Arthroplasty Today 22 (2023) 101166

Contents lists available at ScienceDirect

Arthroplasty Today

journal homepage: http://www.arthroplastytoday.org/

Original research

Preoperative Prediction and Risk Factor Identification of Hospital Length of Stay for Total Joint Arthroplasty Patients Using Machine Learning

Jaeyoung Park, PhD [a], Xiang Zhong, PhD [b, *], Emilie N. Miley, DAT, ATC [c], Chancellor F. Gray, MD [c]

[a] Booth School of Business, University of Chicago, Chicago, IL, USA
[b] Department of Industrial and Systems Engineering, University of Florida, Gainesville, FL, USA
[c] Department of Orthopaedic Surgery and Sports Medicine, University of Florida, Gainesville, FL, USA

A R T I C L E I N F O

Article history:
Received 31 March 2023
Accepted 24 May 2023
Available online xxx

Keywords:
Machine learning
Total joint arthroplasty
Hospital length of stay
Health policy

A B S T R A C T

Background: The aim of this study was to improve understanding of hospital length of stay (LOS) in patients undergoing total joint arthroplasty (TJA) in a high-efficiency, hospital-based pathway.
Methods: We retrospectively reviewed 1401 consecutive primary and revision TJA patients across 67 patient and preoperative care characteristics from 2016 to 2019 from the institutional electronic health records. A machine learning approach, testing multiple models, was used to assess predictors of LOS.
Results: The median LOS was 1 day; outpatients accounted for 16.5%, 1-day inpatient stays for 38.0%, 2-day stays for 26.4%, and 3-days or more for 19.1%. Patients characteristically fell into 1 of 3 broad categories that contained relatively similar characteristics: outpatient (0-day LOS), short stay (1- to 2-day LOS), and prolonged stay (3 days or greater). The random forest models suggested that a lower Risk Assessment and Prediction Tool score, unplanned admission or hospital transfer, and a medical history of cardiovascular disease were associated with an increased LOS. Documented narcotic use for surgery preparation prior to hospitalization and preoperative corticosteroid use were factors independently associated with a decreased LOS.
Conclusions: After TJA, most patients have either an outpatient or short-stay hospital episode. Patients who stay 2 days do not differ substantially from patients who stay 1 day, while there is a distinct group that requires prolonged admission. Our machine learning models support a better understanding of the patient factors associated with different hospital LOS categories for TJA, demonstrating the potential for improved health policy decisions and risk stratification for centers caring for complex patients.

Introduction

Total joint arthroplasty (TJA) is one of the most successful and highest-value surgical procedures in any area of medicine [1] in terms of improvement in quality-adjusted life years for healthcare dollars spent. The growth in TJA surgical volume continues despite the recent novel coronavirus pandemic [2], and TJA remains one of the highest expenditures for the Centers for Medicare & Medicaid Services (CMS). With the high spending impact of TJA, CMS introduced alternative payment models to promote incentives for cost-effective and high-quality care [3]. Additionally, many relevant healthcare policies have encouraged arthroplasty surgeons to reduce hospital stays and postdischarge adverse events (eg, emergency room visits, readmissions, and acute care resource utilization). Specifically, length of stay (LOS) has been substantially decreased, leading to notable cost savings [4-6]. Meanwhile, quality of care for shortened LOS is debatable: 1 study did not identify an association between a shorter stay with the risk of readmission [7],

* Corresponding author. 482 Weil Hall, P.O. Box 116595, Gainesville, FL 32611-6595, USA. Tel.: +1 352 294 7716.
E-mail address: xiang.zhong@ise.ufl.edu

https://doi.org/10.1016/j.artd.2023.101166

while another recent study argued that excessive efforts in the LOS reduction might cause insufficient care and an increased readmission rate [8]. A growing interest in shifting arthroplasty care to low-cost sites of care has also increased the need to understand the effects of this streamlined process on outpatient and short stays. For instance, what portion of patients require (and what patient factors necessitate) a prolonged LOS after TJA and may be best served with care in a true "hospital" setting? And, among higher-functioning patients, who is suitable for an ambulatory surgical center?

In 2018, CMS altered the inpatient only list, removing total knee arthroplasty (TKA) and allowing the CMS patients to be billed under either an inpatient admission or in the outpatient setting. CMS followed this initial move with removal of total hip arthroplasty (THA) in 2020. This change created considerable confusion for healthcare providers [9-11], with no clear standard on who should be billed in which setting, which gave an impetus to identify which patients are more likely to require an inpatient level of care. From a risk-adjustment and the CMS audit perspective, an improved understanding of patient care factors that drive the LOS could improve perioperative decision-making on all sides of the care experience.

To assist in evidence-based decision-making, strategic healthcare policy, and improved clinical outcomes, machine learning approaches can be applied. Several previous studies have utilized machine learning models to predict and identify factors associated with LOS [12-15]. However, these studies typically exploit a dichotomized outcome to predict a prolonged hospital stay (eg, 4-day or 5-day LOS) or same-day discharge. Since improved surgical techniques, anesthetics, and care algorithms have reduced the hospital LOS [4-6], these models might not reflect current practice and may provide limited insights. Therefore, new models considering the practice change are needed to assist providers in guiding whether a patient should be outpatient vs inpatient and who needs prolonged care.

In this study, we hypothesized that machine learning approaches can create a useful predictive model for LOS categories in patients receiving a primary or revision TJA. We first sought to understand how LOS categories might be categorized (ie, 0-, 1-, 2-, 3-day LOS, or short vs long stay). We included various patient factors obtained from patients' electronic health record (EHR) for machine learning model development and identified associated prediction factors to support evidence-based health policy decisions.

Material and methods

Patient population

This retrospective observational study was approved by the institutional review board of the participating institution. The study participants were required to 1) experience elective THA or TKA procedures, which included primary, revision, and removal TJA (excluding trauma cases) between 2016 and 2019 at our institution, and 2) live in the primary service areas during the study period to ensure that adequate preoperational care was documented in the EHR. For the first criterion, the procedures were identified using current procedural terminology and International Classification of Diseases-10 codes (see Table A.1 in Appendix A). We did not differentiate THA and TKA patients because health policy associated with hip and knee arthroplasty tends to move in unison [6,16] and the major determinants of care do not have a significant distinction based on literature [17]. We also did not differentiate procedure types, as in our data set, over 90% of patients received a primary procedure, and our model-based investigation suggested that "procedure type" was not identified as a significant predictor for LOS. For the second criterion, in addition to the zip code, we checked for a history of encounters with primary care, internal medicine clinics, and orthopaedic clinics at the institution before their procedures. We also excluded patients discharged to a skilled-nursing facility among Medicare patients based on the 3-day rule [18] to reduce the model bias. As a result, 1168 unique patients and their 1401 hospitalizations were identified. Each hospitalization was considered as an index hospitalization due to a long time span between 2 consecutive hospitalizations (ie, an average of 261.5 days; interquartile range of 89.5-364.0).

Outcomes

Patients' hospital LOS was recorded as a discrete variable in the EHR. One- or two-day LOS is defined as patients staying either 1- or 2-midnights, respectively. We assumed the LOS outcomes have ordinal characteristics, meaning a longer hospital stay implies a greater burden of disease or disability. Because there is no gold standard for LOS categorization in the literature, we built multiple predictive machine learning models using different LOS categorizations: 1) 0, 1 or 2, and 3 days or more; 2) 0, 1, and 2 days or more; and 3) 0, 1, 2, and 3 days or more. Based on the variable selection results and the model performances, the LOS categorization 0, 1-2, and 3 days or more was chosen for further analysis. These categories were denominated as outpatient, a short stay, and a prolonged stay, respectively. The outpatient category included patients booked as inpatients but then sent home on the same day.

