Machine Learning Model Identifies Increased Operative Time and Greater BMI as Predictors for Overnight Admission After Outpatient Hip Arthroscopy

Bryant M. Song, M.S., Yining Lu, M.D., Ryan R. Wilbur, B.A., Ophelie Lavoie-Gagne, M.D., Ayoosh Pareek, M.D., Brian Forsythe, M.D., and Aaron J. Krych, M.D.

Purpose: The purposes of this study were to identify patient characteristics and risk factors for overnight admission following outpatient hip arthroscopy and to develop a machine learning algorithm that can effectively identify patients requiring admission following elective hip arthroscopy. **Methods:** A retrospective review of a prospectively collected national surgical outcomes database was performed to identify patients who underwent elective outpatient hip arthroscopy from 2006 to 2018. Patients admitted overnight postoperatively were identified as those with length of stay of 1 or more days. Models were generated using random forest (RF), extreme gradient boosting (XGBoost), adaptive boosting (AdaBoost), elastic net penalized logistic regression, and an additional model was produced as a weighted ensemble of the four final algorithms. **Results:** Overall, 1,276 patients were included. The median age was 43 years, and 64.2% (819) were female. Of the included patients, 109 (8.5%) required an overnight stay following elective outpatient hip arthroscopy. The most important factors for inpatient admission were increasing operative time, general anesthesia, age extremes, male gender, greater body mass index (BMI), American Society of Anesthesiologists classification >1, and the following preoperative lab values outside of normal ranges: sodium, platelet count, hematocrit, and leukocyte count. The ensemble model achieved the best performance based on discrimination assessed via internal validation (area under the curve = .71), calibration, and decision curve analysis. The model was integrated into a Web-based open-access application able to provide both personalized predictions and explanations. **Conclusion:** A machine learning algorithm developed based on preoperative features identified increasing operative time, age extremes, greater BMI, sodium, hematocrit, platelets, and leukocyte count as the most important variables associated with inpatient admission with fair validity.

Introduction

Over the past decade, surgeons have increasingly relied on hip arthroscopy for the treatment of intra- and extra-articular hip pathologies.[1] Compared to open surgical hip dislocations, complication rates of hip arthroscopy are low, while surgical outcomes are at least equivalent.[2] Furthermore, greater exposure to hip arthroscopic procedures during specialty training and increased overall case volume contribute to this substantial rise.[3] Although the upward trend in the

incidence of femoroacetabular impingement (FAI) diagnoses and number of hip arthroscopies performed is multifactorial, overall awareness and understanding of hip pathology has increased, contributing to better clinician education for recognition and proper diagnosis.

Hip arthroscopy is one of the fastest growing fields in orthopaedic surgery.[4,5] Patients with labral pathology and concomitant FAI syndrome who are managed with hip arthroscopy are most commonly in the third decade of life, just as their economic contribution to society becomes most productive.[1] Cost-effectiveness analysis of outpatient hip arthroscopy has demonstrated a reduced 10-year societal cost of $67,418 per patient compared to nonoperative treatment.[6] Emphasis on the economic implications of hip arthroscopy is essential, as healthcare delivery continues to transition toward value-based care.

Surgery in the outpatient setting has multiple advantages, including fewer scheduling delays, decreased operating room turnover time, greater physician autonomy, and increased cost efficiency.[7] While hip arthroscopy is generally viewed as a safe procedure in the outpatient setting, studies have identified risk factors for perioperative and postoperative adverse events, including greater body mass index (BMI), increased surgical duration, and chronic corticosteroid *use*.[8,9] Overnight admission following elective outpatient hip arthroscopy has implications on both clinical outcomes as well as cost, particularly due to fees from inpatient services. Identifying patients that may require inpatient admission after hip arthroscopy allows surgeons to counsel patients regarding their postoperative course, and perhaps work on future strategies to mitigate overnight admission. Therefore, the purpose of this study was twofold: to identify patient characteristics and risk factors for overnight admission following outpatient hip arthroscopy and to develop a machine learning algorithm that can effectively identify patients requiring admission following elective hip arthroscopy. We hypothesized that increased operative time, increased age, and greater BMI would be identified by our models as contributing risk factors for overnight admission.

Materials and Methods

The present study was exempt from review by our institutional review board due to the use of an anonymized national database. The analysis was adherent to the transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) guidelines and the Guidelines for Developing and Reporting Machine Learning Models in Biomedical Research.[10,11]

Prospectively collected data from the American College of Surgeons National Surgical Quality Improvement Program (NSQIP) database were used. The database was reviewed for patients undergoing elective outpatient hip arthroscopy between 2006 and 2018. Patients were identified using Current Procedural Terminology (CPT) codes and International Classification of Diseases, Ninth Revision (2005-2015) and Tenth Revision (2015-2016) (ICD-9 and ICD-10, respectively). The CPT codes used were 29860, 29861, 29862, 29863, 29914, 29915, and 29916. Unspecified arthroscopic procedures identified using CPT code 29999 were reviewed manually and included if they had an ICD-9 or ICD10 code indicating hip pathology. The ICD-9 and ICD-10 codes used may be found in Supplemental Table S1.

Overnight admission was defined as patients with a total length of stay (LOS) $\geq$1 day, and those with LOS <1 day were considered to have been discharged the same day. Exclusion criteria included nonelective surgery, concurrent open hip procedures, surgery scheduled as inpatient, patients with infection, emergency cases, and trauma cases. Features that were missing in >35% of the cohort and, therefore, cannot be imputed, were excluded from the analysis.

