Journal of Advanced Research 79 (2026) 633–659

Contents lists available at ScienceDirect

Journal of Advanced Research

journal homepage: www.elsevier.com/locate/jare

The Application of artificial intelligence in periprosthetic joint infection

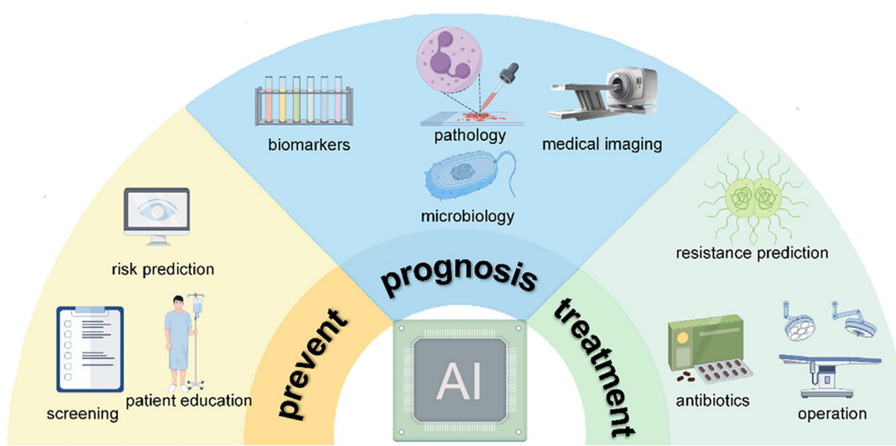
Pengcheng Li[a,1], Yan Wang[a,1], Runkai Zhao[a,1], Lin Hao[a,1], Wei Chai[a], Chen Jiying[a], Zeyu Feng[a], Quanbo Ji[a,b,c,*], Guoqiang Zhang[a,*]

[a] Department of Orthopaedics, General Hospital of Chinese People's Liberation Army, Beijing 100853, China
[b] Beijing National Research Center for Information Science and Technology (BNRist), Beijing, China
[c] Department of Automation, Tsinghua University, Beijing, China

HIGHLIGHTS

- For the first time, we systematically and comprehensively elaborated on the relevant progress of AI technology in the prevention, diagnosis, and treatment of PJI.
- We focused on reviewing the applications and prospects of AI technologies such as machine learning and robotic technology in the risk prediction of PJI, the diagnosis combined with various existing technologies, and the field of treatment.
- This paper summarized and expounds the current challenges and future development directions of AI in PJI field.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 6 January 2025
Revised 6 March 2025
Accepted 20 March 2025
Available online 28 March 2025

Keywords:
Periprosthetic joint infection
Artificial intelligence
Joint
Arthroplasty

ABSTRACT

Periprosthetic joint infection (PJI) represents one of the most devastating complications following total joint arthroplasty, often necessitating additional surgeries and antimicrobial therapy, and potentially leading to disability. This significantly increases the burden on both patients and the healthcare system. Given the considerable suffering caused by PJI, its prevention and treatment have long been focal points of concern. However, challenges remain in accurately assessing individual risk, preventing the infection, improving diagnostic methods, and enhancing treatment outcomes. The development and application of artificial intelligence (AI) technologies have introduced new, more efficient possibilities for the management of many diseases. In this article, we review the applications of AI in the prevention, diagnosis, and treatment of PJI, and explore how AI methodologies might achieve individualized risk prediction, improve diagnostic algorithms through biomarkers and pathology, and enhance the efficacy of antimicrobial and surgical treatments. We hope that through multimodal AI applications, intelligent management of PJI can be realized in the future.

* Corresponding authors.
E-mail addresses: quanbo301@163.com (Q. Ji), gqzhang301orth@163.com (G. Zhang).

[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.jare.2025.03.039
2090-1232/

Contents

Introduction

Total hip and knee arthroplasty, hailed as one of the most successful surgical innovations of the 20th century, has become the gold standard for treating end-stage osteoarthritis, rheumatoid arthritis, and traumatic joint injuries [1]. By replacing damaged joints with prosthetic implants, it effectively alleviates pain, restores joint function, and improves quality of life [2]. However, postoperative complications such as infections, aseptic loosening, and periprosthetic fractures persistently compromise long-term outcomes [3–5]. Among these, periprosthetic joint infection (PJI) is termed a "catastrophic complication" due to its high disability rates and therapeutic complexity [6,7]. The incidence of PJI ranges from 1 % to 3 % depending on patient demographics and surgical techniques, but escalates to 5 %-7% in high-risk populations (e.g., diabetics, obese individuals, or immunocompromised patients) [8,9]. PJI often necessitates prolonged antimicrobial therapy, repeated debridement or revision surgeries, and may result in permanent disability. The 5-year survival rates for hip and knee PJI (67.2 % and 71.7 %, respectively) parallel those of prostate cancer, imposing immense burdens on patients and healthcare systems while severely diminishing quality of life [10,11]. Given these devastating consequences, early prevention, accurate diagnosis, and effective treatment of PJI are critically urgent.

However, identifying high-risk PJI patients in advance remains challenging. The pathogenesis of PJI involves complex interactions among the host, pathogen, and medical interventions. Risk factors encompass patient-related elements such as advanced age, diabetes mellitus, and osteoporosis, as well as treatment-related factors including prolonged surgical duration, substantial intraoperative blood loss, and the propensity for biofilm formation on prostheses [12,13]. Existing risk calculators lack generalizability and are underutilized clinically [14,15]. Diagnostically, PJI has no universally accepted gold standard. Widely used criteria (e.g., ICM 2018, MSIS 2011) rely on composite scoring systems integrating microbiological cultures, serum/synovial biomarkers, and imaging findings. However, rigid scoring frameworks hinder clinical practicality, and studies reveal suboptimal diagnostic accuracy and sensitivity [16]. Therapeutically, biofilm formation and rising antibiotic resistance delay culture results and diminish empirical antibiotic efficacy [17]. Surgical interventions—whether debridement, antibiotics, and implant retention (DAIR), one-stage, or two-stage revision—face challenges such as incomplete debridement, severe bone loss, and recurrent infections complicating soft-tissue reconstruction [18,19]. These limitations underscore the pressing need for innovative technologies to enhance PJI prevention and treatment.

Artificial intelligence (AI), first applied in medicine through Stanford University's MYCIN system for infectious disease diagnosis in the 1970s, has increasingly integrated with healthcare [20]. Modern AI technologies—including machine learning (ML), deep learning (DL), natural language processing (NLP), and computer vision (CV)—offer transformative potential (Table 1). ML algorithms, such as Random Forest (RF), utilize ensemble decision trees for classification/regression tasks, enabling robust risk stratification and factor identification [21,33]. DL, a subset of ML employing multi-layered neural networks, excels in analyzing high-dimensional data; convolutional neural networks (CNNs) autonomously extract features from medical images (e.g., MRI, CT, histopathology), aiding diagnostic precision [24,34,35]. NLP mod-

Table 1
The application of AI tools in medicine.

Classification	Algorithm	Definition	Advantages	Limitations	Clinical Scenarios
ML	RF [21]	An ensemble learning algorithm that builds multiple decision trees and aggregates results through voting or averaging	High accuracy (especially for structured data); Good interpretability; Handles high-dimensional data	Computationally expensive; Poor performance on high-dimensional sparse data (e.g., text)	Disease risk prediction; Patient stratification; Drug efficacy evaluation
	SVM [22]	A classifier based on margin maximization, using kernel functions for nonlinear problems	Strong performance on small datasets; Robust to noise; Memory efficient	Slow training on large datasets; Requires strategies for multi-class tasks	Tumor classification; Survival prediction
	XGBoost [23]	An optimized gradient boosting framework that iteratively trains weak learners to minimize loss	Suitableful for structured data; Handles missing values automatically; Feature importance analysis	Complex hyperparameter tuning; Sensitive to class imbalance; High cost	Personalized medication; Postoperative complication prediction; Multi-omics integration
DL	CNN [24]	A deep learning model using convolutional kernels to extract spatial features, ideal for image processing	Suitableful for recognition; Automatic feature extraction; Translation invariance	Requires large labeled datasets; High computational cost; Poor interpretability	Medical imaging diagnosis; Pathological analysis; Endoscopic detection
	RNN [25]	Processes sequential data through cyclic connections to retain temporal dependencies	Excels in text/speech processing; Supports variable-length inputs	Gradient vanishing/explosion issues; Difficult to parallelize; Limited long-sequence memory	EHR temporal analysis; Pharmacokinetics modeling
	GAN [26]	A framework with a generator (creates data) and discriminator (evaluates authenticity) trained adversarially	Generates high-quality synthetic data;; Enhances model robustness	Unstable training (mode collapse); Subjective evaluation metrics; Limited control over output diversity	Medical data augmentation; Image denoising; Virtual patient generation
NLP	BERT [27]	A bidirectional Transformer-based language model pre-trained via masked language modeling	Deep semantic understanding; Strong cross-task transferability	Extremely resource-intensive; Inefficient for long texts	EHR structuring; Literature mining; Clinical decision support
	GPT series [28]	Autoregressive language models using unidirectional Transformers to generate coherent text	High-quality text generation; Multitask versatility	Uncontrollable outputs; Slow inference	Patient education; Automated note generation; Virtual consultation
	NER [29]	Identifies named entities (e.g., persons, locations) in text via sequence labeling	Core tool for information extraction; Boosts data processing efficiency	Requires domain-specific labeled data; Struggles with entity ambiguity	Clinical entity annotation; Biomedical literature analysis
CV	U-Net [30]	Symmetric encoder-decoder CNN for image segmentation, using skip connections to preserve spatial details	Benchmark for medical image segmentation; Effective with small datasets	Computationally heavy; Limited scalability to 3D data	Medical image segmentation; 3D organ reconstruction
	YOLO [31]	Single-stage object detector framing detection as a regression problem for real-time performance	Real-time performance; Balances speed and accuracy	Poor small object detection; High miss rate for dense objects	Surgical navigation; Real-time endoscopic detection; Ultrasound analysis
	ViT [32]	Applies Transformer architecture to image classification via patch-based self-attention	Global attention mechanism; Parallel computation efficiency	Prone to overfitting on small datasets; Relies heavily on positional embeddings; High computational cost	Multimodal data fusion; Rare disease diagnosis

ML, Machine Learning; DL, Deep Learning; NLP, Natural Language Processing; CV, Computer Vision; RF, Random Forest; SVM, Support Vector Machine; CNN, Convolutional Neural Networks; RNN, Recurrent Neural Network; GAN, Generative Adversarial Network; NER, Named Entity Recognition Model; ViT, Vision Transformer; EHR, Electronic Health Records.

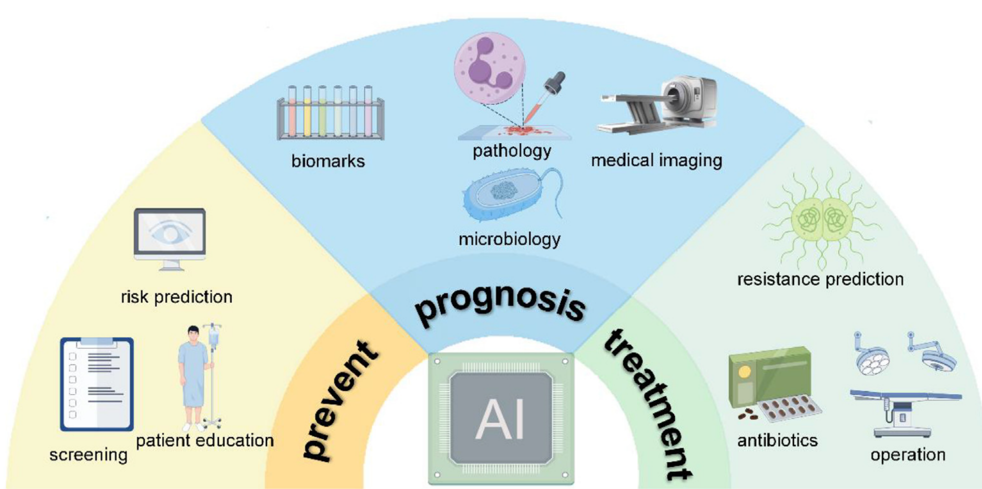


Fig. 1. Application of AI in PJI. This scheme mainly describes the application progress of AI in the prevention, diagnosis and treatment of PJI. Specifically, AI technology is currently used in the field of prevention for the prediction of PJI risk in various clinical scenarios, patient education and screening. In terms of diagnosis, current research mainly combines AI with biomarkers, pathology, medical imaging and microbiology to improve the diagnostic efficacy of PJI. In addition, many studies have shown that machine learning, robotics and other AI technologies have been applied to PJI-related treatments such as drug resistance prediction, antibiotic development and surgery, improving treatment outcomes.

Table 2
Summary of risk factors for PJI.

Country	Application	Patient Number	Risk factors	Time
UK [40]	Recurrence after 2-stage revision for PJI	210	Age > 70; BMI > 30; Presence of sinus; Staphylococcus aureus; CN-PJI	2003–2018
Finnish [37]	PJI after primary THA	33,337	ASA class > 2; Bleeding > 500 mL; Operation time > 120 min; BMI > 31; Anesthesia (epidural);Anesthesia (general); Antithrombotic prophylaxis not used; Bilateral operation; Male sex; avascular necrosis; Femoral head size (reference 32 mm)28/36	2014–2018
Canada [8]	PJI after primary THA	100,674	Male sex; Type-2 diabetes mellitus; Discharged to convalescent care;	2002–2016
USA [41]	Recurrence after 2-stage revision for PJI	198	Both elevated ESR and CRP (ESR 75.2 ± 29.0; CRP 42.6 ± 26.8);Only elevated ESR (65.7 ± 26.1); Only elevated CRP (35.1 ± 16.8).	–
USA [42]	PJI after primary TJA	290	Injection drug	2000–2020
USA [43]	Recurrence of PJI after irrigation and debridement	199	Hematogenous infection; Prior revision Surgery; History of malignancy; diabetes; COPD; Polymicrobial infection; Intraoperative purulence; Elevated systolic blood pressure; Tachycardia; Higher CRP (18.0 vs. 13)	2005–2016
UK [44]	Recurrence of PJI after DAIR	60	ASA > 2, BMI > 35, CRP > 200 mg/L; Staphylococcus aureus growth	2010–2018

PJI, periprosthetic joint infection; BMI, body mass index; ASA, American society of anesthesiologists; THA, total hip arthroplasty; ESR, erythrocyte sedimentation rate; CRP, c-reactive protein; TJA, total joint arthroplasty; COPD, chronic obstructive pulmonary disease; DAIR, debridement, antibiotics, and implant retention.

els (e.g., GPT series) interpret clinical queries about diseases, medications, and treatments, while CV algorithms like YOLO enable real-time object detection for surgical navigation. Currently, several ML models have been developed for predicting PJI risk in various scenarios, diagnosing PJI based on different test results, robotic systems for PJI revision surgeries, and natural language processing (NLP) tools that can answer PJI-related questions. These applications demonstrate the vast potential of AI technologies in the prevention and treatment of PJI (Fig 1).

Search strategy

We conducted searches in the PubMed and Web of Science databases using the following terms to identify literature on AI applications in PJI diagnosis and treatment:"periprosthetic joint infection", "Prosthesis-Related Infections", "implant-associated infection", "Replacement", "Arthroplasty", "Artificial Intelligence", "Machine Learning", "Deep Learning", "Neural Networks", "Algorithm", "diagnos", "predict model", "image analysis", "treatment", "AI", "ML", "DL", "PJI", "periprosthetic knee infections", "prosthetic hip infection". The inclusion timeframe was restricted to studies published between January 2010 and September 2024. Initial screening was performed by reviewing titles and abstracts to assess relevance to the review topic. Full-text articles of selected studies were then evaluated for eligibility. Studies were excluded based on the following criteria: unavailability of full text, non-clinical research, non-English publications, review articles, and *meta*-analyses.

Application in PJI prevention

Disease prevention encompasses reducing exposure to risk factors, seeking timely medical care and treatment, and preventing disease progression. This requires not only the efforts of healthcare professionals but also increased patient awareness and the adoption of healthier lifestyle habits to reduce the risk of disease onset. For different individuals, early identification of high-risk patients and modification of alterable risk factors can effectively reduce disease incidence. However, patients often have limited access to medical knowledge, and even the best physicians are not always effective at public health education, leading to a general lack of disease-related expertise among the population. Additionally, the relative scarcity of medical resources per capita means that many diseases are not promptly addressed, and patients often seek medical care only in the middle or late stages of illness, significantly increasing their burden. AI technology offers a promising solution to these challenges.

PJI risk prediction

Given the substantial burden that PJI places on patients, early identification of high-risk patients following total joint arthroplasty (TJA) is crucial for optimizing treatment strategies and implementing targeted prevention measures [36]. Over the past decade, efforts have been made to identify risk factors for PJI in various scenarios, including primary arthroplasty, aseptic revision surgery, and DAIR (Debridement, Antibiotics, and Implant Retention) procedures [37–39]. These factors include general conditions, medical history, comorbidities, and biomarkers (Table 2). Although previous research has identified modifiable and non-modifiable risk factors such as obesity (BMI > 30 kg/m^2), anesthesia type (epidural or general anesthesia), American Society of Anesthesiologists (ASA) class, history of intravenous drug use, comorbid diabetes, and age (>70), few studies have successfully integrated these factors to quantitatively assess the risk of PJI and the relative importance of each factor in its occurrence.

Previously, several PJI risk calculators were developed using statistical methods such as multivariable analysis [45–48]. However, these calculators often have limited accuracy, and the modeling process is cumbersome, requiring significant costs. In contrast, ML models can efficiently and automatically analyze the complex and nonlinear relationships between numerous patient and surgical variables. This allows them to analyze large, complex datasets with high precision within seconds, predicting the risk of disease onset [49]. As more data are introduced into the algorithm, it continuously learns and improves, becoming a more accurate and predictive model. In recent years, researchers have combined AI technology with PJI risk prediction and risk factor analysis, developing various algorithms for predicting PJI risk in different scenarios. These algorithms aid in the stratified management of patients and the formulation of treatment plans before surgery, and have also changed the understanding of PJI risk factors. We have summarized and compared the performance of these algorithms and the corresponding risk calculators (Table 3 and 4).

The risk of developing PJI after primary arthroplasty has long been a focus of attention, with multiple studies attempting to develop risk calculators to predict this risk. For instance, a risk calculator developed by Tan TL and colleagues using a large cohort of

Table 3
Comparison of risk prediction algorithm and risk calculator for PJI.

Study	Method	Country	Patient number	Variate number	Statistical performance
Predicting PJI after primary TJA[14]	Risk calculator	USA	27,717	42	AUC = 0.83
					AUC = 0.84
Predicting infection after primary TKA[50]	ML	USA	10,021	28	AUC = 0.84
Predicting recurrence after revision of PJI[46]	Risk calculator	USA	1081	56	AUC = 0.75
Predicting recurrence after rTKA of PJI [51]	ML	USA	618	44	AUC = 0.84
Predicting PJI after DAIR[52,53]	Risk calculator	Spain	222	55	AUC = 0.84
					AUC = 0.64
Predicting PJI after I&D[54,55]	ML	USA	1174	52	AUC = 0.74 AUC = 0.69

PJI, periprosthetic joint infection; TJA, total joint arthroplasty; AUC, area under the curve; TKA, total knee arthroplasty; ML, machine learning; rTKA, revision total knee arthroplasty; DAIR, debridement, antibiotics, and implant retention; I&D, irrigation and debridement.
AUC: Measures the model's ability to distinguish classes, where 1.0 indicates perfect discrimination and 0.5 equals random guessing.

Table 4
Summary of risk prediction algorithms for PJI.

First author	Application	Algorithm	AUC	Intercept	Slope	Brier score	Top 5 Predict factors	Limitations
Yeo [50]	Preoperative prediction of the risk of PJI following primary TKA	ANN SGB SVM RF ENP	0.84 0.79 0.78 0.80 0.80	0.09 0.18 −0.21 −0.17 0.18	1.06 1.27 1.15 0.90 1.10	0.054 0.055 0.056 0.055 0.054	CCI, Obesity, Smoking, Diabetes, Medicare insurance	Retrospective study; No external validation; Small cohort size
Klemt [61]	Prediction of PJI Following Aseptic rTKA	ANN SVM k-Nearest	0.78 0.76 0.74	0.10 −0.21 −0.25	1.09 0.20 1.29	0.055 0.058 0.058	>1 open procedure prior to revision TKA, Drug abuse, obesity, Extremity status grade 3, Diabetes	Retrospective study; No external validation; Relatively short follow-up time
Klemt [51]	Predict recurrent infection following rTKA for PJI	ANN RF ENP	0.84 0.81 0.83	0.05 −0.07 0.10	1.05 0.83 1.14	0.053 0.056 0.054	Irrigation and debridement with or without modular component exchange during revision surgery; >4 prior open surgeries, metastatic disease, Drug abuse	Retrospective study; No external validation; Relatively short follow-up time
Shohat [54]	Predict failure of irrigation and debridement for PJI	RF	0.74	–	–	–	Higher CRP levels, Positive blood cultures, Indication for index surgery other than osteoarthritis, Not exchanging the Modular component, Use of immunosuppressive medication	Retrospective study; No external validation

PJI, periprosthetic joint infection; ANN, artificial neural networks; SGB, stochastic gradient boosting; SVM, support vector machines; RF, random forests; ENP, elastic net penalization; CCI, Charlson Comorbidity Index; rTKA, revision total knee arthroplasty; CRP, c-reactive protein.
AUC: Measures the model's ability to distinguish classes, where 1.0 indicates perfect discrimination and 0.5 equals random guessing.
Intercept: Reflects overall bias in predicted probabilities (0 = unbiased, >0 = underprediction, <0 = overprediction).
Slope: Indicates calibration confidence (1.0 = perfect, <1.0 = overconfident predictions, >1.0 = underconfident predictions).
Brier Score: Quantifies probabilistic prediction error (0 = perfect accuracy, 0.25 = random performance for binary tasks).

