

RESEARCH

Open Access



Development and internal validation of machine learning algorithms to predict patient satisfaction after total hip arthroplasty

Siyuan Zhang¹, Jerry Yongqiang Chen², Hee Nee Pang², Ngai Nung Lo², Seng Jin Yeo² and Ming Han Lincoln Liow^{2*} 

Abstract

Background: Patient satisfaction is a unique and important measure of success after total hip arthroplasty (THA). Our study aimed to evaluate the use of machine learning (ML) algorithms to predict patient satisfaction after THA.

Methods: Prospectively collected data of 1508 primary THAs performed between 2006 and 2018 were extracted from our joint replacement registry and split into training (80%) and test (20%) sets. Supervised ML algorithms (Random Forest, Extreme Gradient Boosting, Support Vector Machines, Logistic LASSO) were developed with the training set, using patient demographics, comorbidities and preoperative patient reported outcome measures (PROMs) (Short Form-36 [SF-36], physical component summary [PCS] and mental component summary [MCS], Western Ontario and McMaster's Universities Osteoarthritis Index [WOMAC] and Oxford Hip Score [OHS]) to predict patient satisfaction at 2 years postoperatively. Predictive performance was evaluated using the independent test set.

Results: Preoperative models demonstrated fair discriminative ability in predicting patient satisfaction, with the LASSO model achieving a maximum AUC of 0.76. Permutation importance revealed that the most important predictors of dissatisfaction were (1) patient's age, (2) preoperative WOMAC, (3) number of comorbidities, (4) preoperative MCS, (5) previous lumbar spine surgery, and (6) low BMI (< 18.5).

Conclusion: Machine learning algorithms demonstrated fair discriminative ability in predicting patient satisfaction after THA. We have identified modifiable and non-modifiable predictors of postoperative satisfaction which could enhance preoperative counselling and improve health optimization prior to THA.

Keywords: Machine learning, Artificial intelligence, Total hip arthroplasty, Satisfaction, Patient-reported outcome measures

Introduction

Total hip arthroplasty (THA) is a commonly performed and effective operation for symptomatic hip osteoarthritis, leading to significant improvements in pain,

function and quality of life [1]. To assess a patient's response after THA, patient-reported outcome measures (PROMs) and patient satisfaction are often used. While PROMs predominantly measure improvements in pain and function, patient satisfaction is a unique and holistic outcome that reflects the individual's subjective quality of life improvements in relation to preoperative expectations [2, 3].

* Correspondence: lincoln.liow.m.h@singhealth.com.sg

²Department of Orthopaedic Surgery, Singapore General Hospital, 20 College Road, Academia, Level 4, Singapore 169856, Singapore
Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

This is especially important since the primary aim of elective joint arthroplasty is to improve the patient's quality of life. Prior research has shown that patient satisfaction is a complex and multifactorial phenomenon, with postoperative satisfaction influenced by various determinants such as age, sex, mental health, preoperative expectations, as well as postoperative functional improvements [4–7]. Despite the efficacy of THA, recent literature has shown that around 10–20% of patients remain dissatisfied after surgery [3, 8–10].

As healthcare systems worldwide transition to a value-based, patient-centered model, patient satisfaction will become increasingly important as a measure of success in total joint arthroplasty (TJA) [11, 12]. Being able to predict patients who may be dissatisfied enables surgeons to provide individualized preoperative counselling which could help address unrealistic patient expectations regarding THA [13, 14]. This information may also allow for early intervention and optimization of the patient's physical and mental health. To do so, accurate prediction models are required. One way such models can be developed is using machine learning (ML) algorithms. ML is a subset of artificial intelligence (AI) that uses computer algorithms capable of learning from real-world data and using these insights to predict an outcome without being explicitly programmed [15, 16]. In recent years, the use of ML and AI in medicine have gained traction due to their ability to accurately predict medical outcomes, ranging from heart failure to cancer prognosis [17–19]. In orthopedic surgery, prior studies by Fontana et al, Huber et al and Kunze et al have demonstrated the feasibility of ML algorithms in predicting PROM improvements after TJA [8, 9, 20]. More recently, Kunze et al and Farooq et al have also tried to predict patient satisfaction after TKA, achieving an AUC of 0.77 and 0.81 respectively [21, 22]. However, there have been no prior studies using ML algorithms to predict patient satisfaction after THA.

Thus, the primary aim of our study was to evaluate whether ML algorithms can predict patient satisfaction after THA. Our secondary aim was to identify the underlying variables which drive prediction in these models.

Materials and methods

Data

Following ethics approval by the institutional ethics review board (CIRB 2020/2843), we reviewed prospectively collected data of consecutive patients who underwent elective primary THA from a single institution's joint replacement registry between 2006 and 2018. We identified 1996 adult patients who had undergone unilateral primary THA during this period, of which 1508 (75.6%) had completed their 2 year postoperative follow-up for

patient satisfaction and PROMs (SF-36 PCS/MCS, WOMAC, OHS).

