


# A machine learning-based model for “In-time” prediction of periprosthetic joint infection

DIGITAL HEALTH  
Volume 10: 1–15  
© The Author(s) 2024  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/20552076241253531  
journals.sagepub.com/home/dhj



Weishen Chen<sup>1,2,\*</sup>, Xuantao Hu<sup>1,2,\*</sup>, Chen Gu<sup>3,\*</sup>, Zhaohui Zhang<sup>4</sup>,  
Linli Zheng<sup>1,2</sup>, Baiqi Pan<sup>1,2</sup>, Xiaoyu Wu<sup>1,2</sup>, Wei Sun<sup>3</sup> and Puyi Sheng<sup>1,2</sup> 

## Abstract

**Background:** Previous criteria had limited value in early diagnosis of periprosthetic joint infection (PJI). Here, we constructed a novel machine learning (ML)-derived, “in-time” diagnostic system for PJI and proved its validity.

**Methods:** We filtered “in-time” diagnostic indicators reported in the literature based on our continuous retrospective cohort of PJI and aseptic prosthetic loosening patients. With the indicators, we developed a two-level ML model with six base learners including Elastic Net, Linear Support Vector Machine, Kernel Support Vector Machine, Extra Trees, Light Gradient Boosting Machine and Multilayer Perceptron), and one meta-learner, Ensemble Learning of Weighted Voting. The prediction performance of this model was compared with those of previous diagnostic criteria (International Consensus Meeting in 2018 (ICM 2018), etc.). Another prospective cohort was used for internal validation. Based on our ML model, a user-friendly web tool was developed for swift PJI diagnosis in clinical practice.

**Results:** A total of 254 patients (199 for development and 55 for validation cohort) were included in this study with 38.2% of them diagnosed as PJI. We included 21 widely accessible features including imaging indicators (X-ray and CT) in the model. The sensitivity and accuracy of our ML model were significantly higher than ICM 2018 in development cohort (90.6% vs. 76.1%,  $P=0.032$ ; 94.5% vs. 86.7%,  $P=0.020$ ), which was supported by internal validation cohort (84.2% vs. 78.6%; 94.6% vs. 81.8%).

**Conclusions:** Our novel ML-derived PJI “in-time” diagnostic system demonstrated significantly improved diagnostic potency for surgical decision-making compared with the commonly used criteria. Moreover, our web-based tool greatly assisted surgeons in distinguishing PJI patients comprehensively.

**Level of evidence:** Diagnostic Level III.

## Keywords

periprosthetic joint infection, machine learning, diagnosis, aseptic loosening of prosthesis, web tool, decision-making

Submission date: 17 August 2023; Acceptance date: 22 April 2024

<sup>1</sup>Department of Joint Surgery, The First Affiliated Hospital of Sun Yat-Sen University, Guangzhou, China

<sup>2</sup>Guangdong Provincial Key Laboratory of Orthopaedics and Traumatology, The First Affiliated Hospital of Sun Yat-sen University, Guangzhou, China

<sup>3</sup>School of Electronics and Information Technology, Sun Yat-sen University, Guangzhou, China

<sup>4</sup>Department of Diagnostic Radiology, The First Affiliated Hospital of Sun Yat-Sen University, Guangzhou, China

\*Weishen Chen, Xuantao Hu and Chen Gu contributed equally to this work

## Corresponding authors:

Wei Sun, School of Electronics and Information Technology, Sun Yat-sen University, No.132, Waihuan East Road, Guangzhou Higher Education Mega Center, Guangzhou 510006, Guangdong Province, China.

Email: sunwei@mail.sysu.edu.cn

Puyi Sheng, Department of Joint Surgery, The First Affiliated Hospital of Sun Yat-Sen University, Zhongshan 2nd Road, Guangzhou 510080, Guangdong Province, China.

Email: shengpy@mail.sysu.edu.cn



## Introduction

Periprosthetic joint infection (PJI) is one of the major causes of the failure of total joint arthroplasty (TJA). It demonstrates incidence only second to aseptic prosthetic loosening (APL) but often suggests a more catastrophic prognosis.<sup>1</sup> Unfortunately, it is difficult to make a differential diagnosis between PJI and APL for the atypical clinical signs, especially in the comparison of APL and low-grade PJI.<sup>2</sup>

The true “gold standard” for PJI diagnosis remains elusive.<sup>3</sup> In recent years, organizations including 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) gained significant progress in the diagnosis of PJI by developing diagnostic criteria, like MSIS 2011 and EBJIS 2021, ICM 2018 definition.<sup>4–8</sup> In terms of providing an in-time reference for surgical decision-making, the performance of these definitions was impaired due to the lack of post-operative results. Although EBJIS 2021 demonstrate superior pre-operative diagnostic performance compared with ICM 2018, there is still about 15% patients was misdiagnosed or diagnosed with intermediate result.<sup>9</sup>

Machine learning (ML) is a branch subject of artificial intelligence that provides data analysis methods to automatically optimize the prediction performance of algorithm-based model.<sup>1</sup> It has brought novel solutions to issues in orthopedics, from accurate diagnosis to prognosis prediction.<sup>10,11</sup>

Herein, we attempted to develop an ML-derived diagnostic system using widely accessible indicators obtained before surgical prosthesis management (“in-time”) for diagnosis of PJI and then make a comparison in potency among the present system and commonly used ones, including MSIS 2011, ICM 2018, and EBJIS 2021 definitions). The authors assumed that this ML-derived system could provide “in-time” diagnosis for PJI patients with favorable accuracy.

## Patients and methods

### Patient data

This study was designed as a retrospective cohort study and the data for analysis was de-identified, therefore exempted from the Institutional Review Board. We performed a retrospective review of 668 post-TJA complication cases from the clinical database of the institution between 1995 and 2021. All patients provided written informed consent for using their clinical data in the current study. Patients underwent arthroplasty for bone cancer, megaprosthesis implantation, simple prosthetic wear, joint dislocation, or malalignment, and patients with current other inflammatory arthropathies (like rheumatoid arthritis, crystal arthropathy, evident metallosis, etc.) were excluded from the cohort.

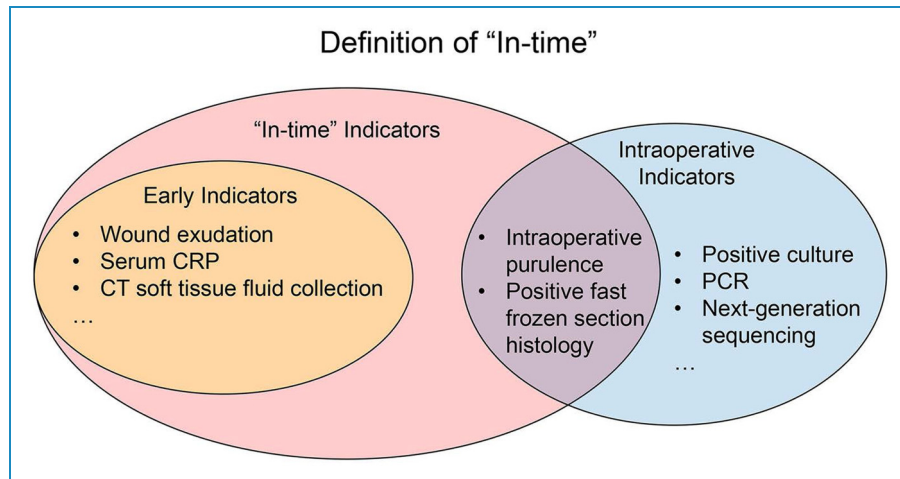
Then, we filtered a set of clinical indicators that can be obtained before surgical management of prosthesis (defined as “in-time” indicators) (Figure 1) according to literature about PJI risk or diagnosis.<sup>4–7,10,12–19</sup> Patients who lack of indicators required by MSIS 2011, ICM 2018, or EBJIS 2021 definition were excluded. Indicators with a missing rate over 25% were also discarded, such as joint function score, MRI, three-phase isotope bone scan, synovial fluid tests, alpha-defensin, etc. (Figure 2).<sup>20</sup> Then, Multiple Imputation by Chained Equations was used to treat missing data. Here, statistical information of the indicators is summarized in Table 1. Finally, a cohort of 199 patients (PJI : APL = 78 : 121) with 46 clinical indicators was obtained (Cohort I).