20,000 individuals achieved high predictive performance [14]. However, ML can achieve similar predictive capabilities without the need for such large cohorts and complex development processes [50]. Similarly, in studies predicting the risk of PJI recurrence after revision surgery, ML models have demonstrated superior predictive capabilities compared to traditional risk calculators [46,51]. In the prediction of PJI recurrence risk after DAIR, limited research indicates that although the predictive performance of ML is slightly weaker than that of risk calculators, it exhibits more stable predictive ability during external validation [52,54].

Overall, the performance of currently developed risk prediction models generally surpasses that of risk calculators in similar scenarios. Additionally, these models are more personalized, capable of making predictions based on the specific circumstances of each patient, and can rank the importance of risk factors, thereby helping guide patients in reducing their PJI risk. However, there are still many challenges to address. For example, the application scenarios covered by these models do not yet encompass all clinical contexts, and similar studies remain scarce. Most models have not undergone external validation, limiting their generalizability, and the predictive performance of the models is heavily influenced by the selection of included indicators, leading to potential biases. Despite these difficulties, we look forward to the emergence of an AI application capable of providing individualized risk predictions for patients across all clinical contexts.

Primary PJI

Previous studies have attempted to use AI technologies to predict the risk of major complications, including PJI, following TJA, but with limited success [56–58]. For example, Bülow and colleagues employed a logistic least absolute shrinkage and selection operator (LASSO) regression algorithm to predict the risk of PJI within 90 days after primary total hip arthroplasty (THA) [57]. The model achieved an AUC of 0.68 in an internal Swedish cohort of 88,830 THA patients and 0.66 in an external Danish cohort of 18,854 THA patients (Fig. 2a,b), demonstrating limited accuracy in predicting the occurrence of PJI within 90 days post-THA.

Similarly, Kunze and colleagues developed a random gradient boosting model to predict the risk of complications, including PJI, following primary THA [59]. Although the model performed well (AUC of 0.88) (Fig. 2c-e), it was only able to predict the overall risk of various complications such as pulmonary embolism, deep vein thrombosis, PJI, and aseptic loosening, rather than the specific probability of PJI occurring. In response, Yeo and colleagues compared the performance of five ML algorithms—artificial neural networks (ANN), stochastic gradient boosting (SGB), support vector machines (SVM), RF, and elastic net penalization (ENP)—in predict-

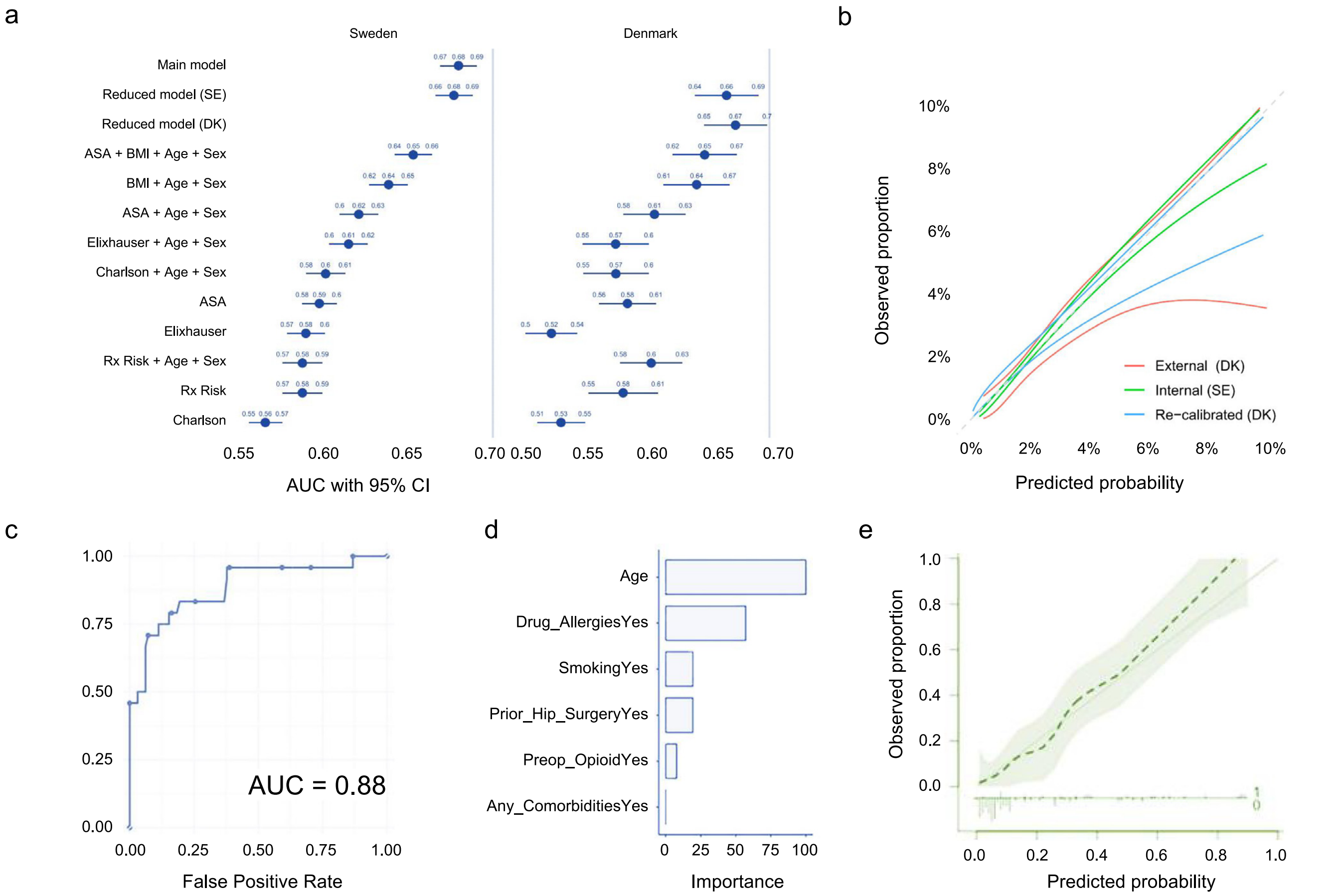


Fig. 2. Application of AI in primary PJI risk prediction. a. AUC value comparison plot of Bülow et al. 's model [57]. b. Calibration plot of Bülow et al. 's model [57]. Copyright 2017 Informa PLC. c. Receiver operative curve of Kunze et al. 's model [59]. d. Global variable importance plot of Kunze et al. 's model. e. Calibration plot of Kunze et al. 's model. [59].Copyright 2022 Springer Nature.

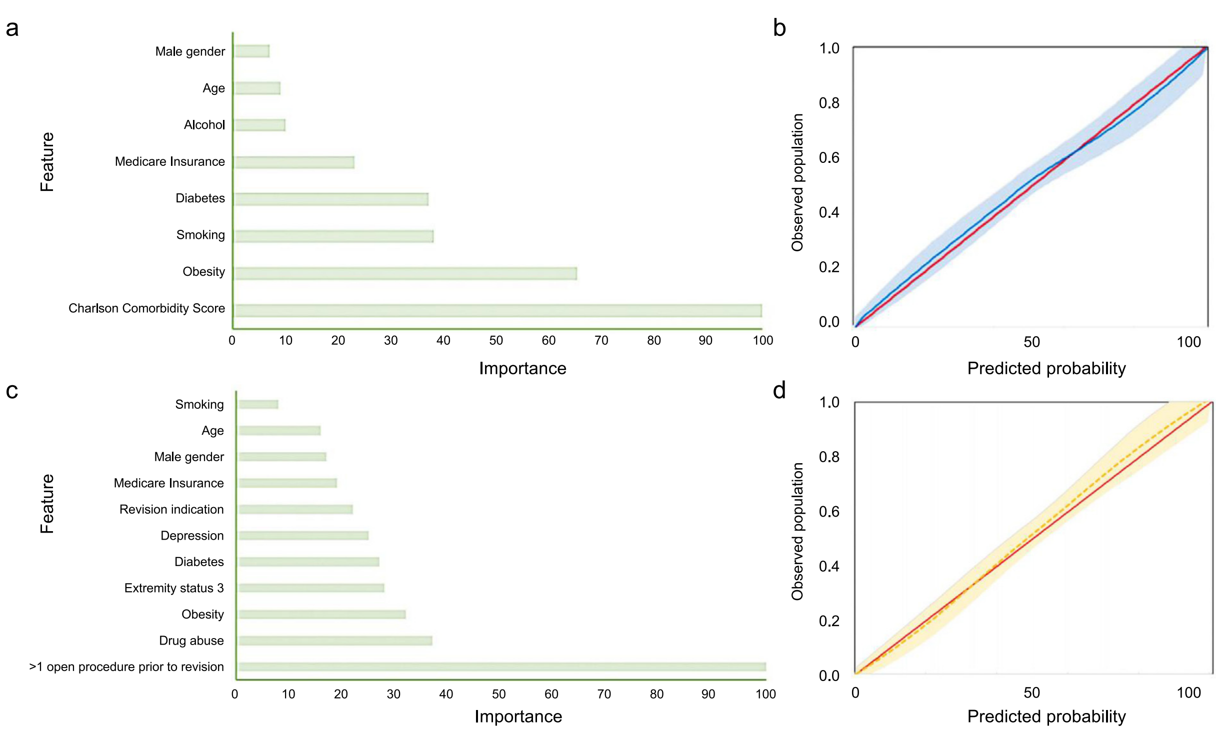


Fig. 3. Application of AI in primary PJI risk prediction. a. Global variable importance plot of Yeo et al. 's model [50]. b. Calibration plot of Yeo et al. 's model [50]. Copyright 2022 Thieme. c. Global variable importance plot of Klemt et al. 's model [61]. d. Calibration plot of Klemt et al. 's model [61]. Copyright 2023 Thieme.

ing the risk of superficial surgical site infection (SSI) and PJI after primary total knee arthroplasty (TKA) [50]. All of these algorithms achieved an accuracy rate exceeding 94 %, with ANN demonstrating the best performance. The study identified Charlson Comorbidity Index (CCI), BMI (>30 kg/m^2), and smoking as the strongest predictors of infection (Fig. 3a,b). The application of this model could help identify high-risk patients for infection before TKA, allowing for the optimization of modifiable risk factors to minimize the risk of infection, but the study included only 181 patients with PJI, and the cohort was small.

Previous research has indicated that PJI is the primary cause of aseptic revision TKA failure, with early aseptic revision TKA significantly increasing the risk of PJI within two years post-revision [60]. This can lead to repeated infections, multiple surgeries, and ultimately, failure of the revision surgery. In this context, Klemt and colleagues developed three supervised ML algorithms—ANN, SVM, and k-nearest neighbors (KNN)—to predict the risk of PJI after aseptic revision TKA [61]. They optimized the model parameters through a coarse-grained grid search algorithm and repeated random subsampling validation, with ANN showing the best predictive performance, achieving an AUC of 0.78 in the test set. The risk factors for PJI identified by this model were largely consistent with those identified by previous retrospective studies using risk calculators [62], namely, pre-revision open surgery, drug abuse (intravenous), and obesity as the main risk factors for PJI after aseptic revision (Fig. 3c,d). However, this study also indicated that a high BMI (>35 kg/m^2) is more significant in the prediction than previously thought [63].

PJI recurrence

Each recurrence of PJI significantly increases the burden on patients' health [64], and the failure rate of surgical treatments for PJI remains high. A recent multicenter retrospective cohort study indicated that the reinfection rate within two years after revision TJA to eradicate PJI can be as high as 24.8 % [65]. To predict the risk of recurrent infection after PJI revision surgery, Klemt and colleagues compared the predictive capabilities of ANN, SGB, and elastic net penalized logistic regression [51]. Their results showed that ANN had the best predictive performance, achieving an AUC of 0.84 in the test set (Supplementary Fig. 1a). Meanwhile, the Intercept of the model is 0.05, which indicates that the prediction of the model has little deviation from the actual probability and has high accuracy. Using the recursive feature elimination method within the RF, they identified several dozen risk factors, such as previous irrigation and debridement procedures, and a history of more than four open surgeries, along with the weighted distribution of these factors for both overall and individual risk predictions (Supplementary Fig. 1b). Compared to previously developed risk calculators for recurrent PJI [46], the ML model not only demonstrated higher accuracy (AUC 0.84 vs. 0.75) but also identified MRSA as a significant risk factor for recurrent PJI.

The classic DAIR approach for treating acute PJI has a highly variable success rate, ranging from 30 % to 90 % [66–69]. When DAIR fails, it not only increases the burden on patients and the healthcare system but can also negatively impact the outcomes of subsequent joint revision surgeries [70,71]. Therefore, it is crucial to identify high-risk patients for DAIR failure in advance. Shohat and colleagues analyzed 1,174 patients who underwent irrigation and debridement (I&D) due to acute PJI, using RF to predict the likelihood of I&D failure (PJI recurrence) [54]. The algorithm demonstrated strong classification ability, with cross-validation tests confirming its accuracy and stability. In the vast majority of cases in this retrospective study, the algorithm's predictions closely matched the observed outcomes (Supplementary Fig. 1c). The algorithm identified the most important predictive variables, including high CRP levels, positive blood cultures, failure to exchange modular components, use of immunosuppressive drugs, late-stage acute (hematogenous) infection, polymicrobial infection, and advanced age, which were consistent with recognized risk factors [72] (Supplementary Fig. 1d). The predicted and observed probabilities of failure were similar, with the model achieving an AUC of 0.74, indicating high accuracy.

In summary, this algorithm enables surgeons to identify high-risk patients for I&D failure in advance, allowing timely adjustments to treatment strategies. By examining the relative importance of variables highlighted by the RF model, surgeons can preoperatively improve modifiable risk factors such as CRP levels and wound leakage, potentially reducing the failure rate of DAIR and the risk and burden of multiple surgeries for patients.

PJI patient education

Large language models (LLMs) like ChatGPT are rapidly emerging as powerful tools in the medical field. These models [73] have demonstrated the ability to pass medical exams with high scores (Supplementary Fig. 2a-c), and can provide advice based on patient information and relevant medical knowledge, much like a surgeon would [74]. Recent reports indicate that consent forms generated by chatbots exhibit good readability, completeness, and accuracy compared to those prepared by surgeons [75]. These advancements offer unique opportunities for creating personalized, patient-oriented interventions. In the future, patients may be able to learn about surgical risks and outcomes through chatbots before their medical appointments, reducing their psychological stress and alleviating the documentation burden on physicians.

Traditionally, patients have relied on brief outpatient visits to understand disease progression and treatment plans, including surgery and medication. Now, chatbots powered by precise LLMs offer patients a more detailed opportunity to learn about their condition. These AI systems can provide personalized risk profiles, answer questions about preoperative optimization and postoperative recovery, and guide patients through the surgical process, including follow-up consultations.

In terms of expertise, preliminary studies have shown that ChatGPT's free-text responses about diseases are relatively reliable and useful [76–78]. For example, Alex and colleagues selected 27 questions related to PJI prevention, diagnosis, pathogens, and treatment from the ICM 2018 and tested ChatGPT 3.5′s free-text responses [79]. The responses were evaluated by three orthopedic surgeons using a five-point Likert scale (Supplementary Fig. 2d). The average Likert score was 3.87 ± 0.66, indicating that ChatGPT's free-text responses on hip and knee PJI were generally considered complete, non-misleading, occasionally containing factual errors, and suitable for use by both patients and orthopedic surgeons. This suggests that people can obtain low-cost yet reliable and effective PJI-related knowledge through GPT, significantly aiding in the understanding of PJI and encouraging timely medical consultation for potential patients. In the future, LLM-based chatbots are expected to provide patients with individualized and professional medical guidance and education.

PJI screening

PJI-related diagnostic data elements are scattered across various sections of electronic health records (EHRs). When dealing with a suspected PJI case, experienced clinicians must extract information from these complex records and meticulously follow diagnostic protocols. This demands a deep familiarity with the relevant scoring systems and processes. Fu et al. developed an NLP algorithm that can automatically extract PJI-related diagnostic data elements (such as surgical records, serological markers,

Table 5
Comparison of performance of PJI diagnostic criteria.

First author	Method	Patient number	Diagnostic criteria	Sen	Spe	Acc	PPV	NPV	AUC
Sigmund [16]	Comparison of the preoperative component of the diagnostic criteria	206, PJI:101; Aseptic: 105	IDSA 2013	49.5 %	100 %	75.2 %	100 %	67.3 %	0.748
			ICM 2018	45.5 %	100 %	73.3 %	100 %	65.6 %	0.728
			EBJIS 2021	69.3 %	100 %	85.0 %	100 %	77.2 %	0.847
Chen [90]	Comparison of complete diagnostic criteria	55 , PJI:19; Aseptic:36	MSIS 2011	52.63 %	100 %	83.64 %			
			ICM 2018	78.57 %	100 %	81.82 %			
			EBJIS 2021	78.57 %	100 %	70.91 %			
Parvizi [83]	Comparison of complete diagnostic criteria	1504, PJI:684; Aseptic: 820	MSIS 2011	79.3 %	99.5 %		79.3 %	20.7 %	
			ICM 2013	86.9 %	99.5 %		86.9 %	13.1 %	
			ICM 2018	97.7 %	99.5 %		95.5 %	2.3 %	

PJI, periprosthetic joint infection; IDSA, Infectious Diseases Society of America; ICM, International Consensus Meeting; EBJIS, European Bone and Joint Infection Society; MSIS, Musculoskeletal Infection Society.
Sen: Proportion of true positives correctly identified (closer to 1.0 reduces missed diagnoses).
Spe: Proportion of true negatives correctly identified (closer to 1.0 reduces false alarms).
Acc: Overall correctness of predictions (misleading for imbalanced data; 1.0 = all correct).
PPV: Probability that a positive prediction is truly positive (high value increases diagnostic confidence).
NPV: Probability that a negative prediction is truly negative (high value strengthens rule-out reliability).
AUC: Measures the model's ability to distinguish classes, where 1.0 indicates perfect discrimination and 0.5 equals random guessing.

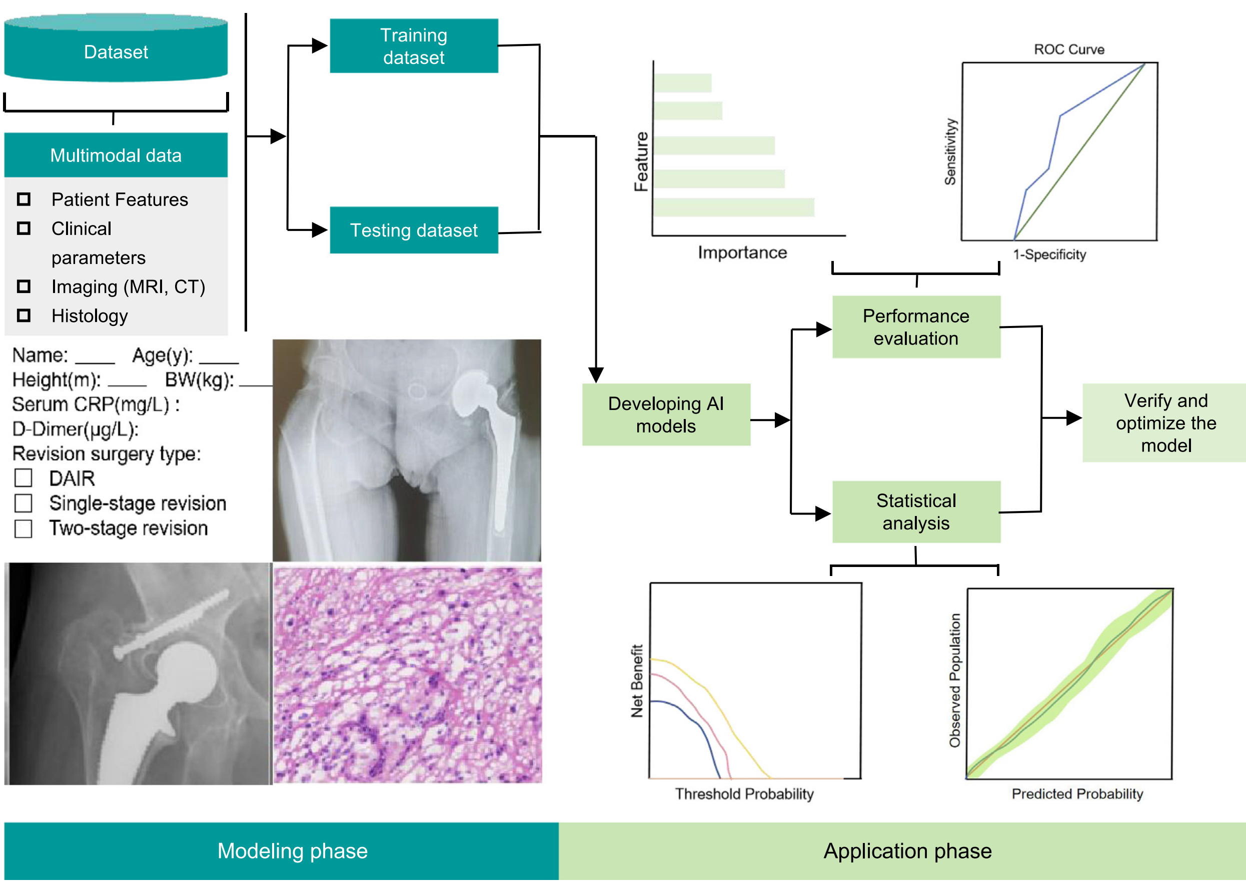


Fig. 4. AI builds a diagnostic algorithm diagram. During the algorithm development phase, researchers collected patients' general conditions, clinical parameters, imaging and histological results, organized them into a dataset, and divided the dataset into a training set and a testing set, which were used to train and develop the model and verify model performance, respectively. In the application phase after the initial establishment of the model, the model performance was evaluated through calibration plot, decision curve analysis, ROC curve analysis, and other methods to optimize the model.

histopathological findings, and microbiological culture results) from EHRs [80]. The algorithm identifies PJI patients based on the MSIS criteria, with high sensitivity (88.7 %) and specificity (99.1 %). This approach significantly reduces the clinical workload by automating the extraction of relevant information from complex EHRs, enabling timely identification and treatment of PJI patients, thus preventing delays in care. Additionally, this algorithm can be integrated with existing PJI risk prediction models, offering the potential for automated PJI risk prediction based on EHRs, and early screening of high-risk populations.