Covariates

We utilized preoperatively obtainable information to support care planning. The 6 categories were incorporated for this study as follows: 1) demographic and socioeconomic characteristics, 2) comorbidities, 3) prior-to-hospitalization care characteristics, 4) preoperative care characteristics during hospitalization, 5) vitals at admission, and 6) presurgical patient-reported outcome measures, including the Patient-Reported Outcomes Measurement Information System-10 [19], the Hip Disability and Osteoarthritis Outcome Score Joint Replacement [20], and the Knee Injury and Osteoarthritis Outcome Score Joint Replacement [21]. A total of 67 variables were identified (Table A.2).

Of these 67 variables, 49 variables had less than 2% missing values, and 18 variables had 16%-29% missing values, including Risk Assessment and Prediction Tool (RAPT) scores [22], body mass index at admission, interval between prelab and procedure, hemoglobin and hematocrit in preoperative laboratory testing, and several patient-reported outcome measures. The absence of RAPT was also inputted into the prediction model as a derived variable to indicate the missingness. The missingness might be related to the LOS (ie, missing not at random) because patients who are emergent cases or transferred from satellite facilities tend to stay longer; these patients do not get RAPT scores collected in our system's workflow. While the RAPT score helps with decision-making and care coordination for a planned elective surgery, it is less useful in such inpatient settings [23]. The missing values were imputed using multivariate imputation by chained equations [24]. The distributions after the missing imputation are depicted in Table A.2.

After the imputation, variables weakly associated with the outcome were filtered out in univariate analyses using a cutoff *P*-value <.1. The associations were examined using the analysis of variance for continuous variables and the chi-square test for categorical variables. A less strict criterion (*P*-value < .1) was employed considering the potential joint associations between multiple variables and the outcome [25]. Furthermore, stepwise variable selection based on Akaike information criterion of the ordinal regression was conducted to reduce multicollinearity [26].

Model construction

Using the selected variables, we built prediction models for the ordinal outcome with 2 different approaches: threshold and ordinal decomposition [27,28]. The threshold approach estimates the thresholds that separate unobserved latent continuous values of the ordinal outcome, such as the ordinal regression. In contrast, the ordinal decomposition approach creates multiple binary prediction models and assigns the category with the maximum likelihood to the sample. Any machine learning model for classification can be applied for this approach, and we chose the support vector machine and the random forest (RF) models to account for the nonlinear effects of the covariates. For the support vector machine model, we used a radial kernel and tuned the cost hyperparameter. For the RF model, the conditional permutation scheme was used to reduce the potential for multicollinearity [29,30]. We tuned the numbers of variables randomly picked at each split, called mtry, with values chosen among 3, 4, and 5. For model evaluation, we assessed all the models based on the prediction accuracy and the Kendall rank correlation coefficient that measures the ordinal association [31]. To avoid overfitting, repeated 4-fold cross-validations were conducted.

Risk factor analysis

We identified important factors from the RF model that provided the best prediction based on the mean decrease in accuracy [32]. This metric measures the decrease in accuracy when a particular variable is excluded, indicating variable importance. In this study, factors that had a positive mean decrease in accuracy for each model were considered significant.

Software

All statistical analyses in this study were performed in the R version 4.0.3 environment [33] and used the following R packages: mice [24], MASS [34], e1071 [35], randomForest [36], and pROC [37].

Results

Patient characteristics

Among the 1401 samples identified, patients were predominantly white (n = 1087, 77.6%), and more than half of patients were women (n = 784, 56.0%). The average age was 66.4, and the interquartile range was from 60 to 73. TKA and THA procedures accounted for 58.6% (n = 821) and 41.4% (n = 580), respectively. TJA procedures consisted of primary (n = 1279, 91.3%) and revision or removal (n = 122, 8.7%).

Outcomes and covariates

The median LOS was 1 day; outpatients accounted for 16.5% (n = 231), 1-day stays for 38.0% (n = 533), 2-day for 26.4% (n = 370), and 3-day or more for 19.1% (n = 267). Table A.2 presents 67 covariates' characteristics by the 4 LOS categories, that is, 0, 1, 2, 3 and more day(s), and the results of the univariate analyses for the 3 different LOS categorizations, as defined in the Outcomes section. The univariate analyses using the analysis of variance and the chi-square test after imputing missing values identified 50 variables significantly associated with the LOS. Many significant care characteristics showed monotonic increase/decrease trends across the LOS categories (see Table A.2). For instance, patients who had lower RAPT scores tended to stay longer (average scores for outpatients and 1-, 2-, and 3-day stays were 10, 9.3, 8.6, and 8.2, respectively). This finding justifies treating the outcome of LOS as ordinal. Additionally, stepwise variable selection was conducted, and 21 variables were considered as the final input covariates for machine learning model development (see Table A.3). Combining 1- and 2-day stays as a short stay was attributed to the monotonicity and similar predictors (eg, preoperative nonsteroidal antiinflammatory drug [NSAID] and RAPT), leading to increased statistical power.

Model performance

Using the identified variables, several machine learning models were constructed and assessed based on accuracy and Kendall rank correlation. As depicted in Table 1, the RF models using the ordinal decomposition approach outperformed any linear models (accuracy: 0.744-0.749 vs 0.710-0.737; Kendall rank correlation: 0.489-0.513 vs 0.445-0.481). We also constructed prediction models for the other LOS categorizations, that is, do not combine 1- and 2-day stays, and combine 2-day stay with 3 or more days, and both had inferior performances (accuracy: 0.617-0.721; Kendall rank correlation: 0.423-0.475; see Table A.4). This suggests that the 2-day-stay patients tend to have similar characteristics to the 1-day-stay ones, compared with the prolonged stay ones.

Risk factor analysis

Important variables were identified from the combination of the 2 RF models, the outpatient (Model 1) and the prolonged hospital LOS (Model 2; Fig. 1b). For Model 1, preoperative NSAID during hospitalization was remarkably important, followed by prior-to-hospitalization narcotic for surgery preparation and preoperative hematocrit to name a few (see Fig. 1a). In contrast, top important variables for Model 2 are listed as follows: RAPT, age, and the absence of RAPT (see Fig. 1b). The 2 models shared some important variables (eg, preoperative NSAID and RAPT, Table 2), whereas other variables were not in common, for example, preoperative hematocrit values and prior-to-hospitalization NSAID for surgery preparation in Model 1, Table 3; age and number of no-shows before hospitalization in Model 2, Table 4.

Discussion

This study highlighted patient features unique to different LOS categories to assist the entire care team in understanding what may drive a need for outpatient, short stay, and prolonged inpatient

Table 1
Model performance of the threshold model and the ordinal decomposition models for the 3 categories (ie, outpatient, short stay, and prolonged stay).

Models	Accuracy (avg., IQR)	Kendall rank correlation (avg., IQR)
Threshold model		
Ordinal regression	0.710 (0.698-0.723)	0.445 (0.417-0.475)
Ordinal decomposition models		
Logistic regression	0.737 (0.723-0.751)	0.481 (0.446-0.511)
SVM with a radial kernel	0.736 (0.721-0.751)	0.479 (0.456-0.512)
Random forest (mtry = 3)	0.749 (0.735-0.763)	0.513 (0.495-0.542)
Random forest (mtry = 4)	0.746 (0.733-0.761)	0.499 (0.472-0.530)
Random forest (mtry = 5)	0.744 (0.731-0.758)	0.489 (0.463-0.515)

Avg, average; LOS, length of stay; mtry, the numbers of variables randomly sampled as candidates at each split; IQR, interquartile range; SVM, support vector machine.

