

Article

Not peer-reviewed version

Preoperative Diagnosis of Periprosthetic Infection in Patients Undergoing Hip or Knee Revision Arthroplasties: Development and Validation of Machine Learning Algorithm

Vincenzo Di Matteo , Pierandrea Morandini , [Victor Savevski](#) , Guido Grappiolo , [Mattia Loppini](#) *

Posted Date: 4 January 2024

doi: 10.20944/preprints202401.0261.v1

Keywords: artificial intelligence; arthroplasty; deep machine learning; hip; knee; periprosthetic joint infection



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Preoperative Diagnosis of Periprosthetic Infection in Patients Undergoing Hip or Knee Revision Arthroplasties: Development and Validation of Machine Learning Algorithm

Vincenzo Di Matteo ^{1,2,3}, Pierandrea Morandini ⁴, Victor Savevski ⁴, Guido Grappiolo ^{3,5} and Mattia Loppini ^{1,3,5,*}

¹ Department of Biomedical Sciences, Humanitas University, Pieve Emanuele, 20090, Milan, Italy

² Adult Reconstruction and Joint Replacement Service, Division of Orthopaedics and Traumatology, Fondazione Policlinico Universitario Agostino Gemelli IRCCS, Largo Agostino Gemelli 8, Roma, RM 00168, Italy; Catholic University of Sacred Heart, Largo Francesco Vito 1, 00168 Rome, Italy

³ Fondazione Livio Sciuotto Onlus, Campus Savona, Università degli Studi di Genova, 17100 Savona, Italy

⁴ Artificial Intelligence Center, IRCCS Humanitas Research Hospital, Via Manzoni 56, 20089 Rozzano – Milan, Italy

⁵ IRCCS Humanitas Research Hospital, Rozzano, 20089 Milan, Italy

* Correspondence: mattia.loppini@hunimed.eu

Abstract: Background: Periprosthetic joint infection (PJI) following total hip and knee arthroplasty remains an extremely challenging and relatively high complication. This study aims to develop, validate and evaluate the use of machine learning (ML) algorithm to predict PJI in patients undergoing revision arthroplasties. Methods: A comprehensive review of patients undergoing hip or knee revision arthroplasty from 1 January 2015 to 31 March 2021 was conducted. Clinical data coming from preoperative patients history, laboratory analysis and demographic characteristics of patients were screened. Final data have been used to train a Logistic Regression model with the aim of predicting PJI preoperatively. Results: 1360 patients were enrolled, 1141 in the aseptic cohort and 219 in the infected cohort were included. ML demonstrated good discriminatory performance in predicting PJI in the selected patients (area under the curve 0.770 ± 0.006 in the training set and 0.720 ± 0.057 in the test set), and identified 3 significant predictors of PJI. Conclusion: ML algorithm trained using preoperative clinical data accurately predicted PJI. The incorporation of ML models into preoperative assessment of patients undergoing prosthetic revision procedures are useful in providing specific risk assessment to aid individualised counselling, shared decision making and presurgical optimization.

Keywords: Artificial intelligence; Arthroplasty; Deep machine learning; Hip; Knee; Periprosthetic joint infection

1. Introduction

The Periprosthetic joint infection (PJI) accounts for 15% of failed total hip arthroplasties (THAs) and for 25% of failed total knee arthroplasties (TKAs) [1]. PJI occurs after 2% of primary THA or TKA [2], which in absolute terms represents approximately 3000 cases in Italy annually [3]. This value is destined to increase following the expected raises in primary procedures performed [2,3]. Preoperative PJI diagnosis remains a clinical challenge, as there is no uniformly accepted gold standard, particularly when secondary to low virulence organisms or when the individual is immunocompromised. There is no efficient system for the preoperative PJI diagnosis, in our clinical practice both pre-operative and intra-operative investigations are involved [4]. Several systems have been developed for the diagnosis such as Musculoskeletal Infection Society (MSIS), the International Consensus Meeting (ICM), the Infectious Disease Society of America (IDSA), and the European Bone and Joint Infection Society (EBJIS) [4]. The diagnosis requires postoperative confirmation provided by the isolation of the pathogen from the result of the intraoperative cultures. Furthermore, in some

cases the diagnosis is achieved following microbiological tests routinely performed during the revision procedures without any suspicion of PJI, therefore it is believed that incidence of PJI is greater than what is reported in the official registries. In addition, biofilms production provided by bacteria prevents the detection of infection during the microbiological analysis, causing 15% of false-negative and tests, which were also impaired by lack of sensitivity [5]. Various microbiological detections were developed such as sonication, synovial sampling and incubation period prolongation [6,7]. Although these technological innovations increased the sensitivity of microbiological analysis they were still too expensive to be used in our daily clinical practice. Hence the rationale for the development of a system to diagnose PJI preoperatively without any intraoperative investigations and postoperative confirmations.