PJI diagnosis

PJI is a common cause of failure in TKA and THA, accounting for 20.4 % of all knee revisions and 12.8 % of all hip revisions. The economic burden of treating this complication is substantial, with

Table 6
Summary of new diagnostic biomarkers for PJI.

Study	Sample type	Sample part	Biomarker	Ref standard
Cobra[94]	Synovial fluid	Knee	cf-DNA	ICM
Fröschen.et al[95]		Hip, Knee	Cytokine(IL-1b, IL-2, IL-4, IL-5, IL-6, IL-8, IL-10, IL-12, IL-17, GM-CSF, TNF-α, and INF-γ)	MSIS
Fuchs et al.[96]		Knee	Soluble Pecam-1	EBJIS
Ikeda et al.[97]		Hip, Knee	MPO	ICM
Grassi et al[98]		Knee	Calprotectin ELISA immunoassay	ICM
Sallai et al[99]		Hip, Knee	EVs	MSIS
Karbysheya et al.[100]		Hip, Knee	D-lactate	MSIS
Chisari et al[101]	Blood	Hip, Knee	Zonulin, sCD14	ICM
Echeverria et al[102]		Hip, Knee	cf-DNA	MSIS
Dartus et al[103]		Hip, Knee	Bacterial antigen	MSIS
Omar et al[104]	Urine	Hip, Knee	Urinary peptide	MSIS

PJI, periprosthetic joint infection; cf-DNA, cell free-DNA; ICM, International Consensus Meeting; MSIS, Musculoskeletal Infection Society; EBJIS, European Bone and Joint Infection Society; ELISA, Enzyme Linked Immunosorbent Assay.

costs per patient ranging from $27,874 to $77,851, along with prolonged hospital stays, and increased morbidity and mortality rates [81,82]. Currently, there are four main diagnostic criteria for PJI: the Musculoskeletal Infection Society (MSIS), the Infectious Diseases Society of America (IDSA), the International Consensus Meeting (ICM), and the European Bone and Joint Infection Society (EBJIS). However, none of these criteria possess sufficient sensitivity and specificity, and there is still no standardized diagnostic algorithm for PJI [83–86].

Researchers are actively exploring new diagnostic algorithms for PJI, with AI playing a significant role. Initially, AI was used to assist in the development of diagnostic algorithms. For instance, during the creation of the ICM 2018 diagnostic criteria, RF was employed to assess the relative importance and weight of the variables included, aiding in the establishment of a scoring system [83]. Similarly, Shohat and colleagues developed a new evidence-based stepwise PJI diagnostic algorithm through a multicenter retrospective study [87]. This algorithm incorporated hematological markers like ESR, D-dimer, and CRP, synovial fluid (SF) markers such as WBC and AD, and intraoperative indicators like purulence and tissue culture. AI techniques like RF were used to assist in determining the weight of variables, enabling the algorithm to interpret the interactions between individual or combined diagnostic test results and their impact on the probability of infection. The algorithm achieved high sensitivity (96.9 %) and specificity (99.5 %).

Although the diagnostic performance has improved with the continual updates of diagnostic algorithms (Table 5), the factors and standards which constitute the scoring systems are predefined and remain fixed for all patients. However, the risk factors leading to PJI vary according to individual patient circumstances [88]. The fixed processes and scoring standards significantly reduce the sensitivity and specificity of diagnostic algorithms for individual cases. Furthermore, the complex processes of standardized diagnostic algorithms increase the likelihood of human error during test ordering, data interpretation, and score calculation [89]. AI can help overcome these issues by offering more individualized and precise diagnoses through learning and training on databases (Fig. 4). Consequently, some researchers have utilized various ML methods and PJI-related indicators to directly develop PJI diagnostic algorithms/models.

Biomarker-based diagnosis

Biomarkers are crucial objective indicators in the diagnostic decision-making process for PJI, aiding in early identification of PJI risk and enhancing diagnostic accuracy [91]. In the diagnosis of PJI, clinically common biomarkers are derived from both serum and SF [92]. Serum biomarkers such as ESR, D-dimer, and CRP, and SF biomarkers like WBC count, PMN%, CRP, leukocyte esterase, and α-defensin, have already been incorporated into widely used diagnostic standards like the MSIS 2011, IDSA 2013, ICM 2018, and EBJIS 2021. Beyond these, researchers are actively exploring additional biomarkers to enable easier and more accurate PJI diagnosis (Table 6). For instance, urinary peptides, calprotectin in SF, and cf-DNA in serum have been shown to assist in PJI diagnosis [93]. AI can aid in discovering new biomarkers and integrating existing ones into new diagnostic models, thereby improving the sensitivity and specificity of PJI diagnosis and helping to determine the relative importance of various biomarkers, leading to more efficient biomarker applications (Table 7).

As a physiological fluid present in joints, SF is a critical component in PJI-related biomarker sources. Both preoperative SF culture and SF biomarkers are significant indicators for PJI [105]. Paranjape et al. developed an ML algorithm that combines unsupervised and supervised techniques, using methods like principal component analysis (PCA), unsupervised clustering (K-Means), and decision tree models [106]. This algorithm, which relies solely on SF test results, assists in PJI diagnosis by selecting the most informative features to distinguish PJI from non-PJI cases (Supplementary Fig. 3a), demonstrating good overall predictive performance with an AUC of 0.993. Additionally, the uncertainty rate of this algorithm in the validation cohort was only 1.3 % (Supplementary Fig. 3c), lower than the 7.4 % uncertainty rate of the ICM 2018 standard. SHAP analysis for feature importance revealed that among SF biomarkers—WBC count, PMN%, CRP, α-defensin (AD), neutrophil elastase, and microbial antigen (MID)—AD was the most critical feature (Supplementary Fig. 3b), consistent with existing literature, including *meta*-analyses [107,108]. This validates the model's effectiveness and the rationality of its decisions.

In addition to SF, blood biomarkers like CRP and ESR are also key factors in PJI diagnosis, and blood samples can be obtained in a relatively non-invasive manner. Based on 18 preoperative blood biochemical tests, Kawakami et al. developed an ML model using RF and UMAP techniques to differentiate between PJI and non-PJI cases (Supplementary Fig. 3d), showing superior predictive performance compared to a multivariate logistic regression control group [109]. SHAP values were then used to evaluate the importance of preoperative blood biomarkers like CRP, total protein, urea, LDH, and platelets in differentiating PJI, with CRP emerging as the most significant blood factor (Supplementary Fig. 3e). Additionally, the Wilcoxon rank-sum test revealed no significant difference in Hb levels between PJI and non-PJI patients, consistent with the literature [110]. However, the model indicated that CRP levels above 0.3 mg/L and Hb levels below 11.0 g/dL were important predictive factors in the RF model, suggesting a partial dependence between

Table 7
Biomarker-based PJI diagnostic methods.

First author	Application	Algorithm	Statistical performance	Top 5 important feature	limitations
Paranjape [106]	Diagnosis of PJI by preoperative synovial fluid test results	A combination algorithm of principal component analysis, K-means clustering and decision tree	AUC: 0.993 Sen: 99.1 % Spc: 97.1 % PPV: 91.9 % NPV: 99.9	AD; MID Sum; %PMN; WBC; SF-CRP	Retrospective study; No external validation
Kawakami [109]	Diagnosis of PJI by preoperative blood biochemical features	A combination of RF and UMAP	AUC: 0.778	CRP; TP; BUN; Hb; ALT	Retrospective study; Susceptibility artifacts; Small cohort size
Omar [104]	Diagnosis of PJI by Urinary peptide markers	SVM	AUC: 0.960 Sen: 95.0 % Spc: 90.0 %	Fragments of structural extracellular matrix proteins; Blood plasma proteins; proteins released from erythrocytes; Proteins derived from cells; Renal immunomodulatory proteins	Retrospective study; No external validation; Small cohort size
Kuo [114]	Individualised PJI diagnosis through multifactoriality	2-level stacked generalization architecture: Base classifiers: RF, XGBoost, LR, NB Meta-classifier: SVM;	AUC: 0.988 Sen: 98.1 % ACC: 96.4 % F1 score: 0.970	Histology; Two positive cultures; Synovial WBC; Sinus tract; Serum CRP	Retrospective study; No external validation; Small cohort size
		ICM 2018	AUC: 0.958 Sen: 97.4 % ACC: 93.5 % F1 score: 0.946		
Chen [90]	Diagnosis of PJI by several preoperative and intraoperative clinical features	2-level ML model: Base classifiers: EN, LSVM, KSVM, ET, LGBM and MLP Meta-classifier: ELWV	AUC: 0.968 Sen: 90.6 % Spc: 98.3 % ACC: 94.5 % F1 score: 0.925	Wound exudation; Increase of local skin temperature; Sinus tract or prosthesis; Exposure; Sinus tract or prosthesis Exposure; Sinus tract or prosthesis exposure	Retrospective study; Small cohort size; Black box; Lack of several expensive examination;
		MSIS 2011	Sen: 47.6 % Spc: 87.9 % ACC: 75.1 % F1 score: 0.604		
		EBJIS 201	Sen: 69.7 % Spc: 87.6 % ACC: 83.1 % F1 score: 0.795		
		ICM 2018	Sen: 76.1 % Spc: 87.9 % ACC: 86.7 % F1 score: 0.838		

PJI, periprosthetic joint infection; P AUC, area under the curve; PV, positive predictive value; NPV, negative predictive value; AD, alpha-defensin; MID Sum, sum of the mid-range; PMN, polymorphonuclear cells; WBC, white blood cell; SF-CRP, synovial fluid C-reactive protein; RF, random forest; UMAP, Uniform Manifold Approximation and Projection; TP, total protein; BUN, blood urea nitrogen; Hb, hemoglobin; ALT, alanine aminotransferase; ICM, International Consensus Meeting; LSVM, Least Squares support Vector Machine; KSVM, Kernel Support Vector Machine; LGBM, Light Gradient Boosting Machine; MLP, multi-layer perceptron; ELWV, ensemble learning of weighted voting; MSIS, Musculoskeletal Infection Society; EBJIS, European Bone and Joint Infection Society; ACC, accuracy.
Sen: Proportion of true positives correctly identified (closer to 1.0 reduces missed diagnoses).
Spe: Proportion of true negatives correctly identified (closer to 1.0 reduces false alarms).
ACC: Overall correctness of predictions (misleading for imbalanced data; 1.0 = all correct).
PPV: Probability that a positive prediction is truly positive (high value increases diagnostic confidence).
NPV: Probability that a negative prediction is truly negative (high value strengthens rule-out reliability).
F1 Score: Balances precision and recall (1.0 = perfect balance, critical for imbalanced classes).
AUC: Measures the model's ability to distinguish classes, where 1.0 indicates perfect discrimination and 0.5 equals random guessing.

CRP and Hb levels in PJI patients (Supplementary Fig. 3f). Thus, ensemble ML can diagnose PJI through low-invasive blood biochemical tests and explore combinations of factors to identify more sensitive and accurate risk factors.

Urine is a critical diagnostic sample in various infectious diseases [111] but has not received widespread attention in PJI diagnosis. Immunohistochemical analysis has confirmed that PJI can alter the protein expression patterns of periprosthetic membranes, which can be observed through urinary peptides [112]. Omar et al. studied urine samples from 10 patients each with aseptic loosening, acute PJI, and chronic PJI, discovering differences in the excretion of 137 peptides between the PJI and aseptic groups [104] (Supplementary Fig. 3g). Using SVM-based MosaCluster software and leave-one-out cross-validation, they generated a PJI-specific peptide model. This biomarker model, comprising 83 peptides, demonstrated a sensitivity of 95 %, specificity of 90 %, and an AUC of 0.96 for PJI diagnosis (Supplementary Fig. 3h). Despite the small sample size, this study suggests that AI-assisted non-invasive PJI diagnosis is becoming increasingly feasible.

Current PJI diagnostic standards include results from microbial culture, pathology, and biomarkers from various samples like SF, blood, tissue, and microorganisms. Biomarkers from these samples

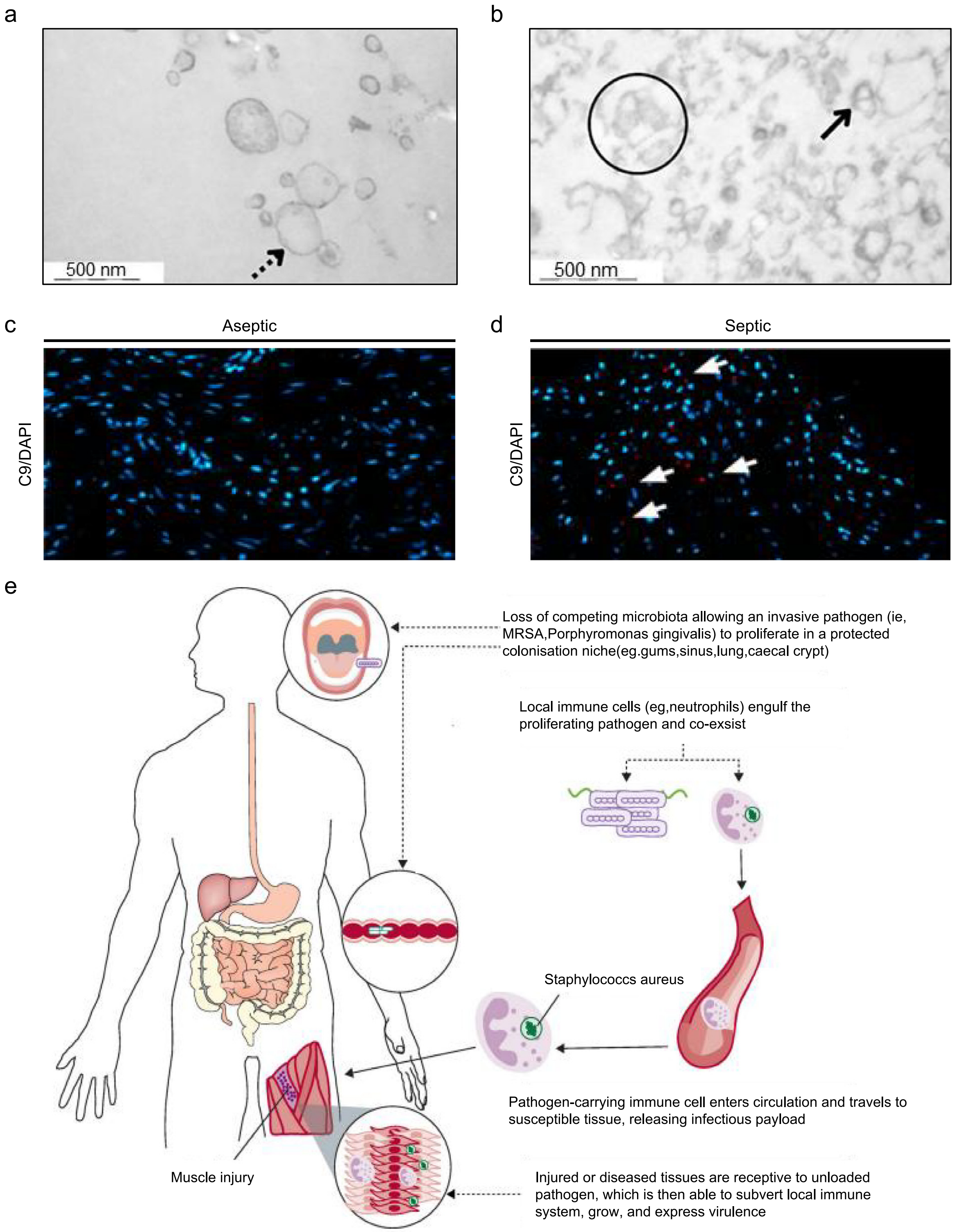


Fig. 5. New biomarkers in PJI. a. Smaller vesicles in aseptic samples [99]. b. EVs in Low-grade septic sample [99]. Copyright 2022 Sallai et al. c. C9 immunostaining in septic samples [115]. d. C9 immunostaining in aseptic samples [115]. Copyright 2023 Meinshausen et al. e. Trojan Horse hypothesis of surgical site infection [119]. Copyright 2020 Elsevier Ltd.

and features extracted from general data help improve the accuracy of PJI diagnosis [113]. Kuo et al. developed a two-tier stacked generalization PJI diagnostic model [114], using RF, XGBoost, Logistic Regression, and Naïve Bayesian as base classifiers, with SVM as the top-level *meta*-classifier. This model made the final diagnosis based on dozens of features, including blood and SF biomarkers, demographic characteristics, and surgery-related features. Unlike the ICM 2018 scoring system [83], this model adapts to existing patient indicators rather than relying on pre-specified and fixed standards, allowing important features to change accordingly for different patients. Compared to the paired ICM, this model's performance metrics, including recall, ACC, F1 score, and AUC, were superior. ML can identify essential features for diagnosis from available data, and PJI diagnostic models trained on these features may be more suitable for patients than the ICM 2018. Similarly, Chen et al. screened 46 clinical indicators, including preoperative markers like ESR and CRP, and intraoperative markers like pus formation and rapid frozen section, to develop a two-tier ML model composed of six base learners (Elastic Net, Linear SVM, Kernel SVM, Extra Trees, LightGBM, and MLP) and a *meta*-learner (Ensemble Learning of Weighted Voting) [90]. When compared to the MSIS 2011, ICM 2018, and EBJIS 2021 diagnostic standards using the same cohort, this model demonstrated superior sensitivity and accuracy. For all the included variables, likely due to the preponderance of acute PJI in the cohort, increased local skin temperature was identified as one of the most important variables. While these diagnostic models still rely on some intraoperative markers and cannot truly achieve preoperative diagnosis, their excellent diagnostic performance suggests that AI models based on multiple feature indicators may be the future direction of diagnostic standard development until perfect biomarkers are discovered.

In recent years, many new biomarkers have been developed, such as activated PMN-derived synovial EVs [99], which are significantly increased in PJI patients compared to aseptic loosening patients, with differences in particle size, morphology, and protein composition (Fig. 5a, b). C9 immunostaining in periprosthetic tissue of PJI patients is also significantly increased [115], with a sensitivity of 89 %, specificity of 75 %, and AUC of 0.84 when used as an independent feature for PJI diagnosis (Fig. 5c, d). Numerous studies [116,117] have shown a close relationship between the gut and PJI (Fig. 5e). In preclinical models of SSI and PJI, Staphylococcus aureus has been shown to migrate from the gut to the prosthetic joint or surgical site [118,119]. NGS has identified a large number of known gut commensals around the joints of PJI patients with high gut permeability and dysbiosis, providing preliminary validation for the "Trojan Horse" hypothesis of PJI [120], which suggests that blood cells like neutrophils and macrophages may mediate bacterial transfer between different sites in PJI patients [121,122]. Zonulin, a marker of increased gut permeability, is also significantly elevated in PJI compared to non-infected individuals [101]. These findings suggest a potential relationship between gut microbiota and PJI, and feces, the most commonly used sample type for studying gut microbiota, has yet to be explored in the field of PJI diagnosis.

Additionally, unconventional biomarkers like those from transcriptomics, proteomics, and metabolomics also warrant attention. The ability of AI to combine multiple omics to develop new biomarkers has been widely validated in other disease fields [123,124]. Extracting specific biomarkers for a particular disease from the vast data of multiple omics manually is challenging, but ML excels at feature extraction and classification from data. For example, Xu et al. identified and validated sepsis-related neutrophil hub genes using multiple ML methods like Lasso, KNN, LDA, LR, RF, and SVM, ultimately identifying five hub genes—APLP, CD 177, GAPDH, S100 A9, and STXBP2 as novel biomarkers targeting neutrophils [124]. Li et al. identified 3,468 proteins from 40 cerebrospinal fluid samples and constructed an RF model based on differential proteins, achieving a diagnostic accuracy of 98.11 % for neurosyphilis (NS) and successfully establishing three biomarkers for NS diagnosis (SEMA7A, SERPINA3, and ITIH4) [125]. These examples demonstrate the vast potential of AI combined with multi-omics technology in the field of biomarker development, which could be applied to PJI diagnosis and enhance its diagnostic efficiency.