Outcomes

Patient satisfaction was recorded at 2 years postoperatively and was rated on a 6-level Likert scale, similar to the approach described by Bourne et al [23]. Patients were asked to grade their level of satisfaction (“terrible”, “poor”, “fair”, “good”, “very good” or “excellent”) and this was dichotomized into patients who were satisfied (“excellent”, “very good” or “good”) and patients who were not satisfied (“terrible”, “poor” or “fair”).

Input variables

Input variables used in this study were prospectively collected by our joint replacement registry. These included patient demographics, comorbidities and PROMs (Table 1). PROMs collected include the Short Form-36 (SF-36), Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) and Oxford Hip Score (OHS).

The SF-36 is a generic health questionnaire that measures an individual's health-related quality of life, and its 8 domains are commonly aggregated into the physical component summary (PCS) and mental component summary (MCS) [24]. On the other hand, the WOMAC is a disease-specific questionnaire for lower limb arthritis, with 24 items grouped into 3 dimensions of pain (5 items), stiffness (2 items) and physical function (17 items) [25]. The total WOMAC score was calculated by summing the aggregate scores for the 3 dimensions and transforming it to a scale of 0–100. Lastly, the OHS is a 12-item questionnaire that assesses hip function and pain. Each item has 5 response options, giving a score between 1 and 5. These scores are summated into a final score ranging between 12 and 60, with a higher score indicating greater disability [26].

Data handling and model training

Data analysis and modelling were performed using Python 3.7 (Python Software Foundation, Wilmington, DE, USA), the Anaconda Distribution (Anaconda, Inc., Austin, TX, USA) and R software, version 4.0.3 (R Foundation for Statistical Computing, Vienna, Austria, 2019). We used four of the most popular supervised ML algorithms (Scikit-learn version 0.24): Random Forest (RF), Extreme Gradient Boosting (XGB), Support Vector Machine (SVM), and Logistic Regression with L1-regularization (LASSO).

The 1508 primary THAs were randomly split into a training set and testing set at an 80–20 ratio. Data in the training set ($n = 1206$) were used to train and fit the various ML algorithms while the independent test ($n = 302$) was reserved for the final evaluation of the model's

Table 1 Input variables for ML models

Demographics	Comorbidities	Preoperative PROMs
Age*	Number of comorbidities*	SF-36 PCS*
Sex	Diabetes	SF-36 MCS*
BMI (numerical)*	Hypertension	WOMAC*
BMI (categorical)*	High cholesterol	OHS*
	Ischemic heart disease,	
	Stroke	
	Renal disease	
	Back pain	
	Depression	
	Previous hip surgery*	
	Previous knee surgery	
	Previous lumbar spine surgery*	

*10 candidate variables selected from recursive feature elimination

performance. The only missing input variable was body mass index (BMI) ($n = 102$, 6.8%). Missing BMI values in the training and test set were imputed using the mean BMI of the training and test sets respectively. Due to class imbalance of the outcomes (majority of patients were satisfied), which could adversely impact the predictive performance of some ML algorithms, we employed a technique called random oversampling during model training. Random oversampling is a simple upsampling technique that randomly resamples the minority class to reach balanced class ratios in the training set [27].

Recursive feature elimination (RFE) using an RF classifier was used to select 10 candidate variables from the larger pool of input variables to train the ML models. RFE selects candidate variables by iteratively calculating the importance scores for each variable and ranking them in order of importance. All ML models were trained using 5-fold stratified cross-validation in the training set to optimize their hyperparameters before final evaluation of their performance on the independent test set.

Model evaluation

All our models were evaluated on an independent test set that was not involved in model training. The main evaluation metric used was the area under the receiver operating characteristic curve (AUC), which is a measure of the model's ability to discriminate between two different classes. A perfect classifier would have an AUC of 1.0 while a completely random classifier (i.e. flipping a coin) would have an AUC of 0.5. Generally speaking, an AUC of 0.7–0.8 is fair, 0.8–0.9 is good and 0.9–1.0 is excellent. Apart from the AUC, other evaluation metrics used include the Brier score, sensitivity and specificity and calibration slope and intercept.

Variable importance

The relative importance of input variables was assessed using permutation importance, a model-agnostic method

that has been shown to generate reliable insights correlating with clinical intuition [28]. The permutation importance of a variable is defined as the decrease in accuracy (AUC in this case) of the trained model on the test set when the said variable is randomly shuffled, thus giving us an estimate of how much the input variable contributes to predictive performance.

Results

Baseline patient characteristics

Baseline patient characteristics of the 1508 THAs in our study are summarized in Table 2. 69.8% ($n = 1052$) of the patients were female and 30.2% ($n = 456$) were male. The mean age was 62.9 (SD:12.1) and mean BMI was 25.8 (SD:4.9).