The field of artificial intelligence (AI) has advanced in the last years with the application of powerful computing and deep learning algorithms in predicting certain surgical outcomes. Machine learning is a form of artificial intelligence that represents a natural extension of the traditional statistical model [8,9]. This scientific discipline focuses on how computers learn from data, imitate human thinking and may overcome human capability [10]. Machine learning is becoming a promising and powerful technology in the prevention of PJI involving developing and using algorithms capable of identifying correlations and patterns of complex risk factors; it may benefit the prediction, diagnosis, treatment and prognosis of PJI [11]. Some machine learning models have been developed for the prevention of PJI. Although a few studies investigated the use of machine learning in the prevention of PJI, to the best of our knowledge, so far, this is the only study that used ML models to predict preoperative PJI in patients undergoing revision THA (r-THA) and revision TKA (r-TKA) in high volume single center.

The purpose of this observational cohort study is to develop, validate and evaluate a machine learning algorithm based prediction tool of preoperative patient-specific clinical objective criteria. The multivariable Logistic Regression analysis predicts preoperative PJI in patients undergoing r-THA and r-TKA in high volume single center.

2. Materials and Methods

The study was conducted in accordance with the Declaration of Helsinki and good clinical practice guidelines. The study protocol for the development of this registry was approved by Ethics Committee protocol code 444/21 of Humanitas Research Hospital IRCCS, in May 2021. According to the retrospective nature of the study all the included patients have already provided a signed written informed consent to be included in the registry of orthopaedic surgical procedures to consent to the use of their clinical data for research purposes.

2.2. Data extraction

The algorithm was developed using clinical data extracted from electronic health records of patients who underwent r-THA and r-TKA from 1 January 2015 to 31 March 2021, at Humanitas Research Hospital IRCCS Italy. The data extraction step consisted in querying the data from Humanitas Data Warehouse (DWH). Oracle SQL™ has been used to gather the interested data of patients admitted at the Orthopaedics Department. The available data consisted of vital parameters and values of laboratory exams conducted during the pre-admission assessments.

2.3. Data selection & Inclusion Criteria

The retrospective observational cohort study included patients undergoing elective r-THA or r-TKA by senior surgeons experienced in joint replacement surgery, from 1 January 2015 to 31 March 2021, at Humanitas Research Hospital IRCCS Italy. Patients were identified from hospital clinical records using International Classification of Diseases, Ninth Revision, Clinical Modification (ICD9-CM) procedure codes (00.70, 00.71, 00.72, 00.73 for revision THA; 80.06, 81.55, 00.80, 00.81, 00.82, 00.83, 00.84 for revision TKA). Eligibility criteria included patients undergoing elective revision THA or TKA in our Orthopaedic Department with all the required predictive features recorded. Patients

undergoing elective primary THA or TKA: International Classification of Diseases, Ninth Revision, Clinical Modification (ICD9-CM) procedure codes (81.51 for THA and 81.54 for TKA), patients aged under 18 years old and who did not have all the required predictive features recorded were excluded. All patients included were further investigated using International Classification of Diseases, Ninth Revision, Clinical Modification (ICD9-CM) diagnostic codes (99666 and 99667 for infection and inflammation reaction due to internal joint prosthesis) and the postoperative culture-positive sampling for prosthetic joint infections. Therefore, the cohort was divided into an aseptic cohort and an infected cohort.

2.4. Feature selection

The first step was to eliminate all variables with a quote of missing values above or equal to 25%. The feature selection phase was performed using a 10-Fold cross-validation: in each fold, a classification tree is used to perform the classification. For all cross-validation iterations, the selected features were annotated. The features used to perform the final classification are those that were selected in all iterations of the cross-validation.

2.5. Analytics

The model used to run the classification is a Logistic Regression implemented in the Python library sklearn [12]. The model is used in a pipeline composed by three stages:

- Missing values imputer: missing values are imputed with the mean in the overall population.
- Feature scaler: The features are scaled using a minimum and maximum scaler.
- Classification model: Logistic Regression.

The hyper parameter selection was performed through a randomized search cross validation, evaluating the models with the logistic loss measure. After the selection of the best model, the final classification has been evaluated with a 10-Fold cross-validation.

