

Development of a multivariable prediction model for early revision of total knee arthroplasty - The effect of including patient-reported outcome measures

Andersen, Jonas Dahl; Hangaard, Stine; Buus, Amanda Agnes Østervig; Laursen, Mogens; Hejlesen, Ole K.; El-Galaly, Anders

Published in:
Journal of Orthopaedics

DOI (link to publication from Publisher):
[10.1016/j.jor.2021.03.001](https://doi.org/10.1016/j.jor.2021.03.001)

Creative Commons License
CC BY 4.0

Publication date:
2021

Document Version
Publisher's PDF, also known as Version of record

[Link to publication from Aalborg University](#)

Citation for published version (APA):

Andersen, J. D., Hangaard, S., Buus, A. A. Ø., Laursen, M., Hejlesen, O. K., & El-Galaly, A. (2021). Development of a multivariable prediction model for early revision of total knee arthroplasty - The effect of including patient-reported outcome measures. *Journal of Orthopaedics*, 24, 216-221. <https://doi.org/10.1016/j.jor.2021.03.001>

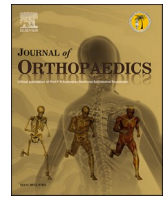
General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy

If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.



Development of a multivariable prediction model for early revision of total knee arthroplasty – The effect of including patient-reported outcome measures

J.D. Andersen^{a,*}, S. Hangaard^{a,b}, A.A.Ø. Buus^c, M. Laursen^{c,d}, O.K. Hejlesen^a, A. El-Galaly^{c,d}

^a Department of Health Science and Technology, Faculty of Medicine, Aalborg University, Denmark

^b Steno Diabetes Center North Denmark, Aalborg University Hospital, Aalborg, Denmark

^c Orthopaedic Research Unit, Aalborg University Hospital, Aalborg, Denmark

^d Department of Clinical Medicine, Aalborg University, Aalborg, Denmark

ARTICLE INFO

Keywords:

TKA
Machine learning
Knee osteoarthritis
Prediction model
Revision

ABSTRACT

Background: Revision TKA is a serious adverse event with substantial consequences for the patient. As revision is becoming increasingly common in patients under 65 years, the need for improved preoperative patient selection is imminently needed. Therefore, this study aimed to identify the most important factors of early revision and to develop a prediction model of early revision including assessment of the effect of incorporating data on patient-reported outcome measures (PROMs).

Material and methods: A cohort of 538 patients undergoing primary TKA was included. Multiple logistic regression using forward selection of variables was applied to identify the best predictors of early revision and to develop a prediction model. The model was internally validated with stratified 5-fold cross-validation. This procedure was repeated without including data on PROMs to develop a model for comparison. The models were evaluated on their discriminative capacity using area under the receiver operating characteristic curve (AUC).

Results: The most important factors of early revision were age (OR 0.63 [0.42, 0.95]; $P = 0.03$), preoperative EQ-5D (OR 0.07 [0.01, 0.51]; $P = 0.01$), and number of comorbidities (OR 1.01 [0.97, 1.25]; $P = 0.15$). The AUCs of the models with and without PROMs were 0.65 and 0.61, respectively. The difference between the AUCs was not statistically significant ($P = 0.32$).

Conclusions: Although more work is needed in order to reach a clinically meaningful quality of the predictions, our results show that the inclusion of PROMs seems to improve the quality of the prediction model.

1. Introduction

Knee osteoarthritis (OA) is one of the most common forms of OA and is estimated to affect 260 million people worldwide.¹ It is estimated that hip and knee OA alone are accounting for 20 million years lived with disability worldwide and thus have considerable impact on the individual patient.^{1–3} Knee OA often results in severe pain, disability and deterioration in health-related quality of life,^{4,5} translating into substantial healthcare costs and productivity losses for individuals and society.^{6,7} There are different treatment options for knee OA aiming to restore function and relieve pain with total knee arthroplasty (TKA) considered the most effective treatment for end-staged multi-compartmental knee OA.^{8–10} TKA is one of the most frequently

performed orthopedic surgeries in Denmark with nearly 7.000 primary TKAs performed in 2018.¹¹ In comparison, more than 700.000 TKAs are performed in USA annually and approximately 100.000 TKAs are performed in the UK.¹² Worldwide rates of both primary and revision TKA are rising, primarily due to the increased longevity of the population and the obesity epidemic.^{4,12,13} Internationally, consistent revision rates have been reported but the number of revisions is expected to rise as an inevitable consequence of the rising number of primary TKAs.^{9,14} In USA, prediction models based on registry data are projecting the demand for primary and revision TKA to increase by 600% by 2030 reaching more than 3.5 million TKAs and more than 270.000 revisions.^{15,16} Revision TKAs are more complicated, more expensive and associated with both reduced implant-survival and inferior

* Corresponding author. Medical Informatics Group, Department of Health Science and Technology, Faculty of Medicine, Aalborg University, Fredrik Bajers Vej 7D - 9220, Aalborg, Denmark.

E-mail addresses: jonas_dahl@live.dk, jsda@hst.aau.dk (J.D. Andersen).

