DR. AFSHIN JAMSHIDI (Orcid ID : 0000-0002-7644-4412)

DR. JEAN-PIERRE PELLETIER (Orcid ID : 0000-0001-9930-6453)

DR. JOHANNE MARTEL-PELLETIER (Orcid ID : 0000-0003-2618-383X)

Article type : Original Article

**Running head: Machine learning-based survival prediction for TKR**

# Machine learning-based individualized survival prediction model for total knee replacement in osteoarthritis: Data from the Osteoarthritis Initiative

Afshin Jamshidi, PhD[1,2], Jean-Pierre Pelletier, MD[1], Aurelie Labbe[3], François Abram, PhD[4], Johanne Martel-Pelletier, PhD[1*], Arnaud Droit, PhD[2*]

[1] Osteoarthritis Research Unit, University of Montreal Hospital Research Centre (CRCHUM), Montreal, Quebec, Canada

[2] Laval University Hospital Research Centre, Quebec, Canada

[3] Department of Decision Sciences, HEC Montreal, Montreal, Quebec, Canada

[4] Medical Imaging Research and Development, ArthroLab Inc., Montreal, Quebec, Canada

* These authors contributed equally as senior investigators.

**Funding**

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article. This work was supported in part by the Osteoarthritis Research Unit of the University of Montreal Hospital Research Centre and the Chair in Osteoarthritis, University of Montreal, Montréal, Canada; CHU de Quebec Research Centre – Laval University, Quebec, Canada; and the Canada First Research Excellence Fund through the TransMedTech Institute, Montreal, Canada.

**Disclosure Statement**

J-P. Pelletier and J. Martel-Pelletier are shareholders in ArthroLab Inc. and François Abram is an employee of ArthroLab Inc. The other authors have no conflicts of interest for this study.

**Direct correspondence to:** Johanne Martel-Pelletier, Osteoarthritis Research Unit, University of Montreal Hospital Research Centre (CRCHUM), 900 Saint-Denis, Suite R11.412, Montreal, Quebec, Canada H2X 0A9; Tel: 514-890-8000; Fax: 514-412-7582; Email: jm@martelppelletier.ca; ORCID ID: https://orcid.org/0000-0003-2618-383X.

**ABSTRACT**

**Objective:** By using machine learning (ML), our study aimed to build a model to predict risk and time to total knee replacement (TKR) of an osteoarthritic knee.

**Methods:** Features were from OAI at baseline. Lasso's Cox identified the ten most important among 1107 features. The prognostic power of the selected features was assessed by Kaplan-Meier and applied to seven ML methods: Cox, DeepSurv, Random Forest, linear/kernel Support Vector Machine (SVM), and linear/neural Multi-Task Logistic Regression. As some of the ten features included similar X-ray measurements, we further looked at using the least features without compromising the accuracy of the model. Prediction performance was assessed by the concordance index (C-index), Brier score, and time-dependent area under the curve (AUC).

**Results:** Identified features included X-rays, the MRI feature bone marrow lesions (BML) in medial condyle, hyaluronic acid injection, performance measure, medical history, and knee symptoms. The methodologies Cox, DeepSurv, and linear SVM demonstrated the highest accuracy (C-index, 0.85, Brier score 0.02, and AUC, 0.87). DeepSurv was chosen to build the prediction model to estimate the time to TKR for a given knee. Moreover, we were able to use only three features (C-index, 0.85, Brier score 0.02, and AUC, 0.86) including BML, KL grade and knee symptoms, to predict risk and time of a TKR event.

**Conclusion:** For the first time, using the OAI cohort, we developed a model to predict with high accuracy if and when a given osteoarthritic knee would require TKR and who would likely progress fast toward this event.

**Significance and Innovations:**

- Individualized prediction time to TKR can be done with high accuracy using only three features.
- Combination of X-ray, bone marrow lesions and symptom demonstrated the most significant impact in prediction time to a TKR event.
- The developed TKR prediction model, once validated, could guide clinicians to appropriate therapeutic strategies.

**Keywords:** osteoarthritis, total knee replacement, survival analysis, prediction, feature selection, machine learning, X-ray, magnetic resonance imaging.

## INTRODUCTION

Knee osteoarthritis (OA) is the most common joint disorder and leading cause of disability across the world. Typically, this disease progresses slowly over many years. However, for many subjects (knees), the disease progresses rapidly. Recent studies documented that on a population with no radiographic knee OA, it is estimated that, over a four-year timeframe, the incidence of "accelerated knee OA" ranged from 0.4% to 22.1% (1, 2). Davis *et al.* (3), further reported that knees with accelerated OA have a higher chance of needing a total knee replacement (TKR) compared to those with a more gradual onset or without accelerated knee OA.

The initiation and rapid progression of primary OA over time for a given knee is generally unknown. Integrating data for uncovering the complex mechanisms of knee OA subjects leading to a TKR event will enable objective-driven analytical leads, improve survival prediction (4-6), and develop better therapeutic strategies for these subjects.

Survival analysis is one of the fundamental tools in the medical domain to identify predictors of time to adverse events and develop systems for clinical decision support. In knee OA research, survival analysis can be used to predict time of pathological events such as TKR. For OA, it is not possible to apply traditional survival analysis methods (7) as this disease involves integrated high-dimensional and nonlinear data structures. For instance, the most commonly used approach for analyzing survival data, the semi-parametric Cox Proportional Hazard (Cox-PH) regression method (8), demonstrated several drawbacks in analyzing non-linear data structures, high dimensions, and low sample size data (9). Machine/deep learning-based survival prediction models have proven to be a better option in the case of complex data with interactions among the features (10) as in OA.

To date, there have been no comprehensive attempts to identify the most important features and to build a survival-based model for predicting the time to TKR for a given OA individual. Recently, Heisinger *et al.* (11), by using a small cohort (n=165), applied 14 factors in a four-year period prior to TKR to predict a patient's need for TKR surgery. In addition, also recently developed was an image-based model with knee radiographs to classify patients with OA at high risk of TKR (12). There are also other works on TKR survival analyses, but all were done for subjects after TKR surgery, not before. For example, the research questions in these studies included how long does a TKR last (13), what is the subject death rate after TKR surgery (14), what are the clinical and radiological outcomes of TKR (15).

Our study was the first to look at TKR as the survival outcome. Two specific research questions were addressed: 1) Which features at baseline are most associated with accelerated knee OA leading to TKR? 2) Can we estimate for a given OA knee, the risk and time to TKR (e.g. remaining useful life [RUL] for a given knee)? To answer these questions, we evaluated the length of time from the date of enrollment until the TKR event (overall survival) by i) applying feature selection methods to find the most important features in survival analysis of a TKR event, and ii) developing a survival prediction model based on the selected features using machine learning (ML)-based methods.