We reviewed cases between 2021 and 2022 and obtained data from another cohort with 55 patients (Cohort II) via an identical workflow with the Cohort I for validation of the model (Figure 2).

Patients were classified into infected and aseptic groups. Due to the lack of a gold standard for PJI, PJI cases were identified based on the criteria of MSIS 2011 in the cohort firstly.<sup>10</sup> A PJI case was confirmed if at least 1 of the major criteria existed, or the following four minor criteria were all met: (1) elevated serum ESR ( $> 28$  mm/h)/CRP ( $> 10$  mg/L), (2) purulence material found in the joint cavity, (3) positive organism culture of one tissue/fluid sample, or (4)  $> 5$  neutrophils per high-power field in 5 high-power field observed in fast frozen section histopathologic analysis. Then, the rest of the patients who met the following criteria were also categorized as PJI (more likely in low-grade): (1) exist  $<$  four of the minor criteria of MSIS 2011 at the initial visit; (2) have been treated with antibiotics but still develop following-up proved readmission as a result of PJI within 6 months after initial visit; and (3) absence of any risk factor of PJI within the 6 months, like invasive operation, bacteremia signs, etc.<sup>21</sup> APL cases were defined as patients who underwent revision with loosening component discovered in surgery, and who did not develop infection on the same joint within 1 year. For cases remaining uncertain after the identification of the above procedures, the final diagnosis was given according to the decision from at least three orthopedic surgeons (at least attending doctor) with over 5 years of working experience.

### Derivation of periprosthetic joint infection diagnostic model

Diagnosis of PJI was treated as a binary classification problem using inductive learning approaches. To increase the robustness and performance of the predictive model, we conceived a two-level ensemble learning architecture with one meta-learner to integrate several base learners. We included eight different ML algorithms in the list of alternative base learners, including two linear models:



**Figure 1.** The concept of “in-time” indicators of PJI demonstrated by Venn diagram. PJI: periprosthetic joint infection.

Elastic Net, Linear Support Vector Machine (LSVM), and six nonlinear models: Kernel Support Vector Machine (KSVM), Decision Tree, Random Forest, Extra Trees (ET), Light Gradient Boosting Machine (LGBM), Multilayer Perceptron (MLP). To evaluate the predictive performance of each classifier (chosen learner for model construction), a stratified fivefold cross-validation method based on data of Cohort I was introduced. The whole train-test process ran iteratively until all the subsets were used for testing (Figure 3). The performance of each model was characterized by the average of 7 metrics produced in each run, including accuracy, Matthews correlation coefficient (MCC), precision, recall (sensitivity), F1 score, specificity, and AUC. Ensemble learning of weighted voting (ELWV) was selected as the top-level meta-classifier.

The importance ranking of the 46 features (clinical indicators included in ML) was estimated according to the weighted feature importance of base classifiers. Then, features demonstrating relatively low importance or high correlation with others were pruned.

The optimal combination of six base learners out of the eight for maximal performance was finally determined as voters for ELWV. A more detailed version of this part can be found in Supplementary Materials and Methods.

### Validation of model

The optimized diagnostic model was used to make PJI “in-time” diagnosis for 55 patients in Cohort II. The statistical information of the 21 indicators is summarized in Supplementary Table 1.

### Development of periprosthetic joint infection diagnostic web tool

We attempted to develop a web tool for PJI diagnosis and clinical information collection based on the established

ML model using a unified modeling language. Detailed technique information can be found in Supplementary Materials and Methods.

### Comparison of model performance and statistics

The predictive performance metrics of selected base learners, meta-learner, and commonly used definitions (including MSIS 2011, ICM 2018, and EBJIS 2021) of PJI were estimated likewise in Cohort I (Figure 3) and Cohort II. The performance of the final ML model and commonly used definitions were compared using paired *t*-test. The overall statistical level was determined as  $P < 0.05$ .

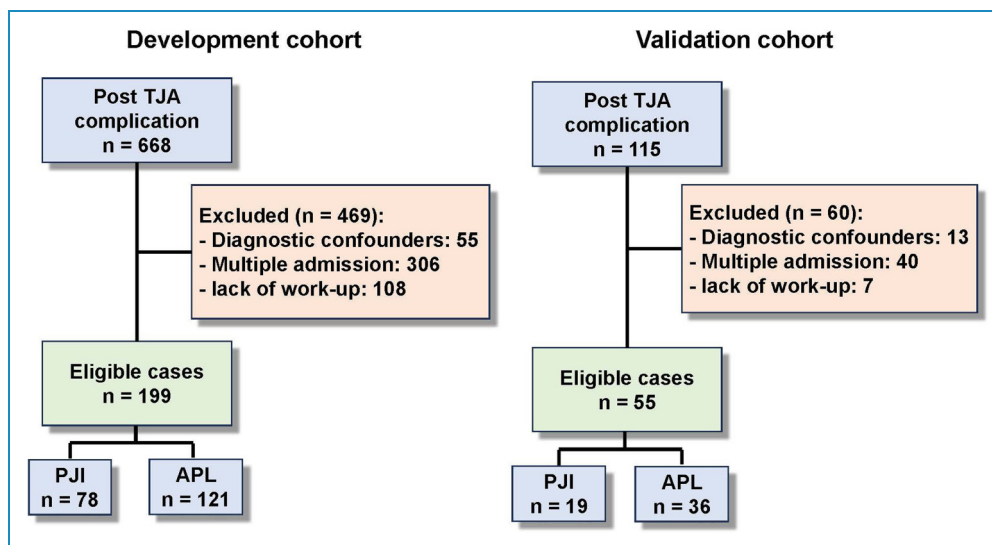
## Results

### Importance ranking of the features

The importance ranking of the 46 features in the first round of inductive learning is summarized in Supplementary Figure 1. Then, features with importance  $\leq 1\%$  were pruned from the list. Also, similar features with high correlation were discarded and the only one with the highest importance among them was kept. The importance ranking of the rest 21 features is demonstrated in Figure 4. As depicted in Figure 4, time to primary arthroplasty (months), hypoalbuminemia (yes/no) and CT soft tissue fluid collection (yes/no) are the top three important features apart from indicators included in commonly used definitions. Table 2 demonstrates the feature list adopted by the final model of ML in comparison to the counterparts in MSIS 2011, ICM 2018, and EBJIS 2021 definitions. After that, the base models were re-trained based on the dataset with pruned feature list.