The variables with less than 35% of missing data were imputed using the Miss Forest multiple imputation method.[12] Exclusion of cases with missing data risks the introduction of biases and a substantial loss of precision and power in the analysis.[13] Multiple imputation is a commonly used and effective method to address missingness in a data set and mitigate these risks.[14]

The primary outcome of interest was overnight admission, defined as LOS $\geq$1 day, which was represented as a binary variable. Following imputation for missing data, feature selection using recursive feature elimination with naïve Bayesian algorithms was used to select preoperative variables that were associated with overnight admission. The variables considered for feature selection and modeling are described in Table 1. After feature selection, modeling was performed using the selected features. The following machine learning algorithms were chosen on the basis of based on prior investigations:[15,16] random forest (RF), extreme gradient boosting (XGBoost), adaptive boosting (Ada-Boost), elastic net penalized logistic regression, and an additional model produced as a weighted ensemble of the four algorithms. Ensemble methods combine the learning techniques of individual models, and, although more memory intensive, ensemble models have been shown to improve predictions, as well as decrease bias and variance.[17] The features included in the final models are described in Table 1. Models were trained and validated via .632 bootstrapping with 1,000 resampled data sets. In brief, model evaluation consists of reiterative partitions of the complete data set into train and test sets. For each combination of train and test set, the model is trained on the train set using 10-fold cross validation repeated 3 times. The performance of this model is then evaluated on the respective

Table 1. Baseline Characteristics of Study Population, $n = 1,276$

Variable	n (%) | Median (IQR)	n (%) Missing	Outpatient Discharge $n = 1167$	Overnight Admission $n = 109$	P Value
Demographics and Intraoperative Variables					
Age, years	43 (34-53)	30 (2.3)	43.5	47.8	.003
Sex		28 (2.2)			.115
Female	819 (64.2)		741 (63.5)	78 (71.6)	
Male	457 (35.8)		426 (36.5)	31 (28.4)	
ASA Class		31 (2.4)			.001
1: No disturbance	269 (21.1)		259 (22.2)	10 (9.2)	
2: Mild disturbance	761 (59.6)		696 (59.6)	65 (59.6)	
3: Severe disturbance	244 (19.1)		210 (18)	34 (31.2)	
4: Life threatening	2 (0.2)		2 (.2)	0	
Body mass index	27.5 (24.2-31.5)	35 (2.7)	28.3	29.9	.005
Dependent functional status	49 (3.8)	28 (2.2)	46 (3.9)	3 (2.8)	.721
Dyspnea	27 (2.1)		22 (1.9)	5 (4.6)	.127
Anesthesia		28 (2.2)			1.000
General	1226 (96.1)		1121 (96.1)	105 (96.3)	
Spinal	50 (3.9)		46 (3.9)	4 (3.7)	
Operative time, minutes	80 (56-109)	28 (2.2)	86.1	94.4	.063
Comorbidities					
Diabetes	84 (6.6)	28 (2.2)	75 (6.4)	9 (8.3)	.593
Smoke	248 (19.4)	28 (2.2)	221 (18.9)	27 (24.8)	.179
Chronic obstructive pulmonary disease	24 (1.9)	28 (2.2)	19 (1.6)	5 (4.6)	.071
Medicated hypertension	306 (24.0)	28 (2.2)	266 (22.8)	40 (36.7)	.002
Chronic steroid use	35 (2.7)		34 (2.9)	1 (0.9)	.361
Coagulopathy	10 (0.8)		7 (0.6)	3 (2.8)	.062
Preoperative laboratory values					
Leukocyte count	6.7 (5.7-7.81)	238 (18.7)	6.6 (5.7-7.8)	6.9 (5.5-8.2)	.093
Platelet	240.1 (212-270)	238 (18.7)	241 (212-269)	236 (207-284)	.058
Hematocrit	41 (39.3-43.6)	199 (15.6)	41 (39.3-43.7)	41 (39-43.1)	.075
Sodium	139.6 (138.8-141)	409 (32.1)	139.51 (138.8-141)	140 (138.86-141)	.145
Overnight admission					
Yes	109 (8.5)				
No	1167 (91.5)				

test set. This sequence of steps is then repeated for 999 more data partitions. The model is, thus, trained and tested on all data points available, and evaluation metrics are summarized with standard distributions of values. Bootstrapping has been shown to improve both model variance and bias when compared to internal validation by train-test split. The optimal model was chosen based on area under the receiver operating characteristics curve (AUROC). Models were compared by discrimination, calibration, and Brier score values. The utility of the ensemble model was compared with a logistic regression model with two separate default strategies. The logistic regression model is fitted with the same inputs and serves as a traditional statistics benchmark against which the machine learning methods are evaluated, while the two default strategies represented changing the management decision for all patients or no patients, respectively.

Discriminative power was assessed via the AUROC. Models that assign the correct label for every output have an AUROC of 1, while completely random predictions have an AUROC of .5. Calibration of the model's predicted probabilities as a function of observed frequencies within the test population are summarized in a calibration plot. An ideal model is a straight line with intercept 0 and slope of 1. Finally, the mean squared difference between predicted probabilities of models and observed outcomes, known as the Brier score, was assessed for each candidate model with smaller values considered more optimal. The candidate algorithm Brier scores are then compared to the Brier score of the null model, which assigns a predicted probability equal to the outcome prevalence in the study population.

Individual explanations for model behavior were provided for transparency into the output using local-interpretable model-agnostic explanations (LIME).[18] Decision curve analysis was used to determine the benefit of implementing the predictive algorithm in practice. The curve plots net benefit against the predicted probabilities of the outcome of interest, in this case, overnight admission, and provides the cost-benefit ratio for every value of the predicted probability. These ratios provide useful guidance for individualized decision-making and accounts for variability in clinician and patient's thresholds for what is considered high risk.