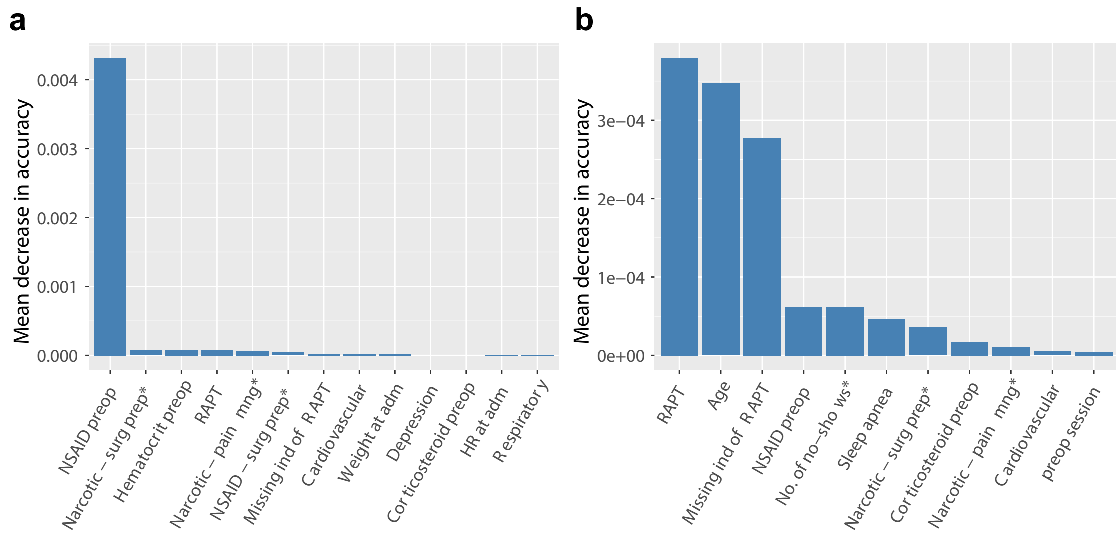

Figure 1. Important variables identified from (a) Model 1 and (b) Model 2. The mean decrease in accuracy measures the accuracy decreases when a particular variable is excluded, indicating variable importance. A higher value is better. Asterisk (*) indicates a covariate before hospitalization; adm, admission; HR, heart rate; missing ind, missing indicator; pain mng, pain management; preop, preoperative; surg prep, surgery preparation.

admission arthroplasty. Such a distinction has profound implications on health policy associated with delivering arthroplasty care. It can also be useful for care planning and management (eg, site of care, discharge planning, and bed utilization).

As there is no gold standard for LOS categorization, this study attempted to investigate the similarity of patients with different LOS categories. The distributions of covariates by the LOS categories in Table A.2 and the model performances in Table 1 suggest that 2-day stay patients had more shared characteristics with 1-day stay patients than prolonged stay patients. This finding reveals that 1- or 2-day LOS patients represent a distinct population, with more in common with each other than with either outpatients or prolonged stay patients. This can be valuable information for patient care decision-making and health policy design; centers that can accommodate outpatients and short-stay patients could cater to the vast majority of even complex arthroplasty patients, as seen in our tertiary care center.

Model performance

We identified that the RF model with the ordinal decomposition had the best performance based on the 2 metrics. The nonparametric model was successful in delineating the nonlinear relationships between the LOS and the covariates, which is hard to be captured by the linear models. In terms of accuracy, the RF model correctly predicted the LOS with a near 3-quarter chance. Furthermore, the model reflected the order of the LOS categories fairly based on the Kendall rank correlation [38]. For the hard-to-distinguish cases, the model would be less likely to incorrectly predict a prolonged stay rather than a short stay for an actual outpatient.

Risk factor analysis

The 2 chosen RF models, Model 1 and Model 2, shared some risk factors while also having unique risk factors (Table 2). Regarding the shared risk factors, preoperative NSAID during hospitalization was a pivotal indicator to determine the 3 categories. The percentage of patients who received NSAIDs preoperatively is significantly higher in the short-stay group (outpatients: 79, 34.2%; short-stay patients: 816, 90.4%; prolonged-stay patients: 200, 74.9%; $P < .001$). However, caution is warranted when attempting to interpret this result, as NSAID prescriptions that are coded in the system might not accurately reflect the actual dose and usage due to many NSAIDs being available over-the-counter.

Table 2
Risk factors identified in both the outpatient prediction model and the prolonged hospital length of stay model and their distributions by the 3 length of stay categories.

Risk factors	Outpatient (n = 231, 16.5%)	Short stay (n = 903, 64.4%)	Prolonged stay (n = 267, 19.1%)	*P*-value
Comorbidities				
Cardiovascular	12 (5.2)	138 (15.3)	61 (22.8)	**<.001**
Prior-to-hospitalization care characteristics				
Narcotic[a] - pain management	37 (16.0)	245 (27.1)	90 (33.7)	**<.001**
Narcotic[a] - surgery preparation	220 (95.2)	728 (80.6)	180 (67.4)	**<.001**
Preoperative care characteristics				
NSAID	79 (34.2)	816 (90.4)	200 (74.9)	**<.001**
Corticosteroid	186 (80.5)	594 (65.8)	143 (53.6)	**<.001**
RAPT score	10 (9-11)	9.1 (8-10.5)	8.4 (7-10)	**<.001**
Absence of RAPT	21 (9.1)	117 (13.0)	89 (33.3)	**<.001**

NSAID, nonsteroidal antiinflammatory drug; RAPT, Risk Assessment and Prediction Tool.
Continuous variables are expressed as mean and interquartile range. The categorical variables are expressed in terms of n (%). A *P*-value less than .05 indicates a significant difference between the 3 groups and is bold.
[a] Events within 3 months prior to hospitalization.

Table 3
Risk factors identified only in the outpatient prediction model and their distributions by the 2 length of stay categories.

Risk factors	Outpatient (n = 231, 16.5%)	LOS 1+ (n = 1170, 83.5%)	*P*-value	Rank of importance in the model
Comorbidities				
Depression	34 (14.7)	305 (26.1)	**<.001**	10
Respiratory	2 (0.9)	59 (5.0)	**.008**	13
Prior-to-hospitalization care characteristics				
NSAID[a] - surgery preparation	116 (50.2)	436 (37.3)	**<.001**	2
Preoperative care characteristics				
Hematocrit			**<.001**	3
Low	59 (25.5)	554 (47.4)		
Normal	171 (74.0)	609 (52.1)		
High	1 (0.4)	7 (0.5)		
Vitals				
Weight at admission (kg)	191.3 (162.0, 218.5)	199.6 (164.5, 230.0)	**.008**	9
Heart rate at admission (mmHg)			.174	12
<60	28 (12.1)	144 (12.3)		
60-100	202 (87.4)	999 (85.4)		
>100	1 (0.4)	27 (2.3)		

LOS, length of stay; NSAID, nonsteroidal antiinflammatory drug.
Continuous variables are expressed as mean and interquartile range. The categorical variables are expressed in terms of n (%). A *P*-value less than .05 indicates a significant difference between the 2 groups and is bold. The rank of importance in the model is obtained from the random forest model.
[a] Events within 3 months prior to hospitalization.

Additionally, a lower RAPT score served as a significant predictive factor of an increased LOS; outpatients had the highest averaged RAPT score, and those with prolonged-stay patients had the lowest RAPT (outpatient: 10.0; short stay: 9.0; prolonged stay: 8.2; $P < .001$), which is consistent with existing studies [39,40]. Also, an unexpected admission or hospital transfer implied by the absence of RAPT was predictive of an increased LOS (outpatient: n = 21, 9.1%; short stay: n = 117, 13.0%; prolonged stay: n = 89, 33.3%; $P < .001$). Furthermore, patients who had experienced cardiovascular disease stayed longer (outpatient: n = 12, 5.2%; short stay: n = 138, 15.3%; prolonged stay: n = 61, 22.8%, $P < .001$).

