Modeling and Predicting Osteoarthritis Progression: Data from the Osteoarthritis Initiative

Eni Halilaj, PhD[*] **[Postdoctoral Fellow]**,
Department of Bioengineering, Stanford University

Ya Le, PhD[*] **[Doctoral Candidate]**,
Department of Statistics, Stanford University

Jennifer L. Hicks, PhD [Senior Research Engineer],
Department of Bioengineering, Stanford University

Trevor J. Hastie, PhD [John A. Overdeck Professor], and
Department of Statistics, Stanford University

Scott L. Delp, PhD [James H. Clark Professor]
Departments of Bioengineering, Mechanical Engineering, and Orthopaedic Surgery, Stanford University

Abstract

Objective: The goal of this study was to model the longitudinal progression of knee osteoarthritis (OA) and build a prognostic tool that uses data collected in one year to predict disease progression over eight years.

Design: To model OA progression, we used a mixed-effects mixture model and eight-year data from the Osteoarthritis Initiative—specifically, joint space width measurements from X-rays and pain scores from the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) questionnaire. We included 1243 subjects who at enrollment were classified as being at high risk of developing OA based on age, body mass index, and medical and occupational histories. After clustering subjects based on radiographic and pain progression, we used clinical variables collected within the first year to build LASSO regression models for predicting the probabilities of belonging to each cluster. Areas under the receiver operating characteristic curve (AUC) represent predictive performance on held-out data.

Please direct correspondence to: Eni Halilaj, PhD, Clark Center Room S-331, 318 Campus Drive, Stanford University, Stanford, CA 94305-5450, Phone: (650) 725-4009/ Fax: (650) 736-0801, ehalilaj@stanford.edu.
[*]equal contribution

Author Manuscript

Results: Based on joint space narrowing, subjects clustered as progressing or non-progressing. Based on pain scores, they clustered as stable, improving, or worsening. Radiographic progression could be predicted with high accuracy (AUC = .86) using data from two visits spanning one year, whereas pain progression could be predicted with high accuracy (AUC = .95) using data from a single visit. Joint space narrowing and pain progression were not associated.

Conclusion: Statistical models for characterizing and predicting OA progression promise to improve clinical trial design and OA prevention efforts in the future.

Keywords

knee osteoarthritis; disease progression; functional data clustering; predictive modeling

1. INTRODUCTION

The current standard of care for knee osteoarthritis (OA) is to manage symptoms with non-steroidal anti-inflammatory drugs, steroidal injections, and physical therapy until these options fail to provide comfort, at which point joint replacement surgery is recommended. The development of robust preventative measures and disease-modifying treatments is partly limited by the current inability to accurately predict OA progression. Although OA is classically described as slowly progressing, recent evidence indicates that in addition to its diverse and multifactorial etiology, the disease is also heterogeneous in how it progresses across the affected population[1]. This heterogeneity presents a challenge in the design of clinical trials due to the likely inclusion of patients with different disease progression trajectories in the treatment and control groups. The ability to characterize and predict OA progression could improve the design and efficacy of studies investigating treatment and prevention strategies. Further, it could facilitate the development of novel treatments, which may be effective for fast, but not slow progressors, or vice versa.

Data from population-based studies such as the Osteoarthritis Initiative (OAI) and the Multicenter Osteoarthritis Study (MOST) suggest that pain trajectories span a spectrum of progression rates, instead of forming clusters. In this context, model-based clustering approaches, which formally model the data as a mixture of distributions, are more appropriate than non-parametric or heuristic approaches, which assume that the data come from distinct subpopulations. Latent Class Growth Analysis (LCGA) has emerged as the method of choice for clustering pain trajectories associated with OA[2–9]. LCGA is a simplified mixture model that represents each cluster by the mean trajectory, modeling fixed effects (i.e., growth), but not random effects (i.e., inter-subject variability within clusters). The assumption that the variance and covariance of trajectories within a cluster are zero results in advantages such as fast convergence, but also several shortcomings, with the most notable one being the inability to overcome the challenge of missing data. A mixed-effects approach, on the other hand, works particularly well with missing data because instead of making assumptions of independence across different time points within a curve, it exploits patient similarity to overcome the challenge of missing data[10–13]. In large-scale studies, missing data is a common problem that affects the reliability of identified clusters.