3. Results

3.1. Dataset

Overall 1360 patients were enrolled (57.7% women, median age 66.6 [56.9 Q1 - 75.1 Q3]). 1141 patients in the aseptic cohort (60.6% women, median age 66.6 [56.8 Q1 - 75.1 Q3]) and 219 observations in the infected cohort (42.9% women, median age 67.1 [57.7 Q1 - 75.0 Q3]) were respectively included. The second surgeries were mainly performed at the hips (1230 (90.5%) total, 1046 (91.8%) aseptic cohort, 184 (84.0%) infected cohort). The selected features are mostly laboratory exams and demographic features. Following Table 1 shows the features distribution in the extracted dataset [13]. Continuous features are described as “median [Q1 Q3]”, and binary variables as “absolute-count (percentage)”.

Table 1. Summary of Patient Features.

Features	Missing Value	Aseptic Cohort	Infected Cohort	Overall	P-value
Patients (number)		1141	219	1360	
Pain	1 (0.10%)	1.0 [0.53 1.05]	1.0 [0.87 1.10]	1.0 [0.57 1.07]	<0.001
Breath Rate (bpm)	69 (5.10%)	15.57 [15.09 16.0]	15.68 [15.24 16.25]	15.6 [15.13 16.01]	0.01
Heart Rate (rpm)	1 (0.10%)	74.18 [69.55 78.85]	74.92 [69.77 79.73]	74.29 [69.56 78.98]	0.46
Systemic Artery Pressure (mmHg)	1 (0.10%)	119.25 [112.79 126.27]	119.22 [113.51 127.56]	119.23 [112.8 126.5]	0.722
Oxygen Saturation Index (%)	1 (0.10%)	98.15 [97.54 98.64]	98.1 [97.52 98.52]	98.14 [97.54 98.61]	0.211
Temperature °C	1 (0.10%)	36.42 [36.29 36.57]	36.39 [36.28 36.49]	36.42 [36.29 36.57]	0.02

Serum Antithrombin (%)	45 (3.30%)	100.0 [92.0 108.0]	96.0 [89.0 103.0]	99.0 [91.0 107.0]	<0.001
Serum Related Eosinophils (%)	0 (0.00%)	2.0 [1.0 3.0]	2.0 [2.0 4.0]	2.0 [1.0 3.0]	0.002
Serum Hematocrit (%)	0 (0.00%)	42.0 [39.3 44.5]	40.6 [38.3 43.65]	41.8 [39.1 44.4]	<0.001
Serum Related Lymphocytes (%)	0 (0.00%)	28.0 [23.0 34.0]	26.0 [20.5 31.0]	28.0 [23.0 33.0]	<0.001
Mean Corpuscular Volume (fL)	0 (0.00%)	89.7 [86.5 92.9]	87.6 [84.35 91.3]	89.4 [86.14 92.7]	<0.001
Mean Platelet Volume (fL)	0 (0.00%)	8.9 [8.2 9.6]	8.5 [7.8 9.0]	8.8 [8.1 9.5]	<0.001
Serum Ironemia (µg/dL)	43 (3.20%)	79.0 [62.0 100.0]	62.0 [43.5 85.5]	76.0 [58.0 98.0]	<0.001
Serum Glucose (mg/dL)	203 (14.90%)	96.0 [89.0 103.0]	98.0 [92.0 108.0]	96.0 [89.0 104.0]	0.004
C-Reactive Protein (mg/dL)	114 (8.40%)	0.3 [0.15 0.59]	0.64 [0.32 1.935]	0.34 [0.17 0.71]	<0.001
Total Protein (g/dL)	0 (0.00%)	11.2 [10.7 11.9]	11.9 [11.2 12.6]	11.3 [10.8 12.0]	<0.001
Serum Sodium (mmol/L)	201 (14.80%)	142.0 [141.0 143.0]	142.0 [140.0 143.0]	142.0 [141.0 143.0]	0.017
Gender (Female)	0 (0.00%)	691 (60.6%)	94 (42.9%)	785 (57.7%)	<0.001
Red Blood Cell Distribution Width (%)	0 (0.00%)	14.0 [13.4 14.8]	14.7 [13.9 15.98]	14.1 [13.5 15.0]	<0.001
Platelets Blood Count (10 ³ /mm ³)	0 (0.00%)	231.0 [196.0 272.0]	254.0 [214.0 308.5]	234.0 [197.0 278.0]	<0.001

