



Development of Explainable Machine Learning Models to Predict Outcomes After Platelet-Rich Plasma Injections for Knee Osteoarthritis

Felix Conrad Oettl,* MD , Antonio Madrazo Ibarra,* MD, Mark A. Fontana,* PhD, Ophelie Loblack,* Mert Marcel Dagli,[†] MD, Miguel Otero,* PhD, Jessica Andres Bergos,* PhD, Brian C. Halpern,* MD, and Scott A. Rodeo,*[‡] MD 

Investigation performed at Hospital for Special Surgery, New York, New York

Background: Platelet-rich plasma (PRP) has been increasingly used to treat knee osteoarthritis, but its efficacy remains unclear due to the variability of outcomes. Machine learning (ML) can improve the ability to predict responses to PRP treatment by identifying specific baseline characteristics of patients who may have greater clinical improvements.

Purpose: To develop and evaluate an ML model predicting clinical outcomes after PRP injection for knee osteoarthritis.

Study design: Cohort study (prognosis); Level of evidence, 2.

Methods: This retrospective study utilized patient demographics and patient-reported outcome measures (PROMs) from 191 patients who received PRP injections for knee osteoarthritis. Patients were randomly split into a training set (80%) and a testing set (20%). The primary outcome was predicting the achievement of the minimal clinically important difference at 6 months after treatment, defined as a ≥ 10 -point increase in the Knee injury and Osteoarthritis Outcome Score for Joint Replacement (KOOS JR) and a $\geq 20\%$ decrease in the numeric pain rating scale for knee pain score. Ten preinjection variables, including demographics and baseline PROMs, were evaluated. Multiple ML algorithms were developed and evaluated on sensitivity, accuracy, precision, area under the receiver operating characteristic curve (AUC)-ROC, and $F1$ score. Feature importance and partial dependency plots were used to explore predictor relationships with the primary outcome.

Results: The Explainable Boosting Machine (EBM) algorithm was determined to be the best algorithm due to its greater explainability (AUC-ROC, 0.81 [95% CI, 0.65-0.94]; $F1$ score, 0.75 [95% CI, 0.57-0.88]; accuracy, 0.74 [95% CI, 0.59-0.90]; sensitivity, 0.71 [95% CI, 0.50-0.90]; precision, 0.79 [95% CI, 0.59-0.96]). The baseline PROMIS (Patient-Reported Outcomes Measurement Information System) Mental score (the higher, the better) and the KOOS JR score (the lower, the better) were the most influential predictors. By excluding the baseline PROMIS health scores, the model's performance significantly deteriorated (AUC-ROC, 0.51 [95% CI, 0.32-0.7]).

Conclusion: ML models effectively predicted a clinically meaningful improvement at 6 months after PRP injection for knee osteoarthritis. The EBM was the algorithm with the best performance, with the PROMIS Mental and Physical scores and the KOOS JR score being the most influential predictors. Additional independent studies are needed to externally validate this model.

Keywords: platelet-rich plasma; knee; osteoarthritis; machine learning; artificial intelligence; minimal clinically important difference

Knee osteoarthritis is the leading cause of lower extremity disability and one of the most common causes of chronic pain in adults, representing an immense personal and

societal burden.¹⁷ As the population ages and obesity rates climb, the prevalence of knee osteoarthritis is projected to continue rising.²⁴ Because of the limited efficacy of current nonoperative treatments, physicians have started to use so-called orthobiologics such as platelet-rich plasma (PRP) to treat osteoarthritis. However, the clinical outcomes after PRP treatments have been highly

The Orthopaedic Journal of Sports Medicine, 13(8), 23259671251349743
DOI: 10.1177/23259671251349743
© The Author(s) 2025

This open-access article is published and distributed under the Creative Commons Attribution - NonCommercial - No Derivatives License (<https://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits the noncommercial use, distribution, and reproduction of the article in any medium, provided the original author and source are credited. You may not alter, transform, or build upon this article without the permission of the Author(s). For article reuse guidelines, please visit SAGE's website at <http://www.sagepub.com/journals-permissions>.

variable,^{2,6,23} and the exact factors underpinning the variable responses to PRP remain poorly understood. Identifying factors that can predict responses to treatment is therefore critical to improve patient care and avoid unnecessary interventions.

Machine learning (ML) models differ from conventional statistics in that they use a bottom-up approach, developing models directly from the data without assuming any predefined model structure. This allows the detection of more complex nonlinear relationships and interactions between a large number of variables. In essence, ML trades interpretability for flexibility and predictive accuracy by learning patterns directly from the data.^{13,18}

Patient outcomes after PRP injections for knee osteoarthritis have been extensively studied.⁶ While numerous studies have assessed patient characteristics associated with positive or negative outcomes, to our knowledge no study has utilized advanced ML to develop predictive models in this population.^{1,4,10,19,20,22} Here, we sought to develop and evaluate an ML model to predict clinical outcomes after PRP injection for knee osteoarthritis.

METHODS

Design

This study was conducted at a large academic tertiary care center. The study was performed according to the TRIPOD (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis) guidelines as well as the Guidelines for Developing and Reporting Machine Learning Predictive Models in Biomedical Research.^{5,14} Clinical data were collected with approval of the institutional review board (No. 2020-0241).

