OPEN ACCESS

Predicting individual knee range of motion, knee pain, and walking limitation outcomes following total knee arthroplasty

Yong-Hao PUA [1], Cheryl Lian-Li POON [1], Felicia Jie-Ting SEAH [2], Julian THUMBOO [3], Ross Allan CLARK [4], Mann-Hong TAN [5], Hwei-Chi CHONG [1], John Wei-Ming TAN [1], Eleanor Shu-Xian CHEW [1], and Seng-Jin YEO [5]

[1] Department of Physiotherapy, Singapore General Hospital, Singapore; [2] Department of Physiotherapy, Sengkang General Hospital, Singapore; [3] Department of Rheumatology and Immunology, Singapore General Hospital, Singapore; [4] Research Health Institute, University of the Sunshine Coast, Sunshine Coast, Australia; [5] Department of Orthopaedic Surgery, Singapore General Hospital, Singapore
Correspondence: pua.yong.hao@sgh.com.sg
Submitted 2018-08-04. Accepted 2018-11-18.

Background and purpose — Up to 20% of patients are dissatisfied after total knee arthroplasty (TKA), mainly because of pain and restricted physical function. We developed a prediction model for 6-month knee range of motion, knee pain, and walking limitations in patients undergoing TKA surgery.

Patients and methods — We performed a prospective cohort study of 4,026 patients who underwent elective, primary TKA between July 2013 and July 2017. Candidate predictors included demographic, clinical, psychosocial, and preoperative outcome measures. The outcomes of interest were (i) knee extension and flexion range of motion, (ii) knee pain rated on a 5-point ordinal scale, and (iii) self-reported maximum walk time at 6 months post TKA. For each outcome, we fitted a multivariable proportional odds regression model with bootstrap internal validation.

Results — At 6 months post TKA, around 5% to 20% of patients had a flexion contracture ≥ 10°, range of motion < 90°, moderate to severe knee pain, or a maximum walk time ≤ 15 minutes. The model c-indices (the probabilities to correctly discriminate between 2 patients with different levels of follow-up TKA outcomes) when evaluating these patients were 0.71, 0.79, 0.65, and 0.76, respectively. Each postoperative outcome was strongly influenced by the same outcome measure obtained preoperatively (all p-values < 0.001). Additional statistically significant predictors were age, sex, race, education level, diabetes mellitus, preoperative use of gait aids, contralateral knee pain, and psychological distress (all p-values < 0.001).

Interpretation — We have developed models to predict, for individual patients, their likely post-TKA levels of knee extension and flexion range of motion, knee pain, and walking limitations. After external validation, they can potentially be used preoperatively to identify at-risk patients and to help patients set more realistic expectations about surgical outcomes.

Some 11% to 20% of patients are dissatisfied following total knee arthroplasty (TKA) mainly because of knee motion limitations (Matsuda et al. 2013, Huang et al. 2017), knee pain (Gunaratne et al. 2017), and functional limitations (Gunaratne et al. 2017). Early and accurate identification of patients at risk for poor post-TKA outcomes would better direct resources toward preventive care for them. Furthermore, to facilitate shared decision-making, providing patients preoperatively (Barlow et al. 2016) with individualized information on expected post-TKA outcomes may help them set more realistic expectations about surgical outcomes (Husain and Lee 2015, Volkmann and FitzGerald 2015), which in turn improves patient satisfaction.

Few studies have systematically combined multiple predictors to generate individualized outcome predictions. These studies have focused singularly on combined pain and physical function outcomes (Dowsey et al. 2016, Sanchez-Santos et al. 2018), health-related quality of life (Gutacker and Street 2017), or patient satisfaction (Van Onsem et al. 2016). Although consensus statements (Singh et al. 2016, Lange et al. 2017) have advocated knee range of motion, knee pain, and physical function as distinct and important post-TKA outcomes, no studies have developed a prediction model to provide individualized predictions on these outcomes separately. Thus, we aimed to create a prediction model for post-TKA knee range- of motion, knee pain, and walking limitations.

Patients and methods

Between July 2013 and July 2017, we identified 5,491 patients aged ≥ 50 years who underwent a unilateral primary TKA for knee osteoarthritis (OA) in Singapore General Hospital—the largest tertiary teaching hospital in Singapore, which performed half of all knee arthroplasties in the nation of 5.6 million people (Singapore Ministry of Health [last accessed Janu-

DOI 10.1080/17453674.2018.1560647

ary 19, 2018]). We excluded patients who underwent revision knee surgery within 6 months post TKA (n = 16). We also excluded patients who had a history of rheumatoid arthritis (n = 58) and patients with stroke or Parkinson's disease (n = 108). For patients with consecutive admissions for TKA (n = 863), only data from their first admission were used. Of the remaining 5,309 patients, we selected a cohort of 4,026 patients with non-missing 6-month follow-up outcomes (Figure 1, Supplementary data). Included patients were similar to those who were excluded because of missing data (Table 1, Supplementary data). All data were collected by physiotherapists and data technicians trained in the testing procedures and entered into an electronic registry database as per routine practice policies of our institution. Study design and reporting was based on the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) Statement (Collins et al. 2015).