Rigorously validated models that follow recommended practices in statistical learning[14], however, have not been commonly reported in the OA literature. It remains to be determined if knee and hip pain progression clusters identified by previous studies generalize well to new data. One study that compared findings from two different cohorts found that not all the clusters identified in one cohort could be reproduced in the other[3]. Similarly, validated prognostic tools for predicting disease progression early have not been commonly reported. While odds ratios associated with specific subject characteristics advance our general understanding of risk factors and inform prevention efforts, they carry little prognostic utility. One of the challenges of developing predictive models is overfitting them to the available data. Ideally, the data must be split into training and validation sets to select the best model and a test set to evaluate its performance[14]. Given a large set of candidate predictive variables, for example, selection of variables to include in the predictive model must be carried out using training and validation data, while model performance must be evaluated using test data. In the case of k-fold cross-validation, two nested rounds of cross-validation would be necessary. Such practices have not been followed by previous studies aiming to build prognostic tools for OA incidence.

The purpose of this study was to identify different clusters of knee OA progression using a mixed-effects mixture model. To characterize OA progression, we used eight-year data from the OAI—specifically, self-reported knee pain and radiographic assessments of joint space narrowing. Additionally, we sought to build cross-validated models that use short-term data to predict long-term disease progression. These tools have been made available in the open-source statistical programing language, R.

2. METHODS

2.1 Subjects

The OAI is a longitudinal observational study on the natural progression of knee OA. Men and women between the ages of 45 and 79 were enrolled at four centers across the United States (Columbus, OH; Baltimore, MD; Pawtucket, RI; Pittsburg, PA) and assessed annually. The collected information includes clinical evaluations, radiological images, nutritional information, physical activity monitoring, and biospecimen samples. Subjects were enrolled as part of a progression (n = 1389), incidence (n = 3285), or control subcohort (n = 122). In this study we included only subjects from the incidence cohort, defined as individuals who at the baseline visit were at high risk of developing OA over the course of the study. The definition of high risk was determined by OAI investigators and included histories of knee pain, aching, or stiffness in a native knee, knee replacement in an ipsilateral knee, family history of OA, high body mass index (BMI), previous knee injury, Herberden's nodes in the hands, history of frequent knee bending, and age—subjects above 70 years were included even in the absence of concomitant risk factors. Subjects with both symptoms and radiographic OA (Kellgren Lawrence grade > 1) were ineligible for enrollment in the incidence subcohort. For subjects who had a knee replacement surgery during the course of the study, we excluded data after the date of the surgery.

2.2 Outcomes

To model OA progression, we used joint space width measurements from X-rays and pain scores from the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) questionnaire, since they are relevant outcomes with respect to structural and symptomatic disease status. OAI participants completed WOMAC questionnaires yearly and X-rays at the baseline and year 1, 2 3, 4, and 6 follow-up visits. We included both the right and left knees in the analysis since OA progression is likely driven by systemic factors that may affect both knees in a similar manner, except in the case of post-traumatic OA or a unilateral deformity. We excluded subjects with less than three data points and focused on the medial compartment because it is the most common site of knee OA initiation and progression. The minimum joint space width in the medial compartment of the knee had been extracted semi-automatically from standardized fixed-flexion X-rays[15,16] and made available by OAI investigators. Briefly, the automated software delineates the edges of the femoral condyle and the tibial plateau, allowing manual correction from the user. The bone-to-bone distance is then determined at fixed intervals across the joint space, and the minimum distance in the medial compartment is used to represent the medial joint space width. The WOMAC questionnaire is the most widely used instrument for assessing knee and hip OA-related symptoms and disability. The pain subscale is based on pain levels during walking, stair-climbing, lying in bed, and standing—each ranging from 0 to 4, for a maximum score of 20. We removed observations that were more than three standard deviations away from the mean for both the pain and joint space data. To ensure that the clustering procedure was representative of disease evolution and independent of baseline status, we expressed joint space narrowing and pain progression as the change from the baseline visit. Given the inherent noise in the data, we estimated pain and joint space width in the baseline visit by fitting linear regression models to the longitudinal data and using the intercept, rather than the observed value (Supplemental Material 1).

2.3 Characterizing Disease Progression

To characterize OA progression, we used a mixed-effects mixture model approach that is designed to work well for sparsely sampled functional data. Thorough descriptions of the clustering methodology[10–12], as well as relevant applications (e.g., clustering of spinal bone mineral density increase in teenagers[13]), have been published earlier and are summarized in the supplementary section (Supplemental Material 2). In addition to model parameters, a final output is the set of posterior probabilities that a subject belongs to any given cluster. To determine the optimal number of clusters, we used the Silhouette approach[17], which is based on intra-cluster cohesion (i.e., how similar curves within a cluster are to each other) and inter-cluster separability (i.e., how different curves within a cluster are from curves in other clusters). The ideal number of clusters is one that maximizes both of these measures, yielding the highest Silhouette.