## MATERIALS AND METHODS

### *Knee selection*

Data used in this study were from the Osteoarthritis Initiative (OAI) cohort (https://data-archive.nimh.nih.gov/oai/). For details on the cohort, refer to Supplementary Methods.

### *Selection of predictor features*

The features were selected at baseline from the OAI database and from the quantitative determination of knee tissues by MRI (Supplementary Tables S1, S2A, S2B). The feature selection was done using the target knee (for the definition, refer to Supplementary Methods) of each subject; 1431 knees were included (Figure 1).

For the feature selection, as the knees with accelerated knee OA (progressors) have a higher chance of needing a TKR compared to those without accelerated knee OA (3), we considered two groups: progressors and non-progressors. The OA knee progressor and non-progressor definition from this cohort was as previously described and discussed (16). The data from the knees of 733 progressors and 698 non-progressors were included (Figure 1). (For further explanation, refer to Supplementary Methods).

Next, we used all the features at baseline in the OAI database which included in addition to standard (clinical, demographics, anthropometry, to name a few) other uncommon features (health-care access, nutrition, knee MRI data, to name a few). After the data cleaning (for the methodology, refer to Supplementary Methods), 1107 features remained and were divided into categories and subcategories as recently described (16) and reported in Supplementary Tables S1, S2.

The feature selection was done using the Lasso's Cox method. The data (1431 knees) were divided into training (80%) and testing (20%) sets to generate the prediction model and data applied for measuring the accuracy of the developed prediction model, respectively. For the extraction of the most important features, the glmnet (17) package was used in R. By using the Lasso's Cox method (18), the ten top baseline features were extracted and served for designing a TKR survival prediction model.

***TKR survival***

For the survival analysis, we considered all the knees to predict the outcome (time to TKR) (Figure 1). Knees (n=7589) with complete data at the final visit (108 months) were included. As a first step, to verify if and which of the most important features described in the literature could individually impact TKR survival, we compared, among the well-known features related to OA, the survival curves of six of them. These included age, gender, race, bone mass index (BMI), Kellgren-Lawrence (KL) (19), and Western Ontario and McMaster Universities Arthritis Index (WOMAC) pain (20). Further, the Kaplan-Meier analysis with log-rank test (P value <0.050) was used on the selected features to compare the prognosis power of each feature for risk and time to TKR.

*ML methods applied for building survival prediction models*

Survival prediction models for TKR events were built with 7589 knees (training, 6071; testing, 1518). The following models in the PySurvival package (21) in Python 3.7 were applied: 1) Cox-PH model (8); 2) DeepSurv/Nonlinear model (10); 3) Linear Multi-Task Logistic Regression (MTLR) (22); 4) Neural MTLR model (23); 5) Random Survival Forest model (24); 6) Linear Support Vector Machines (SVM) model (25); and 7) Kernel SVM model (26) (for details, refer to Supplementary Methods). Figure S1 illustrates the pipeline of the data analysis.

*Hyperparameter tuning*

For the hyperparameter tuning, we applied the GridSearchCV in scikit-learn to determine which values of hyper-parameters perform best in each model (Table S3). We selected the configuration with the largest validation Concordance index (C-index; refer to Supplementary Methods for description).

*Prediction of the performance evaluation*

To compute the prediction performance of the above-mentioned models, we applied three metrics on the test dataset: C-index, Brier score, and the time-dependent AUC (area under the Receiver Operating Characteristic [ROC]) (refer to Supplementary Methods for description).

***Overall and knee-specific predictions of TKR survival***

To compute the prediction performance of the best model, we compared the time series of the actual and predicted number of knees experiencing a TKR event, for each time $t$, by calculating the mean and median absolute error and a root mean square error.

To show representative curves of the TKR survival of different conditions, we used five knees from the OAI that demonstrated a range of values for each selected feature and applied the selected model to predict their specific survival curves.

**RESULTS**

***Association between clinical/demographic features and TKR survival***

Of the 7589 knees included in the analysis, 413 had a TKR event and 7176 survived from a TKR event (right-censored data, which occurs when the TKR event does not happen by the end of study) at the end of follow-up (3320 days or 108 months). Figure S2A illustrates that at 3000 days, 7224 survived from TKR and 365 knees had TKR.

Next, we assessed the association between age, gender, race, BMI, WOMAC pain, and KL grade features with survival probability before a TKR event and illustrated the most important (Figure 2) and less important (Figure S2B) ones.

With regard to race (Figure 2), although the survival curves showed a very small difference between groups in regard to the time to survival, they appeared to be statistically different (P=0.0078). The time for TKR for the Asians and Other non-white groups is slightly higher than the other two groups studied: White/Caucasian and Black/African American. This could be because of the smaller number of patients available in these two groups (105 and 60, respectively). In addition, it is believed that the slight difference (about 3% between the best and worst groups) could not be related to less disease, but to other reasons such as access and preference of some groups to surgery.

Another major risk factor involved in the OA process is the BMI (Figure 2). Data showed knees from morbidly obese, obese, and overweight subjects have a higher chance of TKR in comparison with those from healthy subjects and underweight ($P<0.0001$). In addition, knees from healthy subjects are slightly more at risk of a TKR event than those from underweight subjects.

For WOMAC pain (20) , knees (Figure 2) in the groups having scores 6-10 and 11-15 have a high chance of a TKR event ($P<0.0001$) in comparison with knees from subjects with less pain level (scores 0-5) and with those having high pain level (scores 16-20). The survival at the end of the study of the groups having a score between 6 and 15 is about 87%.

The KL grade, a widely used approach for classifying the severity of knee OA using knee radiographs (19), was also assessed. As illustrated (Figure 2), there is a drastic and significant decrease ($P<0.0001$) in survival chances of OA knees with a KL grade 3 (84% survival at the end of study), but more so with those of KL grade 4, in which the probability is decreasing to around 60% at the end of the study.

For age (Figure S2B), data showed that although there is a statistically significant difference between the two groups ($P<0.0001$), the group over 60 has less chance of survival in comparison with knees from subjects under 60, the probability of survival in both groups is more than 90% at the end of the study.

Regarding the gender (Figure S2B), there is no significant difference between them in survival probability for a TKR event.

When comparing the above six features together (Figures 2 and S2B), it appears that WOMAC pain and more so KL have a more significant impact on TKR survival, while probability of survival in the other groups are still very high (>90%) at the end of the study.

***Building a model for time to TKR event***

*Selection of survival-based features*

Further and to build a model, we used 1431 knees (Figure 1) to identify from 1107 features (Table S1) the most important TKR survival-based features. This was performed with the Lasso's Cox method. The selected top 10 features (Table 1) included five X-ray, MRI assessed bone marrow lesions (BMLs) in the medial condyle, a performance measure, a medical history and a medication feature, and knee symptoms (sometimes swelling, seven days).