Candidate predictors

We selected candidate predictors based on clinical expertise, literature review (Dowsey et al. 2016, Van Onsem et al. 2016, Gutacker and Street 2017), and data availability in our databases. To improve the practicality of the prediction models, we considered variables that were less equipment-dependent and were routinely and easily measured in the clinical setting. We identified 14 predictors, which included demographic, clinical, psychosocial, and preoperative outcome measures (Table 1, Supplementary data). Of interest, the presence of contralateral knee pain was measured by the "Patient Category" item (response choice b) from the Knee Society Clinical Rating System (Insall et al. 1989). For the type of walking aids used preoperatively, we coded the responses into 4 categories: (1) none, (2) walking stick or umbrella, (3) quadstick, and (4) walking frame or 2 canes or crutches. To assess self-reported depression, a single question (Q28) from the SF-36 ("How much of the time during the past four weeks have you felt downhearted and depressed?") was used (Pomeroy et al. 2001).

Outcome measures

The outcomes of interest were the 6-month postoperative knee range of motion, knee pain, and walking limitations. Notably, we have chosen an intermediate (6-month) postoperative timepoint because (i) model prediction accuracy may decrease with a longer time horizon, (ii) knowledge of intermediate-term (6-months) risk for poor TKA outcomes will aid patient education and assist in rehabilitation planning, and (iii) TKA outcomes such as knee range of motion were reportedly nearing their peak at the 6-month timepoint (Stratford et al. 2010).

A long-arm goniometer was used to measure active-assisted knee extension and flexion range of motion with the patients in supine position. To measure knee pain, patients were asked to describe their usual knee pain during the past 4 weeks. This variable, taken from Q1 of the Oxford Knee Questionnaire (Dawson et al. 1998), has 5 categories: (i) none, (ii) very mild, (iii) mild, (iv) moderate, and (v) severe. To measure walking limitations, patients were asked to estimate the time they were able to walk (without a rest) before they had severe difficulty with the operated knee. This variable had 4 categories: (i) > 30 min, (ii) 16–30 min, (iii) 5–15 min, and (iv) around the house only.

Statistics

Data are expressed as means (SD) and medians with quartiles for continuous variables and as counts with percentage for categorical variables. We used proportional-odds ordinal regression models, which examined the multivariable associations of the predictors listed in Table 1 (Supplementary data) with 3 outcomes—namely, knee flexion and extension range of motion, knee pain, and walking limitations. We used proportional-odds ordinal regression because (i) it can handle both ordinal and clumped continuous outcomes (Chang and Pocock 2000, Liu et al. 2017) and (ii) it preserves the information in ordinal outcomes and has greater precision compared with binary logistic regression. To avoid assuming linearity, we modelled all continuous predictors as restricted cubic splines (Durrleman and Simon 1989, Harrell Jr 2015). All other predictors were included in the models as binary or categorical variables. Apart from the "education-level" variable, which was missing in 7.7%, all other predictors were missing at very low levels (0.02% to 0.5%). Thus, we used the transcan function in the R Hmisc package (R Foundation for Statistical Computing) to singly impute missing values. To account for model overfitting, we shrank the odds ratios (ORs) in the models using penalized regression (Moons et al. 2004, Harrell Jr 2015). To account for the clustering of patients within surgeons, which may bias the confidence intervals toward being too narrow, we calculated Huber–White robust estimates of standard errors and confidence intervals (White 1980).

We assessed model performance in 2 ways. First, we measured model discrimination by the concordance index (*c*-index). Similar to an AUC statistic, the generalized *c*-statistic may be interpreted as the probability to correctly discriminate between 2 patients with different levels of follow-up TKA outcomes, where a value of 1 represents perfect discrimination and 0.5 represents no discrimination ("coin flip") (Harrell Jr 2015). We computed the generalized *c*-indices of our ordinal models and the *c*-indices for the predictions of poor post-TKA outcomes (defined as (i) a knee flexion contracture ≥ 10° (Ritter et al. 2008), (ii) a knee flexion range < 90°, (iii) a knee pain rating of "moderate" or "severe," and (iv) a maximum walk time ≤ 15 min). Because a prediction model is expected to perform better ("optimistically") in the development sample than in new (but similar) samples, bootstrap internal validation (Austin and Steyerberg 2017) with cluster sampling (Bouwmeester et al. 2013) was performed to shrink the c-indices for "optimism" (Harrell Jr 2015). Second, we assessed model calibration using loess-smoothed calibration plots.