Partial dependence curves were constructed to demonstrate the independent impact of individual continuous variables to the model's output with all
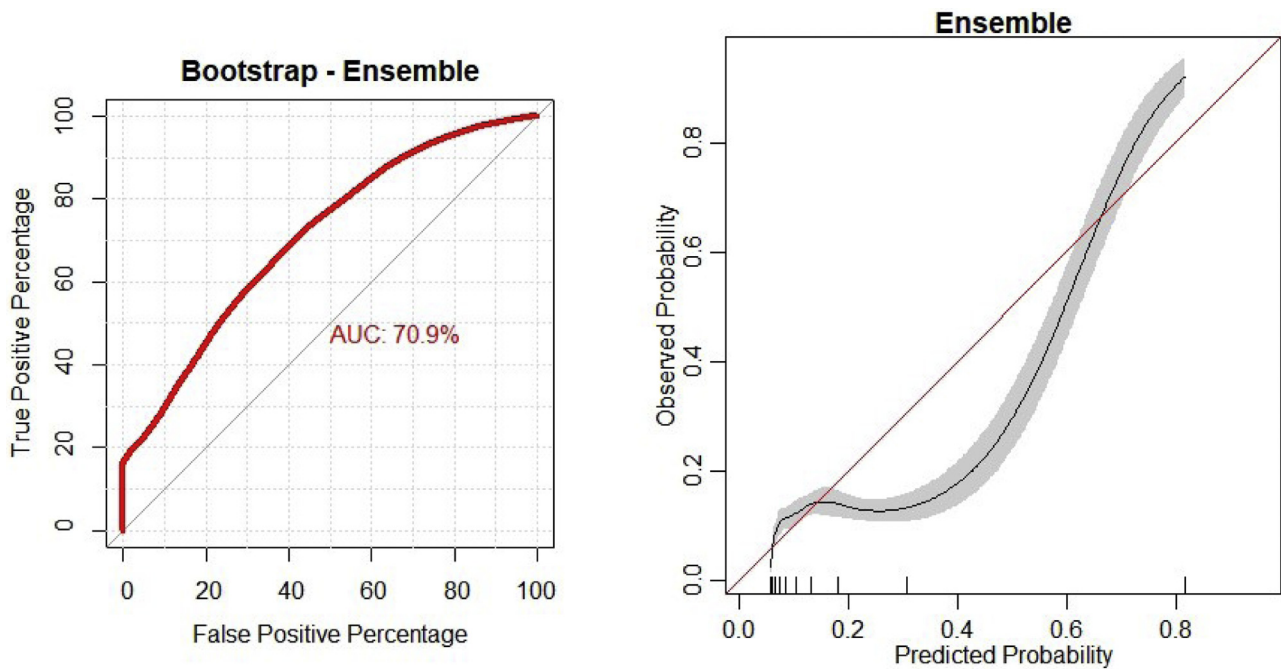
Fig 1. Discrimination and calibration of the ensemble model.

other variables held constant. Dependence on the *y*-axis represents the relative contribution across the range of values of the continuous variable to the predicted probability of the outcome of interest. For each variable, the model tests the range of input values to observe how the prediction changes while holding other input features constant. Reference ranges for normal laboratory values were obtained from the literature.[19]

The candidate algorithm with the best performance was integrated into an interactive, open-access, Web-based application. Clinician input is used to generate outcome predictions (percentage change of overnight admission) with accompanying explanations. All data analysis was performed using RStudio version 1.2.5001 (RStudio, Boston, MA).

Results

Variable Breakdown

A total of 1,276 patients who underwent hip arthroscopy were included after the application of exclusion criteria. The median age was 43 years (IQR 34-53) and 64.2% (819) were female. The most common comorbidities in our population were hypertension (24%) and smoking status (19.4%). The variables identified for model construction, as well as summary statistics and missing data, are presented in Table 1.

In total, 8.5% (109) patients were admitted for at least one night following elective outpatient hip arthroscopy. Feature selection identified the following features as most important globally to the predictive

capabilities of the model: operative time, age, BMI, and the following preoperative lab values: hematocrit, platelet count, sodium, and leukocyte count. Anesthesia type, gender, smoking history, and ASA class were identified as less important variables.

Model Performance

The ensemble model demonstrated the best performance in generating predictions based on internal validation with an area under the curve (AUC) of .71, a calibration intercept of .001, a calibration slope of .993, and a Brier score of .068 (Fig 1). Full details for individual model performance are presented in Table 2.

Decision Curve Analysis

The utility of the ensemble model was compared with a logistic regression model and with two separate default strategies. Upon analysis, the net benefit of applying the ensemble machine learning model was significantly greater than each of the three alternatives (Fig 2).

Explanations

For a given patient, the model generated the probability of admission and specific contributing variables; three examples are provided in Fig 3.