Data Source and Patient Population

This study was performed using data collected in the Center for Regenerative Medicine Registry at our institution. The registry consists of all patients undergoing orthobiological treatments for musculoskeletal pathologies. It contains baseline data (collected before treatment) for each patient, including demographic information and patient-reported outcome measures (PROMs) such as the numeric pain rating scale for knee pain (NPRS) score, the Knee injury and Osteoarthritis Outcome Score for Joint Replacement (KOOS JR), the PROMIS (Patient-Reported Outcomes Measurement Information System) Mental score, the PROMIS Physical score, the Tegner activity score,

and the Single Assessment Numeric Evaluation for knee function (SANE) score. Additionally, the registry prospectively collects PROMs over time (6 weeks, 3 months, 6 months, and yearly) after patients receive treatment. Patients were surveyed at each respective time point after injection, and we utilized the data from the 6-month follow-up survey.

Patients included in this study had to meet the following inclusion criteria: (1) diagnosed with knee osteoarthritis, (2) received intra-articular PRP injections in the knee between June 2020 and October 2022, and (3) had available baseline and 6-month follow-up data. The 6-month time point was selected based on systematic reviews and meta-analyses of patient-reported pain scores of those with knee osteoarthritis at 6 and 12 months after PRP treatment.⁸ The registry included 956 patients who received PRP injections at the time of analysis. A total of 756 patients were excluded because they did not have complete outcome information, which was required to determine if they met the minimal clinically important difference (MCID) thresholds at 6 months after PRP injection. Given the retrospective nature of this study, no attempts were made to reach patients with missing data. Thus, the analyses were conducted in a cohort of 191 patients (Table 1). The demographic distribution of the 191-patient cohort, with a majority (92.1%) of White/Caucasian patients, is also observed in the registry (85.8% White/Caucasian) and the overall hospital patient population (74.1% White/Caucasian).

Outcomes

Our primary outcome was achievement of MCID thresholds for both the NPRS score and the KOOS JR at 6 months after receiving PRP treatment for knee osteoarthritis. The MCID thresholds used were a ≥ 10 -point increase in the KOOS JR and a $\geq 20\%$ decrease in the NPRS score, as previously reported.^{9,21,23}

Data Collection and Input Definitions

Prediction models were developed using 10 preinjection variables included in the registry and collected at our institution as part of standard of care before patients undergo PRP injection of the knee joint. These variables include (1) self-reported sex, (2) age, (3) body mass index (BMI), (4) PROMIS Physical score, (5) PROMIS Mental score, (6) KOOS JR, (7) NPRS score, (8) SANE score, (9) Tegner score, and (10) Kellgren-Lawrence (KL) radiographic

[‡]Address correspondence to Scott A. Rodeo, MD, Hospital for Special Surgery, 535 East 70th Street, New York, NY 10021, USA (email: rodeos@hss.edu).

^{*}Hospital for Special Surgery, New York, New York, USA.

[†]Department of Neurosurgery, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA.

Final revision submitted December 14, 2024; accepted January 10, 2025.

One or more of the authors has declared the following potential conflict of interest or source of funding: Funding was provided by the HSS Center for Regenerative Medicine. M.O. has an ownership interest in Jannu Therapeutics LLC. S.A.R. receives consulting fees from Teladoc Inc, Enovis-DJO, and Novartis Pharmaceuticals. AOSSM checks author disclosures against the Open Payments Database (OPD). AOSSM has not conducted an independent investigation on the OPD and disclaims any liability or responsibility relating thereto.

Ethical approval for this study was obtained from Hospital for Special Surgery (2020-0241-CR4).

TABLE 1
Baseline Characteristics of the Training and Testing Sets^a

Characteristic	Full Set (N = 191)	Training Set (n = 152)	Testing Set (n = 39)
Fulfilled MCID	84 (44.0)	63 (41.4)	21 (53.8)
Age, y	65.23 ± 12.5	65.7 ± 12.6	61.6 ± 12.5
BMI, kg/m ²	26.37 ± 4.5	26.2 ± 4.4	27.1 ± 4.8
Male	103 (53.9)	79 (52.0)	24 (61.5)
Race			
Asian	5 (2.6)	5 (3.3)	0 (0)
Black or African American	5 (2.6)	5 (3.3)	0 (0)
Other	3 (1.6)	3 (2.0)	0 (0)
Patient declined	2 (1.0)	2 (1.3)	0 (0)
White or Caucasian	176 (92.1)	137 (90)	39 (100)
KL osteoarthritis grade			
1	15 (7.9)	11 (7.2)	4 (10.3)
2	54 (28.3)	42 (27.6)	12 (30.8)
3	64 (33.5)	50 (32.9)	14 (35.9)
4	58 (30.4)	49 (32.2)	9 (23.1)
PROMIS Mental T score	54.85 ± 5.7	55.0 ± 5.7	54.4 ± 6.1
PROMIS Physical T score	49.49 ± 5.2	49.8 ± 5.3	48.5 ± 4.8
KOOS JR	62.93 ± 15	63.5 ± 15.1	60.6 ± 14.7
NPRS score	36.11 ± 25.2	34.9 ± 24.9	41.2 ± 25.9
SANE score	51.72 ± 23.2	51.3 ± 23.9	53.1 ± 20.2
Tegner score	3.39 ± 2.1	3.2 ± 2.0	3.5 ± 2.5