3.2. Classification

Since the outcome is heavily imbalanced, the Logistic Regression has been trained using balanced weights for the two classes, so as to penalise more misclassifications errors on class 1 observations. The classification performance is evaluated using classical indexes like precision, recall and F1 Score in a 10-Fold cross-validation. In Table 2 are shown these indexes differentiating for class and train/test sets. As it can be seen, the reported indexes are quite similar between train and test sets, and for both aseptic cohort and infected cohort and 1 (the F1 score for aseptic cohort is 0.815 ± 0.014 in train set and 0.824 ± 0.062 in test set, while for infected cohort is 0.444 ± 0.011 in train set and 0.427 ± 0.067 in test set). This means that the model was able to generalize the classification and that the risk of overfitting is reduced at its minimum.

Figure 1 displays the ROC curves and the distribution of predicted probabilities. This visualization highlights the generalization of the model, with reported AUC values of 0.770 ± 0.006 in the training set and 0.720 ± 0.057 in the test set. Additionally, the predicted probabilities in both the training and test sets exhibit similar distributions.

An Odds Ratio analysis was performed to interpret the modeling. Figure 2 shows the odds ratio obtained using a Logistic Regression fitted with no regularisation. As it can be seen, the measurement of systemic markers of inflammation such as serum C-reactive protein (CRP), is the most predictive feature positively correlated with the outcome PJI. Gender is also informative, as women demonstrate a higher level of protection against PJI. This finding needs to be furthermore explored.

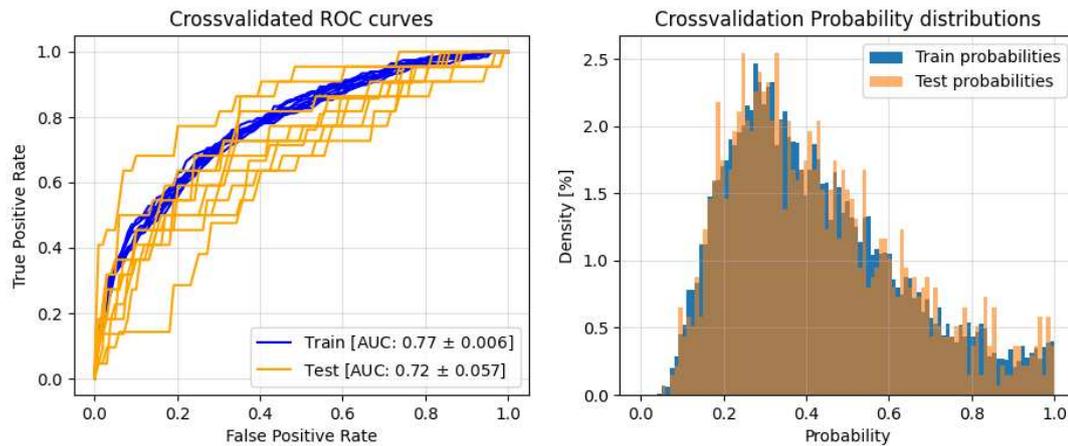


Figure 1. ROC curves and the distribution of the predicted probabilities.

Table 2. Models Classification Scores for the Two Classes.

	Test		Train	
	0	1	0	1
F1	0.824 ± 0.062	0.427 ± 0.067	0.815 ± 0.014	0.444 ± 0.011
Precision	0.911 ± 0.023	0.349 ± 0.092	0.924 ± 0.003	0.329 ± 0.014
Recall	0.761 ± 0.111	0.603 ± 0.143	0.729 ± 0.024	0.688 ± 0.024