Partial Dependence Analysis

Partial dependence curves were constructed for age, BMI, operative time, hematocrit, white blood cell (WBC) count, platelets, and sodium (Fig 4). Higher operative time was associated with increased risk of admission. Greater BMI also increased the risk for overnight

Table 2. Model Assessment on Internal Validation Using .632 Bootstrapping With 1,000 Resampled Datasets, $n = 1,276$

Metric	Area under the Curve		Calibration Slope	Calibration Intercept	Brier Score
	Apparent	Internal Validation			
GLM	.697 (.695-.701)	.678 (.676-.680)	.932 (.920-.943)	.01 (.008-.012)	.12 (.107-.133)
Elastic net	.684 (.680-.687)	.664 (.662-.666)	.993 (.981-1.005)	.001 (−.001-.003)	.188 (.179-.197)
Random forest	.835 (.831-.839)	.707 (.704-.710)	1.03 (1.02-1.04)	−.004 (−.006-.003)	.107 (.095-.119)
XGBoost	.824 (.819-.827)	.708 (.702-.720)	1.01 (1.003-1.02)	−.001 (−.003-0)	.106 (.093-.118)
Adaptive boost	.827 (.823-.831)	.725 (.723-.727)	1.056 (1.049-1.063)	−.008 (−.01-.007)	.107 (.095-.12)
Ensemble	.850 (.846-.854)	.709 (.710-.712)	.993 (.981-1.005)	.001 (−.001-.003)	.068 (.057-.079)

GLM, generalized linear model. Null model Brier score = .085.

admission. Hematocrit (normal range: male 41-53%, female 36-46%), WBC count (4.5-11.0 × 10^3/mm^3), platelet count (150-350 × 10^3/mm^3), and sodium (136-145 mEq/L) outside of normal ranges were associated with increased risk of admission.

The final model was integrated in a Web-based digital application readily accessible on desktops, tablets, and smartphones at this web address: https://sportsmed.shinyapps.io/Hip_Arthroscopy_Admission/. All variables must be entered into the model to generate predictions and explanations, and default values are provided as placeholders.

Discussion

The principal findings of the current investigation include the following: first, feature elimination identified the most important variables in predicting overnight admission following hip arthroscopy: increased operative time, general anesthesia, age extremes, greater BMI, American Society of Anesthesiologists (ASA) classification >1, and the following preoperative lab values: sodium, hematocrit, platelets, and leukocyte count. Second, using an ensemble model achieved the best performance based on discrimination assessed via internal validation (AUC = .71), calibration, and decision curve analysis. Lastly, this model was integrated into a Web-based open-access application capable of providing both predictions and explanations for education and demonstration.

Orthopedic surgeons have witnessed a paradigm shift in the management of FAI and hip labral pathology, with a reported 365% increase in utilization of hip arthroscopy since 2004.[1,5,20] As the prevalence of hip arthroscopy procedures continues to increase, it is important to identify patient characteristics and risk factors related to perioperative and postoperative complications.
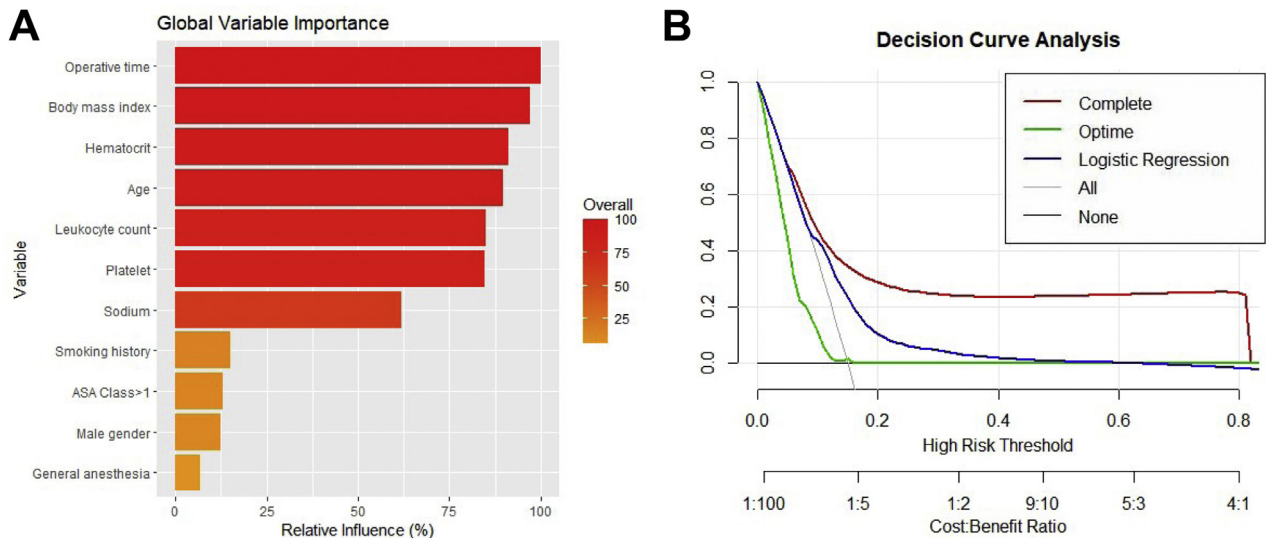