### Result and comparison of diagnostic performance

In the second round of model training, Random Forest and Decision Tree were filtered and Elastic Net (EN), LSVM,



**Figure 2.** Flowchart of patient selection.

**Table 1.** Summary of patient clinical indicators in cohort I.

Features	PJI (n = 78)	APL (n = 121)	p-value	
General information				
Age (y)	64.6 (±12.9)	64.0(±11.0)	.763a	.918b
Age groupd (young/middle-aged/old)	Y (36%), M (45%), O (38%)	Y (64%), M (55%), O (62%)	.573b	
Gender (M/F)	M (37%), F (63%)	M (34%), F (66%)	.635c	
Han ethnic (yes/no)	Y (99%), N (1%)	Y (98%), N (2%)	.834c	
Medical insurance (yes/no)	Y (97%), N (3%)	Y (98%), N (2%)	.970c	
Joint (hip/knee)	H (54%), K (46%)	H (90%), K (10%)	< .001c	
Laterality (left/right)	L (53%), R (47%)	L (46%), R (54%)	.387 <sup>c</sup>	
Body temperature (°C)	36.6 (±0.4)	36.5 (±0.2)	.050a	.041 <sup>b</sup>
Height (cm)	158.3 (±8.4)	159.7 (±7.8)	.222a	.140b
BW (kg)	62.2 (±10.3)	62.5 (±12.0)	.866a	.098b
BMI (kg/m <sup>2</sup> )	24.9 (±3.8)	24.5 (±4.0)	.476a	.427b
BMI classification (underweight/normal/overweight/obesity)e	U (3%), N (46%), Ow (28%), Ob (23%)	U (7%), N (40%), Ow (34%), Ob (18)	.586b	
Clinical history				
Night pain (yes/no)	Y (1%), N (99%)	Y (1%), N (99%)	.760c	

(continued)

Table 1. Continued.

Features	PJI (n = 78)	APL (n = 121)	p-value
Increase of local skin temperature (yes/no)	Y (10%), N (90%)	Y (4%), N (96%)	.088c
Wound exudation (yes/no)	Y (46%), N (54%)	Y (2%), N (98%)	< .001c
LLD (yes/no)	Y (18%), N (82%)	Y (36%), N (64%)	.005c
Time to PA (month)	37.2 ( $\pm$ 54.3)	111.3 ( $\pm$ 68.0)	< .001a < .001b
Symptom duration (month)	9.3 ( $\pm$ 16.8)	18.1 ( $\pm$ 22.3)	.004a < .001b
Number of prior invasive proceduref	0.6 ( $\pm$ 0.9)	0.1 ( $\pm$ 0.5)	< .001a
Prior wound healing problems (yes/no)	Y (37%), N (63%)	Y (7%), N (93%)	< .001c
Recent fever or bacteremia (yes/no)	Y (15%), N (85%)	Y (0%), N (100%)	< .001c
Use of IA (yes/no)	Y (3%), N (97%)	Y (2%), N (98%)	.970c
Obesity (yes/no)	Y (23%), N (77%)	Y (18%), N (82%)	.400c
Diabetes (yes/no)	Y (19%), N (81%)	Y (11%), N (89%)	.093c
Anemia (yes/no)	Y (46%), N (54%)	Y (17%), N (83%)	< .001c
Hypoalbuminemia (yes/no)	Y (53%), N (47%)	Y (10%), N (90%)	< .001c
Hypertension (yes/no)	Y (92%), N (8%)	Y (30%), N (70%)	000c
Cardiac arrhythmia (yes/no)	Y (0%), N (100%)	Y (2%), N (98%)	.254c
CHD (yes/no)	Y (10%), N (90%)	Y (98%), N (2%)	.007c
Thrombosis (yes/no)	Y (0%), N (100%)	Y (1%), N (99%)	.421c
Mild liver disease (yes/no)	Y (1%), N (99%)	Y (0%), N (100%)	.212c
Moderate renal disease (yes/no)	Y (1%), N (99%)	Y (2%), N (98%)	.834c
Cancer (yes/no)	Y (0%), N (100%)	Y (1%), N (99%)	.421c
Other inflammatory arthropathy (yes/no)	Y (3%), N (97%)	Y (6%), N (94%)	.286c
Syphilis (yes/no)	Y (1%), N (99%)	Y (1%), N (99%)	.760c
Allergic history (yes/no)	Y (10%), N (90%)	Y (8%), N (92%)	.632c
Current smoking (yes/no)	Y (10%), N (90%)	Y (4%), N (96%)	.088c
Alcoholism (yes/no)	Y (3%), N (97%)	Y (6%), N (94%)	.286c
Charlson comorbidity index	2.4 ( $\pm$ 1.4)	2.1 ( $\pm$ 1.2)	.235 <sub>a</sub>
ASA	I (23%), II (60%), III (17%), IV (1%)	I (13%), II (77%), III (10%)	.043c

(continued)

Table 1. Continued.

Features	PJI (n = 78)	APL (n = 121)	p-value
Blood tests			
Hb (g/L)	113.0 (±19.8)	127.1 (±20.3)	<.001a <.001b
Anemia grading (normal/mild/moderate/severe/life-threatening) <sup>§</sup>	N (54%), Mi (36%), Mo (9%), LF (1%)	N (83%), Mi (14%), Mo (3%)	<.001b
WBC (10 <sup>9</sup> /L)	7.8 (±2.9)	6.6 (±2.1)	<.001a .002b
Neutrophil (10 <sup>9</sup> /L)	5.4 (±2.6)	3.9 (±9.5)	<.001a <.001b
Neutrophil segment (%)	67.0 (±9.8)	58.4 (±9.5)	<.001a <.001b
Lymphocyte (10 <sup>9</sup> /L)	1.7 (±0.7)	1.9 (±0.6)	.018a .003 <sub>b</sub>
NLR	3.9 (±2.8)	2.5 (±2.6)	<.001a <.001b
Platelet (10 <sup>9</sup> /L)	309.3 (±125.3)	230.6 (±74.5)	<.001a <.001b
Fibrinogen (g/L)	4.86 (±1.5)	3.3 (±1.1)	<.001a <.001b
IL-6 (pg/ml)	26.5 (±29.5)	8.8 (±17.8)	<.001a <.001b
PCT (ng/ml)	0.5 (±1.6)	0.1 (±0.5)	.018a <.001b
Albumin (g/L)	35.1 (±5.1)	39.1 (±3.44)	<.001a <.001b
Globulin (g/L)	35.4 (±6.6)	29.6 (±4.7)	<.001a <.001b
A/G	1.0 (±0.3)	1.4 (±0.2)	<.001a <.001b
FBS (mmol/L)	5.7 (±2.1)	5.1 (±1.1)	.012a .044b
Imaging examinations			
X-ray soft tissue swelling	Y (5%), N (95%)	Y (7%), N (93%)	.520c
X-ray bright line or osteolysis	Y (18%), N (82%)	Y (26%), N (74%)	.207c
CT soft tissue fluid collection	Y (24%), N (76%)	Y (22%), N (78%)	.738c
CT osteolysis	Y (24%), N (76%)	Y (27%), N (73%)	.648c
Classical-criteria-related indicators			
Sinus tract or prosthesis exposure (yes/no)	Y (37%), N (63%)	Y (0%), N (100%)	<.001c
Serum CRP (mg/L)	48.4 (±59.0)	11.4 (±29.1)	<.001a <.001b
Serum ESR (mm/h)	71.4 (±30.1)	33.0 (±28.2)	<.001a <.001b
Intraoperative purulence (yes/no)	Y (40%), N (60%)	Y (0%), N (100%)	<.001c
Positive fast frozen section histology (yes/no)	Y (42%), N (58%)	Y (1%), N (99%)	<.001c

(continued)

Table 1. Continued.