Patient satisfaction and PROM improvements

At 2 years postoperatively, 94.3% ($n = 1422$) of the patients were satisfied and 5.7% ($n = 86$) were dissatisfied (Table 2). Mean PROM improvements were + 20.2 (SD: 12.1) for SF-36 PCS, + 6.9 (SD:12.2) for SF-36 MCS, + 41.5 (SD:21.1) for WOMAC and - 24.1 (SD:9.8) for the OHS.

Model performance – patient satisfaction

Evaluation results for predicting patient satisfaction are given in Table 3. Our ML models demonstrated fair discriminative ability in predicting patient satisfaction, with the LASSO model achieving a maximum AUC of 0.76 (Figs. 1 and 2). This was followed by SVM (AUC:0.74), RF (AUC:0.68) and XGB (AUC:0.66).

Permutation importance (Fig. 3) revealed that the most important predictors of dissatisfaction in the preoperative model were (1) patient's age, (2) preoperative WOMAC, (3) number of comorbidities, (4) preoperative MCS, (5) previous lumbar spine surgery and (6) low BMI (< 18.5).

Table 2 Baseline Patient Characteristics

Variables	Training set (n = 1206)	Test set (n = 302)	P-value
Age	62.8 (12.1)	63.2 (11.8)	0.623
Sex (Female)	841 (69.7%)	211 (69.9%)	0.964
BMI	25.9 (4.8)	25.3 (4.3)	0.075
BMI (categorical)			0.488
< 18.5	39 (3.2%)	14 (4.6%)	–
18.5–29.9	982 (81.4%)	241 (79.8%)	–
≥ 30	185 (15.3%)	47 (15.6%)	–
Comorbidities			
Number of comorbidities	1.0 (1.1)	0.9 (1.1)	0.417
Diabetes	133 (11.0%)	30 (9.9%)	0.584
Hypertension	518 (43.0%)	115 (38.1%)	0.125
High cholesterol	389 (32.3%)	93 (30.8%)	0.626
IHD	52 (4.3%)	16 (5.3%)	0.460
Stroke	17 (1.4%)	7 (2.3%)	0.259
Renal disease	22 (1.8%)	7 (2.3%)	0.576
Back pain	35 (2.9%)	7 (2.3%)	0.581
Depression	7 (0.6%)	2 (0.6%)	0.999
Surgical History			
Previous knee surgery	133 (11.0%)	37 (12.3%)	0.548
Previous hip surgery	236 (19.6%)	64 (21.2%)	0.527
Previous lumbar spine surgery	69 (5.7%)	28 (9.3%)	0.025*
Preop PROMs			
SF-36 PCS	27.4 (8.9)	26.7 (9.8)	0.233
SF-36 MCS	49.1 (12.1)	49.2 (12.2)	0.899
WOMAC	49.5 (20.9)	48.5 (20.9)	0.473
OHS	40.0 (9.2)	40.8 (9.3)	0.190
2-year PROM Improvement			
SF-36 PCS	+ 20.2 (12.0)	+ 19.8 (12.4)	0.601
SF-36 MCS	+ 6.9 (12.1)	+ 6.7 (12.6)	0.804
WOMAC	+ 41.4 (20.9)	+ 41.6 (22.1)	0.921
OHS	–24.1 (9.7)	–24.4 (10.2)	0.658
2-year Satisfaction			
Satisfied	69 (5.7%)	17 (5.6%)	0.951

Continuous outcomes are reported as mean (standard deviation) while categorical outcomes are presented as number (percentage)
P-values are calculated using two-sample t-tests for continuous variables and Chi-squared test/Fisher's exact test for categorical variables
*: P-value < 0.05

Table 3 Model performance for predicting patient satisfaction on test set (n = 302)

	LASSO	SVM	RF	XGB
AUC	0.76 (0.67–0.86)	0.74 (0.63–0.85)	0.68 (0.56–0.80)	0.66 (0.50–0.78)
Brier score	0.23	0.23	0.23	0.21
Threshold	0.50	0.52	0.53	0.58
Sensitivity	65.3%	57.9%	53.0%	50.5%
Specificity	82.4%	76.5%	76.5%	76.5%
Calibration slope	1.29	0.44	1.06	0.29
Calibration intercept	0.13	0.70	0.40	0.77