Patients who had documented narcotic use to prepare for the TJA within 30 days prior to hospitalization contributed to a decreased LOS. Most outpatients received the medications during that period, whereas just two-thirds of prolonged-stay patients did (outpatient: n = 220, 95.2%; short stay: n = 728, 80.6%; prolonged stay: 180, 67.4%; $P < .001$). This result is likely associated with higher postacute care utilization in the prolonged-stay group, which would not receive postoperative pain prescriptions at a presurgical visit. In addition, preoperative corticosteroid use factored in a decreased LOS (outpatient: n = 186, 80.5%; short stay: n = 594, 65.8%; prolonged stay: n = 143, 53.6%; $P < .001$).

Model 1 uniquely stressed the association between the outpatient group and the following covariates: normal preoperative hematocrit values, NSAIDs for surgery preparation before hospitalization, less weight, and no depression (Table 3). Specifically, 3-quarters of the outpatients had normal hematocrit values, while half of the 1+ day stay patients had low hematocrit values (outpatient: n = 171, 74.0%; short or prolonged stay: n = 554, 47.4%; $P < .001$). Use of NSAIDs prior to surgery preparation before hospitalization was also predictive of the outpatients (outpatient: n = 116, 50.2%; short or prolonged stay: n = 436, 37.3%; $P < .001$).

Model 2 distinctively identified the following covariates as predictive of the prolonged stay: younger age, many no-shows before hospitalization, no preoperative corticosteroid, sleep apnea, and no preoperative session attendance (Table 4). Younger patients (age ≤61) stayed longer and might be more likely to be admitted due to other underlying conditions such as trauma injuries, rheumatoid or other inflammatory arthritides rather than osteoarthritis, and osteonecrosis. Additionally, the exclusion criteria related to the 3-day rule may have resulted in identifying a younger age as one of the risk factors. More investigation into the effect of age on LOS needs to be performed. Furthermore, having many no-shows prior to hospitalization not only worsened clinical resource utilization but was also associated with a prolonged stay (outpatient or short stay: n = 163, 14.3%; prolonged stay: n = 81, 30.3%; $P < .001$).

Table 4
Risk factors identified only in the prolonged hospital length of stay model and their distributions by the 2 length of stay categories.

Characteristics	LOS 0-2 (n = 1134, 80.9%)	Prolonged stay (n = 267, 19.1%)	*P*-value	Rank of importance in the model
Demographics and socioeconomic characteristics				
Age, y			.474	2
≤61	305 (26.9)	84 (31.5)		
62-67	318 (28.0)	74 (27.7)		
68-73	276 (24.3)	59 (22.1)		
74+	235 (20.7)	50 (18.7)		
Comorbidities				
Sleep apnea	23 (2.0)	25 (9.4)	**<.001**	6
Prior-to-hospitalization care characteristics				
No. of no-show >1	163 (14.3)	81 (30.3)	**<.001**	5
Preoperative care characteristics				
Attended preoperative sessions	141 (12.4)	14 (5.2)	**.001**	11

LOS, length of stay.
Continuous variables are expressed as mean and interquartile range. The categorical variables are expressed in terms of n (%). A *P*-value less than .05 indicates a significant difference between the 2 groups and is bold. The rank of importance in the model is obtained from the random forest model.

Limitations

First, as this study was a single-center retrospective study, the patient cohort might not represent the national patient population. Also, EHRs possess various care and patient characteristics but can only offer limited samples compared to all-payer claims databases [41]. Using data from different institutions or distributed research networks could improve model generalizability [42]. Additionally, the data only contained discretized LOS. If the actual admission and discharge time can be provided, different modeling choices (eg, survival analysis) can be possible. In this study, we tested several hypotheses regarding LOS categorization options; the selected categorization may be prone to change as care practices continue to evolve and care efficiencies improve. Finally, this study did not differentiate primary, revision, or removal TJA procedures and did not differentiate knee and hip procedures. Studies have shown that different types of procedures shared similar risk factors for the increased LOS (eg, comorbidity) for both TKAs and THAs [43]. Our study also revealed that by including the indicator of procedure type as an additional predictor, the prediction performance was not improved (see Table A.5). The future work could separate primary vs revision and removal TJA procedures and separating TKA and THA procedures, upon obtaining sufficient samples to ensure statistical power.

Conclusions

In this study, we thoroughly examined the relationship between TJA patients' LOS and various covariates obtained from the EHR using machine learning models. The 3-category-based RF model (distinguishing between true "outpatients," short-stay—1-2 days—admissions, and prolonged inpatient stays of 3 or more days) provided the patient factors associated with different levels of care requirement for TJA. As joint arthroplasty care continues to evolve, patients may benefit from receiving highly efficient and cost-favorable care in designed "short-stay" arthroplasty centers, with only a select few requiring care in the hospital setting (patients with poor social support as evidenced by low RAPT, for instance, or patients with prolonged preoperative narcotic use).

Conflicts of interest

C. F. Gray received royalties from Adler Orthopaedics, is a paid consultant for Adler Orthopaedics and Smith & Nephew, owns Stock or stock options in ROMTech and DeBogy Molecular, and is an AAOS Board/Committee Member; all other authors declare no potential conflicts of interest.

For full disclosure statements refer to https://doi.org/10.1016/j.artd.2023.101166.