<https://doi.org/10.1016/j.jor.2021.03.001>

Received 29 January 2021; Accepted 7 March 2021

Available online 11 March 2021

0972-978X/© 2021 The Authors. Published by Elsevier B.V. on behalf of Professor P K Surendran Memorial Education Foundation. This is an open access article

under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

patient-reported outcomes (PRO) when compared to primary TKA.^{17,18} An increasing number of patients under 65 years are estimated to have a TKA in the future. These patients will be more likely to have revision and maybe even re-revision due to their life expectancy.^{19–21} A previous study has shown that the risk of revision TKA is 2.5 times greater in those under 65 years than in those above 65 years at the time of surgery.⁸ More importantly, patients under 55 years at the time of surgery have a 5-fold increased risk of revision three years postoperatively compared with those above 75 years.^{8,21} Machine learning (ML) and prediction modeling is increasingly used within the field of orthopedic surgery and earlier attempts to predict potential outcomes of TKA have been made in terms of patient-satisfaction, range of movement and length of stay.^{22–25} Additionally, a recent study by El-Galaly et al. developed different prediction models for early revision TKA using different ML-algorithms based on national registry data. However, this study did not include any patient-reported outcome measures (PROMs) or information on comorbidity, which has recently been suggested to potentially improve the performance of these models.^{26,27} Therefore, it is yet to be investigated whether it is possible to develop a clinical meaningful prediction model for early revision in patients undergoing TKA while including data on PROMs and comorbidity. Moreover, it is yet to be explored whether including PROMs data yield a better performing model than one without PROMs. Furthermore, because of primary TKA being increasingly considered for patients under 65 years, there is a strong need to identify predictor variables and define risk groups leading to improved preoperative patient selection in clinical practice. Therefore, this study aimed to 1) identify the most important factors of early revision in patients undergoing TKA and 2) develop a prediction model for early revision in patients undergoing TKA including assessment of the effect of incorporating data on PROMs and comorbidity in the model.

2. Material and methods

The study is reported in accordance with the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) guidelines²⁸ and the Strengthening the Reporting of Observational studies in Epidemiology (STROBE) statement.²⁹ The study was approved by the Danish Data Protection Agency before data collection (entry no. 2008-58-0028).

2.1. Study design and cohort

The study was a descriptive correlational and predictive study based on a cohort of patients undergoing primary TKA between November 11, 2014 and December 20, 2016 at the Department of Orthopedic Surgery at Aalborg University Hospital, Denmark. Eligible patients were required to 1) have underwent primary TKA for primary or secondary knee OA 2) be aged ≥ 30 at the time of surgery, and 3) have completed the PROM preoperatively. In cases where a patient had bilateral surgery during this period, only data from the first surgery was included in the analysis. All patients were followed for two years with substitution, removal or addition of an implant or part of an implant considered as revision.

2.2. Data source

We used combined registry data obtained from the Danish Knee Arthroplasty Registry (DKR) and from the Department of Orthopedic Surgery at Aalborg University Hospital's administrative outcome database "Jointbase". The Jointbase was developed in 2013 to monitor the quality of treatment and care of patients and prospectively collects PRO-data on patients undergoing TKA.³⁰ The DKR is a national registry that prospectively collects information on all knee arthroplasties inserted in Denmark, and the completeness of the database has been above 96% for primary arthroplasties and 93% for revision arthroplasties since 2011.^{11,31}

2.3. Predictor variables

All predictor variables were selected from patient demographics and characteristics, PROMs and clinical data on surgery-related factors based on previous studies and clinical reasoning.

Patient demographics and characteristics included in the analysis were age (grouped into decades), gender, body mass index (BMI), pain at rest and activity measured by visual analogue scale (VAS), diabetes mellitus, and number of comorbidities. PROMs included in the analysis were preoperative Oxford Knee Score (OKS)³² and preoperative EuroQol-5D (EQ-5D).³³ OKS is a validated joint-specific instrument consisting of 12 patient-administered questions evaluating pain and function in relation to the knee. The final score is summed up on a scale ranging from 0 to 48 (worst-best). An overall score was calculated between 0 and 48 and subsequently converted to an overall score between 0 and 100, which was used as a continuous variable.^{32,34} EQ-5D is a validated instrument for measuring generic health-related quality of life. The instrument consists of two components: 1) the EQ VAS by which patients assess their perceived health status ranging from 0 (worst imaginable health status) to 100 (best imaginable health status), and 2) the EQ-5D descriptive system which evaluates perceived health status within five dimensions: Mobility, self-care, usual activities, pain/discomfort, and anxiety/depression.^{33,34} The EQ-5D provides an overall index score (EQ-5D index) between 0 and 1 (worst-best) for a patient's health status, which was used as a continuous variable. Clinical data on surgery-related factors included in the analysis were previous surgery in the same knee, duration of surgery, and length of stay.

2.4. Missing data

In total <1% of total values in the dataset were missing with the highest percentage of missing values in previous surgery in the same knee (1.3%). Complete-case analysis was performed and cases with missing data were excluded from the dataset prior to further data analysis.

2.5. Study outcome

The outcome for the prediction model was early revision TKA. This was characterized as revision for any indication within two years of the primary TKA. In accordance with the DKR revision was defined as removal, exchange, or addition of an implant or part of an implant.¹¹

2.6. Prediction model

Multiple logistic regression (LR) with forward selection of variables was used as a prediction model. To internally validate the model, stratified 5-fold cross-validation was performed. The procedure was as follows: Firstly, simple LR analysis was performed to identify potentially important variables. All potential predictor variables with a $P < 0.25$ were included in further analysis. Secondly, multiple LR with forward selection of variables was applied to the dataset to identify the most important features of the model. Thirdly, the dataset was shuffled randomly and then split into five equally sized samples with each sample containing the same amount of outcome observations. Fourthly, the model was trained with four of the samples working as training set and the remaining sample working as test set. This procedure was repeated five times until each sample of the five folds had been used as test set once and to train the model four times. Finally, the overall evaluation score was retained and used as an estimate of the performance of the model. Additionally, the exact same procedure was performed but without the data on PROMs (i.e., EQ-5D and OKS) to build a model for comparison. The performance of the model was evaluated on discrimination and calibration using receiver operating characteristic (ROC) curve, area under the curve (AUC)^{35–38} and the Hosmer-Lemeshow goodness-of-fit score. Difference between the AUCs was compared as

described by Hanley and McNeil.³⁹ The discriminative capacity of prediction models are ranked by their AUCs as follows: Excellent (0.9–1.0), good (0.8–0.89), fair (0.7–0.79), poor (0.6–0.69), and failed (0.5–0.59).⁴⁰ An AUC ≥ 0.70 was chosen as cut-off to define a clinical meaningful model.