Pathology-based diagnosis

Histopathology has long been an essential component in the diagnosis of PJI and is incorporated into various diagnostic criteria [16,91]. For instance, current standards such as the ICM 2018 and EBJIS 2021 stipulate that histopathological criteria require more than 5 neutrophils per high-power field (400x), with examination of at least 5 high-power fields and at least 2 tissue samples [83,84]. Despite continuous revisions aimed at improving diagnostic accuracy by increasing the number of neutrophils observed under the microscope, traditional manual pathological methods face challenges including strong subjectivity, inability to cover all areas of the slice, potential confusion with other inflammatory cells, lengthy processing times, and high costs [126,127]. The advent of AI offers a novel approach to pathological diagnosis (Supplementary Fig. 4), optimizing the entire process from slide preparation and interpretation to addressing related pathological queries, thus enhancing pathologists' consistency, accuracy, and efficiency.

The traditional process of slide preparation and preservation is labor-intensive and resource-demanding. Recent advancements in digital pathology address issues such as the need for fixed locations and equipment, as well as problems with slide breakage and fading, and facilitate a close integration of AI with pathology [128]. However, defocused regions during slide scanning can severely affect the quality of digital slides and subsequent image analysis. To tackle this, Senaras's team developed DeepFocus, a convolutional neural network-based system [129]. It was trained using 216,000 samples from 16 different H&E and immunohistochemically (IHC) slides with varying degrees of blurriness, capable of automatically detecting and segmenting blurred areas in digital slide images for re-scanning or refocusing (Supplementary Fig. 5a), achieving an average accuracy of 93.2 % (±9.6 %), a 23.8 % improvement over existing methods. Additionally, variations in staining reagents, staining protocols, and tissue slice thickness can cause color changes in digital slides. AI has introduced new methods for color normalization. Zanjani and colleagues used CNN and Generative Adversarial Networks (GANs) to develop a universal color normalization model [130], which generates color-standardized images while preserving the crucial structural details of the source image (Supplementary Fig. 5b), with an average standard deviation of 0.022 and 0.0191 in hematoxylin and eosin regions, outperforming many other models. This model can be applied to new slide images without pre-computing the statistical properties of source and template images. Furthermore, to bypass the complex process of tissue preparation, Jin and others developed a rapid slide-free histology platform, DeepDOF-SE [131]. By simply staining 3–4 mm of intraoperative tissue with Rhodamine B and DAPI, DeepDOF-SE uses U-Net and phase mask technology to achieve focused imaging of irregular tissue surfaces and then employs semi-supervised GAN CycleGAN for virtual staining, completing staining and virtual imaging in just 5–10 min, providing pathology information comparable to traditional slides (Supplementary Fig. 5c).

AI techniques, such as convolutional networks, are advantageous for image processing due to their high throughput, reproducibility, and accuracy. They are now widely used in cell counting and annotation for pathological slide image analysis [132,133]. For example, Dave and colleagues developed an adaptive segmentation algorithm, ASA-DL [134], for cell counting in complex dual-stained slides using CNN and U-Net (Supplementary Fig. 5e). This model showed an average error rate of 8.51 % and 6.26 % in neuron and microglia datasets, half or less than that of manual counting. Dy's team [135] evaluated UV-Net, an IHC quantification AI tool based on U-Net and CNN developed by Toronto City University, for calculating the Ki-67 proliferation index to assist in breast cancer prognosis (Supplementary Fig. 5d). The

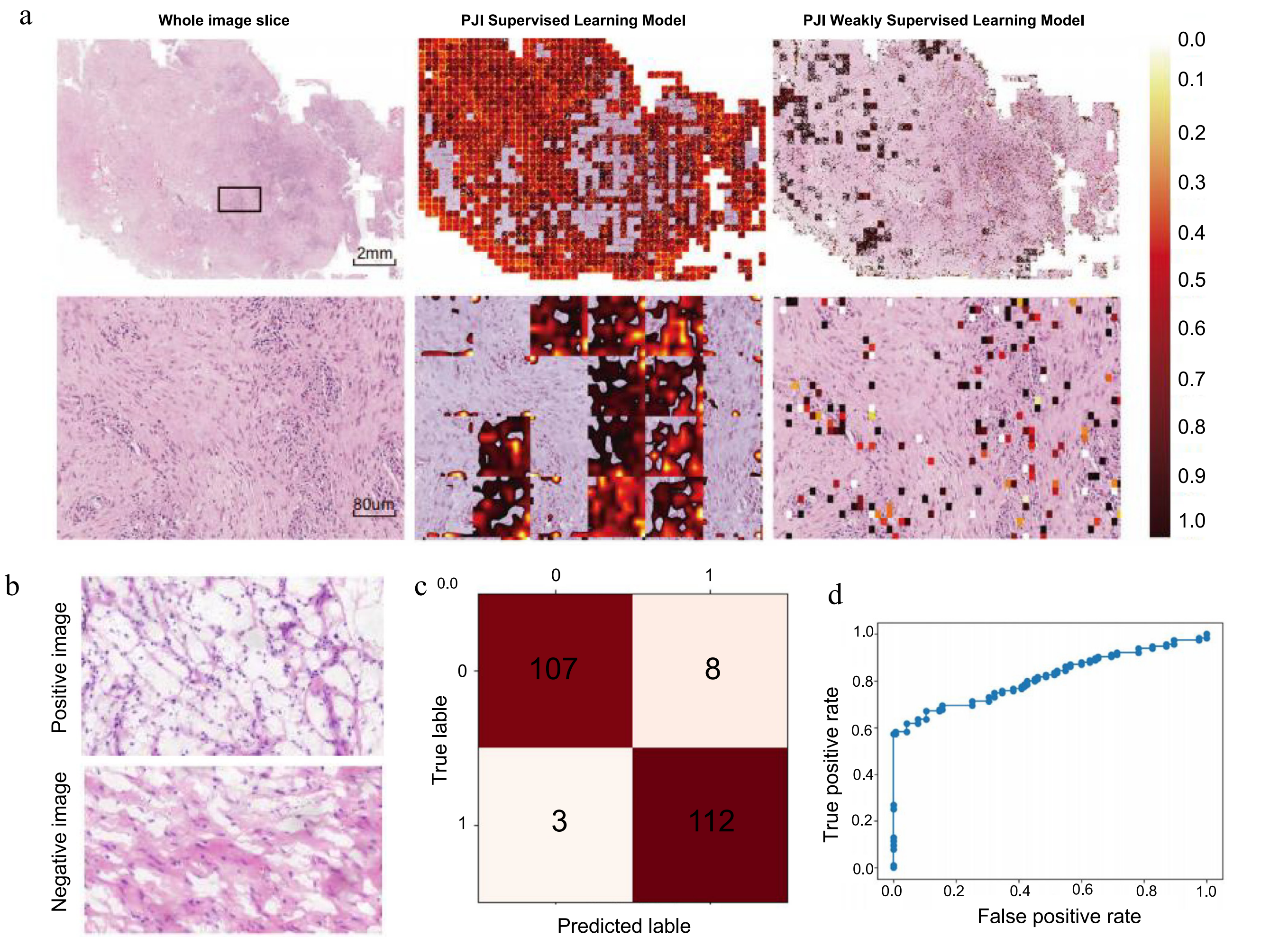


Fig. 6. Application of AI in pathological diagnosis. a. The visual outcomes of the PJI intelligent pathological diagnosis model [18]. b. Our models are consistent with international consensus standards [136] c. Confusion matrix of our model [136] d. ROC curve of our model [136].

results indicated that AI assistance significantly reduced PI scoring errors (2.1 % with AI vs. 5.9 % without AI, $p < 0.001$) and improved scoring consistency (ICC: 0.70 vs. 0.92; Krippendorff's α: 0.63 vs. 0.89; Fleiss' Kappa: 0.40 vs. 0.86), while also accelerating PI scoring (18.3 s without AI vs. 16.8 s with AI, $p < 0.001$), saving a median of 11.9 % of time.

To date, there has been limited research on intelligent slide reading for PJI. Our team [136] has developed a deep learning convolutional network model, ResNet, for intelligent interpretation of PJI pathological slides (Fig. 6b-d). External testing results show an average accuracy of 93.33 %, with an average recall rate of 97.39 % and an AUC of 0.8136, demonstrating high accuracy and excellent classification ability. Unlike manual reading, which involves selecting high-infection areas before scanning at high magnification, the convolutional network analyzes the entire slide, avoiding omissions and misclassification of cell morphology. For PJI diagnosis, in addition to neutrophil aggregation, other features such as soft tissue edema, varying degrees of muscle necrosis, and infiltration by various inflammatory cells are present [137]. These features are sensitive but non-specific and difficult to quantify for pathologists. However, convolutional networks can enhance and quantify these details to assist in diagnosis, reducing reliance on cell counting and the impact of complex cell morphology on diagnosis. Additionally, our external testing results exceed internal validation, indicating that our model has improved with training, achieving a diagnostic accuracy of 93.33 %. However, we were not satisfied with such accuracy levels. To further enhance model performance, we decided to reconstruct the framework by integrating self-supervised learning with supervised/weakly-supervised modeling approaches [138](Fig. 6a). After scanning and segmenting pathological slides, we obtained approximately 1.6 million images. The majority of these data were automatically annotated using the self-supervised model DINO v2, with the results subsequently serving as training labels. Through comparative analysis of six models, EfficientNet and CAMEL2 demonstrated optimal performance at the image differentiation and patient diagnostic levels, achieving AUC values of 0.9652 and 0.9460, respectively. Visual comparisons further revealed that the weakly-supervised CAMEL2 outperformed the s-model in accuracy, completeness, and reliability.

AI is increasingly integrated into pathology, moving beyond single tasks like cell counting or image segmentation to encompass various aspects of pathological diagnosis. For instance, Lu and colleagues developed a multimodal generative AI copilot, PathChat, based on a large multimodal language model, capable of understanding both pathological images and clinical background texts, and responding to complex pathological queries [139]. PathChat outperformed other natural language models such as LLaVA 1.5, LLaVA-Med, and GPT-4 V in diagnosing based on pathological images and clinical background information, and in answering

Table 8
Application of AI in imaging diagnosis.

First author	Application	Algorithm	Statistical performance	Strengths	Limitations
Li [144]	Diagnosis of PJI by CT images and numerical text data	A multimodal fusion network combining USA and GCN based on a hierarchical transformer architecture: HGT	ACC: 0.914; AUC: 0.959; Parameters: 68 M	Fusion of image and text features; utilize few parameters	Retrospective study; no external validation
Albano [145]	Diagnosis of PJI by preoperative MRI features	SVM	Sen: 92 %, Spc: 79 %, PPV: 89 %, NPV: 83 %, ACC: 88 %, AUC: 89 %	Identify Bone edema as the best predictor; avoid the misinterpretation of images	Retrospective study; susceptibility artifacts; small cohort size
Nie [149]	Diagnosis of PHI by 99mTc-MDP DBS.	A neural network based on deep learning and DBS: DBS-eNet	TKA: Acc:87.74 %; Spec:86.25 %; Sen:89.33 %; F1:0.8772; PPV:86.60 %; NPV;89.66 %; AUC:0.957 THA: Acc:86.36 %; Spec:87.86 %; Sen:83.75 %; F1:0.8178; PPV:80.25 %; NPV:90.50 %; AUC:0.906	Diagnostic performance superior to other classification models	Retrospective study; no external validation; small cohort size

AI, artificial intelligence; PJI, periprosthetic joint infection; ACC, accuracy; AUC, area under the curve; MRI, magnetic resonance imaging; SVM, support vector machine; PPV, positive predictive value; NPV, negative predictive value; PHI, periprosthetic hip infection; DBS, dynamic bone scintigraphy; TKA, total knee arthroplasty; THA, total hip arthroplasty.
Sen: Proportion of true positives correctly identified (closer to 1.0 reduces missed diagnoses).
Spe: Proportion of true negatives correctly identified (closer to 1.0 reduces false alarms).
ACC: Overall correctness of predictions (misleading for imbalanced data; 1.0 = all correct).
PPV: Probability that a positive prediction is truly positive (high value increases diagnostic confidence).
NPV: Probability that a negative prediction is truly negative (high value strengthens rule-out reliability).
F1 Score: Balances precision and recall (1.0 = perfect balance, critical for imbalanced classes).
AUC: Measures the model's ability to distinguish classes, where 1.0 indicates perfect discrimination and 0.5 equals random guessing.

open-ended pathological questions, and without the need to combine text data can reach more than 0.75 accuracy (Supplementary Fig. 6d, e). Dolezal and others integrated various deep learning tools into a pathology software, Slideflow (Supplementary Fig. 6a-c), featuring efficient staining normalization, weakly supervised whole-slide classification, uncertainty quantification, image feature extraction and analysis, GANs, and model interpretability [140]. Slideflow can extract and process entire slides at 40 × magnification in just 2.5 s, enabling analysis without switching software or reprocessing data, and the final model had a test-set accuracy of 90.2 %, sensitivity of 77.0 %, and specificity of 92.0 % within high-confidence predictions.

Currently, research on AI applications in PJI pathology is relatively sparse. However, slide-free microscopes like DeepDOF-SE for intraoperative diagnostics hold significant potential for PJI diagnosis. Additionally, AI platforms like PathChat and Slideflow, which integrate multiple AI tools and connect various aspects of pathological diagnosis, are expected to enhance our understanding of PJI pathology and improve diagnostic accuracy.

Imaging-based diagnosis

To date, imaging has not been included in mainstream diagnostic criteria for PJI, but non-invasive imaging techniques remain crucial tools for diagnosing, assessing, and monitoring PJI [141]. When evaluating suspected PJI with CT or MRI, metal-associated artifacts from prosthetics can limit diagnostic value. However, various methods such as parameter adjustment, projection completion techniques, angling, proprietary metal artifact reduction sequences (MARS), and AI can mitigate the impact of these artifacts [142,143]. Additionally, several studies have combined imaging techniques with AI to improve diagnostic performance for PJI (Table 8).

Li and colleagues developed a five-stage modality fusion network based on a transformer architecture, known as HGT [144] (Supplementary Fig. 7a). This structure uses unidirectional selective attention (USA) to integrate important features from text data and CT images, employing a graph convolutional network (GCN) based feature fusion network to reduce the complexity of CT images and integrate features captured by GCN. Ultimately, this model combines CT image features with digital text data to diagnose PJI. It achieves superior performance with fewer parameters compared to other multimodal fusion models, potentially enhancing clinical diagnostic efficiency and accuracy.

Albano and others utilized pre-THA MRI features from PJI patients to train, validate, and test an SVM-based ML algorithm [145]. This classifier demonstrated high diagnostic performance in an external test cohort, identifying bone edema, pericapsular edema, and synovitis as the best predictors of THA infection (Supplementary Fig. 7c).

Dynamic bone scintigraphy (DBS) with 99mTc-methylene diphosphonate (99mTc-MDP) is a widely used, reliable, and straightforward nuclear medicine method for diagnosing PJI. Previous studies have shown that DBS generally exhibits high sensitivity (80 %-94.3 %) for diagnosing PJI but lower specificity [146–148]. Nie and colleagues [149] constructed a PJI diagnostic model based on DBS, named Dynamic Bone Scintigraphy Effective Neural Network (DBS-eNet) (Supplementary Fig. 7b). Compared to mainstream CNNs (VGG, ResNet, DenseNet, and ConvNeXt) and a group of three nuclear medicine physicians, DBS-eNet outperformed all other classification models. It achieved higher diagnostic accuracy in an independent PKI test set and performed comparably to nuclear medicine physicians for PHI, making it a valuable tool for effectively and accurately diagnosing PKI and PHI. Visual analysis revealed that the lesion areas in both the blood flow and blood pool phases of DBS images were areas that contributed the most to DBS-eNet, meaning that DBSeNet could correctly focus on the lesion area when using these images.

In addition to MRI, CT, and bone scintigraphy, other imaging modalities have been explored for improving PJI diagnostic efficacy [150,151]. Molecular imaging techniques, similar to bone scintigraphy, show potential in the field. Systematic reviews of lower limb PJI suggest that 18F-FDG-PET has a total sensitivity and specificity

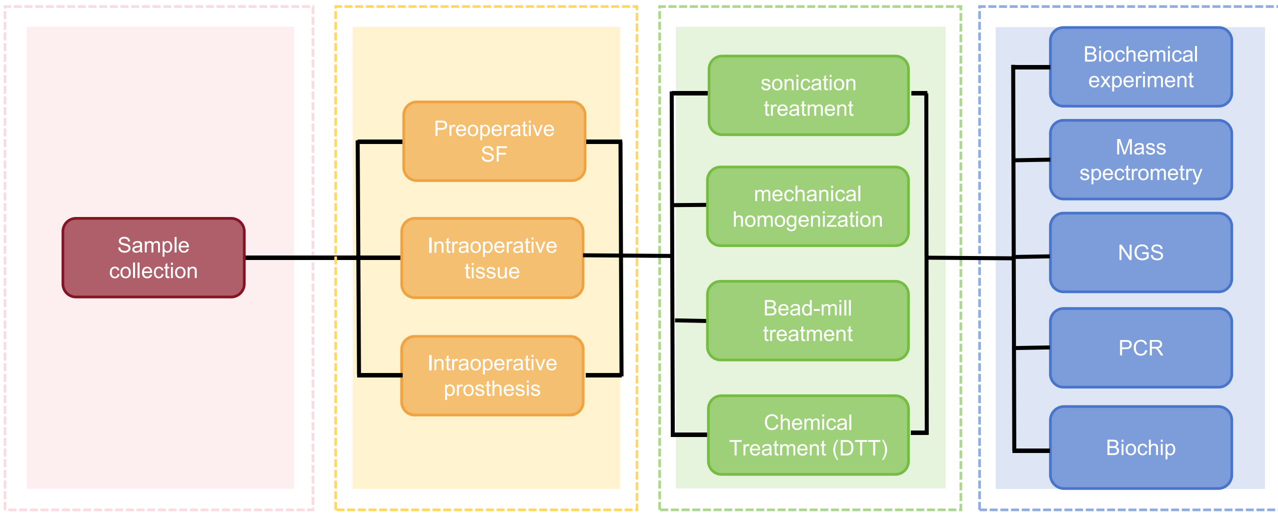


Fig. 7. Common strains identification methods for PJI. The samples for bacterial species identification in PJI usually include synovial fluid, intraoperative periprosthetic tissue and prosthesis. Low-frequency ultrasonic lysis, mechanical homogenization, bead milling and chemical treatment such as DTT can reduce the impact of biofilm on bacterial species identification. The treated samples can be used for bacterial identification from multiple perspectives such as bacterial phenotype and genetic information by means of biochemical experiments, mass spectrometry analysis, PCR, NGS, biochips, etc.

for PJI exceeding 80 % [146,152]. Yama and colleagues found significant differences in the SUVmax and SUVpeak of blood pool and skeletal SPECT/CT between PJI and non-PJI patients [153]. The DBS-eNet model has demonstrated that AI application can significantly enhance the sensitivity and specificity of 99mTc-MDP-DBS diagnostics, suggesting that other molecular imaging techniques may also improve diagnostic performance.

X-ray is typically the primary examination for postoperative complications, but previous studies have indicated that most early PJI X-ray findings appear normal. Decreased mineralization, periosteal reaction, heterotopic ossification, and patchy bone dissolution are only visible with more than 30 % bone loss [154,155]. Current research on X-ray diagnostic capabilities for PJI is limited and involves small sample sizes. With advancements in digital radiography and AI technology, new imaging signs and more effective sign combinations may be discovered, potentially providing simpler imaging evidence for PJI diagnosis.

Microbiology-based diagnosis

PJI is one of the most destructive complications following joint replacement surgery, with its pathogenesis closely related to pathogenic microorganisms. Thus, positive cultures for pathogenic microorganisms have remained a primary diagnostic criterion in scoring systems such as the ISDA 2013, ICM 2018, and EBJIS 2021 [83,84,86]. Typically, the diagnosis of PJI requires at least two positive samples with the same microorganism. However, issues such as preoperative antibiotic use, the presence of biofilms, inadequate culture time, and media selection can lead to many cases of PJI being culture-negative or misdiagnosed as aseptic failure, affecting treatment decisions [156]. As a result, numerous studies focus on improving the positivity rate of cultures from samples such as SF and periprosthetic tissues. This includes techniques like bead milling, mechanical grinding, and dithiothreitol (DTT) pre-treatment of periprosthetic tissues [157–161], as well as ultrasonic treatment of SF, periprosthetic tissues, and prosthetic components, and employing various identification methods such as mass spectrometry, PCR, and NGS (Fig. 7). Despite these efforts, traditional cultures face challenges such as prolonged culture times, the possibility of negative results, and difficulties in culturing atypical pathogens.

In addition to microbial cultures, molecular detection methods targeting pathogen genetic material, such as PCR and NGS, have been widely applied to PJI diagnostics. These methods often demonstrate higher sensitivity and specificity compared to cultures and significantly reduce diagnostic time. However, these non-culture microbial detection techniques also face challenges, including high costs, sample contamination, and difficulties distinguishing between live and dead bacteria. AI technology holds promise in enhancing current microbial culture and sequencing processes, thereby improving diagnostic accuracy and efficiency.