We then assessed the association of the selected features with survival probability before a TKR event (Figure 3 most important and Figure S3 less important features). Of note, we did not include the KL graph as it was already in Figure 2 and data described above; however, this feature was kept for the model. Data showed that in addition to KL grade 4 (Figure 2), four other features (Figure 3) demonstrated a high impact on the TKR event. For three of them (composite OA grade 4, medial condyle BMLs >0.2, and osteophytes and joint space narrowing (JSN) both with a score of 2), the survival probability at the end of the study was around 65%; for the fourth, having received a hyaluronic acid (HA) injection in the past 6 months, survival probability was 75%.

For the other features (Figure S3), although statistical difference was obtained, except for the 400-meter walk feature, the survival probability at the end of study was higher than 80%.

*Development of survival prediction models based on the selected features*

Seven ML methods were applied. As mentioned above, the 400-meter walk feature was not statistically different (Figure S3), the survival prediction models were then developed without this feature. By using the nine remaining features (Table 1) and the seven ML methods, data revealed a very low Brier score for all of them (0.02), indicating that all models have very good predictive abilities. Cox, DeepSurv, and linear SVM models demonstrated the highest C-index (0.85) compared to the Random Forest (C-index, 0.82), Kernel SVM (C-index, 0.83), and linear/neural Multi-Task Logistic Regression (C-index, 0.80). However, as nonlinear analysis outperforms when analyzing the huge amount of data in ML for finding important patterns and predicting diseases, we eliminated the linear SVM method. Between Cox and DeepSurv methods, we chose DeepSurv as it can consider nonlinear interactions between features and was shown to better handle complex data interactions (as

OA dataset) among features and to outperform other models in general and Cox in particular (7, 10). Indeed, the assumption of the linear log-risk function in Cox might by too simplistic when dealing with personalized survival predictions. Therefore, further analysis was performed with the DeepSurv model to better fit OAI survival data with nonlinear log-risk functions.As some of the selected X-ray features including the composite OA; osteophytes and JSN; baseline radiographic knee OA; #7 and KL grade 4 (Table 1) are based on a similar measurement, we further analyzed if one or some of them could be removed from the model without impairing the prediction model. To this end, we consider, in addition to the other five features (Table 1, #2, 6, 8, 9, 10), only one of them each time in the DeepSurv model. Further, to explore if eliminating one or more of the above-mentioned five features could yield similar statistical indexes, we removed each of them in a recursive manner from the model. Data showed that we were able to achieve an identical C-index (0.85) and Brier score (0.02) using the three following features; BML in medial condyle, KL grade, and knee symptoms: sometimes swelling, last seven days (Table 1). Therefore, for the next steps we considered only these three features.

***Overall predictions of TKR survival***

We then compared, by using the DeepSurv model, the time series of the actual vs. predicted number of knees experiencing a TKR and the risk of TKR, for each time point. Data showed (Figure 4A) that the time series of actual and predicted number of knees experiencing a TKR are quite similar in which the predicted values fall in the confidence interval of actual values, with a very low mean, 5.64, and median, 5.10 absolute error; and a root mean square error, 6.55. This indicates that the average prediction error of the model is very low throughout the entire timeline, in which the average error is about five knees in all of 3320 days.

***Knee-specific predictions of TKR survival***

To plot the knee-specific survival curves and to estimate individually the TKR event time using the DeepSurv model, we considered five knees from the OAI dataset which were selected according to a range of values for each selected feature. Figure 4B indicates the values for the three selected features and Figure 4C, the TKR survival curves for the five knees. Comparison of the survival curve (Figure 4C) indicates that the model could perfectly predict the TKR event time. Hence, when the survival curve reaches 0% of survival probability, it indicates the approximate TKR event time. In addition, this could identify the RUL of the knee before TKR. RUL could be calculated as the difference between the enrollment date and the date that the survival curve reaches 0% of survival. Of note, all the curves start from the day 357 (the first TKR event in this study), however, the RUL should be calculated from the day 0 till the 0% of survival. For example, in the case of knee #5, the RUL is around 1000 days. Moreover, knees #1 and 2 showed that the time to TKR will be longer than the cohort time limit. This is not unexpected as their baseline values for the selected features are low (Figure 4B).

From these data, one could classify the knees into different groups of TKR event. For example (Figure 4C), knees reaching 0% of survival before 1000 days could be considered as high risk (knees #4 and 5), those between1000 to 3000 days as medium risk (knee #3), and more than 3000 days as low risk (knees #1 and 2).

### *Time-dependent AUC of the developed model*

As TKR is a time-to-event outcome in which the status changes over time, we further looked at the time-dependent AUC. Figure S4 shows an average AUC of 0.86. Of note the average AUC was 0.87 when considering nine features (data are not shown). The highest performance (up to 0.99) is achieved at the beginning of events from day 357 till day 450, after it slightly fluctuates until 1700 days, then remains stable until almost 2400 days, and slightly decreases (maximum 2%) until the end of study (3320 days [108 months]). These data demonstrated that the prediction model is effective in long-term prediction of TKR until 3320 days and very effective in predicting TKR until 2400 days. Therefore, it is possible to predict the TKR event using the three selected features with high accuracy and virtually stable AUC score for a long-term period at each time point up to 3320 days after enrollment.

## DISCUSSION

In this study, we first considered the best-known risk factors of OA for identifying the most important ones leading to a TKR event. Data revealed that although age, gender, BMI, WOMAC pain, and KL grade were all significant, the KL grade 4 has the highest impact on TKR survival (low survival probability). The importance of KL was further confirmed by the feature selection using ML methodology. This is not surprising, as the importance of KL grade in the prediction of knee OA severity has been known for a long time (19, 27). Although these findings provide insightful information, they did not suffice to build a powerful survival prediction model as there could be unfamiliar features impacting the risk of a TKR event. We then, for the first time, further used 7589 knees and integrated at the same time 1107 features including standard and uncommon ones. The ten most influential features were identified and nine were used to develop a survival prediction model. Further analysis revealed that three of these features were the most influential and included BMLs in the medial condyle, KL, and knee symptoms (sometimes swelling, last seven days). The Lasso's Cox methodology was used for feature selection as it can perfectly deal with multicollinearity (28) issues occurring with OA features and particularly those from MRIs and X-rays.

Of the selected features, the most important included X-rays and also BMLs in the medial condyle. The BML finding is not unexpected and reaffirms this altered structure prediction of the occurrence of TKR (29-34). Moreover, BMLs in the medial condyle as an indication of the likelihood of a TKR event is well in line with findings showing that this is the area in knee OA where both BMLs and cartilage degeneration are the most frequently affected (35-37). There have been medications that were shown to prevent or reduce the severity of BMLs (38, 39), but to the best of our knowledge, there is only one study showing that a bone remodelling therapeutic agent, bisphosphonates, was associated with about 25% lower risk of TKR (40).