References

[1] Etkin CD, Springer BD. The American Joint Replacement Registry—the first 5 years. Arthroplasty Today 2017;3:67—9. https://doi.org/10.1016/j.artd.2017.02.002.
[2] American Joint Replacement Registry releases 2021 annual report, showing increase in number of hip and knee procedures despite pause due to COVID-19. https://aaos.org/aaos-home/newsroom/press-releases/american-joint-replacement-registry-releases-2021-annual-report/. [Accessed 24 August 2022].
[3] McLawhorn AS, Buller LT. Bundled payments in total joint replacement: keeping our care affordable and high in quality. Curr Rev Musculoskelet Med 2017;10:370—7. https://doi.org/10.1007/s12178-017-9423-6.
[4] Rankin KA, Freedman IG, Rubin LE, Grauer JN. Centers for Medicare & Medicaid Services' 2018 removal of total knee arthroplasty from the inpatient-only list led to broad changes in hospital length of stays. J Am Acad Orthop Surg 2021;29(24):1061—7.
[5] Gray CF, Prieto HA, Deen JT, Parvataneni HK. Bundled payment "Creep": institutional redesign for primary arthroplasty positively affects revision arthroplasty. J Arthroplasty 2019;34:206—10. https://doi.org/10.1016/j.arth.2018.10.025.
[6] Gray CF, Prieto HA, Duncan AT, Parvataneni HK. Arthroplasty care redesign related to the comprehensive care for joint replacement model: results at a tertiary academic medical center. Arthroplasty Today 2018;4:221—6. https://doi.org/10.1016/j.artd.2018.02.002.
[7] Barad SJ, Howell SM, Tom J. Is a shortened length of stay and increased rate of discharge to home associated with a low readmission rate and cost-effectiveness after primary total knee arthroplasty? Arthroplasty Today 2018;4:107—12. https://doi.org/10.1016/j.artd.2015.08.003.
[8] Benito J, Stafford J, Judd H, Ng M, Corces A, Roche MW. Length of stay increases 90-day readmission rates in patients undergoing primary total joint arthroplasty. JAAOS Glob Res Rev 2022;6:e21.00271. https://doi.org/10.5435/JAAOSGlobal-D-21-00271.
[9] Curtin BM, Odum SM. Unintended bundled payments for care improvement consequences after removal of total knee arthroplasty from inpatient-only list. J Arthroplasty 2019;34(Supplement):S121—4. https://doi.org/10.1016/j.arth.2019.02.053.
[10] Yates AJ, Kerr JM, Froimson MI, Della Valle CJ, Huddleston JI. The unintended impact of the removal of total knee arthroplasty from the Center for Medicare and Medicaid Services inpatient-only list. J Arthroplasty 2018;33:3602—6. https://doi.org/10.1016/j.arth.2018.09.043.
[11] Krueger CA, Kerr JM, Bolognesi MP, Courtney PM, Huddleston JI. The removal of total hip and total knee arthroplasty from the inpatient-only list increases the administrative burden of surgeons and continues to cause confusion. J Arthroplasty 2020;35:2772—8. https://doi.org/10.1016/j.arth.2020.04.079.
[12] Trautner BW, Bhimani RD, Amspoker AB, Hysong SJ, Garza A, Kelly PA, et al. Development and validation of an algorithm to recalibrate mental models and reduce diagnostic errors associated with catheter-associated bacteriuria. BMC Med Inform Decis Mak 2013;13:48. https://doi.org/10.1186/1472-6947-13-48.
[13] Skura BW, Goubeaux C, Passias BJ, Schuette HB, Melaragno AJ, Glazier MT, et al. Increased length of stay following total joint arthroplasty based on insurance type. Arch Orthop Trauma Surg 2022. https://doi.org/10.1007/s00402-022-04417-1.
[14] Prohaska MG, Keeney BJ, Beg HA, Swarup I, Moschetti WE, Kantor SR, et al. Preoperative body mass index and physical function are associated with length of stay and facility discharge after total knee arthroplasty. Knee 2017;24:634—40. https://doi.org/10.1016/j.knee.2017.02.005.
[15] Cizmic Z, Feng JE, Anoushiravani AA, Borzio RW, Schwarzkopf R, Slover JD. The risk assessment and prediction tool is less accurate in extended length of stay patients following total joint arthroplasty. J Arthroplasty 2019;34:418—21. https://doi.org/10.1016/j.arth.2018.11.008.
[16] Comprehensive care for joint replacement model. https://innovation.cms.gov/innovation-models/cjr. [Accessed 23 June 2023].
[17] Matar WY, Jafari SM, Restrepo C, Austin M, Purtill JJ, Parvizi J. Preventing infection in total joint arthroplasty. JBJS 2010;92(Supplement_2):36—46. https://doi.org/10.2106/JBJS.J.01046.
[18] Skilled nursing facility (SNF) situations | Medicare. https://www.medicare.gov/what-medicare-covers/skilled-nursing-facility-snf-situations. [Accessed 24 August 2022].
[19] Shim J, Hamilton DF. Comparative responsiveness of the PROMIS-10 global health and EQ-5D questionnaires in patients undergoing total knee arthroplasty. Bone Joint J 2019;101-B:832—7. https://doi.org/10.1302/0301-620X.101B7.BJJ-2018-1543.R1.
[20] Nilsdotter AK, Lohmander LS, Klässbo M, Roos EM. Hip disability and osteoarthritis outcome score (HOOS) — validity and responsiveness in total hip replacement. BMC Musculoskelet Disord 2003;4:10. https://doi.org/10.1186/1471-2474-4-10.
[21] Roos EM, Lohmander LS. The Knee injury and Osteoarthritis Outcome Score (KOOS): from joint injury to osteoarthritis. Health Qual Life Outcomes 2003;1:64. https://doi.org/10.1186/1477-7525-1-64.
[22] Sconza C, Respizzi S, Grappiolo G, Monticone M. The risk assessment and prediction tool (RAPT) after hip and knee replacement: a systematic review. Joints 2019;7:41—5. https://doi.org/10.1055/s-0039-1693459.
[23] Oeding JF, Bosco JA, Carmody M, Lajam CM. RAPT scores predict inpatient versus outpatient status and readmission rates after IPO changes for total joint arthroplasty: an analysis of 12,348 cases. J Arthroplasty 2022;37:2140—8. https://doi.org/10.1016/j.arth.2022.05.037.
[24] van Buuren S, Groothuis-Oudshoorn K. MICE: multivariate imputation by chained equations in R. J Stat Softw 2011;45:1—68.
[25] Park J, Zhong X, Babaie Sarijaloo F, Wokhlu A. Tailored risk assessment of 90-day acute heart failure readmission or all-cause death to heart failure with preserved versus reduced ejection fraction. Clin Cardiol 2022;45:370—8. https://doi.org/10.1002/clc.23780.
[26] Bozdogan H. Model selection and Akaike's Information Criterion (AIC): the general theory and its analytical extensions. Psychometrika 1987;52:345—70. https://doi.org/10.1007/BF02294361.
[27] Gutiérrez PA, Pérez-Ortiz M, Sánchez-Monedero J, Fernández-Navarro F, Hervás-Martínez C. Ordinal regression methods: survey and experimental study. IEEE Trans Knowl Data Eng 2016;28:127—46. https://doi.org/10.1109/TKDE.2015.2457911.
[28] Lin HT, Li L. Reduction from cost-sensitive ordinal ranking to weighted binary classification. Neural Comput 2012;24:1329—67. https://doi.org/10.1162/NECO_a_00265.

[29] Strobl C, Boulesteix AL, Kneib T, Augustin T, Zeileis A. Conditional variable importance for random forests. BMC Bioinformatics 2008;9:307. https://doi.org/10.1186/1471-2105-9-307.
[30] Debeer D, Strobl C. Conditional permutation importance revisited. BMC Bioinformatics 2020;21:307. https://doi.org/10.1186/s12859-020-03622-2.
[31] Kendall MG. A new measure of rank correlation. Biometrika 1938;30:81–93. https://doi.org/10.1093/biomet/30.1-2.81.
[32] Han H, Guo X, Yu H. In: Variable selection using mean decrease accuracy and mean decrease Gini based on random forest. 2016 7th IEEE international Conference on Software Engineering and Service Science (ICSESS); New York City, NY: IEEE; 2016. p. 219–24.
[33] R Core Team. R: a language and environment for statistical Computing. R Foundation for Statistical Computing; 2020. https://www.R-project.org/. [Accessed 23 June 2023].
[34] Venables WN, Ripley BD. Modern applied Statistics with S. Fourth. Springer; 2002. https://www.stats.ox.ac.uk/pub/MASS4/. [Accessed 23 June 2023].
[35] Meyer D, Dimitriadou E, Hornik K, Meyer D. e1071: misc functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien. https://CRAN.R-project.org/package=e1071; 2021. [Accessed 25 January 2022].
[36] Liaw A, Wiener M. Classification and regression by randomForest. R News 2002;2:18–22.
[37] Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 2011;12:77. https://doi.org/10.1186/1471-2105-12-77.
[38] Akoglu H. User's guide to correlation coefficients. Turk J Emerg Med 2018;18:91–3. https://doi.org/10.1016/j.tjem.2018.08.001.
[39] Tan C, Loo G, Pua YH, Chong HC, Yeo W, Ong PH, et al. Predicting discharge outcomes after total knee replacement using the risk assessment and predictor tool. Physiotherapy 2014;100:176–81. https://doi.org/10.1016/j.physio.2013.02.003.
[40] Alshahwani AA, Dungey M, Lillie C, Krikler S, Plakogiannis C. Predictive value of the risk assessment and prediction tool (RAPT) score for primary hip and knee arthroplasty patients: a single-center study. Cureus 2021;13:e14112. https://doi.org/10.7759/cureus.14112.
[41] All-payer claims databases. https://www.ahrq.gov/data/apcd/index.html. [Accessed 24 August 2022].
[42] OneFlorida Clinical Research Consortium » Clinical and Translational Science Institute » University of Florida. https://www.ctsi.ufl.edu/ctsa-consortium-projects/oneflorida/. [Accessed 25 January 2022].
[43] Sloan M, Sheth NP. Length of stay and inpatient mortality trends in primary and revision total joint arthroplasty in the United States, 2000–2014. J Orthop 2018;15:645–9. https://doi.org/10.1016/j.jor.2018.05.021.