Fig 2. (A) Global variable importance plot demonstrating the relative contribution of an input variable on the overall predictive performance of the ensemble model. (B) Decision curve analysis comparing the complete ensemble algorithm with conventional logistic regression and a simplified ensemble model using only preoperative use as a predictor. The downsloping line marked by "All" plots the net benefit from the default strategy of changing management for all patients, while the horizontal line marked "None" represents the strategy of changing management for none of the patients (net benefit is zero at all thresholds). The "all" line slopes down because at a threshold of zero, false positives are given no weight relative to true positives; as the threshold increased, false positives gain increased weight relative to true positives and the net benefit for the default strategy of changing management for all patients decreases.
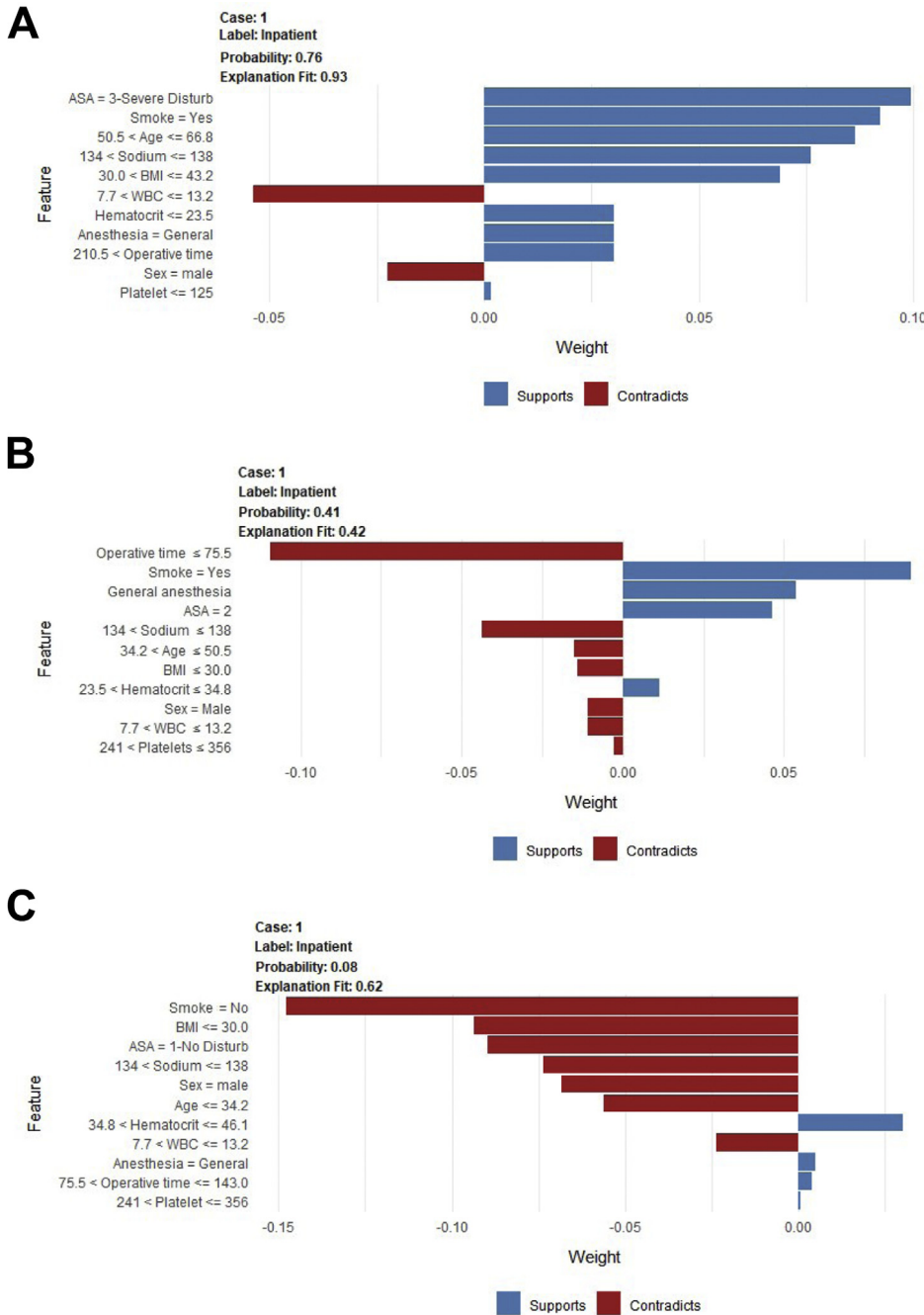
A

Case: 1
Label: Inpatient
Probability: 0.76
Explanation Fit: 0.93

B

Case: 1
Label: Inpatient
Probability: 0.41
Explanation Fit: 0.42

C

Case: 1
Label: Inpatient
Probability: 0.08
Explanation Fit: 0.62

Fig 3. The example patients presented here represent high (patient A), median (patient B), and low (patient C) probability of inpatient admission. The label for each case is set to in-patient admission, and the probability indicates how likely it is that the patient will need in-patient admission. Features are noted as supporting (blue bars) or contradicting (red bars) the prediction, and the scale on the x-axis represents each feature's relative contribution to the prediction for each individual patient. Patient A has a probability of .76 for overnight admission, which is most supported by an American Society of Anesthesiologists (ASA) class of 3, a positive smoking history, and age between 50.5 and 66.8 years; patient A's normal white blood cell count and male sex contradict the prediction. Patient B has a .41 probability of admission, which is most contradicted by an operative time of less than 75 minutes but supported by a positive smoking history, the use of general anesthesia, and an ASA class of 2. Patient C has a .08 probability of admission, which is contradicted by a negative smoking history, body mass index (BMI) less than 30 kg/m^2, an ASA class of 1, normal sodium levels, and male sex. As presented here, patients can be counseled that they can adjust modifiable risk factors (BMI, smoking history, etc.) to reduce their risk of postoperative hospitalization.