Features	PJI (n = 78)	APL (n = 121)	p-value
MSIS 2011 (infected/uncertain/aseptic)	I (35%), A (65%)	I (7%), A (93%)	< .001c
ICM 2018 (infected/uncertain/aseptic)	I (32%), U (13%), A (55%)	I (21%), U (17%), A (62%)	.178c
EBJIS 2020 (infected/uncertain/aseptic)	I (35%), U (18%), A (47%)	I (20%), U (25%), A (55%)	.060c

BW: body weight; PA: primary arthroplasty; IA: immunosuppressive agents; CHD: coronary atherosclerotic heart disease; LLD: lower limb discrepancy; NLR: neutrophil to lymphocyte ratio; PCT: procalcitonin; FBS: fasting blood sugar; A/G: albumin to globulin ratio; PJI: periprosthetic joint infection; APL: aseptic prosthetic loosening; MSIS: Musculoskeletal Infection Society; ICM: International Consensus Meeting; EBJIS: European Bone and Joint Infection Society. The distribution of each indicator was expressed in form of average  $\pm$  standard deviation or percentage.

<sup>a</sup>One-way analysis of variance.

<sup>b</sup>Mann-Whitney U test.

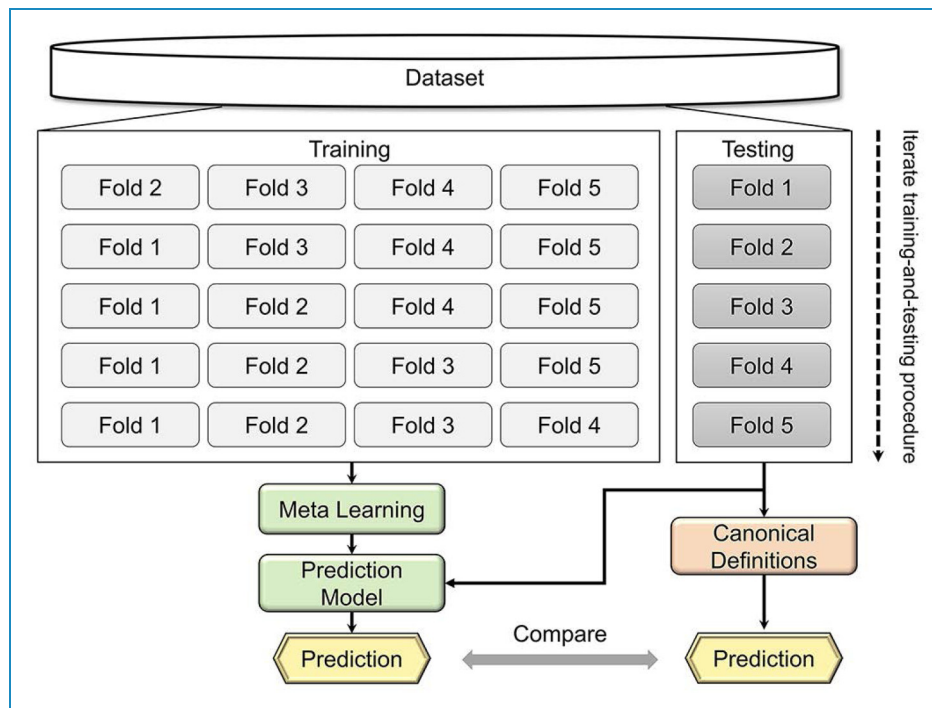
<sup>c</sup>Chi-squared test.

<sup>d</sup>WHO 2015 age group classification.

<sup>e</sup>WHO 2005 BMI classification.

<sup>f</sup>Primary arthroplasty is excluded.

<sup>g</sup>Anemia grading of National Cancer Institute.



**Figure 3.** Schematic diagram showing fivefold cross-validation of each algorithm. One fold was selected as the testing set and the remaining four training sets in an iteration. The same training-and-testing is iterated on each fold.

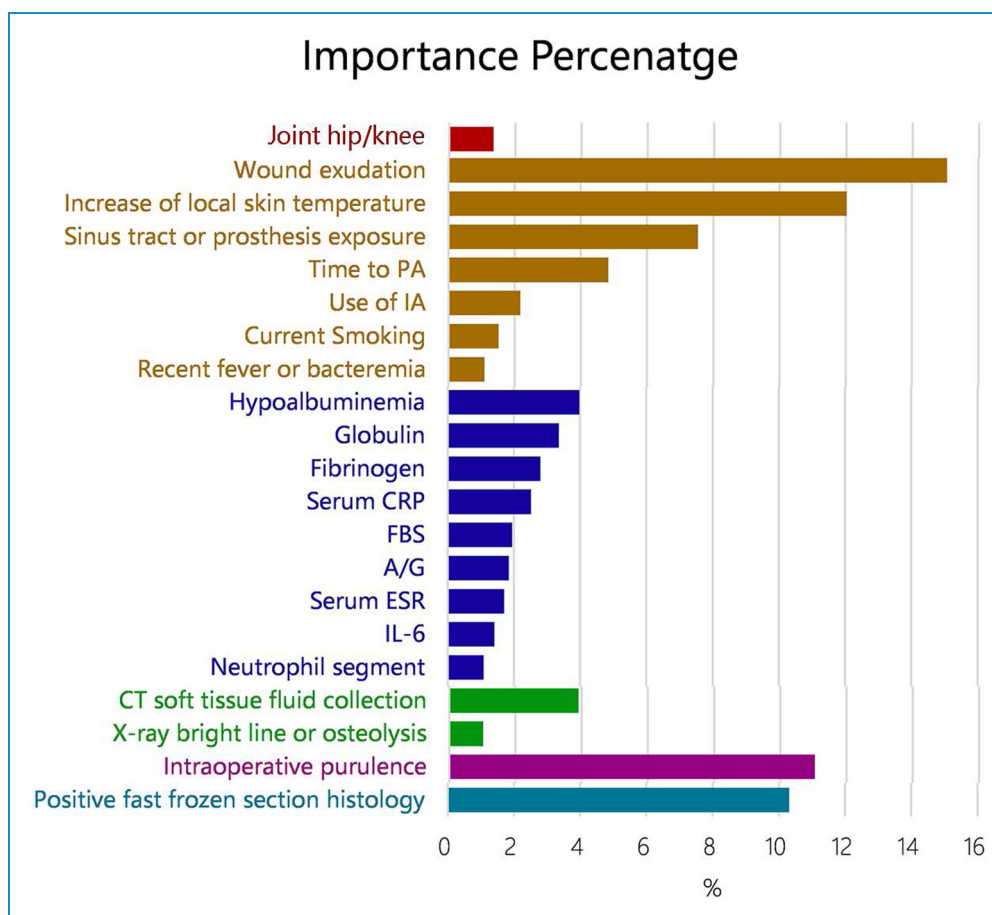
KSVM, ET, LGBM, and MLP were determined as base classifiers for ELWV. The panorama of the final ML model is shown in Figure 5.

Performance metrics of each base classifier for PJI prediction in Cohort I are shown in Table 3. The ROC plot of each classifier used in our model is shown in Supplementary Figure 2. We demonstrated that the LGBM algorithm outperformed all base classifiers with the highest accuracy, MCC, precision, recall/sensitivity,

and F1 score at 0.950 ( $\pm 0.047$ ), 0.897 ( $\pm 0.095$ ), 0.975 ( $\pm 0.056$ ), 0.898 ( $\pm 0.094$ ) and 0.933 ( $\pm 0.063$ ), respectively.

As the meta-classifier, ELWV (i.e. the final ML model) yielded favorable performance in terms of AUC (0.968  $\pm$  0.029), with sensitivity of 0.906 ( $\pm 0.075$ ) and specificity of 0.983 ( $\pm 0.037$ ) in Cohort I (Table 3).

Further, we compared the PJI “in-time” diagnostic performance of the ML model with commonly used criteria (MSIS 2011, ICM 2018, and EBJIS 2021) in Cohort I



**Figure 4.** Importance percentage bar plot of the final feature list. Different colors indicate different categories of the features. Features were sorted in the order of the importance within the categories. FBS: fasting blood sugar; A/G: albumin to globulin ratio.