2.7. Statistical analysis

Continuous data are presented with mean and standard deviation (SD) and categorical data with frequencies (n) and percentages (%). Comparison of continuous variables were conducted with unpaired *t*-tests and categorical variables with chi-square or Fisher's exact test. Multiple LR analyses were performed using forward selection of variables and the stepwise criteria of the probability-of-F to enter and remove variables were set at 0.05 and 0.10, respectively. The regression beta coefficients (B), odds ratios (OR), 95% confidence intervals (CI), and *p*-values (*P*) were calculated. LR diagnostics were performed to assure basic test assumptions were met. To assure absence of multicollinearity, analysis of potential collinear variables was made from the observed variance inflation factor. A *p*-value of <0.05 was considered statistically significant. All statistical analyses were computed using SPSS Statistics (version 27; IBM Corp., Armonk, New York) and MATLAB (version R2020a, MathWorks, Massachusetts, USA).

3. Results

3.1. Sample size

A total of 568 patients were identified through the DKR and Jointbase to have underwent primary TKA between November 11, 2014 and December 20, 2016. Due to other diagnosis than primary or secondary knee OA 19 patients were excluded. Due to revision later than two years from primary TKA, four patients were excluded, and seven patients were excluded due to missing data values. Among the 568 patients, a total of 538 patients met the eligibility criteria and were included in the cohort for analysis. Full overview of the exclusion process is presented in Fig. 1.

3.2. Sample characteristics

A full list of characteristics is presented in Table 1. Patients were aged between 37 and 91 (mean age 67.5) and 55% were female. The majority of the patients had 1–2 comorbidities (51%). The percentage of revision TKA was 4.5% with infection and aseptic loosening as the major indications for revision.

3.3. Predictors of early revision

Potential predictor variables and simple LR analyses of predictor variables associated with early revision are presented in Table 2. After excluding predictor variables with $P > 0.25$, seven variables were included in the multiple LR analysis. Using forward selection of variables, the best predictors of early revision were age (OR 0.63 [0.42, 0.95]; $P = 0.03$), preoperative EQ-5D index (OR 0.07 [0.01, 0.51]; $P = 0.01$) and number of comorbidities (OR 1.01 [0.97, 1.25]; $P = 0.15$). These results are presented in Table 3.

3.4. Prediction model

ROC-curve and AUC of the models with and without PROMs are presented in Fig. 2. As depicted in the ROC-curve, the model including PROMs had an AUC = 0.65, while the model without PROMs had an AUC = 0.61. The difference between the AUCs was not statistically significant ($P = 0.32$). The Hosmer-Lemeshow test score for goodness-of-fit was $X^2 = 4.41$, $P = 0.82$ for the model with PROMs and $X^2 = 4.38$, $P = 0.82$ for the model without PROMs.

4. Discussion

This study aimed to identify the most important predictors of early revision, and to develop a prediction model capable of predicting early revision in patients undergoing TKA including assessment of the effect of incorporating data on PROMs. To the authors' knowledge, this is the first attempt to develop such a model including both data on PROMs and

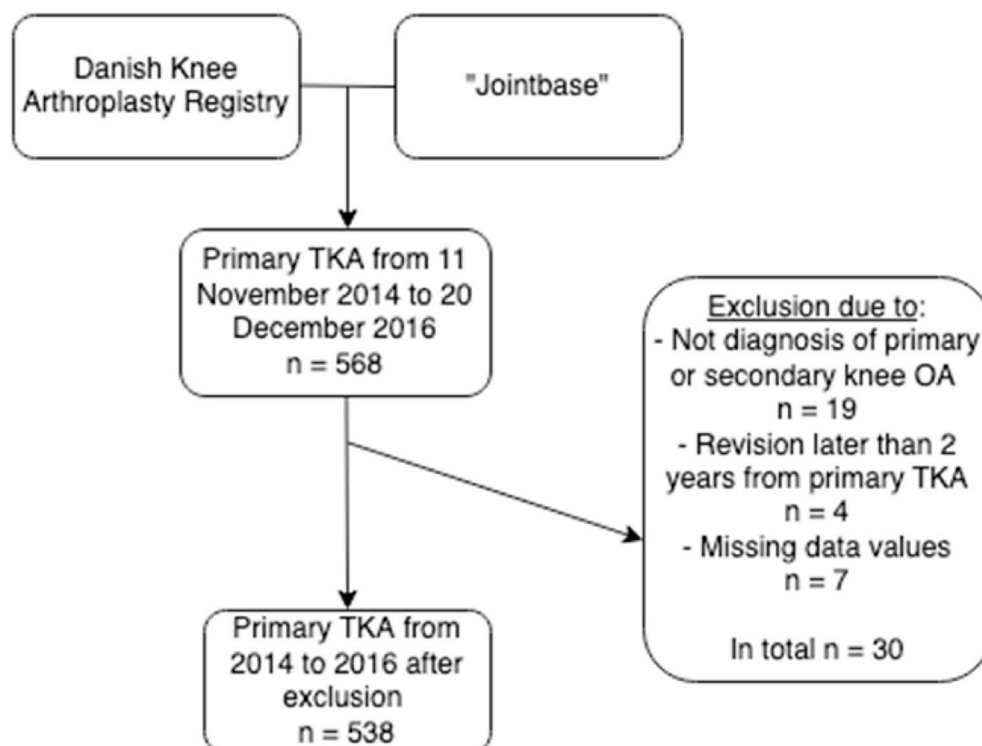


Fig. 1. Flow-chart illustrating the retrieval and exclusion process of the dataset.