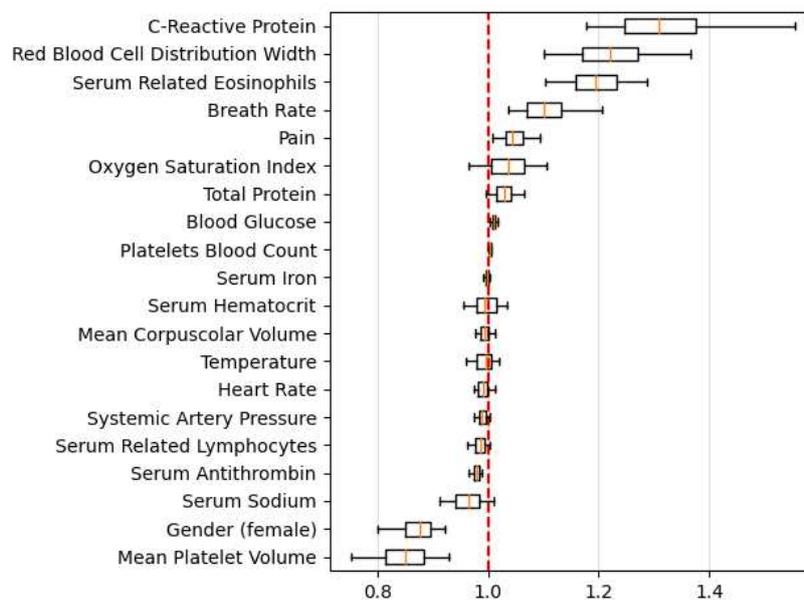


Figure 2. Multivariate Logistic Regression Odds Ratios and CIs.

4. Discussion

Artificial intelligence has shown great potential in improving the accuracy and efficiency of diagnoses, as well as aiding in the identification of risk factors for various diseases. The present study evaluated the preoperative use of ML algorithms to predict PJI in patients undergoing r-THA and r-TKA using available preoperative clinical data. The most pertinent findings of the present study was that the algorithm adopted demonstrated good discriminatory performance for the prediction of PJI in the selected patients (area under the curve 0.770 ± 0.006 in training set and 0.720 ± 0.057 in test set). The measurement of systemic markers of inflammation such as serum C-reactive protein (CRP), serum RDW and serum related eosinophils were the most predictive feature, readily

available preoperative, positively correlated with the outcome PJI. Gender is also informative in the form of reverse association, as women demonstrate a higher level of protection against PJI. The measurement of erythrocyte sedimentation rate (ESR) level, a validated systemic marker of inflammation, was not considered as a feature because the quote of missing values was above to 25%. ML models may be improved by incorporating clinical relevant variables regarding PJI as joint local phenomenon, and not as a systemic process, such as synovial white blood cell (WBC) and synovial polymorphonuclear neutrophils (PMN) which could be obtained through a preoperative joint aspiration. Currently, surgeons and internal medicine physicians seeking to diagnose PJI used a multidisciplinary test battery that included tests to detect joint local inflammation, such as synovial fluid white blood cell (WBC) count and synovial tissue histology [14]. The analysis of synovial fluid obtained by preoperative aspiration, including total cell count and differential leucocyte count and cultures for aerobic and anaerobic organisms, has shown sufficient sensitivity and specificity in multiple studies, it is the most valuable diagnostic tool and should be performed prior to the surgical revision [4]. In our hospital no preoperative synovial fluid analysis was performed, therefore these interesting features were not reported. Regarding features related to formation of cutaneous fistulae and/or drainage of purulent secretions, the present study investigated features related to misdiagnosed "silent" PJI, hence, fistulae and purulent drainage were not considered. Moreover, a previous study reported that the rate of developing a sinus tract in PJI was 21.3%, the presence of sinus tract may be a proxy for other issues such as poor periarticular soft tissue, the poor nutritional status of the host, and multiple prior operations [15]. PJI complication is estimated to occur in 1–3% of patients undergoing primary replacement and in 3–5% of patients undergoing revision [16]. 219 patients were in the infected cohort and 1141 patients were in the aseptic cohort, the patients sample was representative and aligned to the existing literature [1]. With wider application of the joint replacement surgery and related technologies, more and more patients receive the procedure. The incidences of PJI have been on the rise over the recent years and accurate diagnosis of the complication is an urgent task of clinical research. There is currently no universal guideline on the PJI prediction after THA or TKA. In clinical practice, patients at high risk of developing PJI after THA or TKA are identified based on the presence or absence of risk factors and clinicians can only predict the risk of PJI based on their experiences. Whenever a patient's condition is complex or the dataset is incomplete, the difficulty of prediction may increase, requiring more time for an accurate prediction. The machine learning algorithm presented in this study may be a promising alternative to the manual risk prediction method in predicting preoperatively PJI in patients undergoing r-THA or r-TKA. The tool presented here can forecast patients with PJI before the surgery, the strongest predictors of the occurrence of PJI were high levels of serum CRP, serum RDW and serum Related Eosinophils. In recent years, medical researchers are increasingly employing ML applications in the prevention of PJI, the models were divided into four categories: prediction, diagnosis, antibiotic application and prognosis. Yeo et al in 2022 investigated the use of Artificial Neural Networks for the prediction of superficial surgical site infections and PJI following TKA. They retrospectively included a total of 10,021 consecutive primary TKA patients; the average follow-up time lasted for about 2.8 years. SSIs were reported in 404 (4.0%) TKA patients, including 223 superficial surgical site infections and 181 PJI. The patients' demographic and operational variables were collected. The model performance was good, with an AUC of 0.84 and a Brier score of 0.054 (a Brier score close to zero indicates good accuracy of probabilistic prediction). The strongest predictors of the occurrence of surgical site infections following primary TKA, were Charlson comorbidity index, obesity (BMI >30kg/m²), and smoking. The neural network model of the study represented an accurate method to predict patient-specific superficial surgical site infections and PJI following primary TKA [17]. Klemm et al in 2021 used artificial intelligence to evaluate prognostic outcomes. They retrospectively reviewed 618 r-TKA procedures for PJI. They showed a ML model with excellent performance for the prediction of recurrent infections in patients following r-TKA for PJI. The ML models all achieved excellent performance across discrimination (AUC range 0.81–0.84), a Brier score of 0.053 (close to zero indicating good accuracy of the probabilistic prediction). The strongest predictors for recurrent PJI were previous surgeries in patients following r-TKA included irrigation and debridement with or