^aData are presented as n (%) or mean ± SD. BMI, body mass index; KL, Kellgren-Lawrence; KOOS JR, Knee injury and Osteoarthritis Outcome Score for Joint Replacement; MCID, minimal clinically important difference; NPRS, numeric pain rating scale for knee pain; PROMIS, Patient-Reported Outcomes Measurement Information System; SANE, Single Assessment Numeric Evaluation for knee function.

classification of osteoarthritis. The KL grade was assessed on the most recent pre-PRP injection radiograph, evaluated on anteroposterior views by a single experienced reader (F.C.O.). The study sample was predominantly White or Caucasian (92.1%), with higher proportions than in the registry (85.8%) or the hospital (74.1%) data sets, while other racial groups had lower representation. To minimize unwanted noise introduced in the model, race was excluded from the analyses. Variable importances indicate the extent to which each variable contributed to the model's predictions. The 3 variables exhibiting the highest importance were designated as key variables in this study.

Imputation of missing values was performed on the source data set using the missForest package in R (Version 4.3.1; R Foundation for Statistical Computing); the PROMIS score (both Mental and Physical scores) had a 63.4% missing rate before imputation, while the other scales were complete. The data set was first split into a training set (80%) and a test set (20%), with a set seed of 42 to ensure reproducibility. Then, the imputation algorithm was trained and applied on the training set with subsequent application of the model on the test set, to prevent data leakage, while utilizing the same model for imputation.

Model Development

Categorical variables are summarized as absolutes and percentages, while continuous variables are summarized

using means and standard deviations (Table 1). Outcome prediction was conducted utilizing 4 ML algorithms: Explainable Boosting Machine (EBM), random forest (RF), eXtreme Gradient Boosting (XGBoost), and support vector machine (SVM). EBM, RF, and XGBoost are ensemble methods based on decision trees, with XGBoost gaining popularity in recent years because of its high accuracy on complex data sets. SVMs are supervised learning models that classify data by finding the hyperplane that maximizes the margin between the 2 classes in the training data. To evaluate a representative diversity of methods with differing characteristics, we selected these 4 classifiers. The 4 algorithms were derived using the training cohort and subsequently validated on the independent test cohort (Figure 1).

Given the central importance of the PROMIS scores as predictors, which were highly imputed because of missingness, we conducted 2 additional robustness checks that varied how the PROMIS variables were included in the EBM model. These were meant to isolate the effects of (1) including the nonmissing PROMIS predictors and (2) including their imputed values. The first model excluded the PROMIS scores as predictors altogether. The second model included the PROMIS scores, but instead of imputing missing values, we substituted missing values with a generic dummy value (−1). Generally, this strategy for handling missing continuous variables is nonfunctional. However, given the nonparametric nature of the EBM algorithm, this is a valid approach and allows us to visualize, for a given predictor, the mean risk among missing data compared with

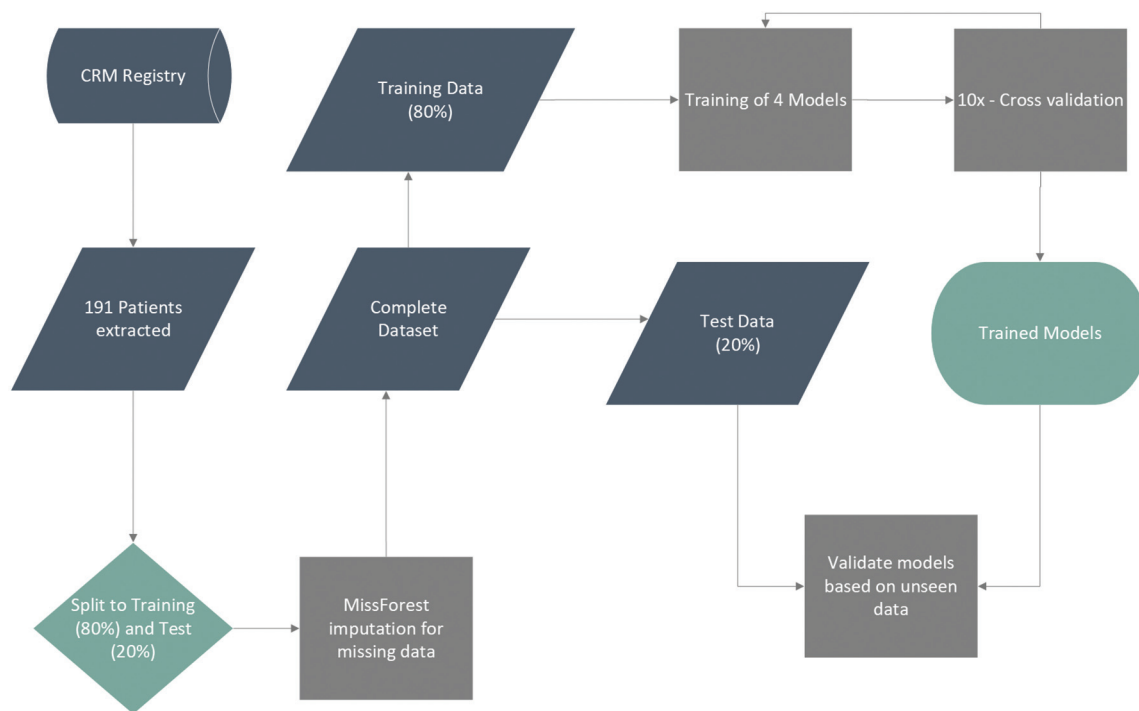


Figure 1. Illustrating the methodology for developing the machine learning models. CRM, Center for Regenerative Medicine.