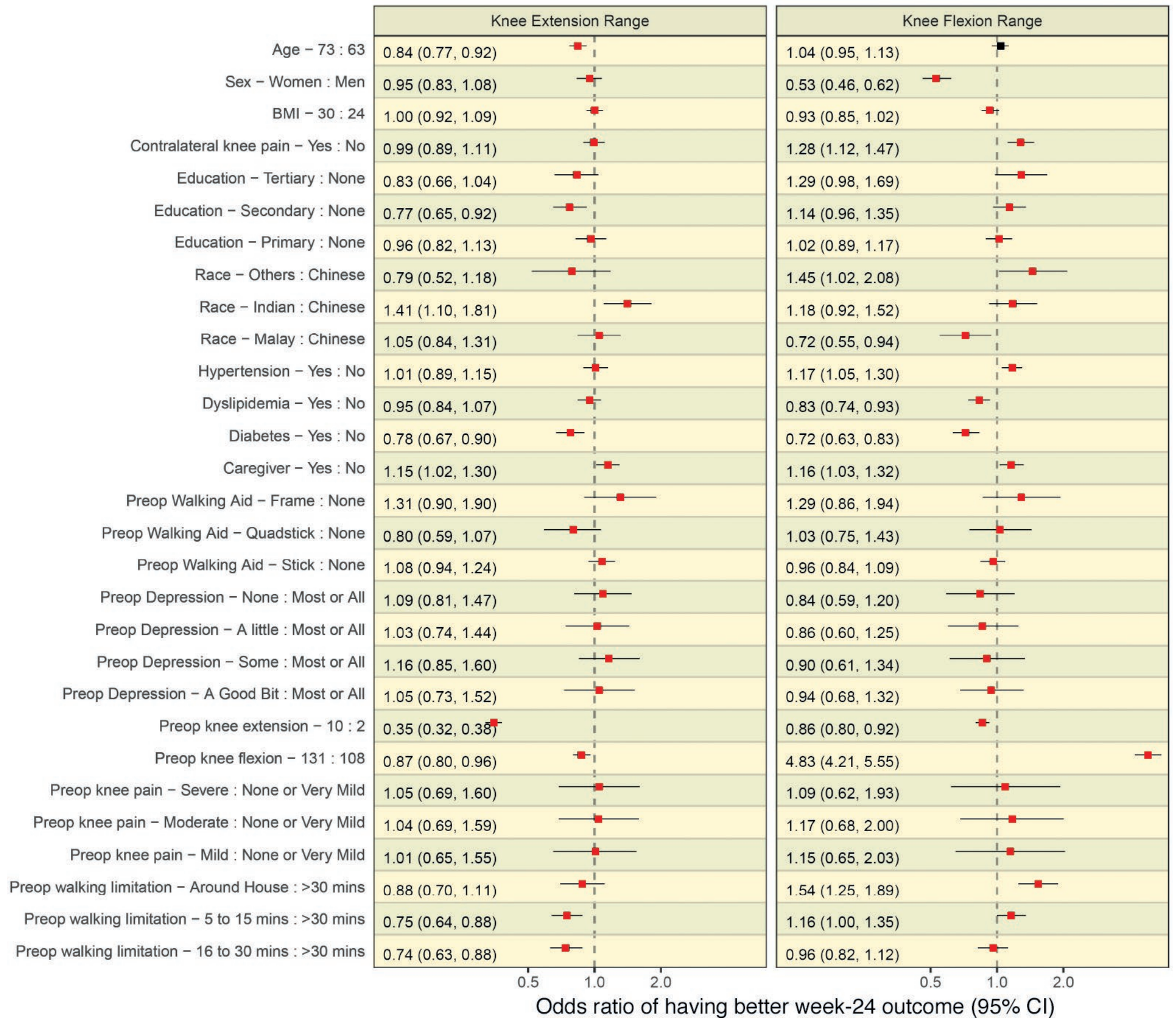

Figure 2. Part 1. Forest plot of odds ratios (ORs) and 95% confidence intervals (CIs) from proportional-odds ordinal regression models predicting (i) knee extension and flexion range of motion, (ii) knee pain intensity, and (iii) walking limitations at 6 months post-TKA. Values to the immediate right of continuous predictors are quartiles, and the ORs estimate the odds of better TKA outcomes (i.e., greater knee range of motion, lower knee pain, and greater walking ability) at the 75th vs. the 25th percentile values. This scaling is done to facilitate the interpretation and comparison of effect sizes of continuous predictors that are often measured on different units. As an illustrative example, other variables being equal, patients with a BMI of 30 (75th percentile) had, on average, 0.89 times (95% CI, 0.80–0.99) the odds of having greater walking ability at 6 months post-TKA, relative to patients with a BMI of 24 (25th percentile). TKA: total knee arthroplasty. Notably, as the goal of this analysis was prediction, the ORs are a measure of predictive effects and they should not be interpreted as causal effects (Steyerberg 2009).

In sensitivity analyses, we examined potential variations in temporal effects (year of knee surgery) and predictor effects over time (Austin et al. 2017), and we found no statistically significant overall time effect (data not shown). We also calculated sensitivity, specificity, positive and negative predictive values (with Wilson 95% confidence intervals [CI]) for various risk thresholds of our prediction models. We assessed and confirmed the appropriateness of our prediction model using residual plots and empirical cumulative logit curves, and all analyses (including the computation of 95% CI) and graphing were done with the rms (Harrell Jr 2017), ROCR (Sing et al. 2005), and ggplot2 (Wickham 2009) R packages (http://www.r-project.org). The web-based application was developed with the R shiny (Chang et al. 2017) package.

Ethics, funding, and potential conflicts of interest

The institutional review board approved the study with a waiver of informed consent (SingHealth CIRB 2014/2027, Singapore). This work was supported by the Singhealth Allied Health Research Publication Grant and the Singapore General Hospital SMART II Centre Grant. No conflicts of interest were declared.

Results

The mean age of all 4,026 patients was 68 years (SD 7.5) and women accounted for three-quarters of the sample (Table 1, Supplementary data). Figure 2 shows the partial effects of each predictor for all outcomes.