without modular component exchange ($p < 0.001$), > 4 prior open surgeries ($p < 0.001$), metastatic disease ($p < 0.001$), drug abuse ($p < 0.001$), HIV/AIDS ($p < 0.01$), presence of *Enterococcus* species ($p < 0.01$) and obesity ($p < 0.01$) [18]. Tao et al in 2022 trained a deep learning model to diagnose PJI. 20 r-TKA were enrolled from Chinese People's Liberation Army General Hospital. PJI infection was based on the 2018 ICM guidelines. Frozen pathological sections collected were converted into electronic images with 461 positive and 461 negative images for model training. The AUC of the model was 0.814 and an average accuracy of 93.3 % [19]. Wu et al developed an accurate machine learning model using administrative and electronic medical records (EMR) to improve the accuracy of surgical site infections (SSI) detection. The study cohort consisted of 16,561 primary TKA and 10,799 primary THA retrospectively included. Their findings suggested ML models derived from administrative data and EMR text data achieved high performance and can be used to automate the detection of complex SSIs: ROC AUC of 0.906 (95% CI 0.835–0.978), PR AUC of 0.637 (95% CI 0.528–0.746) and F1 score of 0.790 (0.670–0.900) [20]. The study findings need to be interpreted in light of its limitations. First, this study design was a retrospective study, which is associated with inherent limitations and reporting bias. Second, it's a single-institution study, which introduced the potential for confounding effects of unmeasured variables and features; our results may not be generalizable to other institutions. Third, the minimum patients follow-up period was 2 years, but that a larger percentage of recurrent infections and re-revisions may occur with longer follow-up time. Lastly, the sample size is relatively small and the label imbalance is quite pronounced. To obtain more consistent results, new studies with larger cohorts are necessary. Future prospective multicenter studies with long-term follow-up should include more features to construct more widely functioning models to improve their accuracy and performance in the prediction of PJI.

5. Conclusions

This study developed and validated a machine learning model to preoperatively predict PJI in patients undergoing r-THA or r-TKA. The study findings illustrate a good performance in predicting PJI based on preoperative clinical data, providing the effectiveness and applicability of artificial intelligence in our clinical practice. The promising results suggest that there is the potential utility to integrate computerized algorithms in electronic health record systems, where they may be employed to assist in clinical decision-making supplementing the ICM criteria for PJI prediction. Further, prospective studies are mandatory to implement our findings and the reliability of this technology to clinical practice.

Author Contributions: Conceptualization, M.L. and G.G.; methodology, M.L.; software, P.M. and V.S.; validation, M.L. and V.D.M.; formal analysis, P.M.; investigation, V.D.M.; resources, G.G.; data curation, P.M.; writing—original draft preparation, V.D.M.; writing—review and editing, P.M. and V.D.M.; visualization, P.M.; supervision, M.L.; project administration, M.L.; funding acquisition, V.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received funding from IRCCS Humanitas Research Hospital (5x1000 funding).

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and good clinical practice guidelines. The study protocol for the development of this registry was approved by Ethics Committee of Humanitas Research Hospital (protocol code 444/21) on May 2021.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study: all individual participants signed a written informed consent before the surgical procedure and a written informed consent to be included in the study.

Data Availability Statement: The data supporting reported results can be found in a repository (Zenodo).

Acknowledgments: All the patients who participated in the study are gratefully acknowledged.