**Table 2.** Comparison of important features included by the final machine learning model and the canonical definitions for “in-time” diagnosis of PJI.

Machine Learning Feature <sup>a</sup>	MSIS 2011 Feature <sup>b</sup>	ICM 2018 Feature <sup>b</sup>	EBJIS 2021 Feature <sup>b</sup>
Joint	Sinus tract or prosthesis exposure	Sinus tract or prosthesis exposure	Sinus tract or prosthesis exposure
Time to PA	Serum CRP	Blood D-dimer	Serum CRP
Recent fever or bacteremia	Serum ESR	Serum CRP	Positive fast frozen section histology
Use of IA	Intraoperative purulence	Serum ESR	
Current smoking	Positive fast frozen section histology	Intraoperative purulence	
Increase of local skin temperature		Positive fast frozen section histology	

(continued)



Table 2. Continued.

Machine Learning Feature <sup>a</sup>	MSIS 2011 Feature <sup>b</sup>	ICM 2018 Feature <sup>b</sup>	EBJIS 2021 Feature <sup>b</sup>
Wound exudation			
Neutrophil segment			
Fibrinogen			
IL-6			
Hypoalbuminemia			
Globulin			
A/G			
FBS			
X-ray bright line or osteolysis			
CT soft tissue fluid collection			
Sinus tract or prosthesis exposure			
Serum CRP			
Serum ESR			
Intraoperative purulence			
Positive fast frozen section histology			

PA: primary arthroplasty; IA: immunosuppressive agents; FBS: fasting blood sugar; A/G: albumin to globulin ratio; PJI: periprosthetic joint infection; MSIS: Musculoskeletal Infection Society; ICM: International Consensus Meeting; EBJIS: European Bone and Joint Infection Society.

<sup>a</sup>Features also considered in the canonical definitions are boldface.

<sup>b</sup>Features also considered in the machine learning model are boldface.

and II, respectively (Table 3 and 4). The result suggested that our ML model showed extra values in accuracy, MMC, recall/sensitivity, and F1 score, and non-inferiority in other metrics in the training cohort (Cohort I). Accordingly, the ML model functioned well in the internal validation cohort (Cohort II) with a sensitivity of 84.21%, specificity of 100.00%, and accuracy of 94.55% (Table 4).

### Accessible PJI diagnostic web tool

We established the PJI “in-time” diagnostic web tool derived from the ML model, which is available at <http://183.6.101.54:48080/forecast-vue/> (see **User Instruction of the web tool** in Supplemental Digital Content). With user-friendly approaches of indicators data input, the PJI “in-time” diagnosis report is presented in a view that

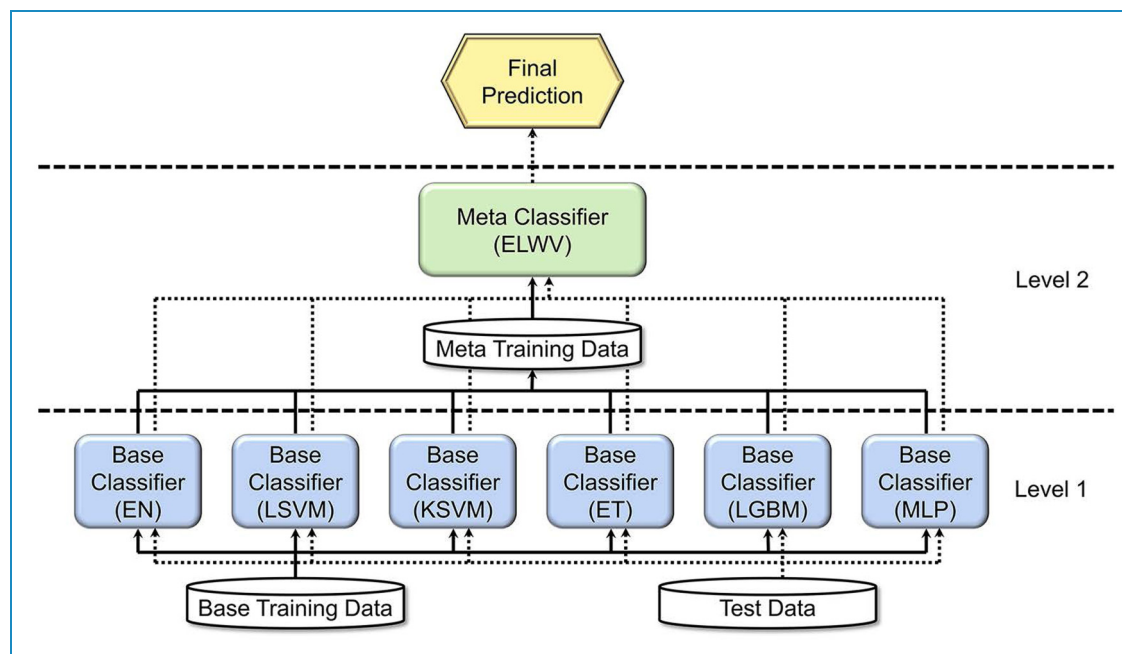
contains the result (Infection/Not) and degree of confidence of ML models, respectively.

### Case presentation

To demonstrate the application of the ML model, we presented the clinical indicator input and prediction result from the web tool in three cases here (Supplementary Table 2 and Supplementary Figure 3). We also made preliminary interpretation for the result of each case.

First, Case 1 was correctly diagnosed as aseptic case, with high confidence level (> 80%) in most classifiers.

Second, Case 2 was correctly diagnosed as infected case, with high confidence level in most classifiers. When using MSIS 2011 or EBJIS 2021 based on the “in-time” clinical data, Case 2 was incorrectly or ambiguous diagnosed.



**Figure 5.** Two-level structure of the ML model with elastic net (EN), linear support vector machine (LSVM), kernel support vector machine (KSVM), extra trees (ET), light gradient boosting machine (LGBM), and multilayer perceptron (MLP) as base classifiers and ensemble learning of weighted voting (ELWV) as meta-classifier. The data flows of the training-and-testing procedures were indicated using the solid and the dashed lines, respectively. ML: machine learning.

Finally, Case 3 is a supposed low-grade PJI case. Although with lower confidence level in each classifier compared with Case 2, Case 3 was still correctly diagnosed as infected case.

## Discussion

Diagnosis of PJI diagnosis remains a critical topic mainly because of the growing patient population and difficulty in differential diagnosis with APL. According to domestic statistics, the amount of TJA operations exceeded 952,000 cases in 2019 with an annual growth rate of 19.96%, which gave rise to a proportional growth in the number of PJI patients.<sup>22</sup> Moreover, up to 20% of patients diagnosed and treated as APL cases are eventually proven to suffer from low-grade PJI instead.<sup>21</sup>

APL cases usually undertake prosthetic revision, while PJI requires adequate control of infection before prosthesis management like retention, single- or two-stage revision. Given the complete difference in the treatment procedures for PJI and APL, it may hinder the proper timing or increase the patient's burden once PJI is misdiagnosed.<sup>23,24</sup> Therefore, early and accurate diagnosis of PJI is of great significance for this rapidly growing patient population to undertake reasonable treatment.

The most important finding of the present study was that the ML-derived system showed excellent performance in the in-time diagnosis of PJI using widely accessible

indicators, which could provide a valuable reference for surgical decision-making.