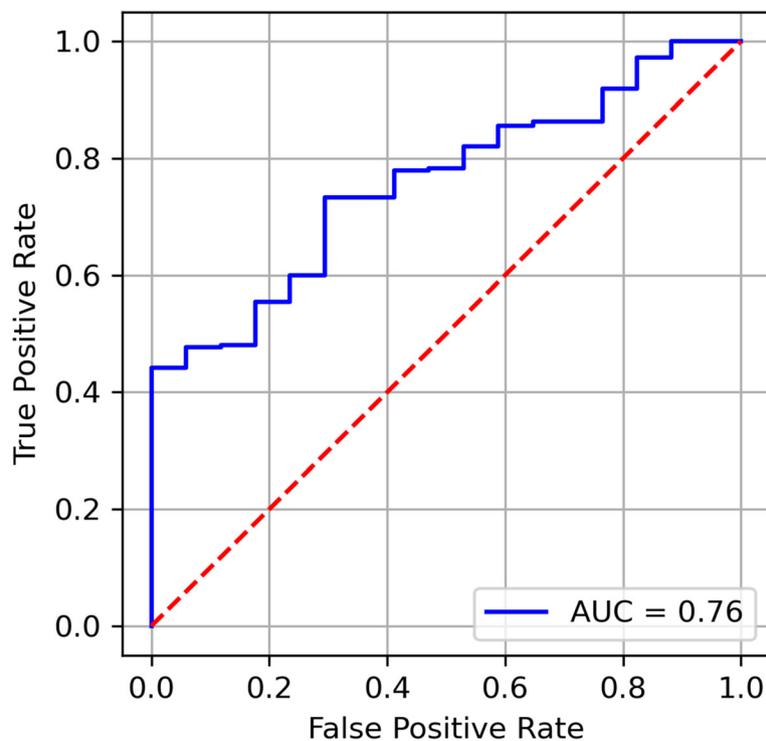


Fig. 1 ROC curve for patient satisfaction using the LASSO model, achieving an AUC of 0.76

Discussion

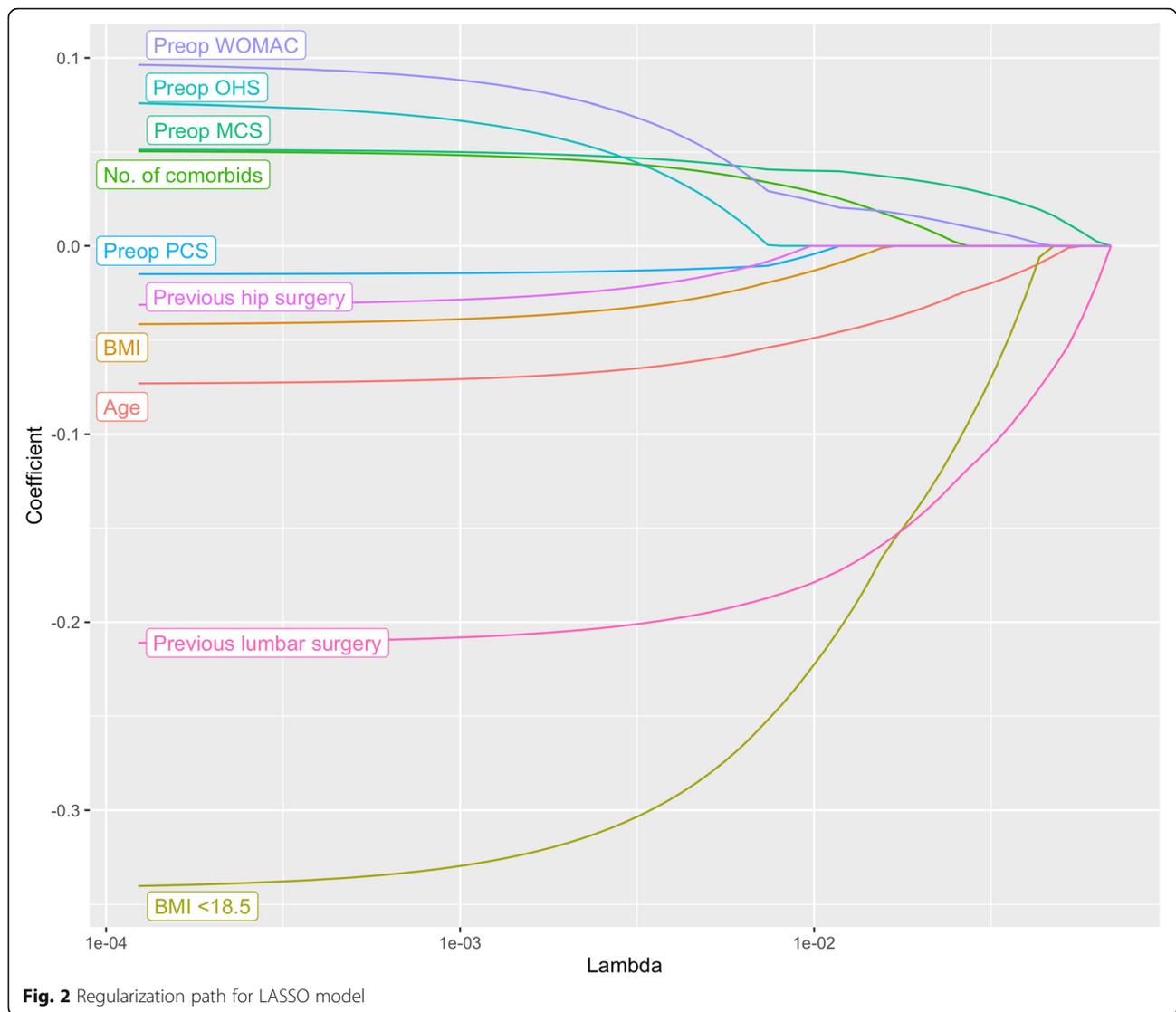
Patient satisfaction is an important outcome measure in THA due to its ability to reflect subjective improvements in quality of life and the fulfillment of preoperative expectations [6, 29]. Our study was the first to use ML algorithms to predict patient satisfaction after THA. While prior studies by Kunze et al and Farooq et al have demonstrated the ability of ML algorithms to predict patient satisfaction after TKA, achieving an AUC of 0.77 and 0.81 respectively, their findings might not be generalizable to patients undergoing THA [21, 22]. Notably, several of the predictors identified in their studies (i.e. condylar-stabilizing implant, preservation of the posterior cruciate ligament, preoperative knee society score) were specific to TKA and not applicable for THA.