Table 1
Sample characteristics.

Characteristic	Data
Total number of patients, <i>n</i> (%)	538 (100%)
Age in years, mean (SD)	67.5 (9.6)
Gender, <i>n</i> (%)	
Male	244 (45%)
Female	294 (55%)
BMI, mean (SD)	29.9 (5.5)
Duration of surgery, mean (SD)	64 min (19.4)
Length of stay, mean (SD)	1.9 (1.3)
Revision within 2 years, <i>n</i> (%)	24 (4.5%)
Revision cause, <i>n</i> (%)	
Infection	9 (38%)
Aseptic loosening	6 (25%)
Instability	3 (13%)
Pain	1 (4%)
Unknown	5 (20%)
Previous surgery in same knee, <i>n</i> (%)	
Yes	143 (27%)
No	395 (73%)
Pain preoperative, mean (SD)	
VAS rest	48.4 (25.3)
VAS activity	68.4 (21.4)
Diabetes mellitus, <i>n</i> (%)	
Yes	174 (32%)
No	364 (68%)
Number of comorbidities, <i>n</i> (%)	
None	203 (38%)
1–2	273 (51%)
3+	62 (11%)
Total	538 (100%)
OKS 0–100 preoperative, mean (SD)	43.4 (14.7)
EQ-5D index preoperative, mean (SD)	0.62 (0.18)

Table 1 - Overview of characteristics. Categorical variables are presented as frequencies (*n*) and percentages (%). Continuous variables are presented as mean and standard deviation (SD). BMI = Body mass index, VAS = Visual analogue scale, OKS = Oxford knee score.

Table 2
Simple Logistic Regression Analysis for identifying predictors of early revision.

Predictors	B	OR	P	OR 95% CI
Gender ^a	0.19	1.21	0.64	[0.53, 2.75]
BMI	0.01	1.01	0.89	[0.93, 1.09]
Age	−0.52	0.59	0.01*	[0.39, 0.88]
Pain rest (VAS)	0.01	1.01	0.35	[0.99, 1.03]
Pain activity (VAS)	0.01	1.01	0.52	[0.99, 1.03]
Diabetes mellitus ^b	−0.15	0.85	0.73	[0.34, 2.10]
Previous surgery in same knee ^c	0.53	1.70	0.22*	[0.73, 3.98]
OKS score 0–100	−0.04	0.96	0.02*	[0.93, 0.99]
EQ-5D index	−2.53	0.08	0.01*	[0.01, 0.55]
Length of stay	0.22	1.25	0.05*	[1.00, 1.57]
Duration of surgery	0.02	1.02	0.01*	[1.00, 1.03]
Number of comorbidities	0.08	1.09	0.16*	[0.97, 1.22]

Table 2 – Simple Logistic Regression analysis for identifying predictors of early revision. B = regression beta coefficients; OR = odds ratio; CI = confidence intervals. P = p-value. *P < 0.25.

^a Male = 1; female = 0.

^b Yes = 1; no = 0.

^c No = 1; yes = 0.

Table 3
The Best Predictors of early revision.

Predictors	B	OR	P	OR 95% CI
Age	−0.46	0.63	0.03	[0.42, 0.95]
EQ-5D index	−2.70	0.07	0.01	[0.01, 0.51]
Number of comorbidities	0.01	1.01	0.15	[0.97, 1.25]

Table 3 – The best predictors of early revision. B = regression beta coefficients; OR = odds ratio; CI = confidence intervals; P = p-value.

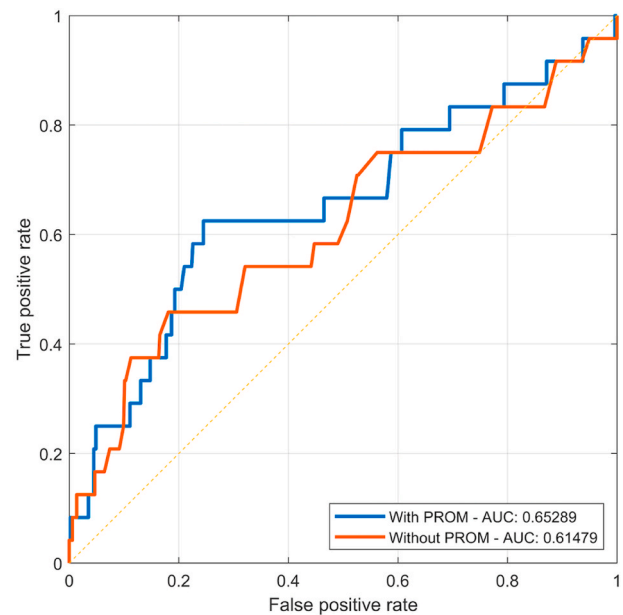


Fig. 2. Receiver Operating Characteristic (ROC)-curve. Area Under the Curve (AUC) with and without PROMs. With PROMs AUC = 0.65. Without PROMs AUC = 0.61. Blue line = with PROMs, red line = without PROMs, yellow line = reference line (no model).

comorbidity. This study demonstrated that the best predictors of early revision were age, preoperative EQ-5D, and number of comorbidities. As depicted in the ROC-curve, neither of the prediction models did reach the predefined threshold for a clinical meaningful model (AUC = 0.65 and 0.61).