Though EBJIS 2021 criteria improved the potency of pre-operative PJI diagnosis, the performances of these commonly used criteria were not satisfactory when using the widely indicators determined in the present study, as is shown in Table 3.<sup>7,9</sup> This may be attributed to the following reasons. Firstly, the indicators were selected according to the panel opinion derived from literature review and clinical experience, which means they were predetermined, not “evidence-based” in the real sense.<sup>6,7,15</sup> The limitation of this “a priori” may hinder the access of potential valuable indicators to the establishment of these criteria. Secondly, commonly used criteria put more emphasis on accurate diagnosis based on complete clinical information including those from post-operative examinations. Moreover, the trade-off between diagnostic potency and complexity of examinations drives the panel to prune the examinations with less potency.<sup>7,15</sup> Therefore, the diagnostic potencies of these criteria were impaired when only early indicators were available, and the contribution of pruned examinations was lost. Thirdly, criteria like ICM 2013, ICM 2018, and EBJIS 2021 could produce ambiguous diagnoses of “infection likely,” which inevitably leave the problem unsolved.<sup>7,14,15</sup>

Although ICM 2018 adopted Random Forest and Multivariate Logistic Regression analysis to develop a weight-balanced scoring system with favorable accuracy,

**Table 3.** Comparison of PJJ “in-time” diagnosis performance results among the inclusive ML classifiers and the canonical definitions in cohort I.

Method	Acc.	MCC	Precision	Recall/Se.	F1 score	Sp.	AUC
EN	0.940 ( $\pm 0.049$ )	0.876 ( $\pm 0.100$ )	0.973 ( $\pm 0.060$ )	0.873 ( $\pm 0.087$ )	0.919 ( $\pm 0.065$ )	0.983 ( $\pm 0.037$ )	0.962 ( $\pm 0.036$ )
LSVM	0.940 ( $\pm 0.055$ )	0.874 ( $\pm 0.114$ )	0.936 ( $\pm 0.077$ )	<b>0.911 (<math>\pm 0.071</math>)</b>	0.923 ( $\pm 0.068$ )	0.959 ( $\pm 0.051$ )	0.958 ( $\pm 0.045$ )
KSVM	0.940 ( $\pm 0.037$ )	0.876 ( $\pm 0.076$ )	0.972 ( $\pm 0.038$ )	0.873 ( $\pm 0.087$ )	0.918 ( $\pm 0.053$ )	<b>0.984 (<math>\pm 0.022</math>)</b>	0.961 ( $\pm 0.045$ )
ET	0.945 ( $\pm 0.049$ )	0.885 ( $\pm 0.100$ )	0.958 ( $\pm 0.060$ )	0.898 ( $\pm 0.087$ )	0.926 ( $\pm 0.065$ )	0.975 ( $\pm 0.037$ )	0.965 ( $\pm 0.036$ )
LGBM	<b>0.950 (<math>\pm 0.047</math>)</b>	<b>0.897 (<math>\pm 0.095</math>)</b>	<b>0.975 (<math>\pm 0.056</math>)</b>	0.898 ( $\pm 0.094$ )	<b>0.933 (<math>\pm 0.063</math>)</b>	0.983 ( $\pm 0.037$ )	0.960 ( $\pm 0.037$ )
MLP	0.945 ( $\pm 0.054$ )	0.888 ( $\pm 0.109$ )	0.965 ( $\pm 0.079$ )	0.898 ( $\pm 0.094$ )	0.927 ( $\pm 0.070$ )	0.975 ( $\pm 0.056$ )	<b>0.968 (<math>\pm 0.028</math>)</b>
ELWV (ML model)	<b>0.945 (<math>\pm 0.054</math>)</b>	<b>0.907 (<math>\pm 0.083</math>)</b>	0.973 ( $\pm 0.060$ )	<b>0.906 (<math>\pm 0.075</math>)</b>	<b>0.925 (<math>\pm 0.072</math>)</b>	0.983 ( $\pm 0.037$ )	0.968 ( $\pm 0.029$ )
MSIS 2011	0.751 ( $\pm 0.034$ )	0.451 ( $\pm 0.097$ )	0.926 ( $\pm 0.069$ )	0.476 ( $\pm 0.213$ )	0.604 ( $\pm 0.159$ )	0.879 ( $\pm 0.214$ )	-
<i>p</i> -value*	<b>.001</b>	< <b>.001</b>	.407	<b>.017</b>	<b>.019</b>	.386	-
ICM 2018	0.867 ( $\pm 0.056$ )	0.675 ( $\pm 0.205$ )	0.940 ( $\pm 0.057$ )	0.761 ( $\pm 0.073$ )	0.838 ( $\pm 0.026$ )	0.879 ( $\pm 0.214$ )	-
<i>p</i> -value*	<b>.020</b>	<b>.027</b>	.512	<b>.032</b>	.050	.366	-
EBJIS 2021	0.831 ( $\pm 0.035$ )	0.615 ( $\pm 0.173$ )	0.944 ( $\pm 0.052$ )	0.697 ( $\pm 0.112$ )	0.795 ( $\pm 0.059$ )	0.876 ( $\pm 0.212$ )	-
<i>p</i> -value*	<b>.001</b>	<b>.005</b>	.546	<b>.023</b>	<b>.024</b>	.352	-

ML: machine learning; Acc.: accuracy; MCC: Matthews correlation coefficient; Se: sensitivity; Sp: specificity; EN: Elastic Net; LSVM: Linear Support Vector Machine; KSVM: Kernel Support Vector Machine; ET: Extra Tree; LGBM: Light Gradient Boosting Machine; MLP: Multilayer Perceptron; ELWV: Ensemble Learning of Weighted Voting; PJJ: periprosthetic joint infection; MSIS: Musculoskeletal Infection Society; ICM: International Consensus Meeting; EBJIS: European Bone and Joint Infection Society.

The performance metrics were expressed in form of average  $\pm$  standard deviation.

\*The *p*-value of paired *t*-test between the ML model and indicated canonical definition.

The best results and *p*-value indicating significant difference in comparison were in boldface.

**Table 4.** PJI “in-time” diagnosis performance results of the established ML model and the canonical definitions in cohort II.

Method	MSIS 2011	ICM 2018	EBJIS 2021	ML model
TP	10	11	11	16
TN	36	34	28	36
FP	0	0	0	0
FN	9	3	3	3
Total	55	55	55	55
Acc.	83.64%	81.82%	70.91%	<b>94.55%</b>
MCC	64.89%	84.97%	84.24%	<b>88.17%</b>
Precision	<b>100.00%</b>	<b>100.00%</b>	<b>100.00%</b>	<b>100.00%</b>
Recall/Se.	52.63%	78.57%	78.57%	<b>84.21%</b>
F1 score	68.97%	88.00%	88.00%	<b>91.43%</b>
Sp.	<b>100.00%</b>	<b>100.00%</b>	<b>100.00%</b>	<b>100.00%</b>

ML: machine learning; TP: true positive; TN: true negative; FP: false positive; FN: false negative; Acc.: accuracy; MCC: Matthews correlation coefficient; Se.: sensitivity; Sp.: specificity; PJI: periprosthetic joint infection; MSIS: Musculoskeletal Infection Society; ICM: International Consensus Meeting; EBJIS: European Bone and Joint Infection Society. The best results were in boldface.

the predetermined indicators and the fixed cut-off values reduce the generalization ability in different diagnoses.<sup>7,25</sup> Here, we developed an ML-derived system for “in-time” diagnosis of PJI without relying on predetermined or fixed criteria. We demonstrated significantly improved diagnostic performance of this system by comparing it with commonly used criteria (Table 3 and 4). As illustrated in Table 2, important features approved by this ML-derived system exceeded the range of ICM 2018.