AI technology has been used to improve various aspects of microbial culture. In microbial culture, ML can optimize and develop culture media, outperforming traditional one-factor-at-a-time (OFAT) and response surface method (RSM) [162]. Specifically, ML models are constructed by learning from datasets to represent unknown relationships between cell culture quality and media composition, and then used to predict novel media compositions to fine-tune media components for improved cell growth (Supplementary Fig. 8). For example, Aghajani1 and colleagues optimized the concentrations of K_2HPO_4, $MgSO_4$, $(NH_4)_2SO_4$, and NH_4Cl in media for high-density cultivation of BL21 bacteria using ANN, identifying optimal concentrations for each component [163].

Clever Culture Systems' APAS automatic culture plate reading system interprets microbial growth patterns on urine culture media based on AI, automating the reading, interpretation, and classification of culture plates [164]. It distinguishes between no growth, significant growth, and plates requiring manual review, helping laboratories reduce plate handling time. Compared to manual handling, the APAS system processes 64 urine cultures in half the time (34 min vs. 71 min), significantly speeding up the diagnostic process.

AI also greatly enhances efficiency in subsequent microbial identification. Yu and colleagues developed a new rapid Methicillin-resistant S. aureus (MRSA) identification method using AI [165]. After exposing MRSA and methicillin-susceptible S. aureus (MSSA) to oxacillin sodium to disrupt cell wall, MRSA remains purple while MSSA turns pink in Gram staining. Machine vision (MV) captures these changes, and linear discriminant analysis (LDA) and nonlinear artificial neural network (ANN) models are used for identification (Supplementary Fig. 8c). This method achieved MRSA and MSSA discrimination accuracies of 96.7 %

Table 9
Summary of AI algorithm for AMR prediction.

First author	Algorithm	Target microorganism	Target antibiotic	Strengths	Limitations	Isolated strains/ mass spectra
Gerada [174]	CNN	Gram-negative Enterobacterales bacterial strains	amikacin, chloramphenicol, ciprofloxacin, gentamicin, meropenem, amoxicillin, amoxicillin/clavulanic acid, cefepime, ceftazidime, flomoxef, and tigecycline	increases throughput; standardize and automate	One plate can only correspond to one concentration; less training data	1086 Gram-negative strains
Liu [177]	SVM; SCM	Actinobacillus pleuropneumoniae	Tetracycline, Ampicillin, Sulfisoxazole, Trimethoprim, Enrofloxacin	Strong classification ability; high precision	Incomplete genomic data; rely on the database	96
Ren [178]	LR; SVM; RF; CNN	E.coli	ciprofloxacin, cefotaxime, ceftazidime, gentamicin	Can provide potential genes; No reliance on databases; Importance evaluation of SNPs	single reference genome; Increased model complexity with increase in layers and nodes	987 of Giessen dataset; 1509 of public dataset
Weis [182]	LR; LightGBM59; MLP	Staphylococcus aureus, Escherichia coli and Klebsiella pneumoniae	Oxacillin, ceftriaxone	relatively low cost; Faster than cultivation	Partial loss of spectra; rely on the database	303,195 mass spectra
Yu [183]	tSNE; LDA; RF; HCA; PCA	E. faecium, S. aureus, K. pneumoniae, A. baumannii, P. aeruginosa, Enterobacter	β-lactam	rapid detection(20 Min); convenient operation; relatively low cost	requires the pretreatment; Small sample size	78

AI, artificial intelligence; AMR, antimicrobial resistance; CNN, convolutional neural network; SVM, support vector machines; LR, logistic regression; RF, random forest; MLP, multi-layer perceptron; AUC, area under the curve; tSNE, t-distributed stochastic neighbor embedding; LDA, linear discriminant analysis; HCA, hierarchical cluster analysis; PCA, Principal component analysis.

and 97.3 %, respectively, with the entire process completed within one hour, greatly improving detection efficiency.

Signoroni's team constructed a hierarchical multi-network framework, DeepColony, for colony recognition and analysis on culture plates using various CNN algorithms [166]. The framework's 0–4 level network structure performs tasks such as locating and quantifying colonies, analyzing and identifying easily proliferating "good colonies," species identification, and selecting clinically significant colonies and plates. Validation on urinary tract infection samples demonstrated species identification accuracy of 97.6 % and consistency with experts in plate interpretation at 95.4 % (unweighted Kappa of 0.920).

Genomic sequencing methods can rapidly sequence sample DNA to identify microorganisms and are therefore used in PJI diagnosis [159,167]. AI can develop faster, more efficient, and accurate data processing and analysis algorithms, aiding researchers in complex data handling and extraction of useful information from large datasets [168,169]. This capability contributes to addressing the challenges of sequencing workflows characterized by complexity, high costs, and time-consuming nature [170]. For instance, Randhawa and colleagues developed the ML-DSP, an unmatched genomic classifier combining six supervised ML classifiers (Linear Discriminant, Linear SVM, Quadratic SVM, Fine KNN, Subspace Discriminant, and Subspace KNN) with digital processing techniques [171]. ML-DSP achieves ultra-fast, accurate genomic classification at all taxonomic levels, with an average classification accuracy of > 97 % and significantly faster than alignment-based classifiers (0.2 s vs. 2 min 14 s). It has successfully categorized 4710 bacterial genomes into three phyla: Spirochaetes, Firmicutes, and Proteobacteria (Supplementary Fig. 8b).

Microbiology is crucial in the diagnosis of PJI and closely related to subsequent treatment decisions. However, the lengthy and complex microbial identification processes often result in delayed diagnoses. AI technology has made microbial culture and identification more intelligent and automated, helping to improve the speed and accuracy of PJI microbiological diagnostics.

Application in PJI treatment

Predicting antibiotic resistance

As PJI is an infectious disease, effective antibiotic treatment is crucial for its management [172]. Common pathogens include *Staphylococcus aureus*, *Escherichia coli*, and *Pseudomonas aeruginosa* [173]. However, the development of antibiotic-resistant strains due to the continuous use of antibiotics significantly limits the effectiveness of treatment. Traditional microbiological methods for detecting resistance include antibiotic susceptibility testing, mass spectrometry, PCR, and sequencing. These methods may require prior cultivation of microorganisms, which is time-consuming and complex, or depend on large equipment and skilled technicians, making them less accessible in resource-limited settings. Therefore, there is a pressing need for new, effective, and rapid methods to detect pathogen resistance, with AI offering a promising solution due to its efficiency and accuracy (Table 9).

Based on susceptibility testing

The agar dilution method is one of the traditional antibiotic susceptibility testing methods, determining the minimum inhibitory concentration (MIC) of antibiotics based on bacterial growth on plates with varying antibiotic concentrations. However, the extensive data collection involved can lead to manual measurement errors. The AIgarMIC model, developed by Gerada and colleagues using CNNs, addresses this issue by automatically annotating plates from photographs to determine MIC values [174]. Validated with MIC results for 651 Gram-negative bacteria, the model achieved 98.9 % agreement with manual annotations, enhancing throughput, facilitating large-scale agar plate annotation, and reducing result interpretation time and human errors.

Based on nucleic acids

Resistance genes within genetic material can help accurately predict pathogen resistance. Current whole-genome sequencing (WGS) methods require comparison with relevant databases to identify pathogens and their antibiotic resistance [175]. However, the vast amount of genetic information complicates data processing and comparison. AI provides more effective alternatives, such as learning resistance gene features from known databases to predict new samples or directly extracting features from genomes for prediction. Research has confirmed that combining ML with sequencing and antimicrobial resistance (AMR) phenotype databases can achieve high-precision AMR prediction [176]. For example, Liu and colleagues used SVM, Set Covering Machine (SCM), and databases like European Nucleotide Archive and ResFinder to accurately predict resistance of *Pseudomonas aeruginosa* to tetracycline, ampicillin, sulfamethoxazole, trimethoprim, and enrofloxacin, achieving 100 % accuracy for all drugs except tetracycline [177](Supplementary Fig. 9b). Additionally, Ren and colleagues combined WGS with RF algorithms to predict *E. coli* resistance to ciprofloxacin, ceftazidime, cefotaxime, and gentamicin [178] (Supplementary Fig. 9a). Compared to other studies based on a database of known resistance genes or s single-nucleotide polymorphisms (SNPs) [179], this approach demonstrated effective AMR prediction without relying on known resistance gene or SNP databases and identified new potential resistance genes and SNPs.

Based on phenotypes

Predicting resistance based on microbial-specific proteins or phenotypes is another approach. Matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry can characterize the protein composition of individual bacterial or fungal colonies within minutes [180]. Compared to the high cost and time consumption of culture and sequencing [181], Weis and colleagues demonstrated that MALDI-TOF-based resistance prediction provides accurate and low-cost predictions within 24 h of sample collection [182] (Supplementary Fig. 9d). Three ML classifiers trained on MALDI-TOF data—logistic regression, gradient boosting decision trees (LightGBM), and deep neural networks (multi-layer perceptron, MLP)—achieved an AUC of 0.80 for *S. aureus*, indicating good accuracy.

Additionally, Yu and colleagues [183] developed a detection method based on gold nanoparticle arrays and ML analysis (Supplementary Fig. 8c). This method uses peptide-functionalized gold nanoparticles (P_AuNPs) to interact with the resistance phenotype of ESKAPE pathogens [184], identifying resistance to β-lactam antibiotics. By incubating P_AuNPs with ESKAPE pathogens for 20 min, the nanoparticles bind to the bacteria and produce characteristic resistance/sensitivity fingerprints. Using ML algorithms like t-distributed stochastic neighbor embedding (t-SNE) and linear discriminant analysis (LDA), the method achieves an accuracy of 89.5 % in identifying antibiotic resistance phenotypes.

AI technology has significantly improved resistance detection from various aspects, including susceptibility testing, pathogen protein expression and phenotyping, and pathogen genetic material. The efficiency of exploring potential resistance genes has also improved. However, the application of AI is still in its early stages, and many studies lack extensive validation. Some algorithms remain "black boxes," lacking interpretability. Nevertheless, as AI applications continue to advance and databases expand, algorithms will become increasingly precise and efficient in predicting antibiotic resistance.

Development of new antibiotics

The discovery and introduction of new antibiotics are both time-consuming and expensive. According to the WHO report on antibacterial agents in clinical development, only 18 new antibiotics have been approved since 2014, while global AMR continues to rise and has become a leading cause of death [185]. In 2019 alone, an estimated 4.95 million deaths globally were associated with bacterial AMR [186]. Therefore, the development of new antibiotics is critically important. Historically, most clinically used antibiotics have been derived from natural products produced by microorganisms or their semi-synthetic derivatives. Unfortunately, discovering new antibiotics through this approach is becoming increasingly difficult, with the same molecules often being repeatedly found [187]. As a result, the discovery of new antibiotics is becoming more challenging, and clinical treatment options for resistant bacterial infections are limited. Modern approaches to antibiotic development often involve screening large chemical libraries to identify the desired phenotype, which requires reviewing hundreds of thousands or even millions of molecules. This process is costly, time-consuming, and unable to perform exhaustive screening [188].

Library-based screening

Compared to manual screening, ML methods enable rapid, low-cost, high-throughput virtual screening of compound libraries to identify new drugs or repurpose existing ones [189,190]. Stokes et al. trained a deep neural network model to screen compounds from several independent chemical libraries that inhibit Escherichia coli [191] (Supplementary Fig. 10a). This led to the identification of a novel broad-spectrum small-molecule antibacterial compound: Halicin. In addition to inhibiting E. coli growth, Halicin effectively treated Clostridium difficile and multi-drug-resistant Acinetobacter baumannii infections in mouse models. Furthermore, the model identified eight structurally distinct antibacterial compounds from over 107 million molecules in the ZINC15 database, two of which demonstrated broad-spectrum antibacterial activity. Similarly, Liu et al. trained a message-passing neural network using molecules that inhibit A. baumannii growth in vitro [192], leading to the discovery of a new narrow-spectrum antibacterial compound against A. baumannii (Supplementary Fig. 10b).

Generative AI-based approaches

However, the models mentioned above evaluate molecules in chemical libraries one by one, which is time-consuming for large chemical spaces. Swanson et al. [193] developed a generative model called SyntheMol that designs synthetic pathways (Supplementary Fig. 10c). By constructing property prediction models using Chemprop, Chemprop-RDKit, and RF, SyntheMol used Monte Carlo tree search (MCTS) guided by the property prediction models to design and synthesize 58 molecules from a chemical space containing nearly 30 billion easily synthesized molecules, based on 132,000 molecules and 13 chemical reactions. Experimental validation showed that six of these molecules exhibited strong antibacterial activity against A. baumannii and other ESKAPE pathogens. This study demonstrated the effectiveness of generative AI in designing novel, synthesizable, and effective small-molecule antibiotic candidates, highlighting the tremendous potential of AI in antibiotic development and its promise for addressing the issue of broad-spectrum antibiotic resistance.

Moreover, regarding existing antibiotics, combining two or more antibiotics can enhance treatment success rates for PJI and reduce pathogen resistance to a single drug [194]. However, antibi-

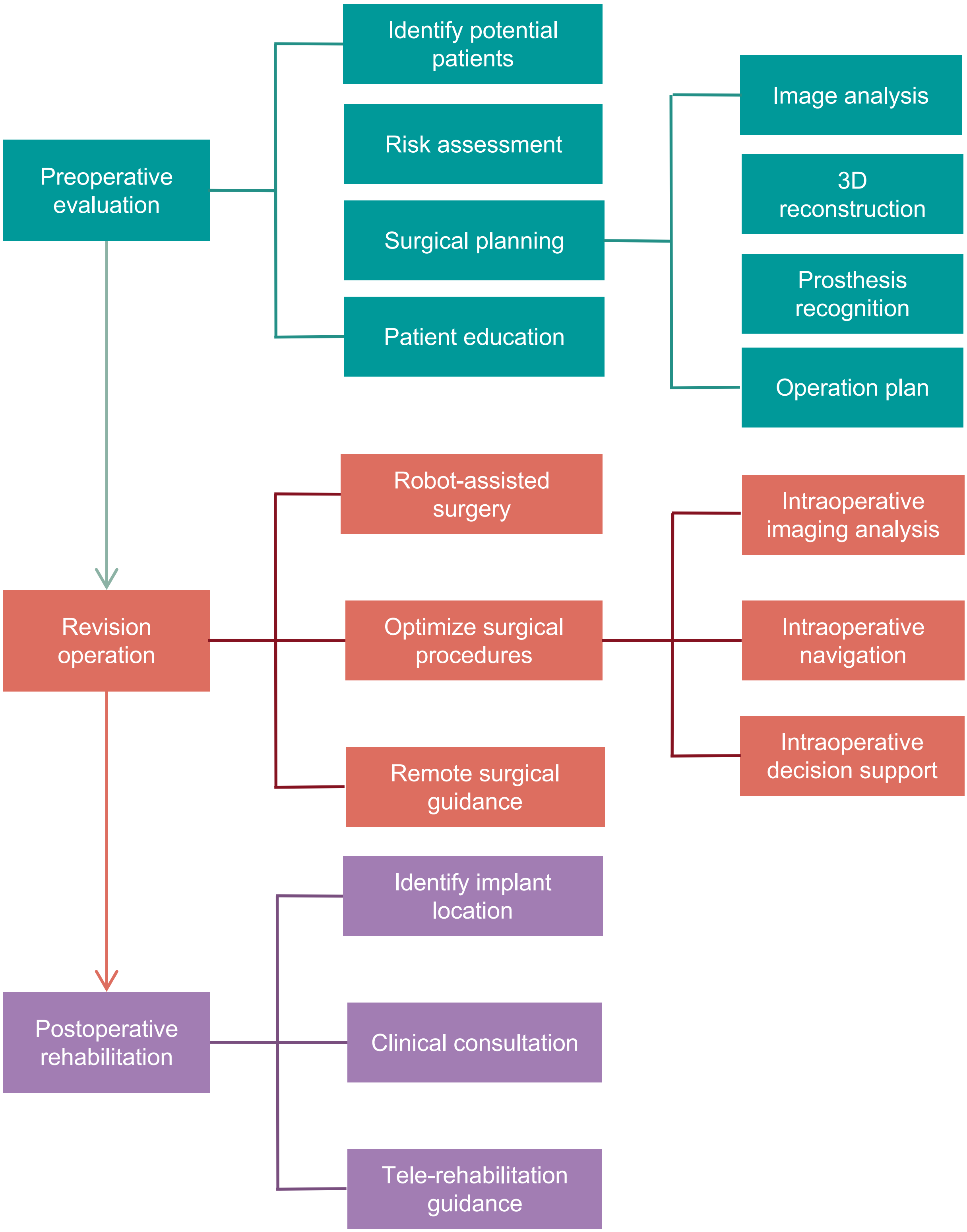


Fig. 8. Application of AI in surgery. It mainly includes perioperative applications such as preoperative evaluation, revision surgery and postoperative rehabilitation. AI can assist clinicians in preoperative planning by assessing surgical risks, identifying high-risk potential patients, popularizing and educating patients, and formulating surgical plans. In the revision surgery of PJI, AI technologies such as robotics and machine vision can also be used to achieve more precise surgical operations and assist in surgical decision-making. After surgery, imaging analysis is used to identify the position of the prosthesis, and smart devices are used to coordinate management and guidance of patients' postoperative rehabilitation training.

otic combinations currently rely largely on the experience of clinicians, with a lack of scientifically precise combination guidelines. Many studies have shown that AI technology can extract features such as drug structure, bioactivity, and gene expression on a large scale, allowing models to be trained to effectively predict drug synergy and efficiently identify new drug combinations [195,196].

Table 10
Summary of AI algorithm for Revision surgery.

First author	Application	Algorithm	Statistical performance	Strength	Limitations
Borjali [199]	Implant design identification by anterior-posterior radiographs	DenseNet-201 CNN	100 % top-3 Acc in identifying seven of nine THR implant designs; Cohen's kappa of 0.78	Short processing time(0.06 sec per radiograph); Visualization of decision process	The size of the dataset; Without consideration of the acetabular component
Klemt [200]	Implant design identification by anterior-posterior radiographs	CNN	Primary THA: AUC: 0.98, Acc: 98.2 %, Sen: 95.8 %, Spe: 98.6 % Primary TKA: AUC: 0.97, Acc: 97.4 %, Sen: 94.9 %, Spe: 97.8 % Revision THA: AUC: 0.98, Acc: 98.0 %, Sen: 94.9 %, Spe: 98.0 % Revision TKA: AUC: 0.96, Acc: 96.3 %, Sen: 94.5 %, Spe: 98.1 %	Relatively large sample size; High accuracy	Single institution sample; Limited implant designs
Lambrechts [201]	Create preoperative plans in a surgeon- and patient-specific manner for total knee arthroplast	SVR; LAD-SVR	Acc: femoral implant size: 82.2 %; tibial implant size: 85.0 %; Number of corrections: 3.76	Accurately predict implant size; Less correction	Black box; Relies on static bone-derived features
Chen [202]	Preoperative planning through CT images	attention U-Net; stacked hourglass network	DSC: 0.973	Short processing time (1.86 ± 0.12 min for each case); Easy accessibility; Highly integration	Small sample size; No statistical significant differences in partial clinical outcome
Rodrigues [208]	Contactless pose estimation through preoperative 3D model and navigation sensor	U-Net	Median rotation error is 3.17°; Median translation error is 6.18 mm	No additional markers; Contactless pose estimation	Small sample size; Relatively high error
Jang [30]	Automatically delineate revision zones and cone placements by postoperative rTKA radiographs	U-Net	DSC of femoral zones: 0.89 ± 0.08 DSC of tibial zones: 0.90 ± 0.05 Acc: identification of femoral cone: 97.8 %; femoral zonal cone placement:95.7 % identification of tibial cone: 100 %; tibial zonal cone placement:89.1 %	Short processing time(8 sec per radiograph)	No external validation; Image artifacts
Prem [217]	Remote acquisition of patient data after TKA	ML	ROM measurements were not different from physician; Uninterrupted passive data collection:100 %	Real-time continuous monitoring; Automatic reminder; portable	small sample size；no control

AI, artificial intelligence; CNN, convolutional neural network; THA, total hip arthroplasty; AUC, area under the curve; TKA, total knee arthroplasty; SVR, support vector regression; LAD-SVR, least absolute deviation support vector machine; DSC, differential scanning calorimetry; ML, machine learning; ROM, range of motion.

While most of the related research currently focuses on anti-cancer drugs, AI technology holds promise for solving the issue of antibiotic combination therapy as well.

Applications in revision surgery treatment

Generally, the treatment of PJI includes either antibiotic therapy alone or in combination with surgery. Surgical options encompass DAIR, arthroplasty, one-stage or two-stage revision surgery, or, in extreme cases, amputation [64]. Two-stage revision remains the preferred method, as it generally offers a higher overall rate of PJI eradication compared to one-stage revision [65]. However, both two-stage and one-stage revisions face challenges such as preoperative planning, intraoperative bone loss, anatomical deformities, and postoperative outcome assessment. Relying solely on clinical experience may not provide accurate assessments, but AI technology offers potential solutions that can enhance the entire perioperative period (Fig. 8, Table 10).