The findings in the present study are predominantly complementing findings from other studies reporting that the risk of revision is increasing with younger age, primarily due to a combination of implant wear and increased level of physical activity compared to older counterparts.^{19,41,42} A low preoperative EQ-5D score has not previously been linked with increased risk of revision but has been reported to be associated with a reduced chance of returning to desired physical activity following TKA.⁴³ Although physical inactivity might not cause wearing of the implant it could instead lead to suboptimal rehabilitation of the knee, which could further lead to a poor outcome following TKA ultimately necessitating revision. The finding that an increasing number of comorbidities was a predictor of revision is generally in accordance with previous studies, which reports on poor outcomes following TKA because of the increased risk of postoperative complications.^{44–46} For example, diabetes is a known risk factor for postoperative infections, and infection was the most common cause of early revision in the present study.⁴⁷ However, diabetes alone was not a predictor of revision in the present study. Surprisingly, predictors such as BMI and previous surgery in the same knee were also not associated with early revision in the present study. These findings are equivocal with some studies reporting that BMI is not a predictor of worse outcome following TKA,^{48–50} while others are suggestive of BMI being associated with increased risk of revision.^{51,52} The finding that previous surgery in the same knee was not associated with a worse outcome following TKA is contrary to what other studies report.^{53,54}

In the present study, the prediction models' discriminative capacity were categorized as "poor" with an AUC of 0.65 and 0.61⁴⁰ and thereby only showed marginally better ability to distinguish between those with and without the outcome than no model. This performance is complementing previous studies attempting to predict orthopedic-related factors, also applying LR models.^{55–57} A recent study by El-Galaly et al. failed to build a clinical meaningful prediction model for early revision despite the use of different ML-algorithms such as random

forest, gradient boosting model and supervised neural network.²⁶ However, future studies might still benefit from combining PROMs data and these more advanced types of models, even though no sophisticated ML-algorithm can fully overcome the fact that revisions are rare.²⁷ No threshold was established in the present study to classify revision and no revision. Therefore, an accurate report of sensitivity and specificity was not possible. It should be noted, however, that for an arbitrarily fixed true-positive rate of 60%, for example, it would result in a false-positive rate at 25% in the model with PROMs and a false-positive rate of 50% in the model without PROMs, thus demonstrating the effect of adding PROMs to the model.

One strength of the present study was its reliance on representative and contemporary data obtained from the DKR, a national clinical quality database and part of the Danish Clinical Quality Program, and from a single center specialized hospital.³¹ Prediction models relying on routinely collected data are more readily implementable in clinical practice.⁵⁸ Another strength was the careful selection of predictor variables that was based on both previous studies and clinical reasoning, and the combination of national registry data and PROMs. Potential limitations include the limited sample size and lack of dealing with the imbalanced dataset (514 unrevised TKAs vs. 24 revised TKAs). A large class imbalance might increase the risk of the model treating all the underrepresented classes as background noise or simply assume that all revised patients were unrevised and thereby remain a high accuracy.⁵⁹ One way to deal with this is by using under- or oversampling techniques to inflate the imbalance combined with increasing the sample size which will make it less prone to imbalance problems.⁵⁹ Many clinical predictors were used in the present study, yet it did not contain information on intraoperative components, surgical technique, or surgeon volume that potentially could be important explanatory variables in terms of revision surgery.⁵⁴ Inclusion of annual volume of arthroplasties on surgeon-level instead of hospital-level might have improved the performance of the model in the present study. However, this information was not accessible through the DKR. Despite this, future studies might benefit from combining patient-related factors with intraoperative or postoperative factors as these possibly have a stronger correlation to revision surgery. Although intraoperative or postoperative information is not of great value during the preoperative patient selection, it might improve the planning and monitoring of the patient's postoperative course. Another limitation is the lack of external validation of the model due to no access to such data. Without external validation it is not possible to estimate the performance of the model on new data, and therefore not possible to evaluate the generalizability and transferability of the model.⁶⁰ To compensate for this, internal validation with stratified 5-fold cross-validation was performed, which have been shown to minimize the risk of bias in datasets with a limited sample size.³⁸ Finally, the study may be limited by the use of stepwise selection of variables at significance level. This approach may cause selection bias and overestimate regression coefficients as a result of overfitting. Overfitting leads to less accurate predictions, especially in small datasets or with weakly predictive variables.^{38,61} However, despite the use of stepwise selection, a relatively high number of variables, and the limited sample size, the present results are not estimated to be a direct consequence of overfitting due to the poor AUC. Instead, it is more likely to be a result of weakly predictive variables with poor correlation to the outcome.

Prediction models as a health technology could have important future applications in clinical practice by assisting clinicians in their pre and postoperative planning and shared decision making.^{24,62,63} The combination of ML and data is essential in terms of predicting the risk of certain outcomes and have already provided predictions regarding opioid use after spinal surgery,⁶⁴ different PROMs following TKA,^{24,55} patient-satisfaction and length of stay following TKA.^{22,25} This combination yields a unique opportunity to improve preoperative patient selection, monitoring, and identification of risk groups. The field of knee arthroplasty have a long tradition of data collecting through national registries, which provide readily accessible data from large

population-based cohorts.⁶⁵ Thus, making it well suited for application of ML to develop prediction models readily useable within Denmark and easily transferable to countries with similar registries.

5. Conclusions

The present study developed and evaluated a prediction model aimed to predict early revision TKA and can be seen as a first step towards determining the most important factors of early revision. Although more work is needed in order to reach a clinically meaningful quality of the predictions, our results show that the inclusion of PROMs seems to improve the quality of the model.

Funding information

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declarations of competing interest

None.

Acknowledgements

We thank the Danish knee surgeons, the Danish Knee Arthroplasty Registry, and the Department of Orthopedic Surgery, Aalborg University Hospital, Denmark for thoroughly collecting, reporting, and providing the data behind this study. We also thank Assistant Professor Thomas Kronborg Larsen for his assistance in the development of the prediction models in the present study.