Our results have shown that ML algorithms using preoperative data had fairly good discriminative ability in predicting patient satisfaction after THA (AUC: 0.76). Although far from ideal, our results demonstrated the feasibility of using ML algorithms to identify high risk patients who may experience dissatisfaction after THA. This information could be used by surgeons to enhance preoperative counselling and manage patient expectations regarding THA, as prior studies have reported that interventions encouraging realistic expectations could improve satisfaction after joint arthroplasty [30]. Identification of at-risk patients may also allow for early intervention and optimization of their physical and mental

health prior to THA, with prior studies showing that some predictors of satisfaction may be fully or partially modifiable [7, 31].

Our study also identified several modifiable and non-modifiable predictors of postoperative satisfaction. Some of our predictors, patient's age (1) and number of comorbidities (3), were similarly observed in Kunze et al's study amongst TKA patients [21]. Despite this, the impact of age on satisfaction remains controversial: While some studies have found that advanced age was associated with dissatisfaction, others have reported that younger patients are more likely to be dissatisfied, possibly due to their higher expectations and functional demands [6, 7]. Regardless, our current findings add to existing literature and reinforces the importance of these variables in a predictive model.

Next, we also identified preoperative WOMAC (2) and SF-36 MCS (4) as important preoperative predictors of dissatisfaction. This is consistent with prior literature which have reported an association between poorer preoperative physical and mental health state and dissatisfaction after TJA [5, 7, 29]. While the underlying mechanisms remain unclear, some authors have suggested that poorer preoperative PROMs may be indicative of poor musculoskeletal health, pathology in other joints (e.g., spine) and even increased pain sensitivity, all of which could potentially impede functional recovery and resolution of symptoms after THA [7, 32]. On the other hand, it is widely recognized that poor mental health



influences a patient's perception of disability and sensitivity to pain which may contribute to poorer satisfaction [33]. Recent studies have also suggested the possibility of mental health optimization, with Geng et al showing that psychological intervention in patients with depression could improve post-TKA satisfaction [31].

We also identified novel predictors of patient satisfaction, such as previous lumbar spine surgery (5). While there is an increasing understanding of the complex relationship between the spine and pelvis, with prior studies identifying lumbar spine pathology and surgery as risk factors for instability and dislocation after THA, much less is understood about how it influences patient satisfaction [34, 35]. Although our results suggest that previous lumbar surgery may be predictive of dissatisfaction, it should be noted that its impact is relatively small and further research is needed to better elucidate the underlying mechanisms. Lastly, we also observed that

patients with low BMI (<18.5) (6) were more likely to be dissatisfied. This could partially be due to poorer nutritional status being associated with poorer physical health, which in turn limits functional recovery and symptom resolution. Prior studies have also reported that patients with low BMI may experience more complications and a longer length of hospital stay after TJA [36, 37]. Although the importance of BMI was observed to be relatively low, it is highly modifiable and highlights the possibility of preoperative nutritional optimization.

Our study had several limitations. First, data used in this study were extracted from a single institution and thus it is unclear to what extent our findings can be extrapolated to other healthcare institutions. Although model performance in our study was assessed on an independent test set, future studies should focus on external validation of ML algorithms using data from other institutions. There was also a significant proportion ($n =$