the risk along the domain of nonmissing values. In this way, and unlike imputation, we can clearly see how risks differ between missing and nonmissing values.¹²

Selection of the Best Model

In selecting which prediction model to prioritize, predictive performance was prioritized along with model interpretability. The following metrics were calculated for performance evaluation: area under the receiver operating characteristic curve (AUC-ROC), *F1* score (harmonic mean of precision and recall), accuracy, sensitivity, and precision. In addition, we calculated confidence intervals for all performance metrics, showcasing the 95% confidence interval for our sample size, which is isomorphic to a power calculation. We performed all analysis and model construction utilizing Python Version 3.11 packages (lightgbm, matplotlib, numpy, pandas, sklearn, xgboost, and InterpretML).

Feature Importance

The feature weights are calculated by the model, representing how important each variable is in predicting the outcome. Furthermore, partial dependency plots were generated for the weighted features to examine the functional form of the relationship between predictors and outcome at the population level. These plots enable investigating the directionality and shape of the associations, providing important information for patients, surgeons, and other stakeholders.

RESULTS

Patient Characteristics and Outcomes

A total of 191 patients were included in the study, of whom 44% achieved our MCID thresholds at 6 months after PRP injection. The 191-patient cohort was randomly divided into a test set (39 patients) and a training set (152 patients). There were no significant differences between the training and test sets (Table 1).

Prediction Model

The EBM model was selected as the best model, exhibiting a comparable AUC-ROC of 0.81 (95% CI, 0.65-0.94) and *F1* score of 0.75 (95% CI, 0.57-0.88) relative to the XGBoost and RF models; however, it is a completely transparent model and therefore explainable. In contrast, the SVM model underperformed relative to the EBM, XGBoost, and RF models based on all metrics (Table 2).

Robustness checks revealed that the PROMIS scores were indispensable predictors. In the first robustness check model, excluding the PROMIS scores altogether from the EBM deteriorated model performance to no better than chance (AUC-ROC, 0.51 [95% CI, 0.32-0.7]). The second robustness check model (substituting missing values of the PROMIS scores with generic dummy values instead of imputed values) resulted in substantially reduced model performance (AUC-ROC, 0.53 [95% CI, 0.34-0.71]). This shows that most of the predictive power of the PROMIS scores was from the actual imputed scores themselves.

TABLE 2
Performance of Machine Learning Models Using Random Set^a

Model	F1 Score (95% CI)	Accuracy (95% CI)	Sensitivity (95% CI)	Precision (95% CI)	AUC-ROC (95% CI)
EBM	0.75 (0.57-0.88)	0.74 (0.59-0.90)	0.71 (0.50-0.90)	0.79 (0.59-0.96)	0.81 (0.65-0.94)
Random forest	0.70 (0.52-0.86)	0.72 (0.56-0.85)	0.62 (0.41-0.82)	0.81 (0.60-1.00)	0.78 (0.63-0.92)
XGBoost	0.65 (0.43-0.8)	0.67 (0.51-0.80)	0.57 (0.35-0.76)	0.75 (0.50-0.95)	0.82 (0.67-0.94)
SVM	0.57 (0.38-0.78)	0.62 (0.49-0.80)	0.48 (0.29-0.73)	0.71 (0.46-0.94)	0.70 (0.54-0.85)

^aAUC-ROC, area under the receiver operating characteristic curve; EBM, Explainable Boosting Machine; SVM, support vector machine; XGBoost, eXtreme Gradient Boosting.

and not from information contained in whether someone completed the PROMIS surveys or not (ie, missingness was less predictive than the imputed scores).

Feature Importance

The features that showed the greatest effect on the EBM model included the baseline PROMIS Mental score, KOOS JR, and PROMIS Physical score (Figures 2 and 3). Age, preoperative activity level, and KL grade were less influential (Figure 4). Examination of partial dependency plots showed that an increased probability of attaining the MCID threshold was associated with greater PROMIS Mental scores, a younger patient age, lower KOOS JR scores, a KL grade of 2, and female sex. Regarding age, risk increases were not uniformly distributed across the age spectrum. The probability of failure to achieve MCID thresholds increased between 79 and 81 years compared with younger patients. An interaction was also detected between KOOS JR scores and attainment of MCID thresholds, indicating a decreased probability of achieving MCID thresholds with higher KOOS JR scores. Additionally, female sex and a KL grade of 2 were associated with improved chances of reaching the MCID.