Conflicts of Interest: G.G. declares royalties and licenses from Zimmer Biomet, Innomed, and Adler Ortho; Financial support for attending symposia and educational programs from Zimmer Biomet. M.L. declares Research grant as principal investigator (2022YME9N3) from the Italian Ministry of University and Research; Research grant as co-principal investigator (GR-2019-12371158) from the Italian Ministry of Health; Research grant as principal investigator (GR-2018-12367275) from the Italian Ministry of Health; Research grants as principal investigator for postmarket study for medical devices from Zimmer Biomet; Financial support for

attending symposia and educational programs from Zimmer Biomet; Scientific Director of Fondazione Livio-Sciotto. V.D.M and P.M and V.S. declare no conflict of interest.

Appendix A

Table A1. LightGBM Scores for the Two Classes.

	Test		Train	
	0	1	0	1
F1	0,785 ± 0,143	0,454 ± 0,109	0,893 ± 0,028	0,622 ± 0,059
Precision	0,291 ± 0,032	0,369 ± 0,148	0,963 ± 0,005	0,501 ± 0,076
Recall	0,706 ± 0,202	0,699 ± 0,184	0,833 ± 0,048	0,835 ± 0,029

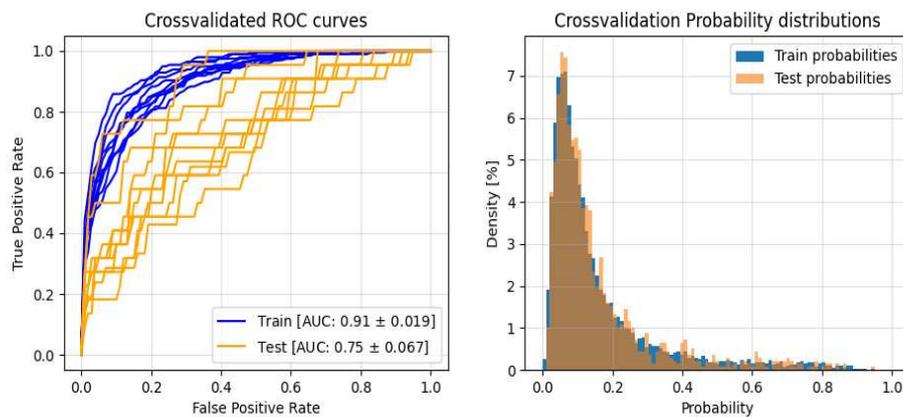


Figure A1. LightGBM ROC curves and the distribution of the predicted probabilities.

As can be seen from both Table A1 and Figure A1, the LightGBM model slightly overfits the data during the training phase, and the performance of the test set drops to a level similar to that reported by Logistic Regression. Also, the performance variance is consistently found to be greater than the measure reported by Logistic Regression. Taking all these factors into account, it is possible to state that the Logistic Regression is more able to generalize the classification showing more reliable results.

References

1. Volpe, L; Indelli, PF; Latella, L; Polli, P; Yakupoglu, J; Marcucci, M. Periprosthetic Joint Infections: A Clinical Practice Algorithm. *Joint* **2014**, *2*, 169-74.
2. Kurtz, S.; Ong, K.; Lau, E.; Mowat, F.; Halpern, M. Projections of Primary and Revision Hip and Knee Arthroplasty in the United States from 2005 to 2030. *J Bone Joint Surg Am* **2007**, *89*, 780–785.
3. Torre, M.; Romanini, E.; Zanolli, G.; Carrani, E.; Luzi, I.; Leone, L.; Bellino, S. Monitoring Outcome of Joint Arthroplasty in Italy: Implementation of the National Registry. *Joints* **2017**, *5*, 70–78.
4. Izakovicova, P.; Borens, O.; Trampuz, A. Periprosthetic Joint Infection: Current Concepts and Outlook. *EFORT Open Reviews* **2019**, *4*, 482–494.
5. Palan, J.; Nolan, C.; Sarantos, K.; Westerman, R.; King, R.; Foguet, P. Culture-Negative Periprosthetic Joint Infections. *EFORT Open Reviews* **2019**, *4*, 585–594.
6. Talsma, D.T.; Ploegmakers, J.J.W.; Jutte, P.C.; Kampinga, G.; Wouthuyzen-Bakker, M. Time to Positivity of Acute and Chronic Periprosthetic Joint Infection Cultures. *Diagnostic Microbiology and Infectious Disease* **2021**, *99*, 115178.