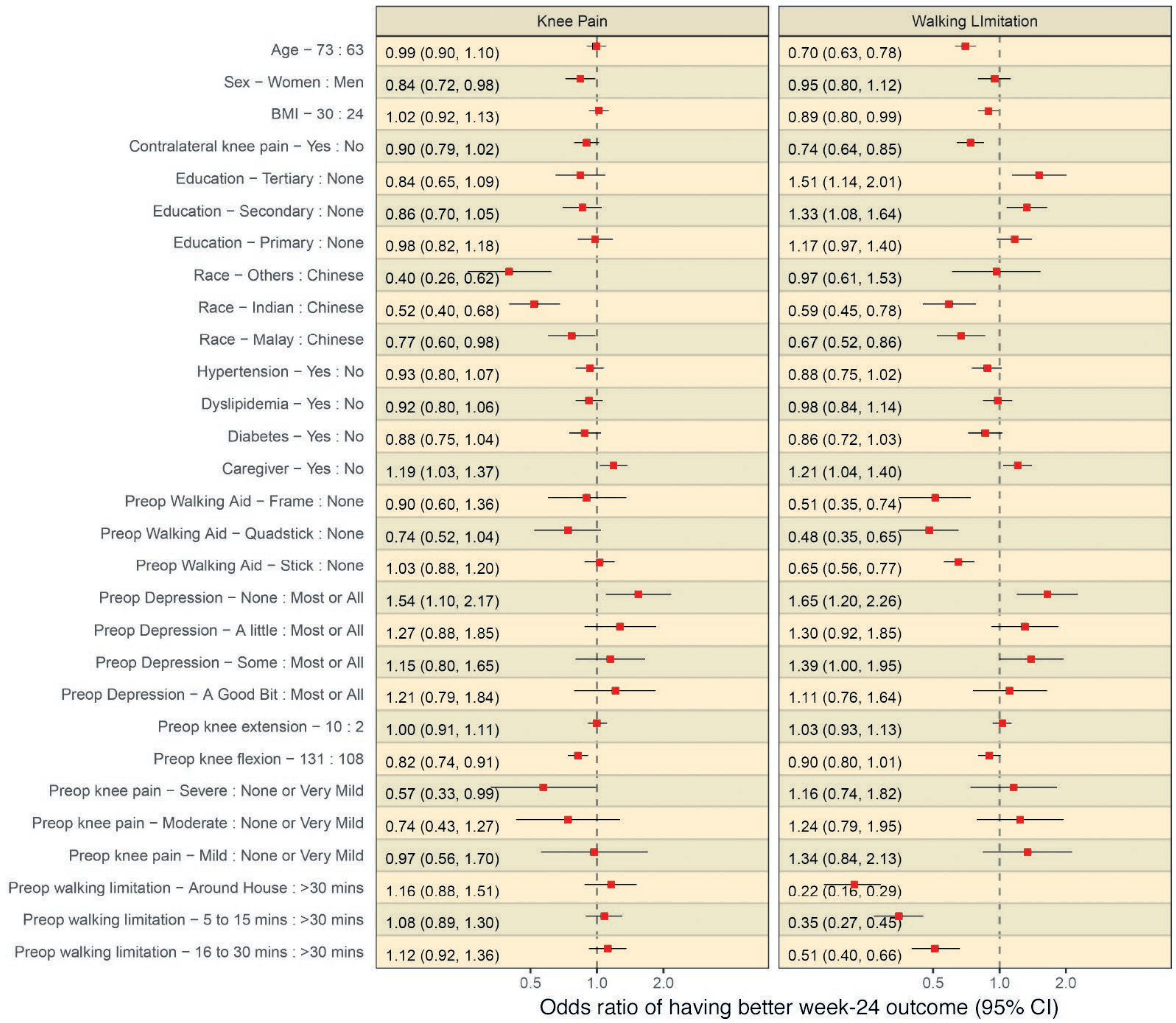

Figure 2. Part 2. See Figure legend on previous page.

Knee range of motion

Preoperatively, mean (SD) knee extension and flexion range were 7° (7°) and 118° (18°), respectively. At 6 months post-surgery, around a fifth of patients (20%; CI, 18–21) had a knee flexion contracture ≥ 10° while 5.6% of the patients (4.9–6.4) had a knee flexion range < 90°. The generalized, optimism-corrected *c*-indices for the knee extension and flexion models were 0.65 and 0.70, respectively. The *c*-indices for the prediction of a postoperative knee flexion contracture ≥ 10° and a knee flexion range < 90° were 0.71 and 0.79. For knee extension range of motion (Figure 2), preoperative knee extension range was the strongest predictor. Additional statistically significant predictors ($p < 0.001$) of greater (better) postoperative knee extension range were younger age, absence of diabetes mellitus, and lower preoperative walking limitations (better walking ability). For knee flexion range of motion, preoperative knee flexion range was the strongest predictor. Additional statistically significant predictors ($p < 0.001$) of greater (better) postoperative knee flexion range were male sex, absence of diabetes mellitus, and greater preoperative knee extension range.

Knee pain

Preoperatively, over four-fifths of patients (83%, CI 82–84) reported at least moderate knee pain; at 6 months post-surgery, the figure was around 1 in every 10 patients (8.7%, CI 7.9–9.7). The generalized, optimism-corrected *c*-index for the knee pain model was 0.58. The *c*-index for the prediction of a postoperative knee pain rating of at least “moderate” was 0.65. Beside lower preoperative knee pain levels, additional statistically significant predictors ($p < 0.001$) of lower (better) postoperative knee pain levels were lower preoperative depression levels, lower preoperative knee flexion range, and Chinese race.