References

- James SL, Abate D, Abate KH, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 354 Diseases and Injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet*. 2018;392:1789–1858. [https://doi.org/10.1016/S0140-6736\(18\)32279-7](https://doi.org/10.1016/S0140-6736(18)32279-7).
- Reinhart L, Knight W, Roberts L, Mendes C. HHS public access global burden of disease study 2010. *Lancet*. 2012;380:2163–2196. [https://doi.org/10.1016/S0140-6736\(12\)61729-2](https://doi.org/10.1016/S0140-6736(12)61729-2).
- Kloppenborg M, Berenbaum F. Osteoarthritis year in review 2019: epidemiology and therapy. *Osteoarthritis Cartilage*. 2020. <https://doi.org/10.1016/j.joca.2020.01.002>. PII: S1063-4584(20)30007-8.
- O'Neill TW, McCabe PS, McBeth J. Update on the epidemiology, risk factors and disease outcomes of osteoarthritis. *Best Pract Res Clin Rheumatol*. 2018;32:312–326. <https://doi.org/10.1016/j.berh.2018.10.007>.
- Zhang W, Doherty M, Peat G, et al. EULAR evidence-based recommendations for the diagnosis of knee osteoarthritis. *Ann Rheum Dis*. 2010;69:483–489. <https://doi.org/10.1136/ard.2009.113100>.
- Kiadaliri AA, Gerhardsson de Verdier M, Turkiewicz A, Lohmander LS, Englund M. Socioeconomic inequalities in knee pain, knee osteoarthritis, and health-related quality of life: a population-based cohort study in southern Sweden. *Scand J Rheumatol*. 2017;46:143–151. <https://doi.org/10.1080/03009742.2016.1181203>.
- Hubertsson J, Petersson IF, Thorstenson CA, Englund M. Risk of sick leave and disability pension in working-age women and men with knee osteoarthritis. *Ann Rheum Dis*. 2013;72:401–405. <https://doi.org/10.1136/annrheumdis-2012-201472>.
- Carr AJ, Robertsson O, Graves S, et al. Knee replacement. *Lancet*. 2012;379:1331–1340. [https://doi.org/10.1016/S0140-6736\(11\)60752-6](https://doi.org/10.1016/S0140-6736(11)60752-6).
- Price AJ, Alvand A, Troelsen A, et al. Knee replacement. *Lancet*. 2018;392:1672–1682. [https://doi.org/10.1016/S0140-6736\(18\)32344-4](https://doi.org/10.1016/S0140-6736(18)32344-4).
- Lützner J, Kasten P, Günther KP, Kirschner S. Surgical options for patients with osteoarthritis of the knee. *Nat Rev Rheumatol*. 2009;5:309–316. <https://doi.org/10.1038/nrrheum.2009.88>.
- Dansk Kneeloplastikregister Årsrapport. 2019.
- Hamilton DF, Howie CR, Burnett R, Simpson AHRW, Patton JT. Dealing with the predicted increase in demand for revision total knee arthroplasty: challenges, risks and opportunities. *Bone Jt J*. 2015;97-B. <https://doi.org/10.1302/0301-620X.97B6.35185>, 723–8.
- Silverwood V, Blagojevic-Bucknall M, Jinks C, Jordan JL, Protheroe J, Jordan KP. Current evidence on risk factors for knee osteoarthritis in older adults: a systematic review and meta-analysis. *Osteoarthritis Cartilage*. 2015;23:507–515. <https://doi.org/10.1016/j.joca.2014.11.019>.
- Labek G, Thaler M, Janda W, Agreiter M, Stöckl B. Revision rates after total joint replacement: cumulative results from worldwide joint register datasets. *J Bone Jt Surg - Ser B*. 2011;93 B:293–297. <https://doi.org/10.1302/0301-620X.93B3.25467>.