2.4 Predicting Disease Progression

After characterizing clusters of joint space narrowing and pain progression, we developed models to predict the probability of belonging to a cluster using patient characteristics collected at the baseline visit, including knee symptoms, medication usage, family history of

OA, general health status and comorbidities, nutritional and mental health information, walking ability, upper leg strength test results, X-ray assessments of knee alignment and evidence of OA (Supplemental Material 3). Given the large number of predictors, we used least absolute shrinkage and selection (LASSO) regression[18]. The LASSO is a shrinkage and variable selection method that mitigates over-fitting when the feature space is large, producing more compact, interpretable, and accurate models than regular regression. We trained predictive models using (1) baseline variables, (2) baseline and year 1 follow-up variables, and (3) baseline, year 1, and year 2 follow-up variables. The rationale for building these three types of models was to determine the time that is necessary to monitor patients before being able to predict long-term disease progression. All the predictive variables were scaled to have zero mean and unit variance.

2.5 Validation and Testing

To ensure that the identified disease progression clusters and predictive models are generalizable to new data, we used k-fold cross-validation both for model selection and performance evaluation, splitting the data into training, validation, and test sets. First, using 10-fold cross-validation, we left aside 10% of the data for model evaluation and used 90% of the data to select the models—this included selecting both the disease progression and predictive models. We repeated this procedure 10 times, each time leaving aside one fold of the data for testing. Given that both the clustering of disease progression and predictive modeling procedures involve hyper-parameters (i.e., number of clusters and the shrinkage parameter for LASSO), we included a nested 10-fold cross-validation step for hyper-parameter tuning, further splitting the data that were allocated for model selection into training and validation sets. For the test data, the true posterior probabilities of belonging to each cluster were estimated using the Bayes rule and cluster parameters from the trained models. We then compared these true posterior probabilities to probabilities predicted by LASSO regression. Here we report areas under the receiving operating characteristic curves (AUC).

3. RESULTS

3.1 Subjects

The number of subjects who satisfied our inclusion criteria was 1243. The mean (± SD) joint space width at the baseline visit was 4.3 ± 1.1 mm, which is lower than that of the normal subcohort (4.8 ± 0.7 mm) and higher than that of the progression subcohort (4.0 ± 1.5 mm) of the OAI, but not different from the excluded sample of the incidence subcohort (Table 1). Average pain on both knees was low ($< 1.7 \pm 2.5$), but significantly higher than zero and not significantly different from the excluded sample of the incidence subcohort (Table 1).

3.2 Characterizing Disease Progression

Analysis of the X-ray data revealed that there are two major clusters of joint space narrowing: a non-progressing and a fast-progressing group (Figure 1). Patterns were similar in the right and left knees. When modeling measurements from both of the knees simultaneously, 71% of the subjects clustered into the non-progressing and 29% clustered into the fast-progressing cluster. On average, over the course of this 8-year study, fast

progressors lost nearly 60% of their baseline joint space width. The model estimated cluster membership in new subjects with high confidence. The mean posterior probabilities for the test data, not used to build the model, were .91 and .84 for the slow and fast progressing clusters respectively. Baseline age, BMI, joint space widths, and WOMAC scores were not significantly different between the two groups, but the non-progressing group had a higher fraction of women (Table 2). Both clusters had a similar proportion of missing data. In the fast-progressing cluster, 32.4% data were missing, whereas in the slow-progressing cluster, 33.2% of the data were missing.

Analysis of the WOMAC pain scores revealed that there are three major clusters of pain progression: 80% of the subjects had stable levels of pain, 14% had worsening pain, and 6% had improving pain. These patterns were similar in the right and left knees. The mean posterior probabilities for the test data were .94, .88, and .88 for the stable, worsening, and improving clusters, respectively. The improving group scored higher on the WOMAC pain and functional limitation scales at baseline, which is indicative of higher levels of pain and disability, whereas the worsening group had a narrower joint space width in the right knee (Table 3). The proportions of missing data were 4.5%, 5.4%, and 7.6% for the stable, worsening, and improving clusters, respectively. Agreement between pain progression and joint space narrowing was low (Figure 2).