Walking limitations

Preoperatively, half of the patients (55%, CI 53–57) reported an inability to walk for more than 15 minutes; at 6 months

Table 2. Operating characteristics of prediction models to predict poor post-TKA outcomes at various risk thresholds

Cutpoint (%)	TN (n)	FP (n)	FN (n)	TP (n)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
Knee extension model:								
0.05	219	3,002	15	773	0.98 (0.97–0.99)	0.07 (0.06–0.08)	0.20 (0.19–0.22)	0.94 (0.90–0.96)
0.10	844	2,377	60	728	0.92 (0.90–0.94)	0.26 (0.25–0.28)	0.23 (0.22–0.25)	0.93 (0.92–0.95)
0.15	1,490	1,731	141	647	0.82 (0.79–0.85)	0.46 (0.45–0.48)	0.27 (0.25–0.29)	0.91 (0.90–0.93)
0.17 [a]	1,790	1,431	183	605	0.77 (0.74–0.80)	0.56 (0.54–0.57)	0.30 (0.28–0.32)	0.91 (0.89–0.92)
0.20	2,070	1,151	261	527	0.67 (0.64–0.70)	0.64 (0.63–0.66)	0.31 (0.29–0.34)	0.89 (0.87–0.90)
Knee flexion model:								
0.05	2,586	1,144	71	207	0.74 (0.69–0.79)	0.69 (0.68–0.71)	0.15 (0.13–0.17)	0.97 (0.97–0.98)
0.06 [a]	2,833	897	84	194	0.70 (0.64–0.75)	0.76 (0.75–0.77)	0.18 (0.16–0.20)	0.97 (0.96–0.98)
0.10	3,213	517	130	148	0.53 (0.47–0.59)	0.86 (0.85–0.87)	0.22 (0.19–0.26)	0.96 (0.95–0.97)
0.15	3,472	258	166	112	0.40 (0.35–0.46)	0.93 (0.92–0.94)	0.30 (0.26–0.35)	0.95 (0.95–0.96)
0.20	3,607	123	196	82	0.29 (0.24–0.35)	0.97 (0.96–0.97)	0.40 (0.34–0.47)	0.95 (0.94–0.96)
Knee pain model:								
0.05	379	3,294	12	340	0.97 (0.94–0.98)	0.10 (0.09–0.11)	0.09 (0.08–0.10)	0.97 (0.95–0.98)
0.08 [a]	1,995	1,678	107	245	0.70 (0.65–0.74)	0.54 (0.53–0.56)	0.13 (0.11–0.14)	0.95 (0.94–0.96)
0.10	2,679	994	185	167	0.47 (0.42–0.53)	0.73 (0.71–0.74)	0.14 (0.12–0.17)	0.94 (0.93–0.94)
0.15	3,455	218	303	49	0.14 (0.11–0.18)	0.94 (0.93–0.95)	0.18 (0.14–0.23)	0.92 (0.91–0.93)
0.20	3,621	52	337	15	0.04 (0.03–0.07)	0.99 (0.98–0.99)	0.22 (0.14–0.34)	0.91 (0.91–0.92)
Walking limitations model:								
0.05	668	2,819	19	514	0.96 (0.94–0.98)	0.19 (0.18–0.20)	0.15 (0.14–0.17)	0.97 (0.96–0.98)
0.10	1,801	1,686	88	445	0.83 (0.80–0.86)	0.52 (0.50–0.53)	0.21 (0.19–0.23)	0.95 (0.94–0.96)
0.15	2,553	934	188	345	0.65 (0.61–0.69)	0.73 (0.72–0.75)	0.27 (0.25–0.29)	0.93 (0.92–0.94)
0.17 [a]	2,705	782	205	328	0.62 (0.57–0.66)	0.78 (0.76–0.79)	0.30 (0.27–0.32)	0.93 (0.92–0.94)
0.20	2,931	556	263	270	0.51 (0.46–0.55)	0.84 (0.83–0.85)	0.33 (0.30–0.36)	0.92 (0.91–0.93)

TN = true negative, FP = false positive, FN = false negative, TP = true positive, PPV = positive predictive value, NPV = negative predictive value.
[a] Optimal cut-off at Youden index.
Poor post-TKA outcomes were defined as (i) a knee flexion contracture ≥ 10°, (ii) a knee flexion range < 90°, (iii) a knee pain rating of "moderate" or "severe," and (iv) a maximum walk time ≤ 15 minutes.

post-surgery, this was just over 1 in every 10 patients (13%, CI 12–14). The generalized, optimism-corrected *c*-index for the walking limitations model was 0.70. The *c*-index for the prediction of a maximum walk time ≤ 15 minutes was 0.76. Besides lower levels of preoperative walking limitations, additional statistically significant predictors ($p < 0.001$) of lower (better) levels of postoperative walking limitations were younger age, the use of a smaller or no gait aid preoperatively, lower preoperative depression levels, the absence of contralateral knee pain, and Chinese race.

The prediction models

Figure 3 (Supplementary data) shows the calibration plots of all prediction models. Table 2 shows the test characteristics of the prediction models at various risk thresholds and suggests that the prediction models tend to be adept in identifying patients at low risk of poor TKA outcomes: its negative predictive values were ≥ 94% for identifying true low-risk patients at a 5% risk threshold.

To facilitate the use of the prediction models in clinical practice, we created a web application (https://sgh-physio.shinyapps.io/predicTKR/) that shows the expected distributions of the TKA outcomes for individual patients based on their preoperative demographic and clinical characteristics. To provide a "bottom-line" prediction of outcome (Barlow et al. 2016) and to avoid setting universal (fixed) outcome thresholds which do not account for patients' baseline (preoperative) levels (Hildon et al. 2012), the app also computes the predicted probabilities of achieving patient-defined acceptable levels of outcomes.