- 15 Kurtz S, Ong K, Lau E, Mowat F, Halpern M. Projections of primary and revision hip and knee arthroplasty in the United States from 2005 to 2030. *J Bone Jt Surg - Ser A*. 2007;89:780–785. <https://doi.org/10.2106/JBJS.F.00222>.
- 16 Kurtz S, Mowat F, Ong K, Chan N, Lau E, Halpern M. Prevalence of primary and revision hip and knee arthroplasty in the United States from 1990 through 2002. *J Bone Joint Surg Am*. 2005;87:1487–1498.
- 17 Greidanus NV, Peterson RC, Masri BA, Garbus DS. Quality of life outcomes in revision versus primary total knee arthroplasty. *J Arthroplasty*. 2011;26:615–620. <https://doi.org/10.1016/j.arth.2010.04.026>.
- 18 Bhandari M, Smith J, Miller LE, Block JE. Clinical and economic burden of revision knee arthroplasty. *Clin Med Insights Arthritis Musculoskelet Disord*. 2012;5:89–94. <https://doi.org/10.4137/CMAMD.S10859>.
- 19 Kurtz SM, Lau E, Ong K, Zhao K, Kelly M, Bozic KJ. Future young patient demand for primary and revision joint replacement: national projections from 2010 to 2030. *Clin Orthop Relat Res*. 2009;467:2606–2612. <https://doi.org/10.1007/s11999-009-0834-6>.
- 20 W-Dahl A, Robertsson O, Lidgren L. Surgery for knee osteoarthritis in younger patients: a Swedish register study. *Acta Orthop*. 2010;81:161–164. <https://doi.org/10.3109/17453670903413186>.
- 21 Stambough JB, Clohisy JC, Barrack RL, Nunley RM, Keeney JA. Increased risk of failure following revision total knee replacement in patients aged 55 years and younger. *Bone Jt J*. 2014;96B:1657–1662. <https://doi.org/10.1302/0301-620X.96B12.34486>.
- 22 Navarro SM, Wang EY, Haeberle HS, et al. Machine learning and primary total knee arthroplasty: patient forecasting for a patient-specific payment model. *J Arthroplasty*. 2018;33:3617–3623. <https://doi.org/10.1016/j.arth.2018.08.028>.
- 23 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;477:1267–1279. <https://doi.org/10.1097/CORR.0000000000000687>.
- 24 Huber M, Kurz C, Leidl R. Predicting patient-reported outcomes following hip and knee replacement surgery using supervised machine learning. *BMC Med Inf Decis Making*. 2019;19:1–13. <https://doi.org/10.1186/s12911-018-0731-6>.
- 25 Scott CEH, Oliver WM, MacDonald D, Wade FA, Moran M, Breusch SJ. Predicting dissatisfaction following total knee arthroplasty in patients under 55 years of age. *Bone Jt J*. 2016;98-B:34–1625. <https://doi.org/10.1302/0301-620X.98B12.BJJ-2016-0375.R1>.
- 26 El-Galaly A, Grazal C, Kappel A, Nielsen PT, Jensen SL, Forsberg JA. Can machine-learning algorithms predict early revision TKA in the Danish knee arthroplasty registry? *Clin Orthop Relat Res*. 2020;1–14. <https://doi.org/10.1097/corr.0000000000001343>. Publish Ah:1–14.
- 27 Fontana MA. CORR Insights®: can machine-learning algorithms predict early revision TKA in the Danish knee arthroplasty registry? *Clin Orthop Relat Res*. 2020;478:2102–2104. <https://doi.org/10.1097/CORR.0000000000000402>.
- 28 Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the tripod statement. *J Clin Epidemiol*. 2015;68:112–121. <https://doi.org/10.1016/j.jclinepi.2014.11.010>.
- 29 Vandembroucke JP, Von Elm E, Altman DG, et al. Strengthening the reporting of observational studies in epidemiology (STROBE): explanation and elaboration. *Epidemiology*. 2007;18:805–835. <https://doi.org/10.1097/EDE.0b013e3181577511>.
- 30 Soerensen NS, Hansen L, Laursen M, Larsen AE, et al. *Udvikling Af PRO Inden for Knæ- Og Hoftealloplastik I Region Nordjylland*. 2019.
- 31 Pedersen AB, Mehner F, Odgaard A, Schröder HM. Existing data sources for clinical epidemiology: the Danish knee Arthroplasty register. *Clin Epidemiol*. 2012;4:125–135. <https://doi.org/10.2147/CLEP.S30050>.
- 32 Dawson J, Fitzpatrick R, Murray D, Carr A. Questionnaire on the perceptions of patients about total knee replacement. *J Bone Joint Surg Br*. 1998;78:856. <https://doi.org/10.1302/0301-620X.80b1.0800063>.
- 33 Rabin R, de Charro F. EQ-SD: a measure of health status from. *Ann Med*. 2001;33:337–343. <https://doi.org/10.1111/j.1461-0248.2011.01689.x>.
- 34 Harris K, Dawson J, Gibbons E, et al. Systematic review of measurement properties of patient-reported outcome measures used in patients undergoing hip and knee arthroplasty. *Patient Relat Outcome Meas*. 2016;7:101–108. <https://doi.org/10.2147/prom.s97774>.
- 35 Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology*. 2010;21:128–138. <https://doi.org/10.1097/EDE.0b013e3181c30fb2>.
- 36 Perkins NJ, Schisterman EF. The inconsistency of “optimal” cutpoints obtained using two criteria based on the receiver operating characteristic curve. *Am J Epidemiol*. 2006;163:670–675. <https://doi.org/10.1093/aje/kwj063>.
- 37 Hosmer DW, Hosmer T, Le Cessie S, Lemeshow S. A comparison of goodness-of-fit tests for the logistic regression model. *Stat Med*. 1997;16:965–980. [https://doi.org/10.1002/\(SICI\)1097-0258\(19970515\)16:9<965::AID-SIM509>3.0.CO;2-O](https://doi.org/10.1002/(SICI)1097-0258(19970515)16:9<965::AID-SIM509>3.0.CO;2-O).
- 38 Royston P, Moons KGM, Altman DG, Vergouwe Y. Prognosis and prognostic research: developing a prognostic model. *BMJ*. 2009;338:1373–1377. <https://doi.org/10.1136/bmj.b604>.
- 39 Hanley A, McNeil J, Ph D. The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve n.d.:29–36.
- 40 Fischer JE, Bachmann LM, Jaeschke R. A readers' guide to the interpretation of diagnostic test properties: clinical example of sepsis. *Intensive Care Med*. 2003;29:1043–1051. <https://doi.org/10.1007/s00134-003-1761-8>.
- 41 Harrysson OLA, Robertsson O, Nayfeh FJ. Higher cumulative revision rate of knee arthroplasties in younger patients with osteoarthritis. *Clin Orthop Relat Res*. 2004;162–8. <https://doi.org/10.1007/01.blo.0000127115.05754.ce>.