3.3 Predicting Disease Progression

Joint space narrowing could not be predicted with high accuracy using baseline data alone (AUC < .6), but adding measurements from the year one and two follow-up visits increased the accuracy of these models (Figure 3). Models built separately for the right and left knee were moderately accurate (AUC < .75), whereas those built to predict progression in both knees simultaneously using joint space width measurements from baseline and year-one follow-up visits were highly accurate (AUC > .86; Figure 4A). Joint space width measurements at baseline and follow-up visits were the only variables selected consistently by LASSO. Other variables fell in the "nuanced features" category (i.e., they are selected by LASSO, but did not significantly affect predictive performance).

Pain progression could be predicted with high accuracy using data from the baseline visit alone (Figure 4B). Given that the clustering procedure identified three clusters, we used a multinomial LASSO model to predict the posterior probabilities associated with each of the three clusters. Both the right and left knee models differentiated worsening from stable subjects with moderate accuracy (AUC > .76) and worsening from improving subjects with high accuracy (AUC > .92). Modeling both knees simultaneously or adding predictive variables from follow-up visits did not increase the accuracy of these models significantly. Important predictive variables selected by LASSO included assessments of pain and function from the KOOS and WOMAC questionnaires, qualitative assessments of X-rays, and assessments of depression and nutrition from the Center for Epidemiologic Studies Depression Scale (CES-D) and Block Brief 2000 questionnaires (Table 4). Generally, worse baseline pain and functional limitation scores were predictive of improvement in the eight-year period. Moderate pain and functional limitation scores were predictive of worsening, whereas good scores were predictive of stability. Evidence of joint space narrowing,

difficulty sleeping, and frequency of certain foods, including dairy and meat, were predictive of worsening pain over time.

4. DISCUSSION

The goal of this study was to characterize different clusters of OA progression and build models to predict these clusters early. We focused on joint space narrowing and pain progression because they are the most widely used surrogates of structural and symptomatic disease status. Our findings, using data from 1243 subjects, indicate that joint space width measurements follow two clusters—progressing and non-progressing—whereas pain scores follow three clusters—stable, improving, and worsening. Eight-year pain progression could be predicted with high accuracy based on data collected in one visit, whereas joint space narrowing could be predicted with high accuracy using data collected in two visits spanning one year.

Our finding that pain progression follows several clusters, or trajectories, is consistent with previous findings; however, the number and shape of these trajectories are different[3,4]. These disparities may stem from the following differences in our approach. First, we modeled change in pain from the baseline visit, rather than absolute pain scores over time because we were interested in clustering subjects based on how they progress, rather than their status at baseline. When modeling the evolution of absolute scores, baseline values can bias clustering, especially in a dataset where inter-subject variability at baseline is, on average, much larger than intra-subject change over time. Second, we allowed each cluster to flexibly and independently model the shape of the underlying data, instead of constraining all the clusters *a priori* to assume one shape. Third, we used a formal approach, the Silhouette, cross-validated on held-out data, to select the number of clusters. Last, another source of incongruity may be the clustering method itself. We chose a mixed-effects mixture model because it is particularly advantageous for sparsely sampled data, using the covariance structure to ensure that curves with missing portions use information from similar curves. Here, we reported mean posterior probabilities associated with test data, not used in the model construction, because they are a better presentation of how well the model can generalize to new data.

Ultimately, the clinical utility of predictive models hinges on their ability to make predictions on new data. Thus, we attempted to develop cross-validated predictive models that can make accurate predictions on held-out data. Previous studies have reported several risk factors associated with OA incidence and progression, including occupational knee bending, BMI, and history of smoking, among others[19]. Although we included these as candidate variables, we did not find them to be strongly predictive of pain progression or joint space narrowing. This, however, does not indicate that factors such as smoking do not elevate the risk of OA, but other variables may be necessary to model the complexity of interacting or confounding factors. For example, individuals who do not have knee malalignment, may never develop OA even if they smoke, while others with high knee malalignment may develop OA despite never having smoked. Our models utilized data from questionnaires and functional tests. Radiologic image assessments and data from biospecimen assays should improve their performance in the future.

The asynchronous progression of pain and joint space narrowing is supported by previous findings that have related pain to structural changes in bone and other soft tissue within the joint capsule, but not cartilage[20–24]. While some studies report that pain and radiographic severity are generally associated[25,26], others have found that nearly 50% of individuals with moderate to severe radiographic evidence of OA are asymptomatic and 10% of individuals with moderate to severe knee pain have normal radiographs[27–29]. It has been suggested that due to the aneural nature of cartilage, pain is mostly related to structural changes in bone and other soft tissue within the joint capsule[30]. Cartilage can produce pro-inflammatory cytokines, but it is likely not the tissue that generates pain. Magnetic resonance imaging (MRI) studies have demonstrated that synovitis, synovial hypertrophy and large synovial effusions, bone marrow edema, subchondral bone sclerosis, and meniscal tears are associated with pain[20–23]. Similarly, early pre-radiographic changes in cartilage microstructure can be captured with T_2-weighted MRI, data that are available for this population, but have not yet been processed. Thus, imaging biomarkers have the potential to improve predictive models both for pain and radiographic progression.