In terms of physical signs, “increase of local skin temperature” was determined as one of the critical variables. This can be attributed to that acute PJI cases with obvious local inflammatory symptoms took up a considerable part of the cohort. Also, clinical history items including “joint type (hip/knee),” “current smoking (yes/no),” and “use of immunosuppressive agents (yes/no)” were considered as significant features of PJI prediction according to ML. Although evidence was limited, the conclusion was supported by several studies.<sup>26–28</sup> This demonstrated that the ML-derived system can be self-improved by learning from rich datasets and making good use of it in model development. Further research will be based on high-quality big data from huge cohorts and improved computational power, which enables the inclusion of more potential

clinical indicators and boosts model performance to another level.

The blood tests including “plasma fibrinogen (g/L)” and “albumin-to-globulin ratio (A/G)” demonstrated novel value in the diagnosis of PJI. Plasma fibrinogen was recently proved to be a potential biomarker for PJI diagnosis. Li et al. reported in a retrospective study with 439 cases that fibrinogen outperformed d-dimer with sensitivity and specificity at 76.3% and 86.2%.<sup>29</sup> As for A/G, it was an index of hepatic function conventionally while recently discovered to be a promising biomarker of inflammation.<sup>17</sup> Both evidence from our previous study and the study of Wang et al. suggested a potential diagnostic value of A/G in PJI.<sup>30,31</sup> Additionally, given the involvement of “hypoalbuminemia (yes/no)” and “globulin (g/L)” in the final model, our study may suggest the potential value of impaired nutritional or immune status in PJI diagnosis, which is supported by the study of Li et al. and Traverso et al.<sup>32,33</sup>

The imaging examinations investigated in the present model also played a part in the “in-time” diagnosis of PJI. “CT soft tissue fluid collection” appeared to be the most important imaging features among them, ranking 8th in the final feature list. “X-ray bright line or osteolysis” also demonstrated potential in PJI diagnosis. These results conform with the previous study by Romanò et al. and Wu et al.<sup>1,34</sup>

Relative studies using ML algorithms to address PJI-related issue were conducted in recent years. Kuo et al. developed a two-level ML system based on a dataset collected during the complete hospitalization, including post-operative microorganism culture. The predictive performances were full beyond ICM 2018.<sup>19</sup> However, the synovial examinations as a group of important features used by their system are not always performed in different institutions for potential infection risk brought by invasive procedures. Moreover, the bottom level of their model was not designed in an optimal approach. Low-performance models like Naïve Bayesian will actually drag down the system performance. Up to the present, the study focusing on early diagnosis of PJI based on state-of-the-art ML prediction using widely accessible indicators remains absent.

Our work developed an ML model with high diagnostic potency, which is feasible in not only well-equipped medical centers but also in local hospitals by using widely accessible “in-time” indicators like clinical history, physical signs, routine blood tests, and imaging examinations. Notably, we have developed the PJI “in-time” diagnostic web tool to improve the usability of our ML-derived system. The web tool enables users with no programming skills to diagnose PJI conveniently.

The findings of the present study should be interpreted in the light of several limitations.

Firstly, the ML-derived system does not perfectly apply to the pre-operative diagnosis of PJI since intraoperative

**Table 5.** Comparison of performance results between the ML models including intraoperative indicators or not.

Model	Acc.	MCC	Precision	Recall/Se.	F1 score	Sp.	AUC
Intraoperative indicators included	<b>0.945 (±0.054)</b>	<b>0.907 (±0.083)</b>	<b>0.973 (±0.060)</b>	<b>0.906 (±0.075)</b>	<b>0.925 (±0.072)</b>	<b>0.983 (±0.037)</b>	<b>0.968 (±0.029)</b>
Not included	0.930 (±0.069)	0.855 (±0.144)	0.947 (±0.087)	0.873 (±0.116)	0.906 (±0.091)	0.967 (±0.054)	0.959
<i>p</i> -value*	<b>.008</b>	.085	<b>.016</b>	<b>.005</b>	<b>.009</b>	<b>.016</b>	.057

Acc.: accuracy; MCC: Matthews correlation coefficient; Se: sensitivity; Sp: specificity; ML: machine learning.

\*The *p*-value of paired *t*-test between the ML models including intraoperative indicators or not.

The better results and *p*-value indicating significant difference in comparison were in boldface.

indicators were included. This is because we found that the addition of intraoperative indicators in features for inductive learning significantly improved the performance of the ML model as shown in Table 5. Therefore, we used “in-time” indicators defined as pre-operative (early) and intraoperative indicators that are available before surgical management of prosthesis (as shown in **Figure 1**) to maximize diagnostic performance. Nevertheless, the prediction made by this ML-derived system still provides a timely reference for decision-making of prosthesis surgical management, like revision for PJI or retention for APL. Secondly, the patient cohort was relatively small and enrolled from a single center, which could impair the reliability, robustness, and generalizability of this system.<sup>35</sup> Thirdly, the raw data was collected retrospectively in nature and the missing data was addressed using Multiple Imputation by Chained Equations. Both of these procedures produced inherent bias in the present study.<sup>36</sup> Fourthly, data from infrequent or expensive examinations was not included in the dataset of model development. The potential diagnostic potency of them contributing to the overall performance of the model was missed. At last, like most ML algorithms, our model inevitably demonstrates method agnostics as a “black box” due to the complexity of base learners, which means it produces binary prediction like “APL” or “PJI” but a concrete interpretation about how it makes the diagnosis. Therefore, we planned to improve this diagnostic system by following methods: (1) enriching the dataset by including more cases and more potential indicators, like synovial examination and (2) performing more reliable external validation via large-scale multicenter clinical study.

## Conclusion

In summary, we developed a novel two-level ensemble ML prediction model for “in-time” diagnosis of PJI and evaluated its performance by comparing it with commonly used definitions in MSIS 2011, ICM 2018, and EBJIS 2021. Using the widely accessible clinical indicators

before the surgical management of prosthesis as input, this diagnostic system produced significantly improved performance. Moreover, we established a user-friendly web tool featuring this system to facilitate a swift diagnosis of PJI.

Our study demonstrated great potential in helping surgeons with timely diagnosis of PJI for surgical decision-making.

**Acknowledgements:** We thank all staff members from the Joint Surgery Department and involved members from the Diagnostic Radiology Department of the institution.

**Declaration of conflicting interests:** The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Funding:** The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Basic and Applied Basic Research Foundation of Guangdong Province (grant number 2023A1515030030), National Natural Science Foundation of China (82172405), National Key Research and Development Program of China (2022YFC2407505), and Science and Technology Institute of National Health Commission (2023HX002112).

**Ethical approval:** This study was exempted from the Institutional Review Board (IRB) since the study design is observational and the data for analysis was de-identified.

**Guarantor:** PS.

**Contributorship:** WC designed the study and participated in the revision of the manuscript. XH constructed the database files, analyzed clinical data, and drafted the manuscript. BP and XW participated in the construction of the database. CG developed the ML model and participated in the revision of the manuscript. ZZ provided the radiological image review. PS and LZ conceived of and participated in the design of the study. WS

participated in the design of the study and the development of the ML model. All authors read and approved the final manuscript.

