term failure.

Implant survivorship and aseptic loosening

Many authors criticize traditional long-term predictive models (e.g. Cox proportional hazards regression) because they assume that the hazard ratio is constant over time. This assumption is often violated in orthopaedic reality (28). To address this gap in the literature, recent studies have employed machine learning-based survival models like Random Survival Forests and DeepSurv. These studies show that ML models can update the probability of aseptic loosening over time by incorporating time-varying covariates, such as progressive osteolysis seen on serial radiographs and specific implant tribology (29).

Periprosthetic fractures and overall revision risk

Periprosthetic femoral fractures (PPFF) are becoming more prevalent, and recent publications have been dedicated to the identification of high-risk patients. ML algorithms using preoperative cortical bone thickness, femoral morphology, and stem design have been studied with promising results in predicting intraoperative and early postoperative fractures (30). Collectively, these reports would suggest that computational tools allow the surgeon to make highly individualized decisions regarding fixation strategies (e.g., cemented vs. uncemented) in frail populations.

Discussion

The literature reviewed is strongly in favor of the claim that machine learning algorithms can predict the risk of complications and revision rates after THA more accurately than traditional statistical methods. ML models demonstrate consistently higher discrimination (higher AUC) and better calibration in several studies based on high-dimensional registry and institutional data [18]. This increased predictive power allows for highly personalized preoperative optimization and tailored postoperative pathways.

However, critical review of the literature suggests significant barriers to routine clinical use. The most frequently mentioned limitation by several authors is the use of retrospective data. Models trained on data from institutions often suffer from algorithmic bias and poor generalizability to other demographics (31). Therefore, the literature strongly supports rigorous external validation prior to clinical deployment.

A further frequent barrier to surgical trust is the “black box” phenomenon. Surgeons are understandably hesitant to alter operative plans based on a black-box algorithm (32). To this end, recent literature points to the necessary shift towards Explainable AI (XAI) techniques such as SHAP (SHapley Additive exPlanations) values, which visually quantify the exact contribution of each patient variable to the final risk prediction (33).

The literature is clear that future studies should focus on prospective, randomized controlled trials comparing ML-augmented clinical decision-making with current standards of care.

Conclusions

The current literature indicates machine learning could revolutionize risk stratification in total hip arthroplasty with complication and revision predictions that are often more accurate than traditional statistical methods. These algorithms assist in personalized patient care, surgical planning, and resource optimization. However, retrospective designs, lack of external validation, and the inherent opaqueness of complex computational models largely limit the current literature. The future of ML in orthopedics hinges on the development of models that are highly transparent and interpretable and the rigorous prospective trials to ensure safe, equitable, and effective integration into daily surgical practice.

Guardian/patient's consent

Not Applicable.

Ethics approval

Compliant with Helsinki.

Consent

Not applicable.

Code availability

Not applicable.

Ethical statement

This study was conducted in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Authors’ contributions

SAH was responsible for the study conception, literature search, data analysis, drafting of the manuscript, critical revision for important intellectual content, and final approval of the version to be published.

Funding statement

No funding has been received for this study.

Conflict of interest statement

On behalf of all authors, the corresponding author states that there is no conflict of interest.

Acknowledgement

None.

Data availability

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

References

1. Rasmussen MB, El-Galaly A, Daugberg L, Nielsen PT, Jakobsen T. Projection of primary and revision hip arthroplasty surgery in Denmark