From the selected features, four (two X-rays, BMLs and HA injection) demonstrated a significant impact on survival analysis by dropping the probability to 75% or less at the end of study, reaffirming the importance of X-ray and MRI features in survival prediction of TKR. For HA, our finding corroborates with recent studies (41, 42) in which injection in OA knees is highly associated with a significant delay in TKR (41). However, caution should be taken in interpreting the role of prior (to TKR) HA injection, as patients who do not want to undergo TKR may use HA as a substitute. This does not apply only to HA, but to other alternative measures taken by a given OA patient for whom surgery is not an option for personal or medical reasons.

As conventional survival models such as Kaplan-Meier analysis are not designed to predict an outcome, we considered ML-based survival models. By comparing seven ML models, data revealed that DeepSurv is the most appropriate. With the DeepSurv method and after further analyses, we were able to reduce the number of features to three without compromising the accuracy of the model. By using them, we further estimated for a given knee the time to TKR event/RUL. Of note, the proposed methodology could also be applied in another articulation, *e.g.* the hip.

As in all studies, there are limitations. First, although we used all the possible features (1107) from one of the most complete databases for OA subjects (OAI) including standard and uncommon ones, other unanticipated features could putatively also influence the TKR survival time. Second, our model was developed using a cohort in which subjects are at a mild-moderate stage of the disease. To ascertain the generalizability of our prediction model, a validation study should be performed with another cohort, preferably from clinical trial OA patients or electronic health research, which will represent more subjects routinely seen by a physician. After validation, the next step of this work will be to develop a predictive tool which could be used to guide clinical decision-making. Finally, one may say that our model requiring an MRI feature to predict time to TKR is not customary in clinical practice. Although MRI may not be commonly used by first-line physicians, it is becoming an increasingly routine part of the investigation of knee OA by subspecialists such as orthopedic surgeons and rheumatologists. Moreover, the use of MRI for the investigation of knee OA is more accessible as availability improves and the cost of the procedure becomes less expensive.

**CONCLUSION**

In this study, we showed that with the use of the OAI cohort it is possible to predict with a high degree of certainty when a TKR event would happen for a given OA knee, and who will progress fast toward this event. To the best of our knowledge, this is the first study in which a survival prediction model for a TKR event was built for OA knees by using ML methods, applying a survival-based feature selection method, and considering a very large number of features. Another important contribution of this work is the development of a prediction model that estimates the time of the risk of a TKR event for a given knee. As at present the time estimate to TKR is arbitrary for clinicians, this developed survival prediction model built with the OAI cohort could, in the future, better guide clinicians to the best therapeutic strategy to improve the survival of a given knee.

**Ethics approval and consent to participate**

Ethics approval was obtained by each OAI clinical site (University of Maryland Baltimore—Institutional Review Board, Ohio State University's Biomedical Sciences Institutional Review Board, University of Pittsburgh Institutional Review Board, and Memorial Hospital of Rhode Island Institutional Review Board) and the OAI coordinating center (Committee on Human Research at University of California, San Francisco, CA, USA). All patients provided written informed consent for participation in the OAI. The Institutional Ethics Committee Board of the University of Montreal Hospital Research Centre approved the study.

**Data availability statement**

All data used in this study is publicly available from the Osteoarthritis Initiative (OAI) cohort (https://data-archive.nimh.nih.gov/oai/). Additional data may be obtained from the corresponding author upon reasonable request, as long as the request is evaluated as scientifically relevant and pertinent.

**Acknowledgements**

The authors would like to thank the OAI participants and Coordinating Center for their work in generating the clinical and radiological data of the OAI cohort and for making them publicly available. None of the authors are part of the OAI investigator team. The authors are grateful to ArthroLab Inc. for having provided the MRI data, and Santa Fiori for preparing the manuscript.

**REFERENCES**


1. Driban JB, Eaton CB, Lo GH, Ward RJ, Lu B, McAlindon TE. Association of knee injuries with accelerated knee osteoarthritis progression: data from the Osteoarthritis Initiative. Arthritis Care Res (Hoboken). 2014;66(11):1673-9.
2. Driban JB, Stout AC, Lo GH, Eaton CB, Price LL, Lu B, et al. Best performing definition of accelerated knee osteoarthritis: data from the Osteoarthritis Initiative. Ther Adv Musculoskelet Dis. 2016;8(5):165-71.
3. Davis JE, Liu SH, Lapane K, Harkey MS, Price LL, Lu B, et al. Adults with incident accelerated knee osteoarthritis are more likely to receive a knee replacement: data from the Osteoarthritis Initiative. Clin Rheumatol. 2018;37(4):1115-8.
4. Lu J, Cowperthwaite MC, Burnett MG, Shpak M. Molecular Predictors of Long-Term Survival in Glioblastoma Multiforme Patients. PLoS One. 2016;11(4):e0154313.
5. Yousefi S, Amrollahi F, Amgad M, Dong C, Lewis JE, Song C, et al. Predicting clinical outcomes from large scale cancer genomic profiles with deep survival models. Sci Rep. 2017;7(1):11707.
6. Zhu B, Song N, Shen R, Arora A, Machiela MJ, Song L, et al. Integrating Clinical and Multiple Omics Data for Prognostic Assessment across Human Cancers. Sci Rep. 2017;7(1):16954.
7. Kim DW, Lee S, Kwon S, Nam W, Cha IH, Kim HJ. Deep learning-based survival prediction of oral cancer patients. Sci Rep. 2019;9(1):6994.
8. Cox DR. Regression models and life tables (with discussion). J R STAT SOC B. 1972;34:187–220.
9. Hao J, Kim Y, Mallavarapu T, Oh JH, Kang M. Interpretable deep neural network for cancer survival analysis by integrating genomic and clinical data. BMC Med Genomics. 2019;12(Suppl 10):189.
10. Katzman JL, Shaham U, Cloninger A, Bates J, Jiang T, Kluger Y. DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network. BMC Med Res Methodol. 2018;18(1):24.