- 42 Meehan JP, Danielsen B, Kim SH, Jamali AA, White RH. Younger age is associated with a higher risk of early periprosthetic joint infection and aseptic mechanical failure after total knee arthroplasty. *J Bone Jt Surg - Ser A*. 2014;96:529–535. <https://doi.org/10.2106/JBJS.M.00545>.
- 43 Harbourne AD, Sanchez-Santos MT, Arden NK, Filbay SR. Predictors of return to desired activity 12 months following unicompartmental and total knee arthroplasty. *Acta Orthop*. 2019;90:74–80. <https://doi.org/10.1080/17453674.2018.1542214>.
- 44 Singh JA, Lewallen DG. Medical and psychological comorbidity predicts poor pain outcomes after total knee arthroplasty. *Rheumatol (United Kingdom)*. 2013;52:916–923. <https://doi.org/10.1093/rheumatology/kes402>.
- 45 Singh JA, Lewallen DG. Depression in primary TKA and higher medical comorbidities in revision TKA are associated with suboptimal subjective improvement in knee function. *BMC Musculoskel Disord*. 2014;15. <https://doi.org/10.1186/1471-2474-15-127>.
- 46 Peter WF, Dekker J, Tilbury C, et al. The association between comorbidities and pain, physical function and quality of life following hip and knee arthroplasty. *Rheumatol Int*. 2015;35:1233–1241. <https://doi.org/10.1007/s00296-015-3211-7>.
- 47 Bolognesi MP, Marchant MH, Viens NA, Cook C, Pietrobon R, Vail TP. The impact of diabetes on perioperative patient outcomes after total hip and total knee arthroplasty in the United States. *J Arthroplasty*. 2008;23:92–98. <https://doi.org/10.1016/j.arth.2008.05.012>.
- 48 Judge A, Arden NK, Cooper C, et al. Predictors of outcomes of total knee replacement surgery. *Rheumatol (United Kingdom)*. 2012;51:1804–1813. <https://doi.org/10.1093/rheumatology/kes075>.
- 49 Kane RL, Saleh KJ, Wilt TJ, Bershadsky B. The functional outcomes of total knee arthroplasty. *J Bone Jt Surg - Ser A*. 2005;87:1719–1724. <https://doi.org/10.2106/JBJS.D.02714>.
- 50 Clement ND, MacDonald D, Burnett R, Breusch SJ. Diabetes does not influence the early outcome of total knee replacement: a prospective study assessing the Oxford knee score, short form 12, and patient satisfaction. *Knee*. 2013;20:437–441. <https://doi.org/10.1016/j.knee.2013.07.009>.
- 51 Jämsen E, Peltola M, Eskelinen A, Lehto MUK. Comorbid diseases as predictors of survival of primary total hip and knee replacements: a nationwide register-based study of 96 754 operations on patients with primary osteoarthritis. *Ann Rheum Dis*. 2013;72:82–1975. <https://doi.org/10.1136/annrheumdis-2012-202064>.
- 52 Namba RS, Cafri G, Khatod M, Inacio MCS, Brox TW, Paxton EW. Risk factors for total knee arthroplasty aseptic revision. *J Arthroplasty*. 2013;28:122–127. <https://doi.org/10.1016/j.arth.2013.04.050>.
- 53 Sanchez-Santos MT, Garriga C, Judge A, et al. Development and validation of a clinical prediction model for patient-reported pain and function after primary total knee replacement surgery article. *Sci Rep*. 2018;8:1–9. <https://doi.org/10.1038/s41598-018-21714-1>.
- 54 Baker PN, Deehan DJ, Lees D, et al. The effect of surgical factors on early patient-reported outcome measures (PROMS) following total knee replacement. *J Bone Jt Surg - Ser B*. 2012;94 B:1058–1066. <https://doi.org/10.1302/0301-620X.94B8.28786>.
- 55 Berliner JL, Brodke DJ, Chan V, SooHoo NF, Bozic KJ. Can preoperative patient-reported outcome measures be used to predict meaningful improvement in function after TKA? *Clin Orthop Relat Res*. 2017;475:149–157. <https://doi.org/10.1007/s11999-016-4770-y>.
- 56 Harris AHS, Kuo AC, Weng Y, Trickey AW, Bowe T, Giori NJ. Can machine learning methods produce accurate and easy-to-use prediction models of 30-day complications and mortality after knee or hip arthroplasty? *Clin Orthop Relat Res*. 2019;477:452–460. <https://doi.org/10.1097/CORR.0000000000000601>.
- 57 Khor S, Lavalley D, Cizik AM, et al. Development and validation of a prediction model for pain and functional outcomes after lumbar spine surgery. *JAMA Surg*. 2018;153:634–642. <https://doi.org/10.1001/jamasurg.2018.0072>.
- 58 Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, May M, Brindle P. Derivation and validation of QRISK, a new cardiovascular disease risk score for the United Kingdom: prospective open cohort study. *Br Med J*. 2007;335:136–141. <https://doi.org/10.1136/bmj.39261.471806.55>.
- 59 Japkowicz N, Stephen S. *The Class Imbalance Problem A Systematic Study* fulltext.Pdf. vol. 6. 2002:429–449.
- 60 Altman DG, Vergouwe Y, Royston P, Moons KGM. Prognosis and prognostic research: validating a prognostic model. *BMJ*. 2009;338:1432–1435. <https://doi.org/10.1136/bmj.b605>.
- 61 Steyerberg EW, Eijkemans MJC, Harrell FE, Habbema JDF. Prognostic modeling with logistic regression analysis: in search of a sensible strategy in small data sets. *Med Decis Making*. 2001;21:45–56. <https://doi.org/10.1177/0272989X0102100106>.
- 62 Haeberle HS, Helm JM, Navarro SM, et al. Artificial intelligence and machine learning in lower extremity arthroplasty: a review. *J Arthroplasty*. 2019;34:2201–2203. <https://doi.org/10.1016/j.arth.2019.05.055>.
- 63 Yu D, Jordan KP, Snell KIE, et al. Development and validation of prediction models to estimate risk of primary total hip and knee replacements using data from the UK: two prospective open cohorts using the UK Clinical Practice Research Datalink. *Ann Rheum Dis*. 2019;78:91–99. <https://doi.org/10.1136/annrheumdis-2018-213894>.
- 64 Karhade AV, Ogink PT, Thio QCB, et al. Development of machine learning algorithms for prediction of prolonged opioid prescription after surgery for lumbar disc herniation. *Spine J*. 2019;19:1764–1771. <https://doi.org/10.1016/j.spinee.2019.06.002>.
- 65 Malchau H, Garellick G, Berry D, et al. Arthroplasty implant registries over the past five decades: development, current, and future impact. *J Orthop Res*. 2018;36:2319–2330. <https://doi.org/10.1002/jor.24014>.