from 2020 to 2050. Acta Orthop. 2022;93:849-53. https://doi.org/10.2340/17453674.2022.5255
2. Amiri F, Haghi S, Khachian A, Hosseini SA, Azadi N. The effect of professional ethics training on the awareness and attitude of operating room nurses. Perioperative Care and Operating Room Management. 2025;41:100564. https://doi.org/10.1016/j.pcorm.2025.100564
3. Egerci OF, Yapar A, Dogruoz F, Selcuk H, Kose O. Preventive strategies to reduce the rate of periprosthetic infections in total joint arthroplasty; a comprehensive review. Arch Orthop Trauma Surg. 2024;144(12):5131-46. https://doi.org/10.1007/s00402-024-05301-w
4. Hosseini SA, Amiri F, Hanani S, Rashidi N, Azadi N. Manual environmental disinfection with hydrogen peroxide and silver ions versus sodium hypochlorite in operating rooms: a multi-centre randomized trial on residual antimicrobial efficacy. Journal of Hospital Infection. 2026;172:150-6. https://doi.org/10.1016/j.jhin.2026.02.007
5. Farash Khayalo H, Bahlgerdi M, Habibi A, Yektay S. Air pollution as a silent driver of osteoporosis: A narrative review of vitamin D impairment and skeletal vulnerability in female populations. Journal of Research in Health and Medical Sciences. 2024;3(4):8-29.
6. Abbasnia A, Bahlgerdi M, Khayalo HF, Bazrafshan E, Moeini Z. Assessment of Carbon Emissions Associated with Artificial Intelligence: A Narrative Review of Data Center Environmental Impacts and Green AI Strategies. Journal of Research in Health and Medical Sciences. 2024;3(1):29-42.
7. Yurtlu DA, Aksun M, Ayvat P, Karahan N, Koroglu L, Aran G. Comparison of Risk Scoring Systems to Predict the Outcome in ASA-PS V Patients Undergoing Surgery: A Retrospective Cohort Study. Medicine (Baltimore). 2016;95(13):e3238. https://doi.org/10.1097/MD.0000000000003238
8. Hosseini SA, Pourdel B, Rajabi E, Alizadeh A. Surgical Time-Outs: Ritual or Real Safety Practice? Perioperative Care and Operating Room Management. 2025:100573. https://doi.org/10.1016/j.pcorm.2025.100573
9. Zhang M, Zheng Y, Maidaiti X, Liang B, Wei Y, Sun F. Integrating Machine Learning into Statistical Methods in Disease Risk Prediction Modeling: A Systematic Review. Health Data Sci. 2024;4:0165. https://doi.org/10.34133/hds.0165
10. Hosseini SA, Rajabi E, Pourdel B. The Empty Chair: OR Nurse Retention as a Patient Safety Imperative. Perioperative Care and Operating Room Management. 2026:100645. https://doi.org/10.1016/j.pcorm.2026.100645
11. Fahim YA, Hasani IW, Kabba S, Ragab WM. Artificial intelligence in healthcare and medicine: clinical applications, therapeutic advances, and future perspectives. Eur J Med Res. 2025;30(1):848. https://doi.org/10.1186/s40001-025-03196-w
12. Pourdel B, Hosseini SA, Rajabi E. Clinical and structural outcomes of hybrid operating rooms in adult cardiac surgery: a systematic review. Perioperative Care and Operating Room Management. 2026;42:100607. https://doi.org/10.1016/j.pcorm.2025.100607
13. Khayalo HF, Yeganeh M, Esrafili A. Critical review on photocatalytic removal of Ciprofloxacin from water and wastewater: Implications for environmental management and wastewater treatment. Critical review. 2024;3(4).
14. Hosseini SA, Pourdel B, Rajabi E, Farash Khayalo H. Efficacy of photocatalytic processes for hospital environmental disinfection: A systematic review. Journal of Research in Health and Medical Sciences. 2025;4(2):11-31.
15. Khan AA, Chaudhari O, Chandra R. A review of ensemble learning and data augmentation models for class imbalanced problems: Combination, implementation and evaluation. Expert Systems with Applications. 2024;244:122778. https://doi.org/10.1016/j.eswa.2023.122778
16. Pourdel B, Hosseini SA, Rajabi E, Alizadeh A. Alarm fatigue in the operating room: A synthesis of the dual threat from device overload and ambient noise. Perioperative Care and Operating Room Management. 2026;42:100624. https://doi.org/10.1016/j.pcorm.2026.100624
17. Wongsak S, Janyawongchot T, Sri-Utenchai N, Owasirikul D, Jaovisidha S, Woratanarat P, et al. Development of a Deep Learning Model for Hip Arthroplasty Templating Using Anteroposterior Hip Radiograph. Journal of Clinical Medicine. 2025;14(24):8689. https://doi.org/10.3390/jcm14248689
18. Sheikhalishahi S, Miotto R, Dudley JT, Lavelli A, Rinaldi F, Osmani V. Natural Language Processing of Clinical Notes on Chronic Diseases: Systematic Review. JMIR Med Inform. 2019;7(2):e12239. https://doi.org/10.2196/12239
19. Khayalo HF, Bahlgerdi M, Habibi A. Efficiency of photocatalytic processes in the removal of the antibiotic ciprofloxacin from aqueous environments: A systematic review.
20. Farash Khayalo H, Bahlgerdi M, Esrafili A. Emerging AI-Driven Methodologies for Rapid Environmental Health Assessment in Crisis