11. Heisinger S, Hitzl W, Hobusch GM, Windhager R, Cotofana S. Predicting Total Knee Replacement from Symptomology and Radiographic Structural Change Using Artificial Neural Networks-Data from the Osteoarthritis Initiative (OAI). J Clin Med. 2020;9(5).
12. Leung K, Zhang B, Tan J, Shen Y, Geras KJ, Babb JS, et al. Prediction of Total Knee Replacement and Diagnosis of Osteoarthritis by Using Deep Learning on Knee Radiographs: Data from the Osteoarthritis Initiative. Radiology. 2020;296(3):584-93.
13. Evans JT, Walker RW, Evans JP, Blom AW, Sayers A, Whitehouse MR. How long does a knee replacement last? A systematic review and meta-analysis of case series and national registry reports with more than 15 years of follow-up. Lancet. 2019;393(10172):655-63.
14. Hunt LP, Ben-Shlomo Y, Whitehouse MR, Porter ML, Blom AW. The Main Cause of Death Following Primary Total Hip and Knee Replacement for Osteoarthritis: A Cohort Study of 26,766 Deaths Following 332,734 Hip Replacements and 29,802 Deaths Following 384,291 Knee Replacements. J Bone Joint Surg Am. 2017;99(7):565-75.
15. Puliero B, Favreau H, Eichler D, Adam P, Bonnomet F, Ehlinger M. Total knee arthroplasty in patients with varus deformities greater than ten degrees: survival analysis at a mean ten year follow-up. Int Orthop. 2019;43(2):333-41.
16. Jamshidi A, Leclercq M, Labbe A, Pelletier JP, Abram F, Droit A, et al. Identification of the most important features of knee osteoarthritis progressors using machine learning methods Ther Adv Musculoskel Dis. 2020;12:1-12.
17. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. J Stat Softw. 2010;33(1):1-22.
18. Tibshirani R. Regression Shrinkage and Selection via the Lasso. J R STAT SOC B. 1996;58(1):267–88.
19. Kohn MD, Sassoon AA, Fernando ND. Classifications in Brief: Kellgren-Lawrence Classification of Osteoarthritis. Clin Orthop Relat Res. 2016;474(8):1886-93.
20. Bellamy N, Buchanan WW. A preliminary evaluation of the dimensionality and clinical importance of pain and disability in osteoarthritis of the hip and knee. Clin Rheumatol. 1986;5(2):231-41.
21. Fotso S. PySurvival: Open source package for Survival Analysis modeling. 2019.

22. Chun-Nam Yu RG, Hsiu-Chin Lin, Vickie Baracos. Learning Patient-Specific Cancer Survival Distributions as a Sequence of Dependent Regressors. In: Shawe-Taylor J, editor. Advances in Neural Information Processing Systems 24 (NIPS 2011); 2011; 2011. p. 1845–53.
23. Fotso S. Deep neural networks for survival analysis based on a multi-task framework. arXiv preprints arXiv:1801.05512; 2018.
24. Hemant Ishwaran UBK, Eugene H. Blackstone, and Michael S. Lauer. Random survival forests. ANN APPL STAT. 2008;2(3):841-60.
25. Sebastian Pölsterl aN, Amin Katouzian. Fast Training of Support Vector Machines for Survival Analysis. Joint European Conference on Machine Learning and Knowledge Discovery in Databases; 2015: Springer, Cham; 2015. p. 243-59.
26. Pölsterl S, Navab, N., Katouzian, A. An Efficient Training Algorithm for Kernel Survival Support Vector Machines. The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases Italy; 2016.
27. Guermazi A, Hayashi D, Roemer F, Felson DT, Wang K, Lynch J, et al. Severe radiographic knee osteoarthritis--does Kellgren and Lawrence grade 4 represent end stage disease?--the MOST study. Osteoarthritis Cartilage. 2015;23(9):1499-505.
28. Jamshidi A, Pelletier JP, Martel-Pelletier J. Machine-learning-based patient-specific prediction models for knee osteoarthritis. Nat Rev Rheumatol. 2019;15(1):49-60.
29. Pelletier JP, Cooper C, Peterfy C, Reginster JY, Brandi ML, Bruyere O, et al. What is the predictive value of MRI for the occurrence of knee replacement surgery in knee osteoarthritis? Ann Rheum Dis. 2013;72(10):1594-604.
30. Raynauld JP, Martel-Pelletier J, Dorais M, Haraoui B, Choquette D, Abram F, et al. Total Knee Replacement as a Knee Osteoarthritis Outcome: Predictors Derived from a 4-Year Long-Term Observation following a Randomized Clinical Trial Using Chondroitin Sulfate. Cartilage. 2013;4(3):219-26.
31. Raynauld JP, Martel-Pelletier J, Haraoui B, Choquette D, Dorais M, Wildi LM, et al. Risk factors predictive of joint replacement in a 2-year multicentre clinical trial in knee osteoarthritis using MRI: results from over 6 years of observation. Ann Rheum Dis. 2011;70(8):1382-8.

32. Dore D, Quinn S, Ding C, Winzenberg T, Zhai G, Cicuttini F, et al. Natural history and clinical significance of MRI-detected bone marrow lesions at the knee: a prospective study in community dwelling older adults. Arthritis Res Ther. 2010;12(6):R223.
33. Tanamas SK, Wluka AE, Pelletier JP, Pelletier JM, Abram F, Berry PA, et al. Bone marrow lesions in people with knee osteoarthritis predict progression of disease and joint replacement: a longitudinal study. Rheumatology (Oxford). 2010;49(12):2413-9.
34. Scher C, Craig J, Nelson F. Bone marrow edema in the knee in osteoarthrosis and association with total knee arthroplasty within a three-year follow-up. Skeletal Radiol. 2008;37(7):609-17.
35. Singh V, Oliashirazi A, Tan T, Fayyad A, Shahi A. Clinical and Pathophysiologic Significance of MRI Identified Bone Marrow Lesions Associated with Knee Osteoarthritis. Arch Bone Jt Surg. 2019;7(3):211-9.
36. Mattap SM, Aitken D, Wills K, Laslett L, Ding C, Pelletier JP, et al. How Do MRI-Detected Subchondral Bone Marrow Lesions (BMLs) on Two Different MRI Sequences Correlate with Clinically Important Outcomes? Calcif Tissue Int. 2018;103(2):131-43.
37. Teichtahl AJ, Cicuttini FM, Abram F, Wang Y, Pelletier JP, Dodin P, et al. Meniscal extrusion and bone marrow lesions are associated with incident and progressive knee osteoarthritis. Osteoarthritis Cartilage . 2017;25(7):1076-83.
38. Laslett LL, Dore DA, Quinn SJ, Boon P, Ryan E, Winzenberg TM, et al. Zoledronic acid reduces knee pain and bone marrow lesions over 1 year: a randomised controlled trial. Ann Rheum Dis. 2012;71(8):1322-8.
39. Martel-Pelletier J, Roubille C, Abram F, Hochberg MC, Dorais M, Delorme P, et al. First-line analysis of the effects of treatment on progression of structural changes in knee osteoarthritis over 24 months: data from the osteoarthritis initiative progression cohort. Ann Rheum Dis. 2015;74(3):547-56.
40. Neogi T, Li S, Peloquin C, Misra D, Zhang Y. Effect of bisphosphonates on knee replacement surgery. Ann Rheum Dis. 2018;77(1):92-7.
41. Altman R, Lim S, Steen RG, Dasa V. Hyaluronic Acid Injections Are Associated with Delay of Total Knee Replacement Surgery in Patients with Knee Osteoarthritis: Evidence from a Large U.S. Health Claims Database. PLoS One. 2015;10(12):e0145776.