DISCUSSION

In this study, we developed an ML model to predict outcomes at 6 months after PRP injections for knee osteoarthritis. We evaluated 4 ML algorithms and found that the EBM algorithm had the best performance with an AUC-ROC of 0.81, indicating that the model accurately discriminates between patients who did and did not achieve MCID thresholds. We found that baseline, preinjection PROMs, including the PROMIS Mental score, KOOS JR, and PROMIS Physical score, were the most important predictors for achieving MCID thresholds at 6 months after PRP administration. Moreover, the associated partial dependency plots of several preinjection parameters showed nonlinear associations with the probability of achieving MCID thresholds, making it possible to potentially establish cutoff points that facilitate decision-making. To our knowledge, this is the first study using ML to predict which patients will respond to PRP injections for knee osteoarthritis.

Previous studies using multivariate regression analysis have explored the association of patient characteristics with the response to PRP treatment for knee osteoarthritis.^{1,4,10,19,20,22} Mild to moderate osteoarthritis and older age have been associated with positive responses to PRP,^{10,22} whereas previous PRP injection, a lower baseline visual analog scale pain score (VAS), bilateral knee injection, severe osteoarthritis, and a greater BMI have been associated with lower clinical improvement after PRP injection.^{1,10,19,20} In the present study, our ML model showed that the most important predictors were the baseline PROMIS Mental score and KOOS JR. To our knowledge, this was the first time the PROMIS Mental score was evaluated as a predictor for PRP response and associated with achievement of MCID thresholds after PRP. This result further suggests the critical importance of mental health assessment for the evaluation of outcomes in orthopaedic interventions, as previously discussed by Ayers et al.³ The KOOS JR showed an opposite association, in agreement with what was described by Prost et al¹⁹ for baseline VAS, with less symptomatic patients having a lower probability of responding to PRP, probably related to a ceiling effect in both scores. We found an association between older age and lower chances of achieving MCID thresholds after PRP injection. Our partial dependency plot showed that around 79 years of age, the probability of achieving MCID thresholds decreased. This observation disagrees with previous studies, which reported that older patients were more likely to respond positively to PRP treatment.¹⁰ These discrepancies could be partly explained because the population evaluated by Korpershoek et al¹⁰ received 3 consecutive injections of PRP (specifically autologous conditioned plasma); the patients studied were, on average, younger than the population analyzed in our study; and there were nuances in the definitions of MCID thresholds.

The definitions of “responders” and “nonresponders” after PRP injections for knee osteoarthritis have varied across the literature, ranging from 40% to 63%.^{4,6,8,11,20,23,24} It is important to note that, like the current study, some of the previous reports did not include a control group and therefore cannot be used to make claims about clinical benefit or efficacy, but rather achievement of predefined MCID thresholds. In the current study, 44% of the analyzed patient cohort met our MCID thresholds (a ≥ 10 -point increase in the KOOS JR and a $\geq 20\%$ decrease in the NPRS score) at 6 months after PRP

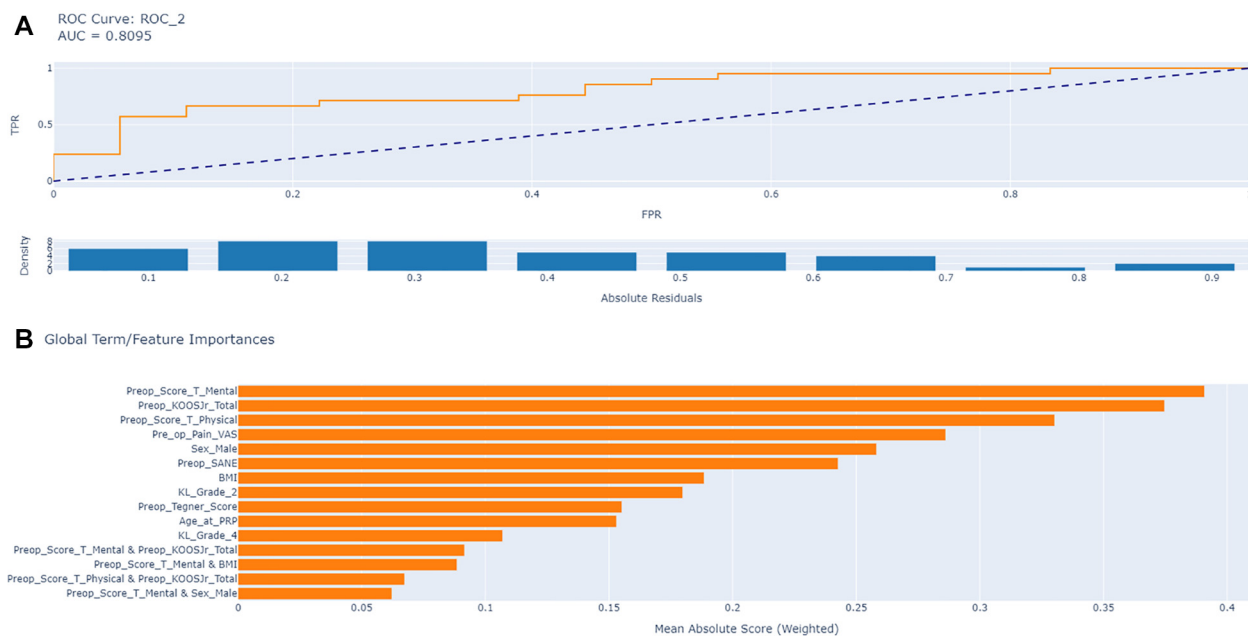


Figure 2. (a) Area under the receiver operating characteristic curve (AUC-ROC) for the machine learning model feature importance analysis. (b) Features ranked by importance for prediction (mean absolute scores). BMI, body mass index; FPR, false-positive rate; KOOS JR, Knee injury and Osteoarthritis Outcome Score for Joint Replacement; PRP, platelet-rich plasma; SANE, Single Assessment Numeric Evaluation for knee function; TPR, true-positive rate; VAS, visual analog scale pain score.