Table A.1
CPT codes and ICD-10 codes for total hip and knee arthroplasty.

Total joint arthroplasty	CPT codes	ICD-10 codes
Total hip arthroplasty	27120, 27125, 27130, 27132, 27134, 27137, 27138, 27090, 27091	Z96.64
Total knee arthroplasty	27440-27447, 27486-27488	Z96.65

CPT, current procedural terminology; ICD-10, International Classification of Diseases-10.

Table A.2
Characteristics and statistical tests based on the LOS categorizations.

Characteristics	LOS 0 (n = 231, 16.5%)	LOS 1 (n = 533, 38.0%)	LOS 2 (n = 370, 26.4%)	LOS 3+ (n = 267, 19.1%)	*P*-value (LOS cat 1: 0/1-2/3+)	*P*-value (LOS cat 2: 0/1/2+)	*P*-value (LOS cat 3: 0/1/2/3+)
Demographics and socioeconomic characteristics							
Age, y					**<.001**	**.004**	**.002**
≤61	86 (37.2)	134 (25.1)	85 (23.0)	84 (31.5)			
62-67	69 (29.9)	151 (28.3)	98 (26.5)	74 (27.7)			
68-73	45 (19.5)	136 (25.5)	95 (25.7)	59 (22.1)			
74+	31 (13.4)	112 (21.0)	92 (24.9)	50 (18.7)			
Sex-male	105 (45.5)	263 (49.3)	135 (36.5)	114 (42.7)	.826	**.002**	.002
Race					**.002**	**.003**	**<.001**
Black	22 (9.5)	66 (12.4)	64 (17.3)	56 (21.0)			
Other	19 (8.2)	38 (7.1)	38 (10.3)	11 (4.1)			
White	190 (82.3)	429 (80.5)	268 (72.4)	200 (74.9)			
Non-Hispanic	217 (93.9)	504 (94.6)	345 (93.2)	260 (97.4)	**.088**	.830	.132
Marital status					**<.001**	**<.001**	**<.001**
Married	165 (71.4)	348 (65.3)	211 (57.0)	147 (55.1)			
Other or unknown	17 (7.4)	33 (6.2)	22 (5.9)	11 (4.1)			
Single/divorced/separated/widowed	49 (21.2)	152 (28.5)	137 (37.0)	109 (40.8)			
Insurance					**<.001**	**<.001**	**<.001**
Medicaid	6 (2.6)	23 (4.3)	26 (7.0)	35 (13.1)			
Medicare	80 (34.6)	288 (54.0)	200 (54.1)	138 (51.7)			
Other	6 (2.6)	18 (3.4)	16 (4.3)	7 (2.6)			
Private or managed care	139 (60.2)	204 (38.3)	128 (34.6)	87 (32.6)			
Income					**.002**	**.002**	**.004**
<40,000	48 (20.8)	161 (30.2)	123 (33.2)	101 (37.8)			
40,000-60,000	88 (38.1)	179 (33.6)	110 (29.7)	78 (29.2)			
>60,000	95 (41.1)	193 (36.2)	137 (37.0)	88 (33.0)			
Rural	17 (7.4)	58 (10.9)	35 (9.5)	31 (11.6)	.269	.313	.374
Smoking status					**.009**	**.043**	**.034**
Current	16 (6.9)	44 (8.3)	32 (8.6)	33 (12.4)			
Former	77 (33.3)	213 (40.0)	150 (40.5)	115 (43.1)			
Never	138 (59.7)	276 (51.8)	188 (50.8)	119 (44.6)			
Alcohol dependence	9 (3.9)	17 (3.2)	12 (3.2)	15 (5.6)	.193	.642	.349
Illicit or cardiotoxic drug use	2 (0.9)	13 (2.4)	10 (2.7)	11 (4.1)	**.073**	.131	.151
Comorbidities							
Renal	1 (0.4)	7 (1.3)	12 (3.2)	17 (6.4)	**<.001**	**<.001**	**<.001**
Cardiovascular	12 (5.2)	81 (15.2)	57 (15.4)	61 (22.8)	**<.001**	**<.001**	**<.001**
Hypertension	112 (48.5)	280 (52.5)	216 (58.4)	164 (61.4)	**.015**	**.004**	**.001**
Sleep apnea	4 (1.7)	9 (1.7)	10 (2.7)	25 (9.4)	**<.001**	**.001**	**<.001**
Diabetes	22 (9.5)	65 (12.2)	62 (16.8)	62 (23.2)	**<.001**	**<.001**	**<.001**
Endocrine	2 (0.9)	3 (0.6)	7 (1.9)	11 (4.1)	**.002**	**.006**	**.002**
Malignancy or cancer	6 (2.6)	13 (2.4)	17 (4.6)	12 (4.5)	.489	.106	.213
Depression	34 (14.7)	120 (22.5)	95 (25.7)	90 (33.7)	**<.001**	**<.001**	**<.001**
Hematologic	0 (0.0)	0 (0.0)	0 (0.0)	11 (4.1)	**<.001**	**.001**	**<.001**
Respiratory	2 (0.9)	16 (3.0)	14 (3.8)	29 (10.9)	**<.001**	**<.001**	**<.001**
Prior-to-hospitalization care characteristics							
NSAID[a] - pain management	55 (23.8)	153 (28.7)	86 (23.2)	67 (25.1)	.685	.143	.244
NSAID[a] - surgery preparation	116 (50.2)	220 (41.3)	133 (35.9)	83 (31.1)	**<.001**	**<.001**	**<.001**
Corticosteroid[a]	42 (18.2)	112 (21.0)	85 (23.0)	62 (23.2)	.362	.283	.470
Narcotic[a] - pain management	37 (16.0)	136 (25.5)	109 (29.5)	90 (33.7)	**<.001**	**<.001**	**<.001**
Narcotic[a] - surgery preparation	220 (95.2)	457 (85.7)	271 (73.2)	180 (67.4)	**<.001**	**<.001**	**<.001**
Chemotherapy[a]	0 (0.0)	0 (0.0)	0 (0.0)	1 (0.4)	.119	.549	.236
Radiation[a]	0 (0.0)	1 (0.2)	1 (0.3)	2 (0.7)	.246	.447	.414
No. of orthopaedics care visits[b]					**.058**	.331	.123
≤3	99 (42.9)	224 (42.0)	167 (45.1)	128 (47.9)			

(*continued on next page*)

Table A.2 (*continued*)