42. Delbarre A, Amor B, Bardoulat I, Tetafort A, Pelletier-Fleury N. Do intra-articular hyaluronic acid injections delay total knee replacement in patients with osteoarthritis - A Cox model analysis. PLoS One. 2017;12(11):e0187227.
43. Guermazi A, Hunter DJ, Li L, Benichou O, Eckstein F, Kwoh CK, et al. Different thresholds for detecting osteophytes and joint space narrowing exist between the site investigators and the centralized reader in a multicenter knee osteoarthritis study--data from the Osteoarthritis Initiative. Skeletal Radiol. 2012;41(2):179-86.
44. Dodin P, Abram F, Pelletier JP, Martel-Pelletier J. A fully automated system for quantification of knee bone marrow lesions using MRI and the osteoarthritis initiative cohort J Biomed Graph Comput. 2013;3(1):51-65.
45. Tormalehto S, Mononen ME, Aarnio E, Arokoski JPA, Korhonen RK, Martikainen J. Health-related quality of life in relation to symptomatic and radiographic definitions of knee osteoarthritis: data from Osteoarthritis Initiative (OAI) 4-year follow-up study. Health Qual Life Outcomes. 2018;16(1):154.

## FIGURE LEGENDS

**Figure 1.** Flow chart of the knees. CVL, cartilage volume loss in the medial tibial plateau; n, number of individuals; MRI, magnetic resonance imaging.

**Figure 2.** Probability of survival for the most important clinical/demographic features before a total knee replacement event. Probability of survival associated with race, bone mass index (BMI), Western Ontario and McMaster Universities Arthritis Index (WOMAC) pain score and Kellgren-Lawrence (KL) grades. In the race graph, Cauc. indicates Caucasian, Black/African A., the A refers to American, and O, Non-white, other non-white. Definitions of the BMIs were as follows in the OAI cohort: healthy, BMI≤24.9 kg/m$^2$; morbidly obese, BMI>40 kg/m$^2$; obese, BMI≤39.9 kg/m$^2$; overweight, BMI≤29.9 kg/m$^2$; and underweight, BMI≤18.5 kg/m$^2$. WOMAC pain was scored as a continued numerical feature (score 0-20 (20), and categorized in four different groups.

**Figure 3.** Probability of survival for the most important selected features before a total knee replacement event. Composite OA grade, composite quasi-Kellgren-Lawrence (KL) grade based on a 0–4 scale (a scale of 4 indicates severe knee baseline X-ray); the quasi-KL OA scale is comprised of evidence of joint space narrowing (JSN) and osteophytes (Ost). Osteophytes and JSN are scored as 0 to 2 for each, corresponding to severe disease (43). BML, bone marrow lesions (data are expressed as BML size in regions/compartments and corresponds to the percentage of BMLs in this region, for example >0.2 corresponds to 20% (37, 44)). Hyaluronic acid (HA) injection, either knee received one HA injection treatment, within the past six months.

**Figure 4.** Predictions of total knee replacement survival before a total knee replacement event. **A)** Comparing actual vs. predicted number of knees. Lower and upper confidence intervals are for the actual curve; both represent the lower and upper 95% confidence interval (CI). **B**) Baseline values of the three selected features. Bone marrow lesions (BML) in the medial condyle data are expressed as BML size in regions and corresponds to the percentage of BMLs in this region (37, 44); KL according to the knee grade based on a 0-4 scale (a scale of 4 indicates severe knee baseline X-ray; knee symptoms: swelling, last seven days according to OAI nomenclature (range 0-4: 0=never swelling and 4=always swelling). **C)** Knee-specific survival curves.

**Table 1. Top ten features selected using Lasso's Cox method for total knee replacement (TKR) survival prediction.**

| Priority | Relative importance | Label | Category |
|---|---|---|---|
| 1 | 1 | Severe knee baseline X-ray: composite OA grade 4 (quasi KL grade [score 0-4]) | X-ray |
| **2** | **0.82** | **BML in the medial condyle (>0.2, data are expressed as BML size in regions).** | **MRI** |
| 3 | 0.62 | Knee baseline X-ray: osteophytes [score 0-2] and joint space narrowing [score 0-2], both with a score of 2 | X-ray |
| 4 | 0.48 | Unable to perform 400-meter walk (excluded for heart rate) | Performance measure |
| 5 | 0.47 | Baseline radiographic knee OA (defined as a KL grade of 2 or greater in left/right knee or both (45) | X-ray |
| 6 | 0.47 | Charlson comorbidity: had stroke, cerebrovascular accident, blood clot or bleeding in brain, or transient ischemic attack | Medical history |
| **7** | **0.37** | **KL grade 4** | **X-ray** |
| 8 | 0.29 | Either knee, received one hyaluronic acid injection treatment, within the past 6 months | Medication |

| | | | |
|---|---|---|---|
| **9** | **0.25** | **Knee symptoms: sometimes swelling, last 7 days** | **Knee symptom** |
| 10 | 0.21 | Baseline symptomatic knee OA (defined as radiographic OA and frequent knee pain in the same knee (45) | X-ray |

Priority indicates the importance of the selected features. Relative importance is calculated as the absolute importance of a variable divided by the absolute importance of the most important variable. The bone marrow lesion (BML) value is between 0 - no BML and 1 - the BML lesion extends into the entire bone region (37, 44) and 0.2 corresponds to 20%. Knee symptoms were scored according to the Osteoarthritis Initiative (OAI) nomenclature (ranges 0-4: 0=never swelling and 4=always swelling), and the Kellgren-Lawrence (KL) grade was based on a 0-4 scale (a scale of 4 indicates severe knee baseline X-ray). For further details about the scoring, refer to Figures 2, 3, and S3.

Data in bold indicate the final three features used for the prediction of TKR survival. OA, osteoarthritis; MRI, magnetic resonance imaging; X-ray, radiographic measurement.