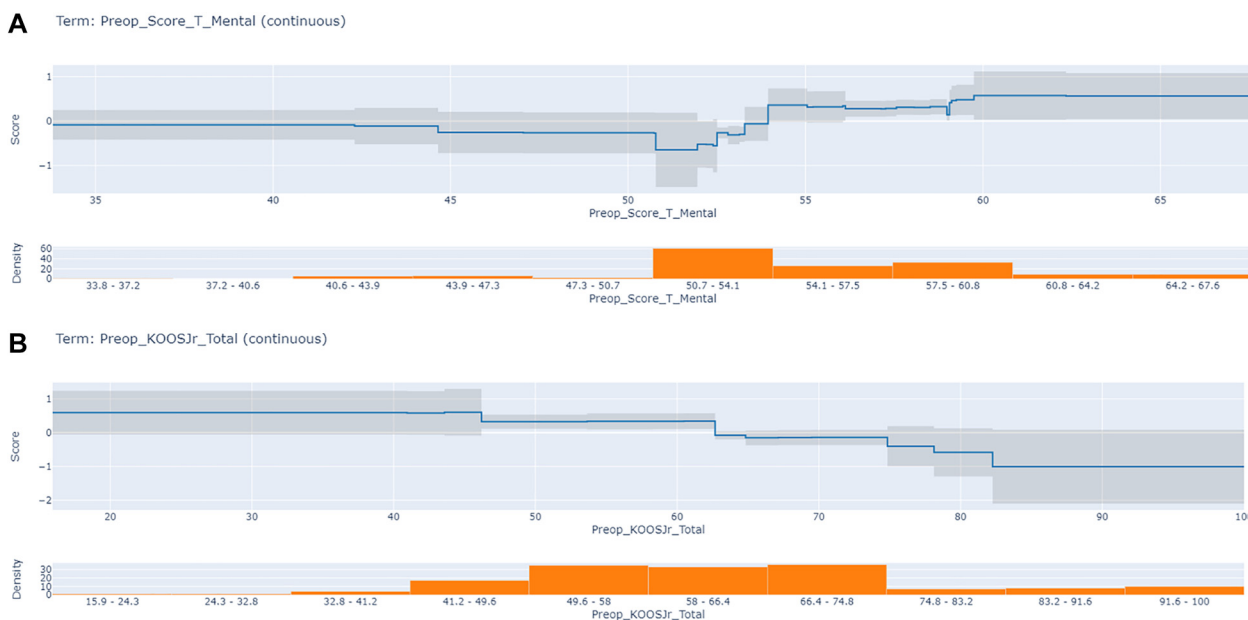


Figure 3. (a) Partial dependency plot of PROMIS (Patient-Reported Outcomes Measurement Information System) Mental score. (b) Partial dependency plot of the Knee injury and Osteoarthritis Outcome Score for Joint Replacement (KOOS JR).

injection. This percentage aligns with previous reports, highlighting that a significant proportion of patients do not achieve MCID thresholds and therefore emphasizing the need to improve treatment allocation by identifying patient candidates who are more likely to respond positively.

Recent studies have reported several accurate predictive models using ML within orthopaedics and sports medicine, forecasting outcomes after common orthopaedic procedures. For example, Martin et al¹⁵ developed a model predicting the risk of anterior cruciate ligament reconstruction (ACLR) revision surgery over 1, 2, and/or 5 years.

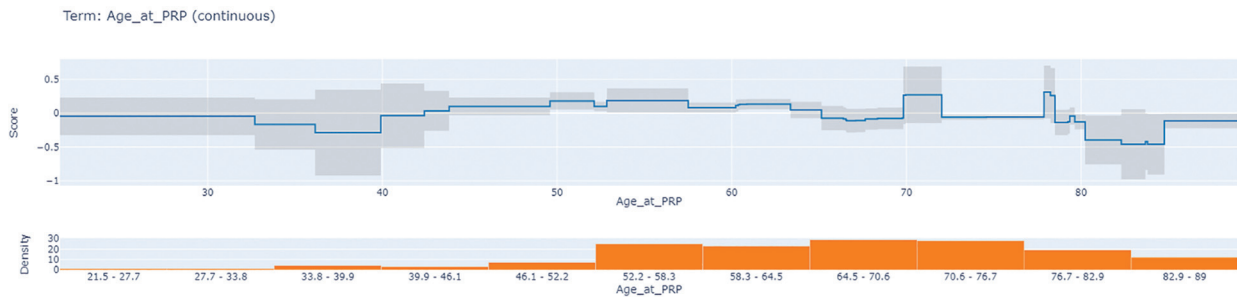


Figure 4. Partial dependency plot of age. PRP, platelet-rich plasma.