Characteristics	LOS 0 (n = 231, 16.5%)	LOS 1 (n = 533, 38.0%)	LOS 2 (n = 370, 26.4%)	LOS 3+ (n = 267, 19.1%)	*P*-value (LOS cat 1: 0/1-2/3+)	*P*-value (LOS cat 2: 0/1/2+)	*P*-value (LOS cat 3: 0/1/2/3+)
4-5	94 (40.7)	212 (39.8)	141 (38.1)	80 (30.0)			
6+	38 (16.5)	97 (18.2)	62 (16.8)	59 (22.1)			
MyChart status-active	146 (63.2)	341 (64.0)	208 (56.2)	136 (50.9)	**.007**	**.001**	**.001**
No. of no-show >1[b]	22 (9.5)	70 (13.1)	71 (19.2)	81 (30.3)	**<.001**	**<.001**	**<.001**
Noncompliance	34 (14.7)	80 (15.0)	58 (15.7)	72 (27.0)	**<.001**	**.026**	**<.001**
Flu vaccine status[b]	75 (32.5)	159 (29.8)	123 (33.2)	92 (34.5)	.604	.355	.537
Preoperative care characteristics							
Days between preoperative labs and operation	5.6 (0, 10)	4.4 (0, 8)	4.7 (0, 8)	3.5 (0, 6)	**<.001**	**.001**	**<.001**
Hemoglobin					**<.001**	**<.001**	**<.001**
Low	56 (24.2)	195 (36.6)	161 (43.5)	147 (55.1)			
Normal	174 (75.3)	336 (63.0)	207 (55.9)	120 (44.9)			
High	1 (0.4)	2 (0.4)	2 (0.5)	0 (0.0)			
Hematocrit					**<.001**	**<.001**	**<.001**
Low	59 (25.5)	229 (43.0)	174 (47.0)	151 (56.6)			
Normal	171 (74.0)	300 (56.3)	193 (52.2)	116 (43.4)			
High	1 (0.4)	4 (0.8)	3 (0.8)	0 (0.0)			
Hemoglobin a1c	4 (1.7)	10 (1.9)	16 (4.3)	12 (4.5)	.185	**.020**	**.049**
Attended preoperative sessions	42 (18.2)	67 (12.6)	32 (8.6)	14 (5.2)	**<.001**	**<.001**	**<.001**
NSAID	79 (34.2)	498 (93.4)	318 (85.9)	200 (74.9)	**<.001**	**<.001**	**<.001**
Corticosteroid	186 (80.5)	393 (73.7)	201 (54.3)	143 (53.6)	**<.001**	**<.001**	**<.001**
Narcotic	231 (100.0)	532 (99.8)	370 (100.0)	267 (100.0)	.759	.443	.653
ASA score >2	117 (50.6)	375 (70.4)	312 (84.3)	227 (85.0)	**<.001**	**<.001**	**<.001**
RAPT score	10.0 (9.0, 11.0)	9.3 (8.0, 11.0)	8.6 (7.0, 10.0)	8.2 (7.0, 10.0)	**<.001**	**<.001**	**<.001**
Absence of RAPT	21 (9.1)	66 (12.4)	51 (13.8)	89 (33.3)	**<.001**	**<.001**	**<.001**
Vitals							
Body mass index at admission (kg/m^2)					.469	.293	.372
<18.5	3 (1.3)	6 (1.1)	2 (0.5)	4 (1.5)			
18.5-<25	39 (16.9)	85 (15.9)	48 (13.0)	39 (14.6)			
25-<30	68 (29.4)	186 (34.9)	121 (32.7)	73 (27.3)			
30-<40	93 (40.3)	205 (38.5)	149 (40.3)	111 (41.6)			
40+	28 (12.1)	51 (9.6)	50 (13.5)	40 (15.0)			
Weight (kg)							
At admission (a)	191.3 (162.0, 218.5)	197.0 (166.0, 222.7)	200.4 (162.3, 234.8)	203.8 (165.0, 239.2)	**.012**	**.011**	**.019**
Within 1 y prior to admission (b)	191.9 (163.0, 216.8)	197.3 (167.8, 224.2)	200.5 (160.0, 235.8)	203.0 (164.5, 235.6)	**.032**	**.024**	**.049**
Difference between (a) and (b)	−0.5 (−4.1, 3.4)	−0.3 (−5.4, 3.7)	−0.1 (−6.5, 4.3)	0.7 (−6.1, 6.8)	.672	.807	.847
Blood pressure (mmHg)							
Systolic BP at admission					.920	.901	.984
<120	59 (25.5)	133 (25.0)	87 (23.5)	61 (22.8)			
120-<130	50 (21.6)	101 (18.9)	72 (19.5)	55 (20.6)			
130-<140	52 (22.5)	122 (22.9)	83 (22.4)	58 (21.7)			
140+	70 (30.3)	177 (33.2)	128 (34.6)	93 (34.8)			
Diastolic BP at admission					.402	**.030**	**.028**
<80	153 (66.2)	363 (68.1)	288 (77.8)	190 (71.2)			
80-<90	54 (23.4)	111 (20.8)	52 (14.1)	48 (18.0)			
90+	24 (10.4)	59 (11.1)	30 (8.1)	29 (10.9)			
Pulse pressure at admission					**.046**	**.001**	**.004**
<40	19 (8.2)	67 (12.6)	38 (10.3)	26 (9.7)			
40-60	124 (53.7)	249 (46.7)	143 (38.6)	114 (42.7)			
>60	88 (38.1)	217 (40.7)	189 (51.1)	127 (47.6)			
MAP at admission					.369	.153	.144
<70	34 (14.7)	84 (15.8)	79 (21.4)	51 (19.1)			
70-100	177 (76.6)	394 (73.9)	261 (70.5)	184 (68.9)			
>100	20 (8.7)	55 (10.3)	30 (8.1)	32 (12.0)			
SPO2 at admission <95%	20 (8.7)	59 (11.1)	30 (8.1)	31 (11.6)	.534	.531	.336
Heart rate at admission (mmHg)					**.001**	**<.001**	**<.001**

<60	28 (12.1)	81 (15.2)	41 (11.1)	22 (8.2)			
60-100	202 (87.4)	448 (84.1)	319 (86.2)	232 (86.9)			
>100	1 (0.4)	4 (0.8)	10 (2.7)	13 (4.9)			
Presurvey responses							
PROMIS-10							
Global health T score-physical	41.7 (37.4, 44.9)	40.5 (34.9, 44.9)	37.5 (32.4, 42.3)	37.8 (32.4, 42.3)	**<.001**	**<.001**	**<.001**
Global health T score-mental	51.2 (45.8, 56.0)	49.5 (43.5, 56.0)	47.0 (41.7, 53.3)	46.5 (41.1, 50.8)	**<.001**	**<.001**	**<.001**
Survey questionnaire (1)[c]	3.5 (3.0, 4.0)	3.2 (3.0, 4.0)	3.0 (2.0, 4.0)	3.0 (2.0, 4.0)	**<.001**	**<.001**	**<.001**
Survey questionnaire (2)	3.3 (2.0, 4.0)	3.1 (2.0, 4.0)	2.7 (2.0, 3.0)	2.7 (2.0, 3.0)	**<.001**	**<.001**	**<.001**
Survey questionnaire (3)	2.7 (2.0, 3.0)	2.6 (2.0, 3.0)	2.3 (2.0, 3.0)	2.4 (2.0, 3.0)	**.006**	**<.001**	**<.001**
Survey questionnaire (4)	3.3 (3.0, 4.0)	3.3 (3.0, 4.0)	3.1 (3.0, 4.0)	3.0 (3.0, 3.0)	**<.001**	**<.001**	**<.001**
Survey questionnaire (5)	3.6 (3.0, 4.0)	3.6 (3.0, 4.0)	3.3 (3.0, 4.0)	3.3 (3.0, 4.0)	**<.001**	**<.001**	**<.001**
Survey questionnaire (6)	3.7 (3.0, 4.0)	3.5 (3.0, 4.0)	3.3 (3.0, 4.0)	3.3 (3.0, 4.0)	**<.001**	**<.001**	**<.001**
Survey questionnaire (7)	3.8 (3.0, 5.0)	3.5 (3.0, 4.0)	3.2 (2.0, 4.0)	3.2 (2.0, 4.0)	**<.001**	**<.001**	**<.001**
Survey questionnaire (8)	3.6 (3.0, 4.0)	3.4 (3.0, 4.0)	3.1 (3.0, 4.0)	3.1 (2.0, 4.0)	**<.001**	**<.001**	**<.001**
Survey questionnaire (9)	4.0 (3.0, 5.0)	3.9 (3.0, 5.0)	3.7 (3.0, 4.0)	3.6 (3.0, 4.0)	**<.001**	**<.001**	**<.001**
Survey questionnaire (10)	3.1 (2.0, 4.0)	2.8 (2.0, 4.0)	2.5 (2.0, 3.0)	2.5 (2.0, 3.0)	**<.001**	**<.001**	**<.001**
Disease specific functional score from HOOS or KOOS[d]	48.4 (42.3, 57.1)	47.8 (39.6, 57.1)	43.9 (34.2, 53.0)	43.9 (34.2, 53.0)	**.005**	**<.001**	**<.001**

LOS, length of stay; cat, category; NSAID, nonsteroidal antiinflammatory drugs; CAM, confusion assessment method; ER, emergency room; SBP, systolic blood pressure; DBP, diastolic blood pressure; PROMIS, patient-reported outcomes measurement information system; KOOS, Knee Injury and Osteoarthritis Outcome Score; HOOS, Hip Injury Disability and Osteoarthritis Outcome Score.