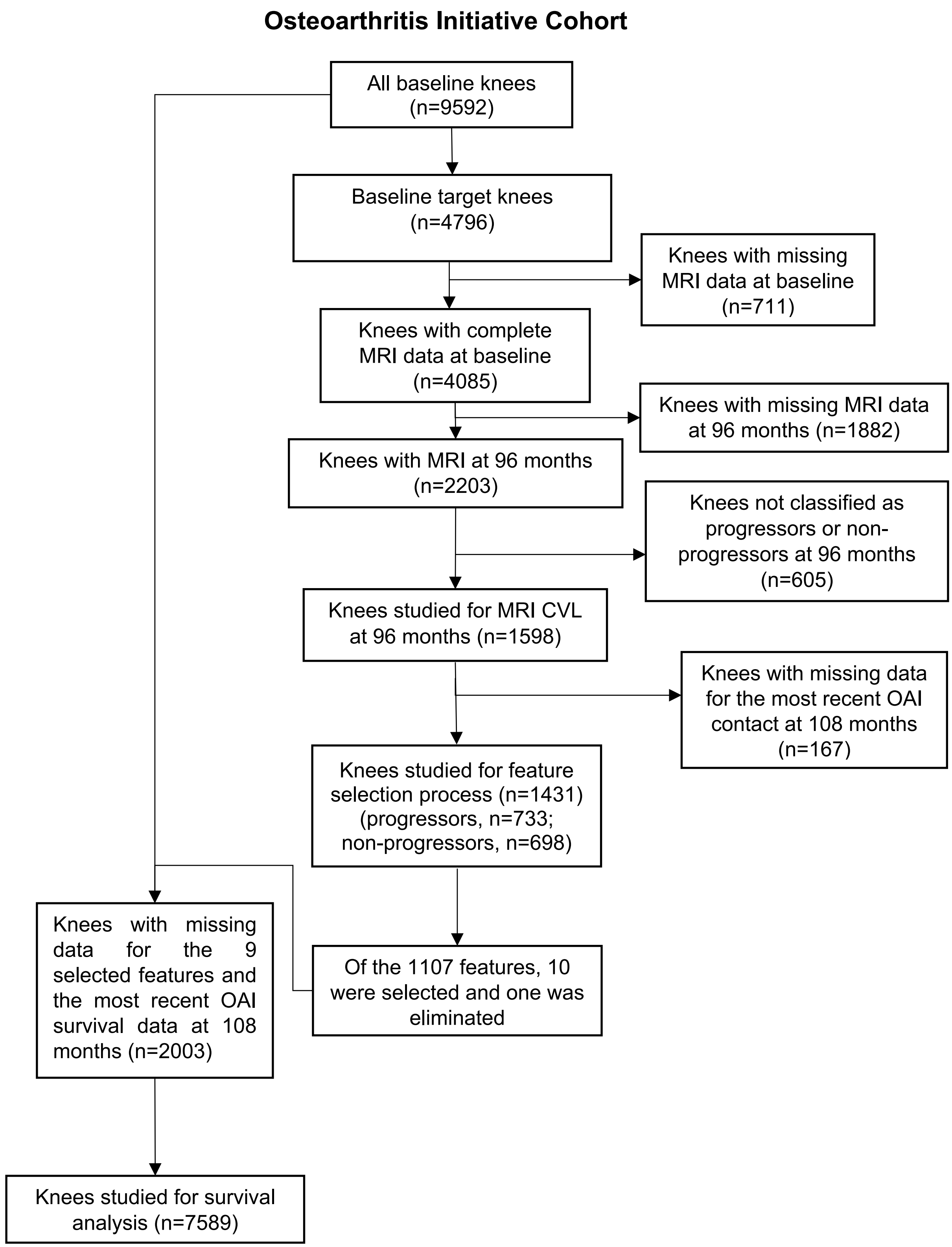





**Figure 1. Flow chart of the knees.**

**Figure 2. Probability of survival for the most important clinical/demographic features before a total knee replacement event.**