Several investigations have used conventional statistical modeling to determine risk factors for inpatient admission after hip arthroscopy.[8,9,21,22] Du et al. identified 1,931 patients who underwent hip arthroscopy and the rate of 30-day readmission following index procedure was estimated to be 0.9%.[8] Multivariate analysis identified increasing BMI, chronic corticosteroid use, and perioperative blood transfusions as independent risk factors associated with unplanned admission within 30 days after surgery.[8] A similar investigation by Hartwell et al. reported that hypertension requiring antihypertensive medication and chronic corticosteroid or immunosuppressant use were significant risk factors for 30-day readmission following hip arthroscopy.[23] Although these two investigations identified patient characteristics related to unplanned admission in the postoperative period, the present investigation focused on comprehensive risk factors for overnight admission following hip arthroscopy. Operative time and greater BMI were identified as the most important variables for predicting overnight admission.
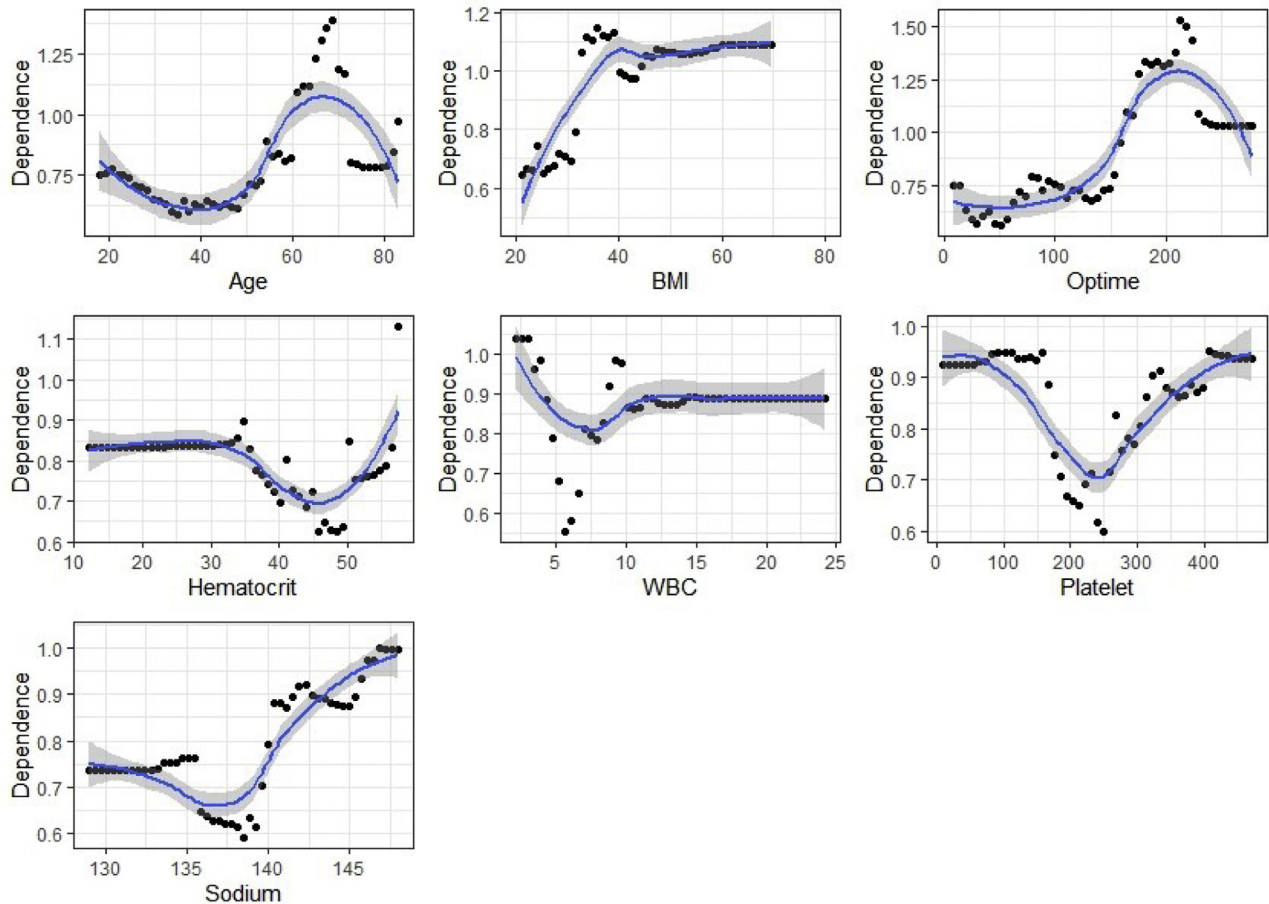
Fig 4. Partial dependence curves demonstrating the relationship between continuous variables and admission risk. The y-axis represents the relative contribution of the variable to the probability of in-patient admission, and the x-axis represents the range of values for the variable.

The present study identified increased operative time as a risk factor for overnight admission following surgery; however, the findings from previous investigations are variable. Sivasundaram et al. reported that increased operating room time was not a significant risk factor for postoperative emergency department use.[22] Two other investigations also using the NSQIP database demonstrated conflicting results regarding the role of operative time.[8,9] Du et al. found that increased operative time was not significantly associated with 30-day unplanned admission.[8] On the other hand, Bovonratwet et al. focused specifically on overnight admission after hip arthroscopy and concluded that operative time of 60-120 min was an independent risk factor for overnight hospital admission.[9] The authors reported an overall readmission rate of 14%, which is comparable to the 9% overnight admission rate from the present study. The findings of our study concur with Bovonratwet et al.: Increased operative time was a strong positive predictor of overnight hospital admission after outpatient hip arthroscopy.

BMI was also identified as an important contributor to the model. Increased BMI is a well-studied risk factor for postoperative complications and poor outcomes following orthopedic surgical procedures, including arthroscopy, which is further reinforced by the current findings.[24-26] Kashanchi et al. used the NSQIP database to identify 18,521 patients undergoing arthroscopic rotator cuff repair and determined that patients with a BMI greater than 30 had significantly higher rates of medical complications, readmissions rates, and non-home discharge.[27] A recent retrospective investigation of 452 patients treated with hip arthroscopy by Seijas et al. determined that BMI above 30 significantly increased the risk of conversion to arthroplasty 5.5-fold.[28] These findings are consistent with large database investigations by Seijas et al., Schairer et al., and Kester et al.[28-30] Similar to the poorer long-term outcomes following hip arthroscopy and the poorer short-term outcomes following shoulder arthroscopy, the present study found higher BMI to be a strong positive predictor of overnight admission following hip arthroscopy.