Preoperative planning

AI can rapidly and accurately identify anatomical landmarks and implant types through CT or MRI, match the appropriate joint prosthesis model, reduce the incidence of complications following joint replacement, and improve surgical outcomes [197]. A survey of orthopedic surgeons in the U.S. indicated that 88 % of respondents view preoperative identification of implant types as a key barrier in total joint revision surgery [198]. When the failed components cannot be identified, all parts, including well-fixed or functional components, must be removed to avoid issues arising from mismatched parts, which increases the burden on patients. Several studies have employed AI to assist in preoperative identification of implants. For example, Borjali and colleagues compared their developed CNN model with the abilities of three orthopedic surgeons to identify implants from X-rays [199] (Supplementary Fig. 11a, b). The CNN demonstrated accuracy equal to or greater than that of the doctors and required only 0.06 s, a significant improvement over the 20 min typically reported for manual identification [198]. Klemt and colleagues trained a CNN-based deep learning model using 11,204 anteroposterior X-ray images [200], achieving accurate identification of 24 THA designs and 14 TKA designs, with AUC values not less than 0.96, and further accuracy improvement with additional data. Accurate preoperative identification also aids surgeons in developing suitable surgical plans. Additionally, the choice of implant size and adjustment of prosthetic degrees of freedom are crucial. Lambrechts and colleagues compared the predictive capabilities of two nonlinear regression models—support vector regression (SVR) and least absolute deviation support vector machine (LAD-SVR)—for predicting implant size and prosthetic degrees of freedom in preoperative planning [201]. AI-provided plans showed higher accuracy compared to manufacturer-provided plans based on imaging and fixed algorithms (femur: 82.2 % vs. 68.4 %; tibia: 85.0 % vs. 73.1 %), with fewer corrections required by surgeons (4.10 vs. 7.13).

With advancements in AI and closer integration with surgical fields, more intelligent and integrated solutions have emerged.

Chen and colleagues integrated automatic image segmentation, image correction, preoperative deformity recognition, and postoperative simulation into an AI preoperative planning system (AIHIP) using U-Net and stacked hourglass networks [202]. AIHIP can complete CT-based preoperative planning in under 2 min and demonstrated greater accuracy in restoring femoral offset compared to X-ray-based planning (Supplementary Fig. 11c). In knee surgery, Li and colleagues [203] developed an AI system—AIJOINT—using 3D-UNet and HRNet neural network structures for TKA preoperative planning and patient-specific instrument design (PSI) (Supplementary Fig. 11d). The system achieved 92.9 % accuracy in predicting femoral and tibial component sizes, significantly outperforming traditional manual methods (42.9 % and 47.6 %). Additionally, AI-based PSI improved the accuracy of the hip-knee-ankle angle and reduced postoperative blood loss.

Intraoperative navigation and robotics

Revision surgeries face challenges such as anatomical changes, bone loss, and soft tissue contractures, leading to a high rate of complications [204,205]. AI-based intraoperative navigation and robotic technologies can improve prosthesis implantation, achieve higher functional scores, and better clinical outcomes [206,207]. With navigation systems, orthopedic surgeons can precisely track and observe anatomical sites and surgical instruments, making real-time adjustments to enhance surgical accuracy. Rodrigues and colleagues developed a video-based contactless navigation system using U-Net and Raposo's registration algorithms [208,209], which showed minimal rotational and displacement errors and can guide intraoperative operations (Supplementary Fig. 12a, b).

Some studies have also confirmed that navigated replacement surgeries yield better radiographic results compared to traditional replacements and reduce malalignment and component misplacement [210,211]. Surgical robots, which integrate machine vision, ML, NLP, and robotics, represent a fusion of AI and surgery [212]. Current robotic systems include passive, semi-autonomous, and autonomous control systems, with semi-autonomous systems being the most widely used due to their combined advantages of passive and autonomous control [213]. These systems typically base surgical plans on preoperative 3D imaging or intraoperative bone surface mapping, with surgeons controlling bone removal through feedback loops while the robot confines manipulations within the surgical plan [214]. This control improves surgeon accuracy and reduces error risk. Among semi-autonomous robots, the MAKO system is the most installed and studied. Research shows that MAKO-assisted placement of acetabular components in THA is more accurately within the Lewinnek's safe zone and improves the precision of femoral component anteversion compared to traditional techniques [215,216]. However, the long-term efficacy of robotic surgery remains unclear, and its application in revision surgery is limited. Whether AI-assisted precise surgery can provide better outcomes for PJI patients requires further investigation.

Postoperative assessment

Proper fixation of prosthetics after revision surgery is crucial for bone ingrowth at the prosthesis-femur interface. Reliable initial fixation promotes integration with surrounding bone tissue, ensuring long-term stability. Deficiencies such as bone loss, deformities, and instability during revision surgery may complicate prosthetic fixation. Therefore, surgeons must use objective methods to assess patient anatomy and prosthetic fixation postoperatively to evaluate surgical outcomes and guide patient rehabilitation. Jang and colleagues [30] trained a U-Net model that can quickly, consistently, and accurately analyze postoperative X-rays of rTKA (Supplementary Fig. 2d). This model automatically delineates rTKA fixation zones and prosthetic placement within 8 s, aiding surgeons in evaluating prosthetic fixation and the potential effectiveness of revisions, thus improving surgical results. Ramkumar and colleagues developed a remote patient monitoring system [217], when paired with wearable devices (knee sleeves) and smartphones, uses ML algorithms to track five indicators reflecting TKA recovery: patient mobility, knee joint range of motion, patient-reported outcome measures, opioid use, and home exercise adherence (Supplementary Fig. 12c). This system provides continuous passive data collection within 3 months postoperatively, enhancing patient compliance with rehabilitation exercises and aiding surgeons in monitoring and guiding postoperative recovery.

Potential applications of AI tools in clinical practice

Numerous studies have demonstrated that AI can play a positive role in various aspects of disease diagnosis and treatment. Specifically regarding PJI, traditional diagnostic processes require clinicians to comprehensively evaluate multiple parameters including local clinical signs, serum and synovial fluid parameters, imaging features, as well as preoperative and intraoperative microbiological culture results, followed by diagnosis through complex scoring systems. Prior to obtaining microbiological culture results, clinicians must rely solely on empirical experience and imaging findings to assess infection status and formulate surgical plans, with inherent uncertainties in this process contributing to the complexity and suboptimal outcomes of PJI management.

Our systematic review of AI applications in PJI reveals that AI systems can automatically extract relevant test results from electronic medical records for PJI diagnosis, achieve accurate risk prediction and diagnosis using partial parameters such as hematological and imaging data, and demonstrate diagnostic accuracy comparable to senior clinicians while surpassing human efficiency in imaging interpretation. Notably, AI demonstrates enhanced capability in quantitatively assessing clinically significant details imperceptible to human observation, such as lesion volume measurements. Furthermore, AI accelerates pathogen identification through rapid genomic analysis, facilitates discovery of novel antimicrobial agents via large-scale compound screening, and enables targeted antibiotic design. In surgical management, AI-assisted robotic systems have achieved clinical implementation with enhanced precision in osteotomy and revision procedures, while showing potential for personalized surgical planning and more thorough debridement strategies. Through integration with wearable devices, AI enables real-time postoperative functional assessment and infection surveillance to optimize rehabilitation.

We anticipate that future multimodal AI platforms integrating comprehensive patient data will establish fully intelligent and personalized PJI management pathways, spanning from initial diagnosis to therapeutic intervention and postoperative monitoring. This technological advancement is expected to significantly improve clinical outcomes in PJI treatment.

Certainly, the current clinical implementation of AI faces multilevel barriers spanning technical, stakeholder-related, and societal dimensions [218]. From a technical perspective, most existing AI applications suffer from inadequate interpretability and lack of external validation, with model outputs remaining insufficiently reliable for clinical confidence [219]. The medical data utilized across studies exhibit significant heterogeneity and fragmentation, necessitating standardized data classification and communication protocols to enhance interoperability [220]. Furthermore, the absence of unified standards and guidelines for clinical validation and performance evaluation further complicates system optimization. At the stakeholder level, widespread AI adoption risks eroding clinical workflows through potential deskilling and workforce displacement of healthcare professionals [221]. Current AI systems lack patient empathy, whereas patients prefer human clinicians [222]. For healthcare institutions, AI deployment demands substantial infrastructure upgrades and specialized technical personnel, with the technology yet to conclusively demonstrate its return on investment [223]. Societally, the establishment of large-scale medical databases raises legitimate privacy concerns among patients. Ambiguities persist regarding liability attribution for AI-driven diagnostic errors, while algorithms trained on fragmented healthcare data risk perpetuating discriminatory outcomes due to inherent data biases [224,225].

Consequently, despite AI's potential to enhance clinical efficiency and therapeutic outcomes, its comprehensive integration into routine clinical practice requires further multidisciplinary deliberation and systemic optimization.

The ethical status of AI applications

Ethical issues surrounding AI in healthcare remain a major focus of concern, spanning multiple dimensions across the entire lifecycle of AI technology applications, including data privacy, algorithmic bias, accountability, informed consent, and ethical review. AI training relies on multimodal data such as electronic health records and medical imaging, yet incomplete data-sharing mechanisms and insufficient anonymization techniques may lead to privacy breaches. In fact, healthcare data leaks have increased across many jurisdictions worldwide. Studies emphasize that even after data deletion or anonymization, re-identification techniques such as linkage attack frameworks can still reconstruct patients' health information, leaving privacy risks unresolved[226–228]. Furthermore, training datasets often fail to represent diverse populations, potentially causing algorithmic bias and fairness deficits. For instance, overrepresentation of White patients in training data may reduce diagnostic sensitivity for minority groups. In the context of PJI, the low incidence rate further limits the availability of comprehensive clinical data for AI training, exacerbating risks of algorithmic bias.

Legally, existing regulations have yet to clarify the legal validity of AI-driven decisions. When AI misdiagnoses result in medical errors, liability may involve multiple parties: developers (algorithmic flaws), hospitals (deployment errors), or clinicians (oversight negligence). While the EU AI Act categorizes obligations for AI systems based on risk levels, its implementation remains imperfect [229,230]. Additionally, patient data usage must safeguard informed consent. AI-driven remote monitoring systems (e.g., wearable devices) require real-time updates of patient authorization, yet most current platforms lack automated workflows to trigger secondary consent.

Therefore, refining ethical review mechanisms and strengthening oversight of AI healthcare applications is critical. Researchers are currently developing a novel assessment tool, CARE-AI (Collaborative Assessment for Responsible and Ethical AI Implementa residual confounding tion), designed to integrate existing guidelines for AI deployment, identify ethical challenges in healthcare AI predictive models, and ensure fairness, trustworthiness, and ethical accountability of AI tools [231]. However, this framework necessitates ongoing refinement to address evolving challenges.

Limitations

Firstly, many current studies are single-center investigations with small cohort sizes and homogeneous patient populations. Demographic characteristics, disease profiles, or clinical workflows in these cohorts may diverge from broader populations, leading to selection bias. Single-center datasets often fail to adequately account for confounding variables (e.g., geographic or socioeconomic factors), resulting in residual confounding bias. Comparative analyses between multicenter and single-center datasets confirm that models trained on multicenter data exhibit lower prediction errors and superior performance [232]. Furthermore, few studies have validated algorithms using external cohorts. Similar to single-center data, models trained and tested solely on internal cohorts may suffer from limited robustness and generalizability, often overestimating predictive accuracy [233]. Thus, existing models require training and validation on more representative populations to enhance predictive performance.

Secondly, most current research is retrospective, where missing data are common. Although researchers employ methods like Multiple Imputation by Chained Equations (MICE) to address missing values, residual bias may persist. This bias is typically associated with the type of missingness (e.g., missing not at random, MNAR) and the imputation approach. Studies indicate that methods such as MICE, RF imputation, and denoising autoencoders exhibit significant bias for MNAR data, with bias magnitude proportional to the extent of missingness [234]. From a methodological perspective, several studies have shown that combining MICE with techniques like RF, KNN, or Classification and Regression Trees (CART) can more effectively handle missing data and reduce bias compared to MICE alone [235]. Therefore, retrospective studies necessitate comprehensive investigations into imputation methods to develop more stable and less biased protocols for mitigating the impact of missing data.

Finally, algorithms such as deep neural networks, convolutional neural networks, and random forests inherently operate as "black boxes," obscuring decision-making pathways and undermining clinician trust in AI tools. Research indicates that explainable AI (XAI) enhances clinician confidence, particularly when explanations are clear, concise, and clinically relevant. Conversely, overly complex or contradictory explanations may erode trust [236]. In clinical practice, physicians rely on established medical knowledge for decision-making, whereas AI models may lack pathophysiological interpretability. For instance, a study using CNNs to diagnose pneumonia revealed that models identified metal tokens indicating body side for diagnosis, which significantly weakened the model's generalization capability on external data [237]. Moreover, when AI outputs conflict with clinical judgment, opaque models demand excessive time for verification, reducing workflow efficiency. Therefore, developing interpretability frameworks remains critical to improving user trust in AI systems.

Concluding remarks

The development of medicine is progressing toward digitalization and intelligence. Over the decades since its emergence, AI has brought disruptive transformations to numerous fields. With the advancement of AI and its deep integration with medical disciplines, its superior capabilities are becoming increasingly evident, particularly demonstrating immense potential in surgical applications and translational research. From disease diagnosis, interpretation of imaging and pathological data, and establishment of novel diagnostic algorithms, to surgical planning, intelligent intraoperative navigation, postoperative rehabilitation guidance, and preventive care, AI is poised to drive profound transformations across all aspects of medical practice.

In this review, we specifically focus on AI applications in PJI, a severe complication following arthroplasty. By synthesizing current advances in AI applications for PJI prevention, diagnosis, and treatment, we observe that in prevention research, most studies concentrate on ML-based risk prediction models for early identification of high-risk populations. These models frequently outperform traditional risk calculators and enable personalized analysis of patient-specific risk factors, facilitating proactive management of modifiable risks to prevent PJI occurrence. In such tasks, ANN may demonstrate enhanced predictive capabilities. Additionally, NLP-enabled large language models and automated medical record screening systems suggest that AI could optimize patient education and enable automated PJI detection through electronic health records. Regarding PJI diagnosis, compared with traditional multimodal diagnostic frameworks, ML algorithms achieve high diagnostic accuracy using limited examinations, significantly improving diagnostic efficiency. Data-driven models not only excel in processing high-dimensional imaging and pathological data but may also identify novel diagnostic features challenging to quantify manually. In microbiological diagnostics—a critical component of PJI diagnosis—AI enhances efficiency through intelligent culture interpretation, sequencing assistance, and antibiotic susceptibility screening. Therapeutically, AI-powered chemical library screening enables high-throughput evaluation of hundreds of millions of compounds with unprecedented efficiency, while generative AI shows promise in designing pathogen-specific antibiotics to optimize antimicrobial therapy. Furthermore, AI-assisted preoperative planning facilitates optimized surgical strategies and implant selection, and AI-driven robotic surgery achieves more precise osteotomy with minimized errors, collectively improving PJI treatment outcomes. Future AI development in PJI should emphasize multimodal data integration encompassing biomarkers, imaging, pathology, and microbiology, combined with intraoperative microbial monitoring and postoperative wearable devices to establish dynamic multimodal surveillance systems. This approach enables personalized treatment adjustments throughout the entire clinical pathway—from preoperative risk assessment and diagnosis to intraoperative navigation and postoperative monitoring. Implementation of federated learning could enhance cross-institutional data sharing and improve model generalizability through multicenter cohort validation.

However, significant challenges remain before widespread clinical implementation of AI. Technically, current PJI-related AI research suffers from limited cohort sizes, lack of external validation, and predominance of retrospective studies, potentially leading to algorithmic bias and compromised generalizability. The black-box nature of many AI algorithms impedes interpretation of decision-making processes, undermining clinical confidence. Moreover, our analysis reveals that numerous studies directly employ existing ML toolkits without sufficient customization for PJI-specific characteristics. Most research focuses on isolated aspects rather than comprehensive optimization across diagnostic workflows. Ethical considerations—including privacy protection, liability, informed consent, and ethical oversight—require substantial refinement, as current legal frameworks and regulatory systems remain inadequate.

Despite these challenges, existing research demonstrates that AI enables cost-effective, rapid, and accurate PJI diagnosis, scientifically optimized preoperative planning, precise surgical execution, personalized risk prediction, and remote rehabilitation guidance—all surpassing conventional approaches. As more data are incorporated into algorithmic training, models will self-optimize to achieve enhanced efficiency and accuracy. Large-scale data analysis may yield novel discoveries to advance PJI management, ultimately improving therapeutic outcomes while reducing morbidity and mortality. Continuous interdisciplinary collaboration and rigorous clinical validation will be essential to realize AI's full potential in revolutionizing PJI care.

Compliance with Ethics Requirements

This article does not contain any studies with human or animal subjects.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

We would like to thank Dr. Javad Parvizi, Dr. Jiying Chen and Dr. Yonggang Zhou for valuable comments. This study was financially supported through grants from the National Natural Science Foundation (82272558, 82072464).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jare.2025.03.039.

References

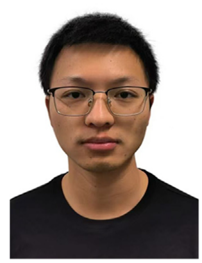
[1] Madry H. Surgical therapy in osteoarthritis. Osteoarthritis Cartilage 2022;30:1019–34.
[2] Portillo ME, Sancho I. Advances in the microbiological diagnosis of prosthetic joint infections. Diagnostics (Basel) 2023;13:809.
[3] Patsiogiannis N, Kanakaris NK, Giannoudis PV. Periprosthetic hip fractures: an update into their management and clinical outcomes. EFORT Open Rev 2021;6:75–92.
[4] Rajput V, Meek RMD, Haddad FS. Periprosthetic joint infection: what next? Bone Joint J 2022;104-B:1193–5.
[5] Elrod RW et al. Could novel radiographic findings help identify aseptic tibial loosening? J Arthroplasty 2023;38:1378–84.
[6] Gazendam A, Wood TJ, Tushinski D, Bali K. Diagnosing periprosthetic joint infection: a scoping review. Curr Rev Musculoskelet Med 2022;15:219–29.
[7] Pannu TS, Villa JM, Higuera CA. Diagnosis and management of infected arthroplasty. SICOT J 2021;7:54.
[8] McMaster Arthroplasty Collaborative (MAC). Risk Factors for Periprosthetic Joint Infection Following Primary Total Hip Arthroplasty: A 15-Year, Population-Based Cohort Study. J Bone Joint Surg Am 102, 503–509 (2020).
[9] Duensing I, Anderson MB, Meeks HD, Curtin K, Gililland JM. Patients with type-1 diabetes are at greater risk of periprosthetic joint infection: a population-based, retrospective, Cohort Study. J Bone Joint Surg Am 2019;101:1860–7.
[10] Premkumar A et al. Projected economic burden of periprosthetic joint infection of the hip and knee in the United States. J Arthroplasty 2021;36:1484–1489.e3.
[11] Kurtz SM et al. Are we winning or losing the battle with periprosthetic joint infection: trends in periprosthetic joint infection and mortality risk for the medicare population. J Arthroplasty 2018;33:3238–45.
[12] Zhang Y-W et al. Prevalence, characteristics, and associated risk factors of the elderly with hip fractures: a cross-sectional analysis of NHANES 2005-2010. Clin Interv Aging 2021;16:177–85.
[13] Tande AJ, Patel R. Prosthetic joint infection. Clin Microbiol Rev 2014;27:302–45.
[14] Tl, T. et al. Development and Evaluation of a Preoperative Risk Calculator for Periprosthetic Joint Infection Following Total Joint Arthroplasty. The Journal of bone and joint surgery. American volume 100, (2018).
[15] Vrancianu CO et al. The challenge of periprosthetic joint infection diagnosis: from current methods to emerging biomarkers. Int J Mol Sci 2023;24:4320.
[16] Sigmund IK, Luger M, Windhager R, McNally MA. Diagnosing periprosthetic joint infections : a comparison of infection definitions: EBJIS 2021, ICM 2018, and IDSA 2013. Bone Joint Res 2022;11:608–18.
[17] Ricciardi BF et al. New developments and future challenges in prevention, diagnosis, and treatment of prosthetic joint infection. J Orthop Res 2020;38:1423–35.
[18] Middleton R, Price A, Alvand A. Contemporary outcomes of debridement, antibiotics and implant retention in knee arthroplasty. Ann Jt 2022;7:9.
[19] Auñón Á et al. Outcomes of the Subsequent periprosthetic joint infection revisions after a failed debridement, antibiotics and implant retention: a multicentric study of 197 patients. Surg Infect (Larchmt) 2024. doi: https://doi.org/10.1089/sur.2024.047.
[20] Sotos JG. MYCIN and NEOMYCIN: two approaches to generating explanations in rule-based expert systems. Aviat Space Environ Med 1990;61:950–4.
[21] Feng H et al. A Random forest model for peptide classification based on virtual docking data. Int J Mol Sci 2023;24:11409.
[22] Ding C, Bao T-Y, Huang H-L. Quantum-inspired support vector machine. IEEE Trans Neural Netw Learn Syst 2022;33:7210–22.
[23] Fan X et al. Prediction of outpatient rehabilitation patient preferences and optimization of graded diagnosis and treatment based on XGBoost machine learning algorithm. Front Artif Intell 2024;7:1473837.