Conditions: A Narrative Review. Journal of Research in Health and Medical Sciences. 2024;3(3):38-48.
21. Khayalo HF, Bahlgerdi M, Esrafili A. Novel artificial intelligence-based approaches for rapid environmental health assessment in crisis situations: A narrative review. health.12:14.
22. Heo KY, Rajan PV, Khawaja S, Barber LA, Yoon ST. Machine learning approach to predict venous thromboembolism among patients undergoing multi-level spinal posterior instrumented fusion. J Spine Surg. 2024;10(2):214-23. https://doi.org/10.21037/jss-24-8
23. Farash Khayalo H, Rajabi E, Yeganeh M. The role of Polycyclic Aromatic Hydrocarbons (PAHs) in urban air as neuroendocrine disruptors in the development of menstrual mood disorders: A narrative review. Journal of Research in Health and Medical Sciences. 2024;3(3):17-27.
24. Farash Khayalo H, Aabbasnia A, Yeganeh M, Rajabi E, Hosseini SA, Pourdel B, et al. Sustainable management of operating room wastes from the perspective of environmental health engineering: A critical review emphasizing strategies for the prevention of healthcare-associated infections. Journal of Research in Health and Medical Sciences. 2025;4(1):14-25.
25. Çetin O, Öztürk A, Avci Ö, Kaya A, Çevik N, Akalin Y. Spinopelvic assessment in preoperative planning of total hip arthroplasty: a comparative cohort analysis. Acta Orthop Traumatol Turc. 2025;59(5):286-91. https://doi.org/10.5152/j.aott.2025.24065
26. Khayalo HF, Bahlgerdi M, Yeganeh M, Esrafili A, Abbasnia A. Systematic review of the efficiency of photocatalytic processes for doxycycline antibiotic removal from aqueous media. 2026.
27. Sharda M, Sharma S, Raikar S, Verhagen N, Wagle J, Mathur R, et al. The Role of Machine Learning in Predicting Hospital Readmissions Among General Internal Medicine Patients: A Systematic Review. Cureus. 2025;17(5):e84761. https://doi.org/10.7759/cureus.84761
28. Austin PC, Giardiello D. The Impact of Violation of the Proportional Hazards Assumption on the Calibration of the Cox Proportional Hazards Model. Stat Med. 2025;44(13-14):e70161. https://doi.org/10.1002/sim.70161
29. Kar İ, Kocaman G, İbrahimov F, Enön S, Coşgun E, Elhan AH. Comparison of deep learning-based recurrence-free survival with random survival forest and Cox proportional hazard models in Stage-I NSCLC patients. Cancer Med. 2023;12(18):19272-8. https://doi.org/10.1002/cam4.6479
30. Foster A, Beeharry MW. Periprosthetic Proximal Femoral Fractures: A Comprehensive Review of Epidemiology, Risk Factors, Classification and Management. Cureus. 2026;18(1):e100589. https://doi.org/10.7759/cureus.100589
31. Suleman MU, Mursaleen M, Khalil U, Saboor A, Bilal M, Khan SA, et al. Assessing the generalizability of artificial intelligence in radiology: a systematic review of performance across different clinical settings. Ann Med Surg (Lond). 2025;87(12):8803-11. https://doi.org/10.1097/MS9.0000000000004166
32. Ahmed MI, Spooner B, Isherwood J, Lane M, Orrock E, Dennison A. A Systematic Review of the Barriers to the Implementation of Artificial Intelligence in Healthcare. Cureus. 2023;15(10):e46454. https://doi.org/10.7759/cureus.46454
33. Sheu RK, Pardeshi MS. A Survey on Medical Explainable AI (XAI): Recent Progress, Explainability Approach, Human Interaction and Scoring System. Sensors (Basel). 2022;22(20). https://doi.org/10.3390/s22208068