7. Evangelopoulos, D.S.; Stathopoulos, I.P.; Morassi, G.P.; Koufos, S.; Albarni, A.; Karampinas, P.K.; Stylianakis, A.; Kohl, S.; Pneumaticsos, S.; Vlamis, J. Sonication: A Valuable Technique for Diagnosis and Treatment of Periprosthetic Joint Infections. *The Scientific World Journal* **2013**, *2013*, 1–5.
8. Fu, S.; Wyles, C.C.; Osmon, D.R.; Carvour, M.L.; Sagheb, E.; Ramazanian, T.; Kremers, W.K.; Lewallen, D.G.; Berry, D.J.; Sohn, S.; et al. Automated Detection of Periprosthetic Joint Infections and Data Elements Using Natural Language Processing. *The Journal of Arthroplasty* **2021**, *36*, 688–692.
9. Bingham, J.S.; Salib, C.G.; McQuivey, K.; Temkit, M.; Spangehl, M.J. An Evidence-Based Clinical Prediction Algorithm for the Musculoskeletal Infection Society Minor Criteria. *The Journal of Arthroplasty* **2018**, *33*, 2993–2996.
10. Deo, R.C. Machine Learning in Medicine. *Circulation* **2015**, *132*, 1920–1930.
11. Chong, Y.Y.; Chan, P.K.; Chan, V.W.K.; Cheung, A.; Luk, M.H.; Cheung, M.H.; Fu, H.; Chiu, K.Y. Application of Machine Learning in the Prevention of Periprosthetic Joint Infection Following Total Knee Arthroplasty: A Systematic Review. *Arthroplasty* **2023**, *5*, 38.
12. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-Learn: Machine Learning in Python. *MACHINE LEARNING IN PYTHON*.
13. Pollard, T.J.; Johnson, A.E.W.; Raffa, J.D.; Mark, R.G. Tableone: An Open Source Python Package for Producing Summary Statistics for Research Papers. *JAMIA Open* **2018**, *1*, 26–31.
14. Bedair, H.; Ting, N.; Jacovides, C.; Saxena, A.; Moric, M.; Parvizi, J.; Della Valle, C.J. The Mark Coventry Award: Diagnosis of Early Postoperative TKA Infection Using Synovial Fluid Analysis. *Clin Orthop Relat Res* **2011**, *469*, 34–40.
15. Xu, C.; Wang, Q.; Kuo, F.-C.; Goswami, K.; Tan, T.L.; Parvizi, J. The Presence of Sinus Tract Adversely Affects the Outcome of Treatment of Periprosthetic Joint Infections. *J Arthroplasty* **2019**, *34*, 1227–1232.e2.
16. Sigmund, I.K.; McNally, M.A.; Luger, M.; Böhler, C.; Windhager, R.; Sulzbacher, I. Diagnostic Accuracy of Neutrophil Counts in Histopathological Tissue Analysis in Periprosthetic Joint Infection Using the ICM, IDSA, and EBJS Criteria. *Bone Joint Res* **2021**, *10*, 536–547.
17. Yeo, I.; Klemm, C.; Robinson, M.G.; Esposito, J.G.; Uzosike, A.C.; Kwon, Y.-M. The Use of Artificial Neural Networks for the Prediction of Surgical Site Infection Following TKA. *J Knee Surg* **2023**, *36*, 637–643.
18. Klemm, C.; Laurencin, S.; Uzosike, A.C.; Burns, J.C.; Costales, T.G.; Yeo, I.; Habibi, Y.; Kwon, Y.-M. Machine Learning Models Accurately Predict Recurrent Infection Following Revision Total Knee Arthroplasty for Periprosthetic Joint Infection. *Knee Surg Sports Traumatol Arthrosc* **2022**, *30*, 2582–2590.
19. Tao, Y.; Hu, H.; Li, J.; Li, M.; Zheng, Q.; Zhang, G.; Ni, M. A Preliminary Study on the Application of Deep Learning Methods Based on Convolutional Network to the Pathological Diagnosis of PJI. *Arthroplasty* **2022**, *4*, 49.
20. Wu, G.; Cheligeer, C.; Southern, D.A.; Martin, E.A.; Xu, Y.; Leal, J.; Ellison, J.; Bush, K.; Williamson, T.; Quan, H.; et al. Development of Machine Learning Models for the Detection of Surgical Site Infections Following Total Hip and Knee Arthroplasty: A Multicenter Cohort Study. *Antimicrob Resist Infect Control* **2023**, *12*, 88.
21. Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; Liu, T.-Y. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. In Proceedings of the Advances in Neural Information Processing Systems; Curran Associates, Inc., 2017; Vol. 30.

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