Continuous variables are expressed as mean and interquartile range. The categorical variables are expressed in terms of n (%). A *P*-value less than .1 indicates a significant difference between the 2 group and is bold.

[a] Events within 3 months prior to hospitalization.

[b] Events within 1 year prior to hospitalization.

[c] Survey questionnaire: (1) In general, how would you rate your physical health; (2) To what extent are you able to carry out your everyday physical activities; (3) How would you rate your pain on average; (4) How would you rate your fatigue on average; (5) How often have you been bothered by emotional problems such as feeling anxious, depressed, or irritable; (6) In general, would you say your health is; (7) In general, how would you rate your satisfaction with your social activities and relationships; (8) In general, would you say your quality of life is; (9) In general, how would you rate your mental health including your mood and your ability to think; (10) In general, please rate how well you carry out your usual social activities and roles.

[d] For an individual sample, only a single disease specific functional score (HOOS or KOOS) was included depending on the type of surgery.

Table A.3
Final model input covariates by the LOS categorization.

Characteristics	LOS cat 1 (0/1-2/3+)	LOS cat 2 (0/1/2+)	LOS cat 3 (0/1/2/3+)
Demographics and socioeconomic characteristics			
Age, y	O		
Sex-male			
Race			
Non-Hispanic	O		
Marital status			
Insurance			O
Income			
Rural			
Smoking status			
Alcohol dependence			
Illicit or cardiotoxic drug use			
Comorbidities			
Renal	O	O	O
Cardiovascular	O		
Hypertension			
Sleep apnea	O	O	O
Diabetes			
Endocrine		O	O
Malignancy or cancer			
Depression	O		
Hematologic		O	O
Respiratory	O		
Prior-to-hospitalization care characteristics			
NSAID[a] - pain management		O	
NSAID[a] - surgery preparation	O		O
Corticosteroid[a]			
Narcotic[a] - pain management	O	O	
Narcotic[a] - surgery preparation	O		O
Chemotherapy[a]			
Radiation[a]			
No. of orthopaedics care visits[b]			
MyChart status-active		O	
No. of no-show >1[b]	O		O
Noncompliance			
Flu vaccine status[b]			
Preoperative care characteristics			
Days between preoperative labs and operation	O		
Hemoglobin			
Hematocrit	O	O	O
Hemoglobin a1c			
Attended preoperative sessions	O	O	O
NSAID	O	O	O
Corticosteroid	O	O	O
Narcotic			
ASA score >2	O	O	O
RAPT Score	O	O	O
Missing indicator of RAPT	O	O	O
Vitals			
Body mass index at admission (kg/m^2)			
Weight (kg)			
At admission (a)	O	O	O
Within 1 y prior to admission (b)			
Difference between (a) and (b)			
Blood pressure (mmHg)			
Systolic BP at admission			
Diastolic BP at admission		O	
Pulse pressure at admission			
MAP at admission			
SPO2 at admission <95%			
Heart rate at admission (mmHg)	O	O	O
Presurvey responses			
PROMIS			
Global health T score-physical		O	
Global health T score-mental		O	O
Survey questionnaire (1)[c]			
Survey questionnaire (2)			O
Survey questionnaire (3)			
Survey questionnaire (4)		O	
Survey questionnaire (5)			
Survey questionnaire (6)		O	O
Survey questionnaire (7)			

Table A.3 (*continued*)

Characteristics	LOS cat 1 (0/1-2/3+)	LOS cat 2 (0/1/2+)	LOS cat 3 (0/1/2/3+)
Survey questionnaire (8)			
Survey questionnaire (9)			
Survey questionnaire (10)			
Disease specific functional score from HOOS or KOOS			

LOS, length of stay; cat, category; NSAID, nonsteroidal antiinflammatory drugs; CAM, confusion assessment method; ER, emergency room; SBP, systolic blood pressure; DBP, diastolic blood pressure; PROMIS, patient-reported outcomes measurement information system; KOOS, Knee Injury and Osteoarthritis Outcome Score; HOOS, Hip Injury Disability and Osteoarthritis Outcome Score.
After the variable selection phases, model input covariates were identified for each outcome setting.
[a] Events within 3 months prior to hospitalization.
[b] Events within 1 year prior to hospitalization.
[c] Survey questionnaire (1) In general, how would you rate your physical health; (2) To what extent are you able to carry out your everyday physical activities; (3) How would you rate your pain on average; (4) How would you rate your fatigue on average; (5) How often have you been bothered by emotional problems such as feeling anxious, depressed, or irritable; (6) In general, would you say your health is; (7) In general, how would you rate your satisfaction with your social activities and relationships; (8) In general, would you say your quality of life is; (9) In general, how would you rate your mental health including your mood and your ability to think; (10) In general, please rate how well you carry out your usual social activities and roles.

Table A.4
Model performance of the threshold model and the ordinal binary decomposition models for the 4-category setting.

Models	Different 3 categories (LOS cat 2: 0/1/2+)		Four categories (LOS cat 3: LOS 0/1/2/3+)	
	Accuracy (avg., sd)	Kendall rank correlation (avg., sd)	Accuracy (avg., sd)	Kendall rank correlation (avg., sd)
Threshold model				
Ordinal regression	0.721 (0.022)	0.475 (0.038)	0.506 (0.023)	0.470 (0.034)
Ordinal decomposition models				
Logistic regression	0.637 (0.022)	0.459 (0.037)	0.535 (0.022)	0.457 (0.037)
SVM with a radial kernel	0.617 (0.025)	0.423 (0.042)	0.508 (0.024)	0.423 (0.043)
Random forest (mtry = 3)	0.632 (0.021)	0.441 (0.038)	0.541 (0.023)	0.453 (0.041)
Random forest (mtry = 4)	0.628 (0.019)	0.432 (0.034)	0.534 (0.021)	0.447 (0.034)
Random forest (mtry = 5)	0.622 (0.020)	0.425 (0.036)	0.530 (0.022)	0.445 (0.039)

avg, average; LOS, length of stay; cat, category; mtry, the numbers of variables randomly sampled as candidates at each split; sd, standard deviation; SVM, support vector machine.

Table A.5
Model performance including the indicator of procedure type for the 3 LOS categorization of interest (ie, outpatient, short stay, and prolonged stay).

Models	Accuracy (avg., IQR)	Kendall rank correlation (avg., IQR)
Threshold model		
Ordinal regression	0.714 (0.696-0.731)	0.450 (0.422-0.471)
Ordinal decomposition models		
Logistic regression	0.734 (0.717-0.749)	0.484 (0.452-0.517)
SVM with a radial kernel	0.730 (0.714-0.746)	0.450 (0.421-0.482)
Random forest (mtry = 3)	0.745 (0.726-0.762)	0.498 (0.464-0.527)
Random forest (mtry = 4)	0.743 (0.729-0.759)	0.494 (0.470-0.519)
Random forest (mtry = 5)	0.739 (0.723-0.759)	0.484 (0.456-0.511)

LOS, length of stay; avg, average; mtry, the numbers of variables randomly sampled as candidates at each split; IQR, interquartile range; SVM, support vector machine.