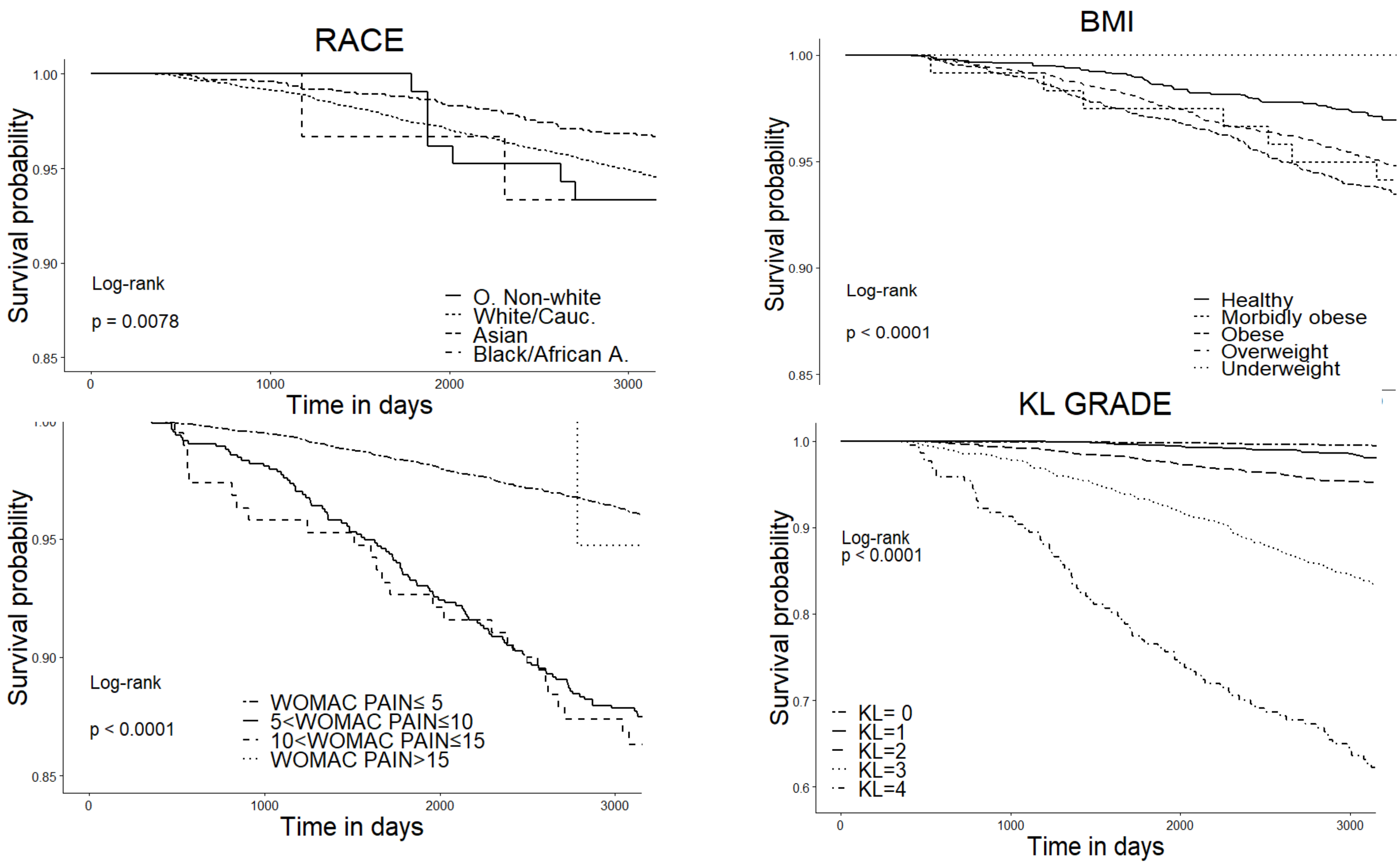


**Figure 3. Probability of survival for the most important selected features before a total knee replacement event.**

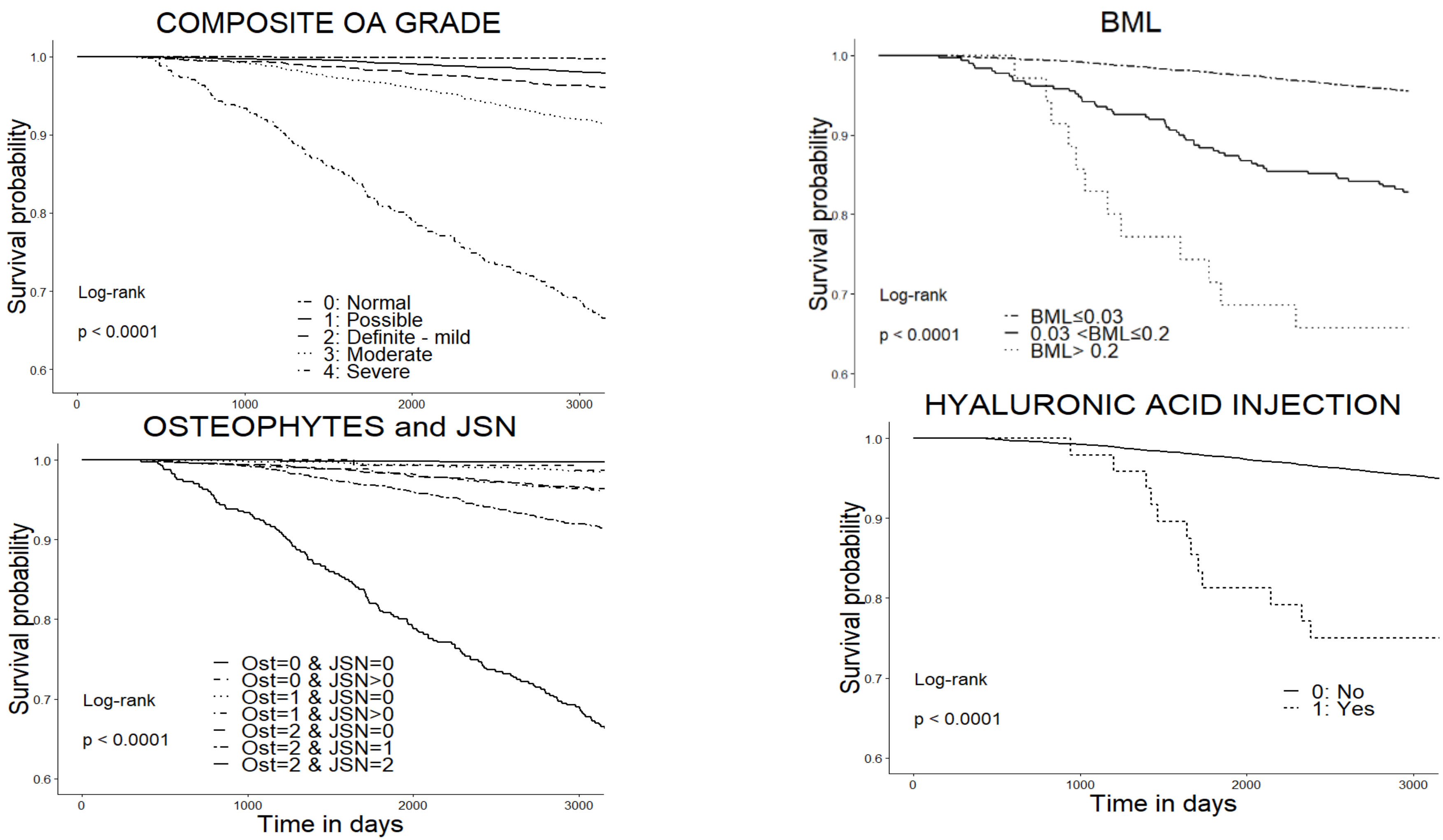
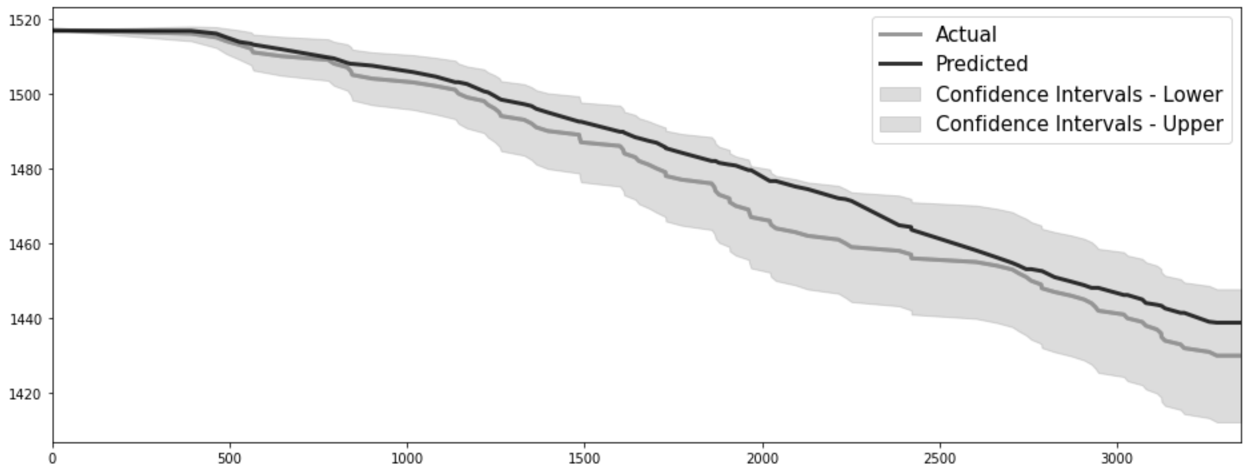


**Figure 4. Predictions of total knee replacement survival before a total knee replacement event.**

**A.**




**B.**

| Knees | BML in the medial condyle | KL grade | Knee symptoms: sometimes swelling, last seven days |
|---|---|---|---|
| 1 | 0 | 2 | 0 |
| 2 | 0.011 | 2 | 2 |
| 3 | 0.043 | 3 | 1 |
| 4 | 0.214 | 4 | 3 |
| 5 | 0.311 | 4 | 4 |

**C.**