The literature on machine learning models for predicting clinical outcomes after hip arthroscopy is limited. Two previous investigations used machine learning algorithms to predict clinically significant improvements in function and satisfaction at 2 years postoperatively in patients undergoing primary hip arthroscopy.[16,31] The current study identified patient factors that increase their risk for overnight admission. By identifying such patients, it allows potential for intervention to improve their current health and overall risk profile before hip arthroscopy, thus mitigating the likelihood of an overnight stay.

Artificial intelligence, namely machine learning, has transformed numerous industries by providing powerful computational tools for improved efficiency and supportive decision making. These computational algorithms analyze training sets for pattern recognition from imputed data to classify and predict outcomes that would otherwise not be elucidated through conventional statistical models. While utilization of machine learning models in orthopedics remains nascent, a 2018 systematic review reported an increase in applications of machine learning models, namely, focusing on osteoarthritis prediction and detection, bone and cartilage imaging, and spinal pathology detection.[32] Particularly in sports medicine, machine learning application has shown promise in automated imaging analysis and risk stratification.[33] A recent primer highlights the potential application of machine learning in sports medicine as a physician aid, emphasizing that the technology should be used to augment physician capabilities rather than replace them.[34] Future implementation of this technology in practice necessitates a fundamental working knowledge of the strengths, limitations, and applications of artificial intelligence-based tools.

Limitations

Several important limitations must be considered when interpreting these findings. First, because of the nature of the NSQIP database, the findings are limited to information that is retrievable within the database. Surgeon-specific information, including surgical volume, and patient-specific information, including severity of comorbidities, indications for surgeries, reasons for admission, psychosocial and socioeconomic factors, and current medications, was not readily available. As a result, these variables could not be included as input variables in our model for analysis. Secondly, while NSQIP-participating hospitals generally have laboratory service routinely available, many surgeons are no longer collecting routine labs for hip arthroscopy patients. The model may, therefore, not be as applicable without this information readily available. Additionally, ambulatory surgery centers were not included in the database. Lastly, the performance of algorithms for both prediction and explanation is dependent on the available training data, and the model underestimates hospitalization at 15-65% risk. Although the model demonstrated fair internal validity, optimization of the model through rigorous external validation must be performed prior to widespread implementation.

Conclusion

A machine learning algorithm developed on the basis of preoperative features identified increasing operative time, age extremes, greater BMI, sodium, hematocrit, platelets, and leukocyte count as the most important variables associated with inpatient admission with fair validity.

Disclosures

The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health.

References

1. Hale RF, Melugin HP, Zhou J, et al. Incidence of femoroacetabular impingement and surgical management trends over time. *Am J Sports Med* 2021;49:35-41.
2. Matsuda DK, Carlisle JC, Arthurs SC, Wierks CH, Philippon MJ. Comparative systematic review of the open dislocation, mini-open, and arthroscopic surgeries for femoroacetabular impingement. *Arthroscopy* 2011;27:252-269.
3. Gordon AM, Flanigan DC, Malik AT, Vasileff W. Orthopaedic surgery sports medicine fellows see substantial increase in hip arthroscopy procedural volume with high variability from 2011 to 2016. *Arthroscopy* 2021;37:521-527.
4. Bonazza NA, Homcha B, Liu G, Leslie DL, Dhawan A. Surgical trends in arthroscopic hip surgery using a large national database. *Arthroscopy* 2018;34:1825-1830.
5. Byrd JWT. Editorial commentary: Hip arthroscopy—A microcosm in the evolution of arthroscopy in sports medicine. *Arthroscopy* 2020;36:773-775.
6. Mather RC 3rd, Nho SJ, Federer A, et al. Effects of arthroscopy for femoroacetabular impingement syndrome on quality of life and economic outcomes. *Am J Sports Med* 2018;46:1205-1213.
7. Crawford DC, Li CS, Sprague S, Bhandari M. Clinical and cost implications of inpatient versus outpatient orthopedic surgeries: A systematic review of the published literature. *Orthop Rev (Pavia)* 2015;7:6177.
8. Du JY, Knapik DM, Trivedi NN, et al. Unplanned admissions following hip arthroscopy: Incidence and risk factors. *Arthroscopy* 2019;35:3271-3277.
9. Bovonratwet P, Boddapati V, Nwachukwu BU, Bohl DD, Fu MC, Nho SJ. Increased hip arthroscopy operative duration is an independent risk factor for overnight hospital admission. *Knee Surg Sports Traumatol Arthrosc May* 2021;29:1385-1391.
10. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement. *Br J Surg* 2015;102:148-158.

11. 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:e323.

12. Stekhoven DJ, Bühlmann P. MissForest—Non-parametric missing value imputation for mixed-type data. *Bioinformatics (Oxford, England)* 2012;28:112-118.

13. Hughes JD, Hughes JL, Bartley JH, Hamilton WP, Brennan KL. Infection rates in arthroscopic versus open rotator cuff repair. *Orthopaed J Sports Med* 2017;5:2325967117715416.

14. Huque MH, Carlin JB, Simpson JA, Lee KJ. A comparison of multiple imputation methods for missing data in longitudinal studies. *BMC Med Res Methodol* 2018;18:168.

15. Karhade AV, Schwab JH, Bedair HS. Development of machine learning algorithms for prediction of sustained postoperative opioid prescriptions after total hip arthroplasty. *J Arthroplasty* 2019;34:2272-2277.e1.