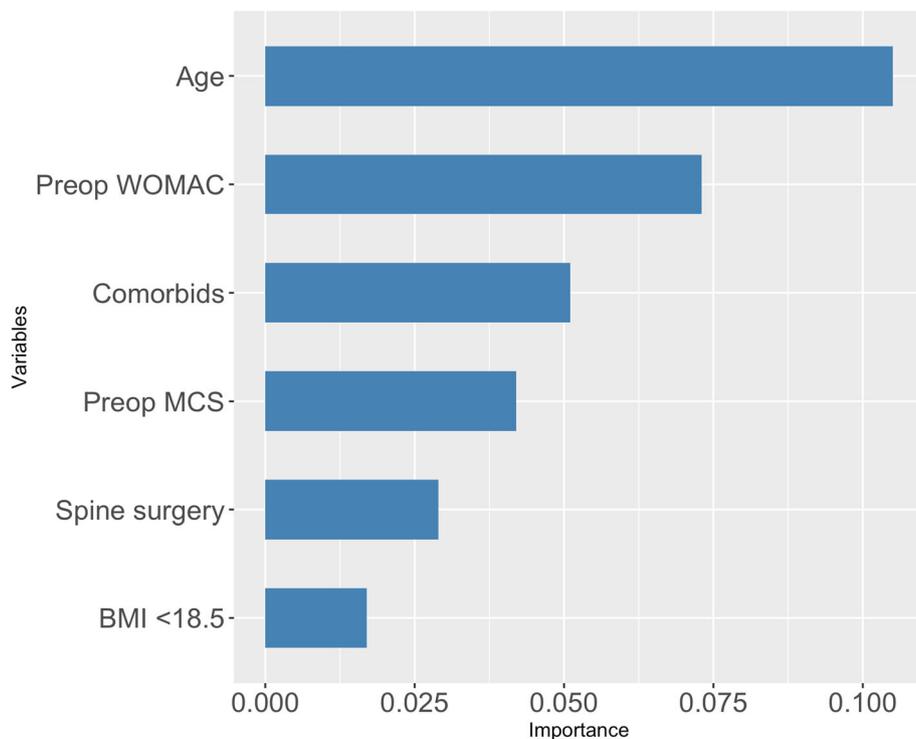


Fig. 3 Most important predictors of patient satisfaction: (1) patient's age, (2) preoperative WOMAC, (3) number of comorbidities, (4) preoperative MCS, (5) previous lumbar spine surgery and (6) low BMI (< 18.5)

488, 24.4%) of patients who were excluded from our study due to incomplete follow-up. This follow-up rate of 75.6%, while not ideal, is expected for studies with intermediate follow-up durations, with prior studies reporting similar follow-up rates of around 70–80% [32, 33]. Next, the performance of our ML models may have been limited by the input variables currently available in our joint replacement registry. Although our variables are largely similar to predictors identified in a systematic review by Gunaratne et al, there may be other factors such as the patient's preoperative expectations, psychological coping mechanisms and other socio-economic factors which may influence postoperative satisfaction [7]. However, these factors are often overlooked by joint replacement registries and remain poorly understood, thus highlighting the need for further research in this area. Lastly, we acknowledge that our models may not be fully optimized in terms of calibration as the primary aim of our study was to demonstrate the ability of ML models to discriminate between the two outcomes.

Conclusion

Machine learning (ML) algorithms demonstrated fair discriminative ability in predicting patient satisfaction after THA. We have identified modifiable and non-modifiable predictors of postoperative satisfaction which

could enhance preoperative counselling and improve health optimization prior to THA.

Authors' contributions

All authors were involved in the conception and design of the study. SZ was involved in statistical analysis and modelling. SZ and MHLL were involved in drafting of the initial manuscript. All authors contributed to the critical revision and approval of the final manuscript. MHLL was in charge of overall project supervision.

Funding

No funding was received for conducting this study.

Availability of data and materials

Data available on request.

Declarations

Conflict of interest

Seng Jin Yeo is a member of the Editorial Board of *Arthroplasty* and other authors declare that they have no competing interests. All authors were not involved in the journal's review of or decisions related to, this manuscript.

Ethics approval and consent to participate

Ethics approval (CIRB: 2020/2843) was obtained prior to initiation of study. Waiver of informed consent was obtained from the ethics board as de-identified data from the joint replacement registry were used.

Consent for publication

Not applicable.

Author details

¹Yong Loo Lin School of Medicine, National University of Singapore, 1E Kent Ridge Road, NUHS Tower Block, Level 11, Singapore 119228, Singapore.

²Department of Orthopaedic Surgery, Singapore General Hospital, 20 College Road, Academia, Level 4, Singapore 169856, Singapore.