Discussion

We developed 4 models to predict, for individual patients, their likely levels of knee extension and flexion range of motion, knee pain, and walking limitations at 6 months post TKA. Our models showed adequate calibration (Figure 2) and modest to moderately good predictive discrimination when evaluating patients with poor postoperative outcomes, with c-indices ranging between 0.65 (knee pain model) and 0.79 (knee flexion model). Across all models, we found that the postoperative outcomes were strongly influenced by the same outcome measure obtained preoperatively—an unsurprising finding that is consistent with the literature (Gandhi et al. 2006, Stratford et al. 2010, Lewis et al. 2015, Harmelink et al. 2017, Sanchez-Santos et al. 2018). Nonetheless, preoperative outcomes are not the only predictor of postoperative outcomes (Figure 2), which supports the need to consider several factors when predicting postoperative outcomes.

Interestingly, we observed that greater preoperative knee flexion range of motion was associated with lower odds of better (lower) postoperative knee pain rating (IQR-OR, 0.82, CI, 0.74–0.91; Figure 1). To our knowledge, this association has not previously been examined and may seem paradoxical at first. However, it is plausibly explained by the emerging evidence of the inverse associations (i) between greater preoperative knee radiographic severity and lower (better) postoperative pain severity (Dowsey et al. 2012, Valdes et al. 2012) and (i) between greater knee flexion range of motion and lower radiographic severity of knee OA (Holla et al. 2011). Accordingly, it is possible that knee pain in patients with mild knee radiographic OA (and good flexion range of motion) is not directly driven by knee joint damage, but rather by chronic pain mechanisms such as higher pain sensitivity and/or central sensitization (Valdes et al. 2012). In our study, knee radiographic severity, indexed by the Kellgren–Lawrence grade in previous studies (Dowsey et al. 2016), was not included because it is not readily available for incorporation into real-time prediction given its semi-quantitative nature (Wright 2014). Thus, it would be of interest for future prediction modelling studies to compare the predictive information provided by Kellgren–Lawrence grade with that provided by knee flexion range of motion.

In our study, patients' preoperative depression level was an important predictor of postoperative 6-month pain and walking limitations. Reviewing the literature, recent systematic reviews have reported that greater preoperative psychological distress is associated with worse pain and physical function post TKA (Lewis et al. 2015, Bletterman et al. 2017). In terms of intervention studies, Riddle et al. (2011) reported that patients who were preoperatively instructed in pain coping skills reported better Month-2 self-reported pain and physical function outcomes. Similarly, Cai et al. (2017) demonstrated, in a recent randomized clinical trial, that providing targeted cognitive behavioral therapy in the early postoperative care resulted in better Month-6 self-reported pain and physical function outcomes. Thus, taken together, our results indicate that at-risk patients may benefit from targeted behavioral interventions during the preoperative and early postoperative periods to modify this risk factor and improve outcomes.

In our multi-racial sample of Chinese, Malay, and Asian-Indian patients, we observed racial variation in post-TKA outcomes: Chinese race was associated with significantly lower (better) postoperative levels of knee pain and walking limitations, and these findings persisted after adjustment for multiple demographic, clinical, psychosocial, and preoperative outcome measures (Figure 1). At a time when race has been generally inadequately studied in orthopedics (Somerson et al. 2014), our findings provide useful timely information and, potentially, differences in perceived pain thresholds (Tan et al. 2008) and sociocultural expressions of pain (Campbell and Edwards 2012) may help explain them. Although we found important variations in outcomes according to patient race, patients belonging to a racial group that is considered as having poor post-TKA outcomes may still achieve a good outcome prognosis if other variables are favorable. We emphasize the need to consider multiple factors when providing individualized outcome predictions. Also, we caution that predictions from our models should serve as a starting point for shared decision-making, and not as a definitive or final recommendation.

Limitations

Our study has limitations. First, our data come from only 1 institution, but it delivers care to a large segment of the nation's population. Having a large and representative population-based sample improves the stability of our model predictions and their applicability to institutions with patients who have similar characteristic to our patients. Second, although our prediction models have satisfactory discrimination, their performance is not optimal. Furthermore, our models could be criticized for not including potentially important predictors such as the severity of radiographic knee OA and a comprehensive list of comorbidities and psychosocial factors. Nevertheless, as we continue to grow our database and refine the variables collected, we will be able to update our prediction models and improve their prediction accuracy. Third, although we have implemented our prediction models in a web application to improve their accessibility in the clinical setting, we acknowledge that prediction models are unlikely to be widely used unless they can be incorporated into electronic medical records systems. As our prediction models comprise routinely and easily measured variables in the clinical setting, it is possible to integrate and externally validate them in electronic medical records systems. Future studies should explore this possibility. Fourth, we studied an Asian sample so the extent to which our results may apply to non-Asians is unknown. Finally, although we believe that knowledge of intermediate-term (6 months) risk for poor TKA outcomes will aid patient education and assist in rehabilitation planning, our models do not predict longer-term outcomes; TKA outcomes such as self-reported knee pain and physical function may require as long as 2 years to reach a plateau (Giesinger et al. 2014, Lim et al. 2015).

In summary, at 6 months post TKA, around 5% to 20% of patients had knee range limitations, moderate to severe knee pain, or walking limitations (maximum walk time ≤ 15 minutes). Using data that are routinely collected and available from our database, we have developed prediction models that can potentially complement clinical and shared decision-making by providing personalized risk estimates of these important outcomes. The next step in translating this work is to perform external testing before evaluating the impact of individualized risk predictions on improving patient outcomes and satisfaction.