16. Kunze KN, Polce EM, Rasio J, Nho SJ. Machine learning algorithms predict clinically significant improvements in satisfaction after hip arthroscopy. *Arthroscopy* 2021;37:1143-1151.

17. Dietterich TG. Ensemble methods in machine learning. *Lect Notes Comput Sc* 2000;1857:1-15.

18. Kumarakulasinghe NBBT, Liu J, Leao AS, Papapetrou P. Evaluating local interpretable model-agnostic explanations on clinical machine learning classification models, *In 2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS)*. Rochester, MN: IEEE, 2020;7-12.

19. Kratz A, Ferraro M, Sluss PM, Lewandrowski KB. Normal reference laboratory values. *New Engl J Med* 2004;351:1548-1563.

20. Montgomery SR, Ngo SS, Hobson T, et al. Trends and demographics in hip arthroscopy in the United States. Arthroscopy 29:661-665. doi:10.1016/j.arthro.2012.11.005.

21. Malviya A, Raza A, Jameson S, James P, Reed MR, Partington PF. Complications and survival analyses of hip arthroscopies performed in the national health service in England: A review of 6,395 cases. *Arthroscopy* 31:836-842. doi:10.1016/j.arthro.2014.12.013.

22. Sivasundaram L, Trivedi NN, Kim CY, et al. Emergency department utilization after elective hip arthroscopy. Arthroscopy 36:1575-1583. doi:10.1016/j.arthro.2020.02.008.

23. Hartwell MJ, Morgan AM, Johnson DJ, et al. Risk factors for 30-day readmission following hip arthroscopy. *Knee Surg Sports Traumatol Arthrosc* 2020;28:1290-1295.

24. Katakam A, Florissi IS, Colon Iban YE, et al. Class III obesity increases risk of failure to achieve the 1-year hip disability and osteoarthritis outcome score—Physical function short form minimal clinically important difference following total hip arthroplasty. *J Arthroplasty* 2021;36:187-192.

25. Katakam A, Collins AK, Sauder N, et al. Obesity increases risk of failure to achieve the 1-year PROMIS PF-10a minimal clinically important difference following total joint arthroplasty. *J Arthroplasty* 2021;36:S184-S191.

26. Macchi M, Spezia M, Elli S, Schiaffini G, Chisari E. Obesity increases the risk of tendinopathy, tendon tear and rupture, and postoperative complications: A systematic review of clinical studies. *Clin Orthop Relat Res Aug* 2020;478:1839-1847.

27. Kashanchi KI, Nazemi AK, Komatsu DE, Wang ED. Level of obesity is directly associated with complications following arthroscopic rotator cuff repair. *J Shoulder Elbow Surg* 2021;30:1581-1587.

28. Seijas R, Barastegui D, Lopez-de-Celis C, et al. Preoperative risk factors in hip arthroscopy. *Knee Surg Sports Traumatol Arthrosc* 2021;29:1502-1509.

29. Schairer WW, Nwachukwu BU, McCormick F, Lyman S, Mayman D. Use of hip arthroscopy and risk of conversion to total hip arthroplasty: A population-based analysis. *Arthroscopy* 2016;32:587-593.

30. Kester BS, Capogna B, Mahure SA, Ryan MK, Mollon B, Youm T. Independent risk factors for revision surgery or conversion to total hip arthroplasty after hip arthroscopy: A review of a large statewide database from 2011 to 2012. *Arthroscopy* 2018;34:464-470.

31. Kunze KN, Polce EM, Nwachukwu BU, Chahla J, Nho SJ. Development and internal validation of supervised machine learning algorithms for predicting clinically significant functional improvement in a mixed population of primary hip arthroscopy. *Arthroscopy* 2021;37:1488-1497.

32. Cabitza F, Locoro A, Banfi G. Machine learning in orthopedics: A literature review. *Front Bioeng Biotechnol* 2018;6:75.

33. Ramkumar PN, Kunze KN, Haeberle HS, et al. Clinical and research medical applications of artificial intelligence. *Arthroscopy* 2021;37:1694-1697.

34. Ramkumar PN, Luu BC, Haeberle HS, Karnuta JM, Nwachukwu BU, Williams RJ. Sports medicine and artificial intelligence: A primer. *Am J Sports Med* In press.

Appendix Table 1. Inputs Considered for Feature Selection

Demographics
  Gender
  Race
  Hispanic ethnicity
  Age
  BMI
Comorbidities
  ASA classification
  Diabetes mellitus w/ oral agents or insulin
  Current smoker within one year
  Dyspnea
  Functional health status prior to surgery
  Hypertension requiring medication
  Acute renal failure
  Disseminated cancer
  Open wound/wound infection
  Steroid use for chronic condition
  >10% loss of body weight in last 6 months
  Bleeding disorders
  History of COPD
  Ascites
  Congestive heart failure 30 days before surgery
Preoperative Lab Values
  Serum sodium
  BUN
  Serum creatinine
  Serum albumin
  Total bilirubin
  SGOT
  Alkaline phosphatase
  WBC
  Hematocrit
  Platelet count
  PTT
  International normalized ratio of PT values
  PT
Operative Variables
  Principal anesthesia technique
  Inpatient/outpatient status
  Transfer status
  Total operation time
  Quarter of admission
  Concomitant loose body removal
  Concomitant synovectomy
  Femoroplasty
  Acetabuloplasty
  Labral repair

BUN, blood, urea, nitrogen; PT, prothrombin time; PTT; partial thromboplastin time; SGOT, serum glutamic oxaloacetic transaminase; WBC, white blood cell.