Received: 2 May 2021 Accepted: 22 July 2021

Published online: 02 September 2021

References

- Ethgen O, Bruyère O, Richey F, Dardennes C, Reginster JY. Health-related quality of life in total hip and total knee arthroplasty. A qualitative and systematic review of the literature. *J Bone Joint Surg Am*. 2004;86(5):963–74. <https://doi.org/10.2106/00004623-200405000-00012>.
- Siljander MP, McQuivey KS, Fahs AM, Galasso LA, Serdahely KJ, Karadsheh MS. Current trends in patient-reported outcome measures in Total joint Arthroplasty: a study of 4 major Orthopaedic journals. *J Arthroplast*. 2018; 33(11):3416–21. <https://doi.org/10.1016/j.arth.2018.06.034>.
- Halawi MJ, Jongbloed W, Baron S, Savoy L, Cote MP, Lieberman JR. Patient-reported outcome measures are not a valid proxy for patient satisfaction in Total joint Arthroplasty. *J Arthroplast*. 2020;35(2):335–9. <https://doi.org/10.1016/j.arth.2019.09.033>.
- Mancuso CA, Salvati EA, Johanson NA, Peterson MG, Charlson ME. Patients' expectations and satisfaction with total hip arthroplasty. *J Arthroplast*. 1997; 12(4):387–96. [https://doi.org/10.1016/s0883-5403\(97\)90194-7](https://doi.org/10.1016/s0883-5403(97)90194-7).
- Anakwe RE, Jenkins PJ, Moran M. Predicting dissatisfaction after total hip arthroplasty: a study of 850 patients. *J Arthroplast*. 2011;26(2):209–13. <https://doi.org/10.1016/j.arth.2010.03.013>.
- Palazzo C, Jourdan C, Descamps S, Nizard R, Hamadouche M, Anract P, et al. Determinants of satisfaction 1 year after total hip arthroplasty: the role of expectations fulfilment. *BMC Musculoskelet Disord*. 2014;15(1):1–9. <https://doi.org/10.1186/1471-2474-15-53>.
- Gunaratne R, Pratt DN, Banda J, Fick DP, Khan RJK, Robertson BW. Patient dissatisfaction following Total knee Arthroplasty: a systematic review of the literature. *J Arthroplast*. 2017;32(12):3854–60. <https://doi.org/10.1016/j.arth.2017.07.021>.
- Fontana MA, Lyman S, Sarker GK, Padgett DE, MacLean CH. Can machine learning algorithms predict which patients will achieve minimally clinically important differences from Total joint Arthroplasty? *Clin Orthop Relat Res*. 2019;477(6):1267–79. <https://doi.org/10.1097/CORR.0000000000000687>.
- Huber M, Kurz C, Leidl R. Predicting patient-reported outcomes following hip and knee replacement surgery using supervised machine learning. *BMC Med Inform Decis Mak*. 2019;19(1):3. <https://doi.org/10.1186/s12911-018-0731-6>.
- Yeo MGHGG, Chen JY, Lo NN, Yeo SJ, Liow MHL. Are Oxford hip score and Western Ontario and McMaster universities osteoarthritis index useful predictors of clinical meaningful improvement and satisfaction after Total hip Arthroplasty? *J Arthroplast*. 2020;35(9):2458–64. <https://doi.org/10.1016/j.arth.2020.04.034>.
- Wei DH, Hawker GA, Jevsevar DS, Bozic KJ. Improving value in musculoskeletal care delivery: AOA critical issues. *J Bone Joint Surg Am*. 2015;97(9):769–74. <https://doi.org/10.2106/JBJS.N.00841>.
- Halawi MJ. Outcome measures in Total joint Arthroplasty: current status, challenges, and future directions. *Orthopedics*. 2015;38(8):e685–9. <https://doi.org/10.3928/01477447-20150804-55>.
- Lee EOE. Shared decision making to improve care and reduce costs. *N Engl J Med*. 2013;368(1):6–8. <https://doi.org/10.1056/NEJMp1209500>.
- Elwyn G, Frosch D, Thomson R, Joseph-Williams N, Lloyd A, Kinnersley P, et al. Shared decision making: a model for clinical practice. *J Gen Intern Med*. 2012;27(10):1361–7. <https://doi.org/10.1007/s11606-012-2077-6>.
- Bini SA. Artificial intelligence, machine learning, deep learning, and cognitive computing: what do these terms mean and how will they impact health care? *J Arthroplast*. 2018;33(8):2358–61. <https://doi.org/10.1016/j.arth.2018.02.067>.
- Wiens J, Shenoy ES. Machine learning for healthcare: on the verge of a major shift in healthcare epidemiology. *Clin Infect Dis*. 2018;66(1):149–53. <https://doi.org/10.1093/cid/cix731>.
- Ngiam KY, Khor IW. Big data and machine learning algorithms for health-care delivery. *Lancet Oncol*. 2019;20(5):e262–73. [https://doi.org/10.1016/S1470-2045\(19\)30149-4](https://doi.org/10.1016/S1470-2045(19)30149-4).
- Dorado-Díaz PI, Sampedro-Gómez J, Vicente-Palacios V, Sánchez PL. Applications of artificial intelligence in cardiology. The future is already here. *Rev Esp Cardiol (Engl Ed)*. 2019;72(12):1065–75. <https://doi.org/10.1016/j.rec.2019.05.014>.
- Beam AL, Kohane IS. Big data and machine learning in health care. *JAMA*. 2018;319(13):1317–8. <https://doi.org/10.1001/jama.2017.18391>.