**ORCID iD:** Puyi Sheng  <https://orcid.org/0000-0002-5898-2161>

**Supplemental material:** Supplemental material for this article is available online.

## References

- Rajput V, Meek RMD, Haddad FS, et al. Periprosthetic joint infection: what next?. *Bone Joint J* 2022; 104-b: 1193–1195. DOI: 10.1302/0301-620x.104b11.Bjj-2022-0944
- Hantouly AT, Alzobi O, Toubasi AA, et al. Higher sensitivity and accuracy of synovial next-generation sequencing in comparison to culture in diagnosing periprosthetic joint infection: a systematic review and meta-analysis. *Knee Surg Sports Traumatol Arthrosc* 2022; 31(9): 3672–3683. DOI: 10.1007/s00167-022-07196-9
- Kloos J, Vander Linden K, Vermote S, et al. Prevalence, interpretation, and management of unexpected positive cultures in revision TKA: a systematic review. *Knee Surg Sports Traumatol Arthrosc* 2022; 30(12): 3998–4009. DOI: 10.1007/s00167-021-06856-6
- Osmon DR, Berbari EF, Berendt AR, 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. 2012/12/12.
- Parvizi J, Gehrke T and Chen AF. Proceedings of the international consensus on periprosthetic joint infection. *Bone Joint J* 2013; 95-b: 1450–1452. 2013/10/24.
- Parvizi J, Jacovides C, Zmistowski B, et al. Definition of periprosthetic joint infection: is there a consensus? *Clin Orthop Relat Res* 2011; 469: 3022–3030. 2011/07/14.
- Parvizi J, Tan TL, Goswami K, et al. The 2018 definition of periprosthetic hip and knee infection: an evidence-based and validated criteria. *J Arthroplasty* 2018; 33: 1309–1314.e1302. 2018/03/20.
- McNally M, Sousa R, Wouthuyzen-Bakker M, et al. The EBJIS definition of periprosthetic joint infection. *Bone Joint J* 2021; 103-b: 18–25. 2021/01/01.
- Sousa R, Ribau A, Alfaro P, et al. The European bone and joint infection society definition of periprosthetic joint infection is meaningful in clinical practice: a multicentric validation study with comparison with previous definitions. *Acta Orthop* 2023; 94: 8–18. 2023/01/23.
- Klemm C, Laurencin S, Uzosike AC, et al. Machine learning models accurately predict recurrent infection following revision total knee arthroplasty for periprosthetic joint infection. *Knee Surg Sports Traumatol Arthrosc* 2021; 30(8): 2582–2590. DOI: 10.1007/s00167-021-06794-3
- Wouthuyzen-Bakker M, Shohat N, Parvizi J, et al. 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. 2021/05/21.
- Parvizi J, Adeli B, Zmistowski B, et al. Management of periprosthetic joint infection: the current knowledge: AAOS exhibit selection. *J Bone Joint Surg Am* 2012; 94: e104. 2012/07/20.
- Parvizi J and Gehrke T. Definition of periprosthetic joint infection. *J Arthroplasty* 2014; 29: 1331. 2014/04/29.
- Parvizi J, Zmistowski B, Berbari EF, et al. New definition for periprosthetic joint infection: from the workgroup of the musculoskeletal infection society. *Clin Orthop Relat Res* 2011; 469: 2992–2994. 2011/09/23.
- Tubb CC, Polkowski GG and Krause B. Diagnosis and prevention of periprosthetic joint infections. *J Am Acad Orthop Surg* 2020; 28: e340–e348. 2020/01/24.
- Ye Y, Chen W, Gu M, et al. Limited value of serum neutrophil-to-lymphocyte ratio in the diagnosis of chronic periprosthetic joint infection. *J Orthop Traumatol* 2021; 22: 37. 2021/09/19.
- Choe H, Kobayashi N, Abe K, et al. Evaluation of serum albumin and globulin in combination with C-reactive protein improves serum diagnostic accuracy for low-grade periprosthetic joint infection. *J Arthroplasty* 2022; 38(3): 555–561. DOI: 10.1016/j.arth.2022.09.011
- Chalmers BP, Kapadia M, Chiu YF, et al. Accuracy of predictive algorithms in total hip and knee arthroplasty acute periprosthetic joint infections treated with debridement, antibiotics, and implant retention (DAIR). *J Arthroplasty* 2021; 36: 2558–2566. 2021/03/23.
- Kuo FC, Hu WH and Hu YJ. Periprosthetic joint infection prediction via machine learning: comprehensible personalized decision support for diagnosis. *J Arthroplasty* 2022; 37: 132–141. 2021/09/21.
- Azur MJ, Stuart EA, Frangakis C, et al. Multiple imputation by chained equations: what is it and how does it work? *Int J Methods Psychiatr Res* 2011; 20: 40–49. 2011/04/19.
- Milandt NR, Gundtoft PH and Overgaard S. A single positive tissue culture increases the risk of rerevision of clinically aseptic THA: a national register study. *Clin Orthop Relat Res* 2019; 477: 1372–1381. 2019/05/29.
- Bian Y, Cheng K, Chang X, et al. Reports and analysis of amount of hip and knee arthroplasty in China from 2011 to 2019. *Chinese J Orthopedics* 2020; 40: 1453–1460.
- Park JW, Won SH, Kim HS, et al. Current incidence and future projection of periprosthetic fractures in South Korea: a study based on national claim database. *Orthop Surg* 2022; 14: 530–535. 2022/02/01.
- Trebbse R and Roskar S. Evaluation and interpretation of prosthetic joint infection diagnostic investigations. *Int Orthop* 2021; 45: 847–855. 2021/02/09.
- Kong L, Cao J, Zhang Y, et al. Risk factors for periprosthetic joint infection following primary total hip or knee arthroplasty: a meta-analysis. *Int Wound J* 2017; 14: 529–536. 2016/07/12.
- Springer BD, Cahue S, Etkin CD, et al. Infection burden in total hip and knee arthroplasties: an international registry-based perspective. *Arthroplast Today* 2017; 3: 137–140. 2017/07/12.
- Xu C, Guo H, Wang Q, et al. Interaction of obesity with smoking and inflammatory arthropathies increases the risk of periprosthetic joint infection: a propensity score matched study in a Chinese Han population. *J Hosp Infect* 2019; 101: 222–228. 2018/07/04.

28. Resende VAC, Neto AC, Nunes C, et al. Higher age, female gender, osteoarthritis and blood transfusion protect against periprosthetic joint infection in total hip or knee arthroplasties: a systematic review and meta-analysis. *Knee Surg Sports Traumatol Arthrosc* 2021; 29: 8–43. 2018/11/11.
  29. Li R, Shao HY, Hao LB, et al. Plasma fibrinogen exhibits better performance than plasma D-dimer in the diagnosis of periprosthetic joint infection: a multicenter retrospective study. *J Bone Joint Surg Am* 2019; 101: 613–619. 2019/04/05.
  30. Wang H, Zhou H, Jiang R, et al. Globulin, the albumin-to-globulin ratio, and fibrinogen perform well in the diagnosis of periprosthetic joint infection. *BMC Musculoskelet Disord* 2021; 22: 583. 2021/06/27.
  31. Ye Y, Chen W, Gu M, et al. Serum globulin and albumin to globulin ratio as potential diagnostic biomarkers for periprosthetic joint infection: a retrospective review. *J Orthop Surg Res* 2020; 15: 459. 2020/10/09.
  32. Li Z, Maimaiti Z, Fu J, et al. The superiority of immune-inflammation summary index for diagnosing periprosthetic joint infection. *Int Immunopharmacol* 2023; 118: 110073. 20230328.
  33. Traverso G, Núñez JH, Gehrke T, et al. Hypoalbuminemia increases the risk of failure following one-stage septic revision for periprosthetic joint infection. *Arch Orthop Trauma Surg* 2023; 143: 5641–5648, 20230418.
  34. Romanò CL, Petrosillo N, Argento G, 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.
  35. Rozycki M, Satterthwaite TD, Koutsouleris N, et al. Multisite machine learning analysis provides a robust structural imaging signature of schizophrenia detectable across diverse patient populations and within individuals. *Schizophr Bull* 2018; 44: 1035–1044. 2017/12/01.
  36. Beesley LJ and Taylor JMG. A stacked approach for chained equations multiple imputation incorporating the substantive model. *Biometrics* 2021; 77: 1342–1354. 2020/09/14.
-