They found that Cox lasso (least absolute shrinkage and selection operator) and the generalized additive model were the best models, with a concordance index of 0.68. Similarly, Kunze et al¹¹ developed an ML model predicting the achievement of the MCID 2 years after undergoing ACLR, finding that the top-performing model was the elastic-net regularized logistic regression obtaining an AUC of 0.82. Haeberle et al⁷ took a different approach, using RF models to predict the risk of 4 different subsequent hip surgeries a minimum of 2 years after hip arthroscopic surgery for femoroacetabular impingement. Performance was strong for projecting the risk of a second hip arthroscopy (AUC, 0.77) and conversion to total hip arthroplasty (AUC, 0.88), while it was modest for resurfacing (AUC, 0.62) and periacetabular osteotomy (AUC, 0.76). Our model developed to predict patient response to PRP for knee osteoarthritis obtained an AUC-ROC of 0.81, which could be considered a good performance, placing it in the upper range of models previously developed in sports medicine.¹⁶ As ML continues to advance and more data become available, predictive models will likely become more accurate.

Limitations

This study has several limitations that should be considered when interpreting the results. Two of the evaluated variables (PROMIS Mental and Physical scores) had a high number of missing data, which were imputed using state-of-the-art methods. We performed 2 robustness checks on these variables, proving the imputation did not bias the model's performance. Although our model effectively predicted achievement of MCID thresholds after PRP treatment, a homogeneous cohort of 191 patients with knee osteoarthritis from a single institution represents a small sample size for developing ML models. Our high exclusion rate is due to a majority of patients not reaching the 6-month threshold at the time of analysis as well as poor survey response. Including patients and data from multiple centers could increase both the sample size and population diversity, thus improving the model's performance, generalizability, and external validity. Our models included 10 predefined variables but did not account for the characteristics of the PRP injected (number of injections, composition or volume injected, kit used, and injection procedure), state-of-the-art imaging metrics, or other


potentially relevant clinical information, as they were not available. Incorporating these variables could help us better define the optimal patient candidates for PRP treatment. It may well be that different PRP formulations have different outcomes across different patient profiles, which we did not account for in our study. Our study did not include a comparator (placebo or other treatment) group, which means that we cannot make claims about the efficacy of PRP treatment or its clinical benefit compared with other interventions or no treatment. Additionally, our analysis was limited to patients who received PRP and completed follow-up questionnaires, which may introduce selection bias if there are systematic differences between those who completed follow-up and those who did not. Finally, the mean BMI in our study population was 26.37; therefore, this model may not be generalizable to an obese population.


Despite these limitations, this study provides a roadmap with valuable methodology and insight into predictors of responses to PRP treatment for knee osteoarthritis that can potentially be extended to other orthobiologics and pathologies. While preliminary, our ML model provides a workflow that, when further developed and validated, may assist in clinical decision-making and treatment allocation, avoiding unnecessary procedures for patients unlikely to achieve MCID thresholds. It is crucial to emphasize that this tool should complement, not replace, clinical judgment, providing an additional evidence-based layer of information to enhance the decision-making process. It is important to note that we also believe that larger multicenter studies, including more patient diversity and a large number of demographic and PRP-specific variables, are needed to independently validate our results, and to improve our ability to allocate treatment by identifying patients with knee osteoarthritis who are more likely to benefit from PRP injections.

CONCLUSION

ML models can effectively predict clinically meaningful improvement at 6 months after PRP injection for knee osteoarthritis. The EBM was the algorithm with the best performance, with the PROMIS Mental and Physical scores and the KOOS Jr score being the most influential predictors. Additional independent studies are needed to externally validate this model.