- Kunze KN, Karhade AV, Sadauskas AJ, Schwab JH, Levine BR. Development of machine learning algorithms to predict clinically meaningful improvement for the patient-reported health state after Total hip Arthroplasty. *J Arthroplast*. 2020;35(8):2119–23. <https://doi.org/10.1016/j.arth.2020.03.019>.
- Kunze KN, Polce EM, Sadauskas AJ, Levine BR. Development of machine learning algorithms to predict patient dissatisfaction after primary Total knee Arthroplasty. *J Arthroplast*. 2020;35(11):3117–22. <https://doi.org/10.1016/j.arth.2020.05.061>.
- Farooq H, Deckard ER, Ziemba-Davis M, Madsen A, Meneghini RM. Predictors of patient satisfaction following primary Total knee Arthroplasty: results from a traditional statistical model and a machine learning algorithm. *J Arthroplast*. 2020;35(11):3123–30. <https://doi.org/10.1016/j.arth.2020.05.077>.
- Bourne RB, Chesworth BM, Davis AM, Mahomed NN, Charron KDJ. Patient satisfaction after total knee arthroplasty: who is satisfied and who is not? *Clin Orthop Relat Res*. 2010;468(1):57–63. <https://doi.org/10.1007/s11999-009-1119-9>.
- Laucis NC, Hays RD, Bhattacharya T. Scoring the SF-36 in Orthopaedics: a brief guide. *J Bone Joint Surg Am*. 2015;97(19):1628–34. <https://doi.org/10.2106/JBJS.O.00030>.
- Walker LC, Clement ND, Deehan DJ. Predicting the outcome of Total knee Arthroplasty using the WOMAC score: a review of the literature. *J Knee Surg*. 2019;32(8):736–41. <https://doi.org/10.1055/s-0038-1666866>.
- Nilsson ABA. Measures of hip function and symptoms: Harris hip score (HHS), hip disability and osteoarthritis outcome score (HOOS), Oxford hip score (OHS), lequesne index of severity for osteoarthritis of the hip (LISOH), and American Academy of orthopedic surgeons (AAOS) hip and knee questionnaire. *Arthritis Care Res*. 2011;63:S200–7.
- Batista GEAPA, Prati RC, Monard MC. A study of the behavior of several methods for balancing machine learning training data. *SIGKDD Explor Newsl*. 2004;6(1):20–9. <https://doi.org/10.1145/1007730.1007735>.
- Cava W, Bauer C, Moore JH, Pendergrass SA. Interpretation of machine learning predictions for patient outcomes in electronic health records. *AMIA Annu Symp Proc*. 2019;2020:572–81.
- Hamilton DF, Lane JV, Gaston P, Patton JT, Macdonald D, Simpson AH, et al. What determines patient satisfaction with surgery? A prospective cohort study of 4709 patients following total joint replacement. *BMJ Open*. 2013; 3(4):e002525. <https://doi.org/10.1136/bmjopen-2012-002525>.
- Tolk JJ, Janssen RPA, Haanstra TM, van der Steen MC, Bierma-Zeinstra SMA, Reijnen M. The influence of expectation modification in knee arthroplasty on satisfaction of patients: a randomized controlled trial. *Bone Joint J*. 2021 Apr; 103-B(4):619–26. <https://doi.org/10.1302/0301-620X.103B4.BJJ-2020-0629.R3>.
- Geng X, Wang X, Zhou G, Li F, Li Y, Zhao M, et al. A randomized controlled trial of psychological intervention to improve satisfaction for patients with depression undergoing TKA: a 2-year follow-up. *J Bone Joint Surg Am*. 2021; 103(7):567–74. <https://doi.org/10.2106/JBJS.20.00169>.
- Kim TK, Chang CB, Kang YG, Kim SJ, Seong SC. Causes and predictors of patient's dissatisfaction after uncomplicated total knee arthroplasty. *J Arthroplast*. 2009;24(2):263–71. <https://doi.org/10.1016/j.arth.2007.11.005>.
- Gandhi R, Davey JR, Mahomed NN. Predicting patient dissatisfaction following joint replacement surgery. *J Rheumatol*. 2008;35(12):2415–8. <https://doi.org/10.3899/jrheum.080295>.
- Eftekhary N, Shimmin A, Lazennec JY, Buckland A, Schwarzkopf R, Dorr LD, et al. A systematic approach to the hip-spine relationship and its applications to total hip arthroplasty. *Bone Joint J*. 2019;101-B(7):808–16. <https://doi.org/10.1302/0301-620X.101B7.BJJ-2018-1188.R1>.
- Ike H, Dorr LD, Trasolini N, Steffl M, McKnight B, Heckmann N. Spine-pelvis-hip relationship in the functioning of a Total hip replacement. *J Bone Joint Surg Am*. 2018;100(18):1606–15. <https://doi.org/10.2106/JBJS.17.00403>.
- Sayeed Z, Anoushiravani AA, Chambers MC, Gilbert TJ, Scaife SL, El-Othmani MM, et al. Comparing in-hospital Total joint Arthroplasty outcomes and resource consumption among underweight and morbidly obese patients. *J Arthroplast*. 2016;31(10):2085–90. <https://doi.org/10.1016/j.arth.2016.03.015>.
- Katakam A, Melnic CM, Bragdon CR, Sauder N, Collins AK, Bedair HS. Low body mass index is a predictor for mortality and increased length of stay following Total joint Arthroplasty. *J Arthroplast*. 2021;36(1):72–7. <https://doi.org/10.1016/j.arth.2020.07.055>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.