Supplementary data

Table 1 and Figures 1 and 3 are available as supplementary data in the online version of this article, http://dx.doi.org/10.1080/17453674.2018.1560647

The authors acknowledge the support from Jennifer Liaw, the head of the Department of Physiotherapy, Singapore General Hospital. They also thank William Yeo from the Orthopaedic Diagnostic Centre, Singapore General Hospital, for his assistance. Finally, they thank Nai-Hong Chan and the therapy assistants (Penny Teh, Hamidah Binti Hanib, Bamgbang Heryanto, and Shela Devi D/o Perumal) for their kind assistance.

YHP wrote the first draft of the manuscript and performed the data analysis. All authors contributed to conception and design of the study, critical analysis of the data, interpretation of the findings, and critical revision of the manuscript.

Acta thanks Daan Nieboer and Annette W-Dahl for help with peer review of this study.

Austin P C, Steyerberg E W. Events per variable (EPV) and the relative performance of different strategies for estimating the out-of-sample validity of logistic regression models. Stat Methods Med Res 2017; 26(2): 796-808.

Austin P C, van Klaveren D, Vergouwe Y, Nieboer D, Lee D S, Steyerberg E W. Validation of prediction models: examining temporal and geographic stability of baseline risk and estimated covariate effects. Diagn Progn Res 2017; 1: 12.

Barlow T, Scott P, Griffin D, Realpe A. How outcome prediction could affect patient decision making in knee replacements: a qualitative study. BMC Musculoskelet Disord 2016; 17: 304.

Bletterman A N, de Geest-Vrolijk M E, Vriezekolk J E, Nijhuis-van der Sanden M W, van Meeteren N L, Hoogeboom T J. Preoperative psychosocial factors predicting patient's functional recovery after total knee or total hip arthroplasty: a systematic review. Clin Rehabil 2017; 32(4): 512-25.

Bouwmeester W, Moons K G, Kappen T H, van Klei W A, Twisk J W, Eijkemans M J, Vergouwe Y. Internal validation of risk models in clustered data: a comparison of bootstrap schemes. Am J Epidemiol 2013; 177(11): 1209-17.

Cai L, Gao H, Xu H, Wang Y, Lyu P, Liu Y. Does a program based on cognitive behavioral therapy affect kinesiophobia in patients following total knee arthroplasty? A randomized, controlled trial with a 6-month follow-up. J Arthroplasty 2017; 33(3): 704-10.

Campbell C M, Edwards R R. Ethnic differences in pain and pain management. Pain Manag 2012; 2(3): 219-30.

Chang B H, Pocock S. Analyzing data with clumping at zero: an example demonstration. J Clin Epidemiol 2000; 53(10): 1036-43.

Chang W, Cheng J, Allaire J J, Xie Y, McPherson J. shiny: web application framework for R. R package version 1.0.5. https://CRAN.R-project.org/package = shiny 2017.

Collins G S, Reitsma J B, Altman D G, Moons K G. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): the TRIPOD statement. Ann Intern Med 2015; 162(1): 55-63.

Dawson J, Fitzpatrick R, Murray D, Carr A. Questionnaire on the perceptions of patients about total knee replacement. J Bone Joint Surg Br 1998; 80(1): 63-9.

Dowsey M M, Nikpour M, Dieppe P, Choong P F. Associations between pre-operative radiographic changes and outcomes after total knee joint replacement for osteoarthritis. Osteoarthritis Cartilage 2012; 20(10): 1095-102.

Dowsey M M, Spelman T, Choong P F. Development of a prognostic nomogram for predicting the probability of nonresponse to total knee arthroplasty 1 year after surgery. J Arthroplasty 2016; 31(8): 1654-60.

Durrleman S, Simon R. Flexible regression models with cubic splines. Stat Med 1989; 8(5): 551-61.

Gandhi R, de Beer J, Leone J, Petruccelli D, Winemaker M, Adili A. Predictive risk factors for stiff knees in total knee arthroplasty. J Arthroplasty 2006; 21(1): 46-52.

Giesinger K, Hamilton D F, Jost B, Holzner B, Giesinger J M. Comparative responsiveness of outcome measures for total knee arthroplasty. Osteoarthritis Cartilage 2014; 22(2): 184-9.

Gunaratne R, Pratt D N, Banda J, Fick D P, Khan R J K, Robertson B W. Patient dissatisfaction following total knee arthroplasty: a systematic review of the literature. J Arthroplasty 2017; 32(12): 3854-60.

Gutacker N, Street A. Use of large-scale HRQoL datasets to generate individualised predictions and inform patients about the likely benefit of surgery. Qual Life Res 2017; 26(9): 2497-505.

Harmelink K E M, Zeegers A, Hullegie W, Hoogeboom T J, Nijhuis-van der Sanden M W G, Staal J B. Are there prognostic factors for one-year outcome after total knee arthroplasty? a systematic review. J Arthroplasty 2017; 32(12): 3840-53 e1.

Harrell Jr F E. Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis. New York: Springer; 2015.

Harrell Jr F E. rms: Regression Modeling Strategies. R package version 5.1-1. http://CRAN.R-project.org/package = rms 2017.

Hildon Z, Neuburger J, Allwood D, van der Meulen J, Black N. Clinicians' and patients' views of metrics of change derived from patient reported outcome measures (PROMs) for comparing providers' performance of surgery. BMC Health Serv Res 2012; 12: 171.

Holla J F, Steultjens M P, van der Leeden M, Roorda L D, Bierma-Zeinstra S M, den Broeder A A, Dekker J. Determinants of range of joint motion in patients with early symptomatic osteoarthritis of the hip and/or knee: an exploratory study in the CHECK cohort. Osteoarthritis Cartilage 2011; 19(4): 411-19.