ORCID iDs

Felix Conrad Oettl  <https://orcid.org/0000-0001-9721-685X>

Scott A. Rodeo  <https://orcid.org/0000-0002-0745-9880>

REFERENCES

- Alessio-Mazzola M, Lovisolo S, Sonzogno B, et al. Clinical outcome and risk factor predictive for failure of autologous PRP injections for low-to-moderate knee osteoarthritis. *J Orthop Surg*. 2021;29(2):230949902110219. doi:10.1177/23094990211021922
- Andia I, Maffulli N. Platelet-rich plasma for managing pain and inflammation in osteoarthritis. *Nat Rev Rheumatol*. 2013;9(12):721-730. doi:10.1038/nrrheum.2013.141
- Ayers DC, Franklin PD, Ring DC. The role of emotional health in functional outcomes after orthopaedic surgery: extending the biopsychosocial model to orthopaedics. *J Bone Joint Surg Am*. 2013;95(21):e165. doi:10.2106/bjbs.l.00799
- Boffa A, Andriolo L, Franceschini M, et al. Minimal clinically important difference and patient acceptable symptom state in patients with knee osteoarthritis treated with PRP injection. *Orthop J Sports Med*. 2021;9(10):232596712110262. doi:10.1177/23259671211026242
- Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): the TRIPOD Statement. *Br J Surg*. 2015;102(3):148-158. doi:10.1002/bjs.9736
- Filardo G, Previtali D, Napoli F, et al. PRP injections for the treatment of knee osteoarthritis: a meta-analysis of randomized controlled trials. *Cartilage*. 2021;13(1 suppl):364S-375S. doi:10.1177/1947603520931170
- Haeberle HS, Ramkumar PN, Karnuta JM, et al. Predicting the risk of subsequent hip surgery before primary hip arthroscopy for femoroacetabular impingement syndrome: a machine learning analysis of preoperative risk factors in hip preservation. *Am J Sports Med*. 2021;49(10):2668-2676. doi:10.1177/03635465211024964
- Hohmann E, Tetsworth K, Glatt V. Is platelet-rich plasma effective for the treatment of knee osteoarthritis? A systematic review and meta-analysis of level 1 and 2 randomized controlled trials. *Eur J Orthop Surg Traumatol*. 2020;30(6):955-967. doi:10.1007/s00590-020-02623-4
- Hung M, Bounsanga J, Voss MW, Saltzman CL. Establishing minimum clinically important difference values for the Patient-Reported Outcomes Measurement Information System Physical Function, hip disability and osteoarthritis outcome score for joint reconstruction, and Knee injury and Osteoarthritis Outcome Score for Joint Reconstruction in orthopaedics. *World J Orthop*. 2018;9(3):41-49. doi:10.5312/wjo.v9.i3.41
- Korpershoek JV, Vonk LA, Filardo G, et al. Effect of autologous conditioned plasma injections in patients with knee osteoarthritis. *Orthop J Sports Med*. 2023;11(7):23259671231184848. doi:10.1177/23259671231184848
- Kunze KN, Polce EM, Ranawat AS, et al. Application of machine learning algorithms to predict clinically meaningful improvement after arthroscopic anterior cruciate ligament reconstruction. *Orthop J Sports Med*. 2021;9(10):232596712110465. doi:10.1177/23259671211046575
- Kunze KN, So MM, Padgett DE, et al. Machine learning on Medicare claims poorly predicts the individual risk of 30-day unplanned readmission after total joint arthroplasty, yet uncovers interesting population-level associations with annual procedure volumes. *Clin Orthop Relat Res*. 2023;481(9):1745-1759. doi:10.1097/CORR.0000000000002705
- Ley C, Martin RK, Pareek A, et al. Machine learning and conventional statistics: making sense of the differences. *Knee Surg Sports Traumatol Arthrosc*. 2022;30(3):753-757. doi:10.1007/s00167-022-06896-6
- Luo W, Phung D, Tran T, et al. Guidelines for developing and reporting machine learning predictive models in biomedical research: a multidisciplinary view. *J Med Internet Res*. 2016;18(12):e323. doi:10.2196/jmir.5870
- Martin RK, Wastvedt S, Pareek A, et al. Ceiling effect of the combined Norwegian and Danish Knee Ligament Registers limits anterior cruciate ligament reconstruction outcome prediction. *Am J Sports Med*. 2023;51(9):2324-2332. doi:10.1177/03635465231177905
- Nahm FS. Receiver operating characteristic curve: overview and practical use for clinicians. *Korean J Anesthesiol*. 2022;75(1):25-36. doi:10.4097/kja.21209
- Neogi T. The epidemiology and impact of pain in osteoarthritis. *Osteoarthritis Cartilage*. 2013;21(9):1145-1153. doi:10.1016/j.joca.2013.03.018
- Oettl FC, Pareek A, Winkler PW, et al. A practical guide to the implementation of AI in orthopaedic research, part 6: how to evaluate the performance of AI research? *J Exp Orthop*. 2024;11(3):e12039. doi:10.1002/jeo2.12039
- Prost D, Bardot T, Baud A, et al. Long term improvement of knee osteoarthritis after injection of single high/very high volume of very pure PRP: a retrospective analysis of patients optimally managed in dedicated centers. *Regen Ther*. 2024;25:203-212. doi:10.1016/j.reth.2023.12.006
- Saita Y, Kobayashi Y, Nishio H, et al. Predictors of effectiveness of platelet-rich plasma therapy for knee osteoarthritis: a retrospective cohort study. *J Clin Med*. 2021;10(19):4514. doi:10.3390/jcm10194514
- Salaffi F, Stancati A, Silvestri CA, Ciapetti A, Grassi W. Minimal clinically important changes in chronic musculoskeletal pain intensity measured on a numerical rating scale. *Eur J Pain*. 2004;8(4):283-291. doi:10.1016/j.ejpain.2003.09.004
- Sánchez M, Jorquera C, De Dicastillo LL, et al. Real-world evidence to assess the effectiveness of platelet-rich plasma in the treatment of knee degenerative pathology: a prospective observational study. *Ther Adv Musculoskelet Dis*. 2022;14:1759720X2211003. doi:10.1177/1759720X221100304
- Zahir H, Dehghani B, Yuan X, et al. In vitro responses to platelet-rich-plasma are associated with variable clinical outcomes in patients with knee osteoarthritis. *Sci Rep*. 2021;11(1):11493. doi:10.1038/s41598-021-90174-x
- Zhang Y, Jordan JM. Epidemiology of osteoarthritis. *Clin Geriatr Med*. 2010;26(3):355-369. doi:10.1016/j.cger.2010.03.001