[24] Herzog NJ, Magoulas GD. Convolutional neural networks-based framework for early identification of dementia using MRI of brain asymmetry. Int J Neural Syst 2022;32:2250053.
[25] Mao S, Sejdic E. A review of recurrent neural network-based methods in computational physiology. IEEE Trans Neural Netw Learn Syst 2023;34:6983–7003.
[26] Wang R et al. Applications of generative adversarial networks in neuroimaging and clinical neuroscience. Neuroimage 2023;269:119898.
[27] Lu W et al. Application of Entity-BERT model based on neuroscience and brain-like cognition in electronic medical record entity recognition. Front Neurosci 2023;17:1259652.
[28] Hou W, Ji Z. Assessing GPT-4 for cell type annotation in single-cell RNA-seq analysis. Nat Methods 2024;21:1462–5.
[29] Nath N, Lee S-H, Lee I. NEAR: Named entity and attribute recognition of clinical concepts. J Biomed Inform 2022;130:104092.
[30] Jang SJ et al. Standardized fixation zones and cone assessments for revision total knee arthroplasty using deep learning. J Arthroplasty 2023;38:S259–S265.e2.
[31] Yao Z et al. Construction and multicenter diagnostic verification of intelligent recognition system for endoscopic images from early gastric cancer based on YOLO-V3 algorithm. Front Oncol 2022;12:815951.
[32] Al-Hammuri K, Gebali F, Kanan A, Chelvan IT. Vision transformer architecture and applications in digital health: a tutorial and survey. Vis Comput Ind Biomed Art 2023;6:14.
[33] Jin Y, Lan A, Dai Y, Jiang L, Liu S. Development and testing of a random forest-based machine learning model for predicting events among breast cancer patients with a poor response to neoadjuvant chemotherapy. Eur J Med Res 2023;28:394.
[34] Wang X et al. Detection and classification of mandibular fracture on CT scan using deep convolutional neural network. Clin Oral Invest 2022;26:4593–601.
[35] Collins T et al. Automatic optical biopsy for colorectal cancer using hyperspectral imaging and artificial neural networks. Surg Endosc 2022;36:8549–59.
[36] Gonzalez MR, Pretell-Mazzini J, Lozano-Calderon SA. Risk factors and management of prosthetic joint infections in megaprostheses-a review of the literature. Antibiotics (Basel) 2023;13:25.
[37] Panula VJ et al. Risk factors for prosthetic joint infections following total hip arthroplasty based on 33,337 hips in the Finnish Arthroplasty Register from 2014 to 2018. Acta Orthop 2021;92:665–72.
[38] Sigurdardottir M et al. Prevalence of modifiable risk factors in primary elective arthroplasty and their association with infections. Acta Orthop 2023;94:38–44.
[39] Shichman I et al. Failed 2-stage revision knee arthroplasty for periprosthetic joint infection-patient characteristics and outcomes. J Arthroplasty 2023;38:2177–82.
[40] Matar HE, Bloch BV, Snape SE, James PJ. Outcomes of single- and two-stage revision total knee arthroplasty for chronic periprosthetic joint infection : long-term outcomes of changing clinical practice in a specialist centre. Bone Joint J 2021;103-B:1373–9.
[41] Klemt C et al. Elevated ESR and CRP Prior to second-stage reimplantation knee revision surgery for periprosthetic joint infection are associated with increased reinfection rates. J Knee Surg 2023;36:354–61.
[42] Humphrey TJ et al. Rates and outcomes of periprosthetic joint infection in persons who inject drugs. J Arthroplasty 2023;38:152–7.
[43] Shohat N, Goswami K, Tan TL, Fillingham Y, Parvizi J. Increased failure after irrigation and debridement for acute hematogenous periprosthetic joint infection. J Bone Joint Surg Am 2019;101:696–703.
[44] Walkay S et al. Outcomes of debridement, antibiotics and implant retention (DAIR) for periprosthetic joint infection in a high-volume arthroplasty centre. Indian J Orthop 2022;56:1449–56.
[45] Wingert NC et al. The ACS NSQIP risk calculator is a fair predictor of acute periprosthetic joint infection. Clin Orthop Relat Res 2016;474:1643–8.
[46] Klemt C, Tirumala V, Smith EJ, Padmanabha A, Kwon Y-M. Development of a preoperative risk calculator for reinfection following revision surgery for periprosthetic joint infection. J Arthroplasty 2021;36:693–9.
[47] Bozic KJ et al. Estimating risk in Medicare patients with THA: an electronic risk calculator for periprosthetic joint infection and mortality. Clin Orthop Relat Res 2013;471:574–83.
[48] Inacio MCS, Pratt NL, Roughead EE, Graves SE. Predicting infections after total joint arthroplasty using a prescription based comorbidity measure. J Arthroplasty 2015;30:1692–8.
[49] Haeberle HS et al. Artificial intelligence and machine learning in lower extremity arthroplasty: a review. J Arthroplasty 2019;34:2201–3.
[50] Yeo I et al. The use of artificial neural networks for the prediction of surgical site infection following TKA. J Knee Surg 2023;36:637–43.
[51] Klemt C et al. Machine learning models accurately predict recurrent infection following revision total knee arthroplasty for periprosthetic joint infection. Knee Surg Sports Traumatol Arthrosc 2022;30:2582–90.
[52] Tornero E et al. KLIC-score for predicting early failure in prosthetic joint infections treated with debridement, implant retention and antibiotics. Clin Microbiol Infect 2015;21(786):e9–786.e17.
[53] Löwik CAM et al. Predicting failure in early acute prosthetic joint infection treated with debridement, antibiotics, and implant retention: external validation of the KLIC score. J Arthroplasty 2018;33:2582–7.
[54] Shohat N et al. 2020 Frank Stinchfield Award: Identifying who will fail following irrigation and debridement for prosthetic joint infection. Bone Joint J 2020;102-B:11–9.
[55] Sancho, I. et al. Accuracy of a Novel Preoperative Failure Risk Model for Debridement Antibiotics and Implant Retention (DAIR) in Acute Prosthetic Joint Infection. Diagnostics (Basel) 12, 2097 (2022).
[56] Howie CM, Mears SC, Barnes CL, Stambough JB. Readmission, complication, and disposition calculators in total joint arthroplasty: a systemic review. J Arthroplasty 2021;36:1823–31.
[57] Bülow E et al. Prediction of early periprosthetic joint infection after total hip arthroplasty. Clin Epidemiol 2022;14:239–53.
[58] Harris AH et al. Prediction models for 30-day mortality and complications after total knee and hip arthroplasties for veteran health administration patients with osteoarthritis. J Arthroplasty 2018;33:1539–45.
[59] Kunze KN, Karhade AV, Polce EM, Schwab JH, Levine BR. Development and internal validation of machine learning algorithms for predicting complications after primary total hip arthroplasty. Arch Orthop Trauma Surg 2023;143:2181–8.
[60] Quinlan ND, Werner BC, Brown TE, Browne JA. Risk of prosthetic joint infection increases following early aseptic revision surgery of total hip and knee arthroplasty. J Arthroplasty 2020;35:3661–7.
[61] Klemt C et al. The use of artificial intelligence for the prediction of periprosthetic joint infection following aseptic revision total knee arthroplasty. J Knee Surg 2024;37:158–66.
[62] Klemt C et al. Outcomes and risk factors associated with 2-stage reimplantation requiring an interim spacer exchange for periprosthetic joint infection. J Arthroplasty 2021;36:1094–100.
[63] Harris AHS et al. Can machine learning methods produce accurate and easy-to-use prediction models of 30-day complications and mortality after knee or hip arthroplasty? Clin Orthop Relat Res 2019;477:452–60.
[64] Citak M et al. Risk factors for failure after 1-stage exchange total knee arthroplasty in the management of periprosthetic joint infection. J Bone Joint Surg Am 2019;101:1061–9.
[65] Kandel CE et al. Predictors of treatment failure for hip and knee prosthetic joint infections in the setting of 1- and 2-stage exchange arthroplasty: a multicenter retrospective cohort. Open Forum Infect Dis 2019;6:ofz452.
[66] Wakabayashi H, Hasegawa M, Naito Y, Tone S, Sudo A. Outcome of irrigation and debridement with topical antibiotic delivery using antibiotic-impregnated calcium hydroxyapatite for the management of periprosthetic hip joint infection. Antibiotics (Basel) 2023;12:938.
[67] Barry JJ et al. Irrigation and debridement with chronic antibiotic suppression is as effective as 2-stage exchange in revision total knee arthroplasty with extensive instrumentation. J Bone Joint Surg Am 2021;103:53–63.
[68] Tai DBG et al. Outcomes and risk factors for failure after débridement, antibiotics, and implant retention for elbow periprosthetic joint infection. J Shoulder Elbow Surg 2023;32:475–9.
[69] Wouthuyzen-Bakker M, Shohat N, Parvizi J, Soriano A. Risk scores and machine learning to identify patients with acute periprosthetic joints infections that will likely fail classical irrigation and debridement. Front Med (Lausanne) 2021;8:550095.
[70] Rajgopal A, Panda I, Rao A, Dahiya V, Gupta H. Does prior failed debridement compromise the outcome of subsequent two-stage revision done for periprosthetic joint infection following total knee arthroplasty? J Arthroplasty 2018;33:2588–94.
[71] Sherrell JC et al. The Chitranjan Ranawat Award: fate of two-stage reimplantation after failed irrigation and débridement for periprosthetic knee infection. Clin Orthop Relat Res 2011;469:18–25.
[72] Bernaus M et al. Risk factors of DAIR Failure and validation of the KLIC score: A Multicenter Study Of Four Hundred Fifty-Five Patients. Surg Infect (Larchmt) 2022;23:280–7.
[73] Kung TH et al. Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. PLOS Digit Health 2023;2:e0000198.
[74] Ayers JW et al. Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. JAMA Intern Med 2023;183:589–96.
[75] Decker H et al. Large language model-based Chatbot vs surgeon-generated informed consent documentation for common procedures. JAMA Netw Open 2023;6:e2336997.
[76] Uz C, Umay E. 'Dr ChatGPT': Is it a reliable and useful source for common rheumatic diseases? Int J Rheum Dis 2023;26:1343–9.
[77] Humar P, Asaad M, Bengur FB, Nguyen V. chatgpt is equivalent to first-year plastic surgery residents: evaluation of ChatGPT on the plastic surgery in-service examination. Aesthet Surg J 2023;43:NP1085–NP1089.
[78] Jung LB et al. ChatGPT passes german state examination in medicine with picture questions omitted. Dtsch Arztebl Int 2023;120:373–4.
[79] A, d. et al. Are ChatGPT's free-text responses on periprosthetic joint infections of the hip and knee reliable and useful? J Clin Med 2023;12.
[80] Fu S et al. Automated detection of periprosthetic joint infections and data elements using natural language processing. J Arthroplasty 2021;36:688–92.
[81] Upfill-Brown A et al. Epidemiology of revision total knee arthroplasty in the United States, 2012 to 2019. Arthroplast Today 2022;15:188–195.e6.
[82] Gwam CU et al. Current epidemiology of revision total hip arthroplasty in the United States: national inpatient sample 2009 to 2013. J Arthroplasty 2017;32:2088–92.

[83] Parvizi J et al. The 2018 definition of periprosthetic hip and knee infection: an evidence-based and validated criteria. J Arthroplasty 2018;33:1309–1314.e2.
[84] McNally M et al. The EBJIS definition of periprosthetic joint infection. Bone Joint J 2021;103-B:18–25.
[85] Workgroup Convened by the Musculoskeletal Infection Society. New definition for periprosthetic joint infection. J Arthroplasty 26, 1136–1138 (2011).
[86] Osmon DR et al. Diagnosis and management of prosthetic joint infection: clinical practice guidelines by the Infectious Diseases Society of America. Clin Infect Dis 2013;56:e1–e25.
[87] Shohat N et al. Development and validation of an evidence-based algorithm for diagnosing periprosthetic joint infection. J Arthroplasty 2019;34:2730–2736.e1.
[88] Kunutsor SK, Whitehouse MR, Blom AW, Beswick AD, INFORM Team. Patient-related risk factors for periprosthetic joint infection after total joint arthroplasty: a systematic review and meta-analysis. PLoS One 11, e0150866 (2016).
[89] Deirmengian C, McLaren A, Higuera C, Levine BR. Physician use of multiple criteria to diagnose periprosthetic joint infection may be less accurate than the use of an individual test. Cureus 2022;14:e31418.
[90] Chen W et al. A machine learning-based model for 'In-time' prediction of periprosthetic joint infection. Digit Health 2024;10:20552076241253531.
[91] McNally M, Sigmund I, Hotchen A, Sousa R. Making the diagnosis in prosthetic joint infection: a European view. EFORT Open Rev 2023;8:253–63.
[92] Yilmaz MK, Abbaszadeh A, Tarabichi S, Azboy I Parvizi J. Diagnosis of periprosthetic joint infection: the utility of biomarkers in 2023. Antibiotics (Basel) 12, 1054 (2023).
[93] Schindler M et al. Novel diagnostic markers for periprosthetic joint infection: a systematic review. Front Cell Infect Microbiol 2023;13:1210345.
[94] Cobra HA. de AB, Mozella AP, da Palma IM, Salim R, Leal AC. Cell-free deoxyribonucleic acid: a potential biomarker of chronic periprosthetic knee joint infection. J Arthroplasty 2022;37:2455–2459.
[95] Fröschen FS et al. Analysis of synovial biomarkers with a multiplex protein microarray in patients with PJI undergoing revision arthroplasty of the hip or knee joint. Arch Orthop Trauma Surg 2020;140:1883–90.
[96] Fuchs M, Trampuz A, Kirschbaum S, Winkler T, Sass FA. Soluble Pecam-1 as a biomarker in periprosthetic joint infection. J Clin Med 2021;10:612.
[97] Ikeda S, Uchiyama K, Minegishi Y, Nakamura M, Takaso M. Evaluation of myeloperoxidase in synovial fluid as a biomarker for chronic periprosthetic joint infection. Int Orthop 2020;44:1915–20.
[98] Grassi M et al. synovial biomarkers to detect chronic periprosthetic joint infection: a pilot study to compare calprotectin rapid test, calprotectin ELISA immunoassay and leukocyte esterase test. J Arthroplasty 2022;37:781–6.
[99] Sallai I et al. Activated polymorphonuclear derived extracellular vesicles are potential biomarkers of periprosthetic joint infection. PLoS One 2022;17:e0268076.
[100] Karbysheva S et al. Synovial fluid d-lactate-A novel pathogen-specific biomarker for the diagnosis of periprosthetic joint infection. J Arthroplasty 2020;35:2223–2229.e2.
[101] Chisari E, Cho J, Wouthuyzen-Bakker M, Parvizi J. Gut permeability may be associated with periprosthetic joint infection after total hip and knee arthroplasty. Sci Rep 2022;12:15094.
[102] Echeverria AP et al. Sequencing of circulating microbial cell-free DNA can identify pathogens in periprosthetic joint infections. J Bone Joint Surg Am 2021;103:1705–12.
[103] Dartus J et al. Diagnostic accuracy of the BJI InoPlexTM (Diaxonhit) immunoassay on blood samples for periprosthetic joint infection in complex microbiological situations. Preliminary results of 24 cases in a French Reference Center for Complex Bone and Joint Infection (CRIOAC). Orthop Traumatol Surg Res 2021;107:102909.
[104] Omar M, Windhagen H, Krettek C, Ettinger M. Noninvasive diagnostic of periprosthetic joint infection by urinary peptide markers: a preliminary study. J Orthop Res 2021;39:339–47.
[105] Toler KO et al. Nationwide results of microorganism antigen testing as a component of preoperative synovial fluid analysis. J Bone Joint Surg Am 2023;105:448–54.
[106] Paranjape PR et al. Achieving high accuracy in predicting the probability of periprosthetic joint infection from synovial fluid in patients undergoing hip or knee arthroplasty: the development and validation of a multivariable machine learning algorithm. Cureus 2023;15:e51036.
[107] Bonanzinga T et al. The role of alpha defensin in prosthetic joint infection (PJI) diagnosis: a literature review. EFORT Open Rev 2019;4:10–3.
[108] Ivy MI et al. Synovial fluid α defensin has comparable accuracy to synovial fluid white blood cell count and polymorphonuclear percentage for periprosthetic joint infection diagnosis. Bone Joint J 2021;103-B:1119–26.
[109] Kawakami E et al. Monitoring of blood biochemical markers for periprosthetic joint infection using ensemble machine learning and UMAP embedding. Arch Orthop Trauma Surg 2023;143:6057–67.
[110] Yaradılmış YU et al. Do low hemoglobin levels affect the healing process of periprosthetic joint infection? Cureus 2021;13:e14393.
[111] Caceres PS et al. High SARS-CoV-2 viral load in urine sediment correlates with acute kidney injury and poor COVID-19 outcome. J Am Soc Nephrol 2021;32:2517–28.
[112] Morawietz L et al. Differential gene expression in the wear particle induced and infectious periprosthetic membrane of loosened knee-endoprostheses. Verh Dtsch Ges Pathol 2003;87:204–14.
[113] Li Z et al. The superiority of immune-inflammation summary index for diagnosing periprosthetic joint infection. Int Immunopharmacol 2023;118:110073.
[114] Kuo F-C, Hu W-H, Hu Y-J. Periprosthetic joint infection prediction via machine learning: comprehensible personalized decision support for diagnosis. J Arthroplasty 2022;37:132–41.
[115] Meinshausen A-K et al. C9 immunostaining as a tissue biomarker for periprosthetic joint infection diagnosis. Front Immunol 2023;14:1112188.
[116] Valtetsiotis K et al. The potential role of gut bacteriome dysbiosis as a leading cause of periprosthetic infection: a comprehensive literature review. Microorganisms 2023;11:1778.
[117] Zhang Y-W et al. Fecal microbiota transplantation ameliorates bone loss in mice with ovariectomy-induced osteoporosis via modulating gut microbiota and metabolic function. J Orthop Translat 2022;37:46–60.
[118] Zhu H, Jin H, Zhang C, Yuan T. Intestinal methicillin-resistant Staphylococcus aureus causes prosthetic infection via 'Trojan Horse' mechanism: evidence from a rat model. Bone Joint Res 2020;9:152–61.
[119] Alverdy JC, Hyman N, Gilbert J. Re-examining causes of surgical site infections following elective surgery in the era of asepsis. Lancet Infect Dis 2020;20:e38–43.
[120] Chisari E, Cho J, Wouthuyzen-Bakker M, Parvizi J. Periprosthetic joint infection and the trojan horse theory: examining the role of gut dysbiosis and epithelial integrity. J Arthroplasty 2022;37:1369–74.
[121] Thwaites GE, Gant V. Are bloodstream leukocytes Trojan Horses for the metastasis of Staphylococcus aureus? Nat Rev Microbiol 2011;9:215–22.
[122] Muraille E, Leo O, Moser M. TH1/TH2 paradigm extended: macrophage polarization as an unappreciated pathogen-driven escape mechanism? Front Immunol 2014;5:603.
[123] Mann M, Kumar C, Zeng W-F, Strauss MT. Artificial intelligence for proteomics and biomarker discovery. Cell Syst 2021;12:759–70.
[124] Xu P, Tao Z, Zhang C. Integrated multi-omics and artificial intelligence to explore new neutrophils clusters and potential biomarkers in sepsis with experimental validation. Front Immunol 2024;15:1377817.
[125] Li J et al. Large-scale proteome profiling identifies biomarkers associated with suspected neurosyphilis diagnosis. Adv Sci (Weinh) 2024;11:e2307744.
[126] Sigmund IK et al. Diagnostic accuracy of neutrophil counts in histopathological tissue analysis in periprosthetic joint infection using the ICM, IDSA, and EBJIS criteria. Bone Joint Res 2021;10:536–47.
[127] Kelly ME et al. The utility of frozen section histology in diagnosing periprosthetic joint infection in revision total joint arthroplasty. J Arthroplasty 2021;36:2137–43.
[128] Niazi MKK, Parwani AV, Gurcan MN. Digital pathology and artificial intelligence. Lancet Oncol 2019;20:e253–61.
[129] Senaras C, Niazi MKK, Lozanski G, Gurcan MN. DeepFocus: Detection of out-of-focus regions in whole slide digital images using deep learning. PLoS One 2018;13:e0205387.
[130] Zanjani F et al. Stain normalization of histopathology images using generative adversarial networks. Eindhoven University of Technology 573–577 (2018).
[131] Jin L et al. DeepDOF-SE: affordable deep-learning microscopy platform for slide-free histology. Nat Commun 2024;15:2935.
[132] Rönnau MM, Lepper TW, Amaral LN, Rados PV, Oliveira MM. A CNN-based approach for joint segmentation and quantification of nuclei and NORs in AgNOR-stained images. Comput Methods Programs Biomed 2023;242:107788.
[133] Bouteldja N et al. Deep learning-based segmentation and quantification in experimental kidney histopathology. J Am Soc Nephrol 2021;32:52–68.
[134] Dave P et al. An adaptive digital stain separation method for deep learning-based automatic cell profile counts. J Neurosci Methods 2021;354:109102.
[135] Dy A et al. AI improves accuracy, agreement and efficiency of pathologists for Ki67 assessments in breast cancer. Sci Rep 2024;14:1283.
[136] Tao Y et al. A preliminary study on the application of deep learning methods based on convolutional network to the pathological diagnosis of PJI. Arthroplasty 2022;4:49.
[137] Li Z et al. Identification of immune infiltration and immune-related biomarkers of periprosthetic joint infection. Heliyon 2024;10:e26062.
[138] Tao Y et al. Clinically applicable optimized periprosthetic joint infection diagnosis via AI based pathology. NPJ Digit Med 2024;7:303.
[139] Lu MY et al. A multimodal generative AI copilot for human pathology. Nature 2024. doi: https://doi.org/10.1038/s41586-024-07618-3.
[140] Dolezal JM et al. Slideflow: deep learning for digital histopathology with real-time whole-slide visualization. BMC Bioinf 2024;25:134.
[141] Jardon M, Fritz J, Samim M. Imaging approach to prosthetic joint infection. Skeletal Radiol 2023. doi: https://doi.org/10.1007/s00256-023-04546-7.
[142] Sacher SE, Koff MF, Tan ET, Burge A, Potter HG. The role of advanced metal artifact reduction MRI in the diagnosis of periprosthetic joint infection. Skeletal Radiol 2023. doi: https://doi.org/10.1007/s00256-023-04483-5.
[143] Shi Z, Wang N, Kong F, Cao H, Cao Q. A semi-supervised learning method of latent features based on convolutional neural networks for CT metal artifact reduction. Med Phys 2022;49:3845–59.
[144] Li R, Yang F, Liu X, Shi HHGT. A hierarchical GCN-based transformer for multimodal periprosthetic joint infection diagnosis using computed tomography images and text. Sensors (Basel) 2023;23:5795.
[145] Albano D et al. MRI-based artificial intelligence to predict infection following total hip arthroplasty failure. Radiol Med 2023;128:340–6.