Huang Y, Lee M, Chong H C, Ning Y, Lo N N, Yeo S J. Reasons and factors behind post-total knee arthroplasty dissatisfaction in an Asian population. Ann Acad Med Singapore 2017; 46(8): 303-9.

Husain A, Lee G C. Establishing realistic patient expectations following total knee arthroplasty. J Am Acad Orthop Surg 2015; 23(12): 707-13.

Insall J N, Dorr L D, Scott R D, Scott W N. Rationale of the Knee Society clinical rating system. Clin Orthop Relat Res 1989; (248): 13-14.

Lange T, Schmitt J, Kopkow C, Rataj E, Gunther K P, Lutzner J. What do patients expect from total knee arthroplasty? A Delphi consensus study on patient treatment goals. J Arthroplasty 2017; 32(7): 2093-9 e1.

Lewis G N, Rice D A, McNair P J, Kluger M. Predictors of persistent pain after total knee arthroplasty: a systematic review and meta-analysis. Br J Anaesth 2015; 114(4): 551-61.

Lim J B, Chi C H, Lo L E, Lo W T, Chia S L, Yeo S J, Chin P L, Tay K J, Lo N N. Gender difference in outcome after total knee replacement. J Orthop Surg (Hong Kong) 2015; 23(2): 194-7.

Liu Q, Shepherd B E, Li C, Harrell F E, Jr. Modeling continuous response variables using ordinal regression. Stat Med 2017; 36(27): 4316-35.

Matsuda S, Kawahara S, Okazaki K, Tashiro Y, Iwamoto Y. Postoperative alignment and ROM affect patient satisfaction after TKA. Clin Orthop Relat Res 2013; 471(1): 127-33.

Moons K G M, Donders A R, Steyerberg E W, Harrell F E. Penalized maximum likelihood estimation to directly adjust diagnostic and prognostic prediction models for overoptimism: a clinical example. J Clin Epidemiol 2004; 57(12): 1262-70.

Pomeroy I M, Clark C R, Philp I. The effectiveness of very short scales for depression screening in elderly medical patients. Int J Geriatr Psychiatry 2001; 16(3): 321-6.

Riddle D L, Keefe F J, Nay W T, McKee D, Attarian D E, Jensen M P. Pain coping skills training for patients with elevated pain catastrophizing who are scheduled for knee arthroplasty: a quasi-experimental study. Arch Phys Med Rehabil 2011; 92(6): 859-65.

Ritter M A, Lutgring J D, Davis K E, Berend M E. The effect of postoperative range of motion on functional activities after posterior cruciate-retaining total knee arthroplasty. J Bone Joint Surg Am 2008; 90(4): 777-84.

Sanchez-Santos M T, Garriga C, Judge A, Batra R N, Price A J, Liddle A D, Javaid M K, Cooper C, Murray D W, Arden N K. Development and validation of a clinical prediction model for patient-reported pain and function after primary total knee replacement surgery. Sci Rep 2018; 8(1): 3381.

Sing T, Sander O, Beerenwinkel N, Lengauer T. ROCR: visualizing classifier performance in R. Bioinformatics 2005; 21(20): 7881.

Singapore Ministry of Health. Knee replacement surgery. https://www.moh.gov.sg/content/moh_web/home/costs_and_financing/hospital-charges/Total-Hospital-Bills-By-condition-procedure/Knee_Knee_replacement_surgery.html (last accessed January 19, 2018).

Singh J A, Dowsey M M, Dohm M, Goodman S M, Leong A L, Scholte Voshaar M, Choong P F. Achieving consensus on total joint replacement trial outcome reporting using the OMERACT filter: endorsement of the final core domain set for total hip and total knee replacement trials for endstage arthritis. J Rheumatol 2016; 44(11): 1723-6.

Somerson J S, Bhandari M, Vaughan C T, Smith C S, Zelle B A. Lack of diversity in orthopaedic trials conducted in the United States. J Bone Joint Surg Am 2014; 96(7): e56.

Steyerberg E. Clinical prediction models: a practical approach to development, validation and updating. New York: Springer; 2009.

Stratford P W, Kennedy D M, Robarts S F. Modelling knee range of motion post arthroplasty: clinical applications. Physiother Can 2010; 62(4): 378-87.

Tan E C, Lim Y, Teo Y Y, Goh R, Law H Y, Sia A T. Ethnic differences in pain perception and patient-controlled analgesia usage for postoperative pain. J Pain 2008; 9(9): 849-55.

Valdes A M, Doherty S A, Zhang W, Muir K R, Maciewicz R A, Doherty M. Inverse relationship between preoperative radiographic severity and postoperative pain in patients with osteoarthritis who have undergone total joint arthroplasty. Semin Arthritis Rheum 2012; 41(4): 568-75.

Van Onsem S, Van Der Straeten C, Arnout N, Deprez P, Van Damme G, Victor J. A new prediction model for patient satisfaction after total knee arthroplasty. J Arthroplasty 2016; 31(12): 2660-7 e1.

Volkmann E R, FitzGerald J D. Reducing gender disparities in post-total knee arthroplasty expectations through a decision aid. BMC Musculoskelet Disord 2015; 16: 16.

White H. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica 1980; 48: 817-30.

Wickham H. ggplot2: elegant graphics for data analysis. New York: Springer; 2009.

Wright R W. Osteoarthritis classification scales: interobserver reliability and arthroscopic correlation. J Bone Joint Surg Am 2014; 96(14): 1145-51.