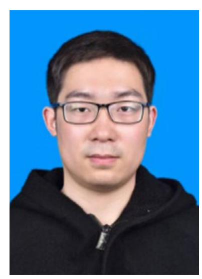
[146] Verberne SJ, Raijmakers PG, Temmerman OPP. The accuracy of imaging techniques in the assessment of periprosthetic hip infection: a systematic review and meta-analysis. J Bone Joint Surg Am 2016;98:1638–45.
[147] Kumar R, Kumar R, Kumar V, Malhotra R. Comparative analysis of dual-phase 18F-fluoride PET/CT and three phase bone scintigraphy in the evaluation of septic (or painful) hip prostheses: a prospective study. J Orthop Sci 2016;21:205–10.
[148] Zhang F et al. The temporal impact of prosthesis implantation and semi-quantitative criteria on the diagnostic efficacy of triple-phase bone scanning for periprosthetic joint infection. Orthop Surg 2022;14:1438–46.
[149] Nie L et al. An artificial intelligence framework for the diagnosis of prosthetic joint infection based on 99mTc-MDP dynamic bone scintigraphy. Eur Radiol 2023;33:6794–803.
[150] Hofmann UK, Eleftherakis G, Migliorini F, Fink B, Mederake M. Diagnostic and prognostic relevance of plain radiographs for periprosthetic joint infections of the hip: a literature review. Eur J Med Res 2024;29:314.
[151] Palestro CJ. Molecular imaging of periprosthetic joint infections. Semin Nucl Med 2023;53:167–74.
[152] Jin H et al. Diagnostic performance of FDG PET or PET/CT in prosthetic infection after arthroplasty: a meta-analysis. Q J Nucl Med Mol Imaging 2014;58:85–93.
[153] Yama N, Nagoya S, Sugita S, Koyama M, Hatakenaka M. Diagnosis of prosthetic joint infection at the hip using the standard uptake value of three-phase 99mTc-hydroxymethylene diphosphonate SPECT/CT. Ann Nucl Med 2022;36:634–42.
[154] Signore A et al. Consensus document for the diagnosis of prosthetic joint infections: a joint paper by the EANM, EBJIS, and ESR (with ESCMID endorsement). Eur J Nucl Med Mol Imaging 2019;46:971–88.
[155] Romanò CL et al. The role of imaging techniques to define a peri-prosthetic hip and knee joint infection: multidisciplinary consensus statements. J Clin Med 2020;9:2548.
[156] Palan J et al. Culture-negative periprosthetic joint infections. EFORT Open Rev 2019;4:585–94.
[157] Kolenda C et al. Experience with the use of the MicroDTTect device for the diagnosis of low-grade chronic prosthetic joint infections in a routine setting. Front Med (Lausanne) 2021;8:565555.
[158] Kuo F-C et al. Rapid diagnosis of periprosthetic joint infection from synovial fluid in blood culture bottles by direct matrix-assisted laser desorption ionization time-of-flight mass spectrometry. PLoS One 2020;15:e0239290.
[159] Indelli PF, Ghirardelli S, Violante B, Amanatullah DF. Next generation sequencing for pathogen detection in periprosthetic joint infections. EFORT Open Rev 2021;6:236–44.
[160] Roux A-L et al. Diagnosis of prosthetic joint infection by beadmill processing of a periprosthetic specimen. Clin Microbiol Infect 2011;17:447–50.
[161] Hameister R, Lim CT, Lohmann CH, Wang W, Singh G. What is the role of diagnostic and therapeutic sonication in periprosthetic joint infections? J Arthroplasty 2018;33:2575–81.
[162] Hashizume T, Ying B-W. Challenges in developing cell culture media using machine learning. Biotechnol Adv 2024;70:108293.
[163] Aghajani MH, Aghagolzadeh haji2 H, Mir, M, Aghajani M. Optimization of medium components using artificial neural networks. Int J Health Stud 2017;3.
[164] Chiu M et al. Comparison of the APAS independence automated plate reader system with the manual standard of care for processing urine culture specimens. Microbiol Spectr 2022;10:e0144222.
[165] Yu M et al. Simple and rapid discrimination of methicillin-resistant staphylococcus aureus based on gram staining and machine vision. Microbiol Spectr 2023;11:e0528222.
[166] Signoroni A et al. Hierarchical AI enables global interpretation of culture plates in the era of digital microbiology. Nat Commun 2023;14:6874.
[167] Wang C-X et al. Preliminary assessment of nanopore-based metagenomic sequencing for the diagnosis of prosthetic joint infection. Int J Infect Dis 2020;97:54–9.
[168] Schmidt B, Hildebrandt A. Deep learning in next-generation sequencing. Drug Discov Today 2021;26:173–80.
[169] Pandey D, Onkara Perumal P. A scoping review on deep learning for next-generation RNA-Seq. data analysis. Funct Integr Genomics 2023;23:134.
[170] Ye SH, Siddle KJ, Park DJ, Sabeti PC. Benchmarking metagenomics tools for taxonomic classification. Cell 2019;178:779–94.
[171] Randhawa GS, Hill KA, Kari L. ML-DSP: machine learning with digital signal processing for ultrafast, accurate, and scalable genome classification at all taxonomic levels. BMC Genomics 2019;20:267.
[172] Le Vavasseur B, Zeller V. Antibiotic therapy for prosthetic joint infections: an overview. Antibiotics (Basel) 2022;11:486.
[173] Hao L et al. Direct detection and identification of periprosthetic joint infection pathogens by metagenomic next-generation sequencing. Sci Rep 2023;13:7897.
[174] Gerada A, Harper N, Howard A, Reza N, Hope W. Determination of minimum inhibitory concentrations using machine-learning-assisted agar dilution. Microbiol Spectr 2024;12:e0420923.
[175] Lüftinger L et al. Predictive antibiotic susceptibility testing by next-generation sequencing for periprosthetic joint infections: potential and limitations. Biomedicines 2021;9:910.
[176] Boolchandani M, D'Souza AW, Dantas G. Sequencing-based methods and resources to study antimicrobial resistance. Nat Rev Genet 2019;20:356–70.
[177] Liu Z et al. Evaluation of machine learning models for predicting antimicrobial resistance of actinobacillus pleuropneumoniae from whole genome sequences. Front Microbiol 2020;11:48.
[178] Ren Y et al. Prediction of antimicrobial resistance based on whole-genome sequencing and machine learning. Bioinformatics 2022;38:325–34.
[179] Shi J et al. Antimicrobial resistance genetic factor identification from whole-genome sequence data using deep feature selection. BMC Bioinf 2019;20:535.
[180] Dierig A, Frei R, Egli A. The fast route to microbe identification: matrix assisted laser desorption/ionization-time of flight mass spectrometry (MALDI-TOF MS). Pediatr Infect Dis J 2015;34:97–9.
[181] Lee K, Kim D-W, Cha C-J. Overview of bioinformatic methods for analysis of antibiotic resistome from genome and metagenome data. J Microbiol 2021;59:270–80.
[182] Weis C et al. Direct antimicrobial resistance prediction from clinical MALDI-TOF mass spectra using machine learning. Nat Med 2022;28:164–74.
[183] Yu T, Fu Y, He J, Zhang J, Xianyu Y. Identification of Antibiotic resistance in ESKAPE pathogens through Plasmonic Nanosensors and machine learning. ACS Nano 2023;17:4551–63.
[184] De Oliveira DMP et al. Antimicrobial resistance in ESKAPE pathogens. Clin Microbiol Rev 2020;33:e00181–219.
[185] Liu G-Y et al. Antimicrobial resistance crisis: could artificial intelligence be the solution? Mil Med Res 2024;11:7.
[186] Antimicrobial Resistance Collaborators. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. Lancet 2022;399:629–55.
[187] Cox G et al. A common platform for antibiotic dereplication and adjuvant discovery. Cell Chem Biol 2017;24:98–109.
[188] Tommasi R, Brown DG, Walkup GK, Manchester JI, Miller AA. ESKAPEing the labyrinth of antibacterial discovery. Nat Rev Drug Discov 2015;14:529–42.
[189] Sarkar C et al. Artificial intelligence and machine learning technology driven modern drug discovery and development. Int J Mol Sci 2023;24:2026.
[190] Paul D et al. Artificial intelligence in drug discovery and development. Drug Discov Today 2021;26:80–93.
[191] Stokes JM et al. A deep learning approach to antibiotic discovery. Cell 2020;180:688–702.e13.
[192] Liu G et al. Deep learning-guided discovery of an antibiotic targeting Acinetobacter baumannii. Nat Chem Biol 2023;19:1342–50.
[193] Swanson K et al. Generative AI for designing and validating easily synthesizable and structurally novel antibiotics. Nat Mach Intell 2024;6:338–53.
[194] Thompson JM et al. Oral-only linezolid-Rifampin is highly effective compared with other antibiotics for periprosthetic joint infection: study of a mouse model. J Bone Joint Surg Am 2017;99:656–65.
[195] Fan K, Cheng L, Li L. Artificial intelligence and machine learning methods in predicting anti-cancer drug combination effects. Brief Bioinform 2021;22: bbab271.
[196] Tsigelny IF. Artificial intelligence in drug combination therapy. Brief Bioinform 2019;20:1434–48.
[197] Velasquez Garcia A et al. Artificial intelligence-based three-dimensional templating for total joint arthroplasty planning: a scoping review. Int Orthop 2024;48:997–1010.
[198] Wilson NA, Jehn M, York S, Davis CM. Revision total hip and knee arthroplasty implant identification: implications for use of Unique Device Identification 2012 AAHKS member survey results. J Arthroplasty 2014;29:251–5.
[199] Borjali A et al. Comparing the performance of a deep convolutional neural network with orthopedic surgeons on the identification of total hip prosthesis design from plain radiographs. Med Phys 2021;48:2327–36.
[200] Klemt C et al. The ability of deep learning models to identify total hip and knee arthroplasty implant design from plain radiographs. J Am Acad Orthop Surg 2022;30:409–15.
[201] Lambrechts A, Wirix-Speetjens R, Maes F, Van Huffel S. Artificial intelligence based patient-specific preoperative planning algorithm for total knee arthroplasty. Front Robot AI 2022;9:840282.
[202] Chen X et al. Development and validation of an artificial intelligence preoperative planning system for total hip arthroplasty. Front Med (Lausanne) 2022;9:841202.
[203] Li S et al. Development and validation of an artificial intelligence preoperative planning and patient-specific instrumentation system for total knee arthroplasty. Bioengineering (Basel) 2023;10:1417.
[204] Rodríguez-Merchán EC, Gómez-Cardero P, Encinas-Ullán CA. Management of bone loss in revision total knee arthroplasty: therapeutic options and results. EFORT Open Rev 2021;6:1073–86.
[205] Wu X-D et al. Robotic-assisted revision total joint arthroplasty: a state-of-the-art scoping review. EFORT Open Rev 2023;8:18–25.
[206] Arshi A et al. Navigated and robot-assisted technology in total knee arthroplasty: do outcome differences achieve minimal clinically important difference? J Arthroplasty 2022;37:1562–9.
[207] Zhang Y-W et al. Efficacy evaluation of three-dimensional printing assisted osteotomy guide plate in accurate osteotomy of adolescent cubitus varus deformity. J Orthop Surg Res 2019;14:353.
[208] Rodrigues P et al. Deep segmentation leverages geometric pose estimation in computer-aided total knee arthroplasty. Healthc Technol Lett 2019;6:226–30.
[209] Raposo C, Barreto JP. Using 2 point+normal sets for fast registration of point clouds with small overlap. In: 2017 IEEE International Conference on

Robotics and Automation (ICRA) 5652–5658 (IEEE, Singapore, Singapore, 2017). doi:10.1109/ICRA.2017.7989664.
[210] Chen JY et al. Less outliers in pinless navigation compared with conventional surgery in total knee arthroplasty. Knee Surg Sports Traumatol Arthrosc 2014;22:1827–32.
[211] Chen JY et al. Radiological outcomes of pinless navigation in total knee arthroplasty: a randomized controlled trial. Knee Surg Sports Traumatol Arthrosc 2015;23:3556–62.
[212] Breda A, Territo A. Virtual reality simulators for robot-assisted surgery. Eur Urol 2016;69:1081–2.
[213] Subramanian P, Wainwright TW, Bahadori S, Middleton RG. A review of the evolution of robotic-assisted total hip arthroplasty. Hip Int 2019;29:232–8.
[214] Batailler C, Swan J, Sappey Marinier E, Servien E, Lustig S. New technologies in knee arthroplasty: current concepts. J Clin Med 2020;10:47.
[215] Sugano N et al. Accuracy of femoral component anteversion in robotic total hip arthroplasty. Bone Joint J 2024;106-B:104–9.
[216] Guo D-H et al. Total hip arthroplasty with robotic arm assistance for precise cup positioning: a case-control study. Orthop Surg 2022;14:1498–505.
[217] Ramkumar PN et al. Remote patient monitoring using mobile health for total knee arthroplasty: validation of a wearable and machine learning-based surveillance platform. J Arthroplasty 2019;34:2253–9.
[218] Li LT, Haley LC, Boyd AK, Bernstam EV. Technical/Algorithm, Stakeholder, and Society (TASS) barriers to the application of artificial intelligence in medicine: a systematic review. J Biomed Inform 2023;147:104531.
[219] Diprose WK et al. Physician understanding, explainability, and trust in a hypothetical machine learning risk calculator. J Am Med Inform Assoc 2020;27:592–600.
[220] Mehta N, Pandit A. Concurrence of big data analytics and healthcare: a systematic review. Int J Med Inf 2018;114:57–65.
[221] S, S. et al. Physician perspectives on integration of artificial intelligence into diagnostic pathology. NPJ digital medicine 2, (2019).
[222] Yokoi R, Eguchi Y, Fujita T, Nakayachi K. Artificial intelligence is trusted less than a doctor in medical treatment decisions: influence of perceived care and value similarity. Int. J. Human–Computer Interact. 2021;37:981–90.
[223] Xiao X, Xue L, Ye L, Li H, He Y. Health care cost and benefits of artificial intelligence-assisted population-based glaucoma screening for the elderly in remote areas of China: a cost-offset analysis. BMC Public Health 2021;21:1065.
[224] K D et al. Public vs physician views of liability for artificial intelligence in health care. JAMIA 2021;28(7):1574–7. doi: https://doi.org/10.1093/jamia/ocab055.
[225] Mc L, MB. Mf, M. Perceptions of artificial intelligence in healthcare: findings from a qualitative survey study among actors in France. J. Translat. Med. 18 (2020).
[226] Murdoch B. Privacy and artificial intelligence: challenges for protecting health information in a new era. BMC Med Ethics 2021;22:122.
[227] Ji S. et al. De-health: all your online health information are belong to us. In: 2020 IEEE 36th International Conference on Data Engineering (ICDE) 1609–1620 (IEEE, Dallas, TX, USA, 2020). doi:10.1109/ICDE48307.2020.00143.
[228] Na L et al. Feasibility of reidentifying individuals in large national physical activity data sets from which protected health information has been removed with use of machine learning. JAMA Netw Open 2018;1:e186040.
[229] De Micco F et al. Robotics and AI into healthcare from the perspective of European regulation: who is responsible for medical malpractice? Front Med (Lausanne) 2024;11:1428504.
[230] Cestonaro C, Delicati A, Marcante B, Caenazzo L, Tozzo P. Defining medical liability when artificial intelligence is applied on diagnostic algorithms: a systematic review. Front Med (Lausanne) 2023;10:1305756.
[231] Y, N. et al. An ethics assessment tool for artificial intelligence implementation in healthcare: CARE-AI. Nat Med 30, (2024).
[232] Samaga D et al. Single-center versus multi-center data sets for molecular prognostic modeling: a simulation study. Radiat Oncol 2020;15:109.
[233] Ho SY, Phua K, Wong L, Bin Goh WW. Extensions of the external validation for checking learned model interpretability and generalizability. Patterns (N Y) 2020;1:100129.
[234] Getz K, Hubbard RA, Linn KA. Performance of multiple imputation using modern machine learning methods in electronic health records data. Epidemiology 2023;34:206–15.
[235] Pridham G, Rockwood K, Rutenberg A. Strategies for handling missing data that improve Frailty Index estimation and predictive power: lessons from the NHANES dataset. Geroscience 2022;44:897–923.
[236] Rosenbacke R, Melhus Å, McKee M, Stuckler D. How explainable artificial intelligence can increase or decrease clinicians' trust in AI applications in health care: systematic review. JMIR AI 2024;3:e53207.
[237] Jr Z et al. Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: a cross-sectional study. PLoS Med 2018;15.



Mr. **Pengcheng Li** is currently pursuing a master's degree at the Academy of Military Medical Sciences of the People's Liberation Army of China. He is currently researching the application of artificial intelligence in the diagnosis and treatment of knee joint diseases. In addition, he is also paying attention to the latest developments in knee and hip joint replacement surgery.

Dr. **Yan Wang** , Chief Physician, Professor, and Doctoral Supervisor in the Department of Orthopedics at the General Hospital of the People's Liberation Army of China, serves as the President of the Orthopedic Physicians Branch of the Chinese Medical Doctor Association, the Director of the Orthopedic Professional Committee of the entire military, the President of the Asia-Pacific Arthroplasty Society (APAS), the Deputy Editor-in-Chief of "Spine", and the Co-Editor-in-Chief of "Journal of Arthroplasty". He has published 578 papers and edited or translated 30 monographs (textbooks). He also co-edited the spinal surgery book "Spinal Osteotomy" with Professor Laurence Lenke and Professor Oheneba Boachie-Adjei, achieving outstanding accomplishments in the fields of joint surgery and spinal surgery.

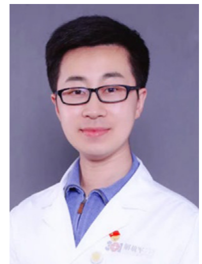


Mr. **Runkai Zhao** is currently pursuing a Master's degree at the PLA MedicalCollege. He has co-authored 8 research papers, 4 of which as the first authoror co-first author. Her current research focuses on osteoarthritis and hip and knee joint replacement.

Mr. **Lin Hao** is currently pursuing a master's degree at the Academy of Military Medical Sciences of the People's Liberation Army of China. he is currently researching the molecular mechanisms related to the pathogenesis of knee osteoarthritis and is also interested in the field of knee joint replacement surgery.

Dr. **Wei Chai**, Director of the Joint Surgery Department at the Department of Orthopedics, General Hospital of the People's Liberation Army of China, Chief Physician, and Doctoral Supervisor. He has published a total of 79 papers, been granted 4 international patents, and edited or translated 8 monographs. His current main research directions include complex hip and knee joint replacement surgeries for conditions such as developmental dysplasia of the hip, revision treatments for prosthetic loosening and infection after surgery, minimally invasive joint replacement and rapid recovery, and the assistance of joint replacement with precision medical technologies such as robotics and artificial intelligence.

Chen Jiying Chief Physician of the Orthopedics Department at the General Hospital of the People's Liberation Army of China, and Doctoral Supervisor. He has published 72 SCI papers as the first or corresponding author, with the highest impact factor of 18.2; he has edited or translated 8 monographs. He has extensive experience and unique insights in complex hip and knee joint replacement and revision surgeries, as well as related joint diseases such as prosthetic joint infections and loosening.



Dr. **Quanbo Ji,** Deputy Chief Physician of the Department of Orthopedics, Medical Division of the General Hospital of the People's Liberation Army of China, and Master's Program Supervisor, is a young editorial board member of Arthroplasty at the Asian Association of Arthroplasty. He has published 23 papers as the first or corresponding author, and has translated and co-edited the publication of 6 monographs. His current main research direction is the optimization of the pathogenesis and treatment strategies of osteoarthritis.

Mr. **Zeyu Feng** is currently pursuing a master's degree at the Academy of Military Medical Sciences of the People's Liberation Army of China. He is currently researching how to rapidly and accurately detect the pathogens of prosthetic joint infections, as well as how to effectively clear the infections. In addition, he is also paying attention to the latest developments in knee and hip joint replacement surgery.

Dr. **Guoqiang Zhang**, Deputy Director of the Joint Surgery Department at the General Hospital of the People's Liberation Army of China, is a Chief Physician/Professor/PhD Supervisor. He has published a total of 75 papers, edited (translated) 6 monographs, and been granted 4 international invention patents. He has extensive clinical and research experience in the field of joint surgery. His current main research directions include the application of artificial intelligence in joint surgery, rapid detection of prosthetic joint infections, related pathogenesis and treatment strategies, and new strategies for the diagnosis and treatment of knee and hip joint diseases.