A few characteristics and limitations of this study must be considered when interpreting the findings presented here and placing them in context with the current and future literature. First, this study is based on a US population, which although diverse, may not encompass all the varying combinations of genetic, demographic, and mechanical factors that increase predisposition to OA and accelerate its rate of progression thereafter. Validating the identified disease progression clusters in a different population, such as the Cohort Hip and Cohort Knee (CHECK) study carried out in the Netherlands[31], is a desirable next step for testing their generalizability to different populations. Second, we found that eight-year joint space narrowing cannot be accurately predicted using data from a single visit, and follow-up data from at least one visit improves model performance by 28%. Data from a second follow-up visit increases predictive performance by an additional 7%, but this incremental improvement should be weighted against longer monitoring time, which may be infeasible for clinical trials. Ultimately, predictive accuracy is a function of effort, increasing with higher quality data collected in the baseline visit or longer-term monitoring. Once processed data from biospecimen assays and MRI scans become available for this population, regularized models (e.g., LASSO models) have the potential to identify new biomarkers of disease progression from a large pool of candidate variables and, subsequently, improve the predictive performance of these models. In the absence of these high quality data from the baseline visit, monitoring subjects for at least one year may be needed to make accurate predictions. Third, given two highly correlated predictors, LASSO may select only one of them. Although this is accomplished at no expense on the predictive performance of a model, the variables selected by LASSO should not be interpreted as the only ones that are associated with the outcomes. Our goal was to develop the most accurate predictive model, rather than identify risk factors.

Studies investigating the effect of disease modifying treatments for osteoarthritis are currently limited since they combine subjects with different disease progression profiles into one group. The heterogeneity of OA progression is a confounding factor that may obscure the positive effect of a treatment, especially if the treated group contains a higher number of fast progressors compared to the untreated group. To make these studies more efficient and

informative, predictive models of disease progression are needed. Our work contributes toward achieving this goal in several ways. First, while most previous studies have modeled progression of symptoms, here we also modeled structural disease progression measured through joint space narrowing, which is the only clinical endpoint recognized by the Food and Drug Administration. Second, we presented a novel statistical modeling approach that utilizes subject similarity to overcome the challenge of missing data. Unlike previous approaches, our approach does not constrain all the clusters to take one shape (e.g., be linear, but have different slopes) and inter-subject variability within clusters to be zero. Fewer assumptions about the underlying structure of the data allow us to model disease progression more accurately. Third, we built prognostic models to predict disease progression early—a necessary step for improving the design of future clinical trials. A common practice in the field is to report odds ratios as a measure of the association between a subject characteristic and an outcome. While useful in understanding risk factors, odds ratios are not well-suited for prediction, as is evidenced by the data presented here. Subjects in the incidence subcohort of the OAI were all selected using a set of risk factors indicating they were at high risk of developing OA. However, as shown here, the majority of them do not progress. Fourth, to ensure that both the disease progression and predictive models are generalizable, we used thorough cross-validation. This is not common practice in the field, but it is a necessary step for building models that translate well to new data. Additionally, we have provided a set of open-source tools in the statistical package R to encourage the refinement of these models as more sensitive biomarkers become available to the community through the OAI repository. Ultimately, identification and early prediction of OA progression trajectories could boost OA prevention efforts by improving the efficacy of clinical trials and accelerating the discovery of new treatments.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

REFERENCES

1. Driban JB, Sitler MR, Barbe MF, Balasubramanian E. Is osteoarthritis a heterogeneous disease that can be stratified into subsets? Clin Rheumatol 2009;29(2):123 10.1007/s10067-009-1301-1. [PubMed: 19924499]

2. Losina E, Collins JE. Forecasting the future pain in hip OA: can we rely on pain trajectories? Osteoarthritis Cartilage 2016;24(5):765–767. 10.1016/j.joca.2016.01.989. [PubMed: 26854793]

3. Nicholls E, Thomas E, van der Windt DA, Croft PR, Peat G. Pain trajectory groups in persons with, or at high risk of, knee osteoarthritis: findings from the Knee Clinical Assessment Study and the Osteoarthritis Initiative. Osteoarthritis Cartilage 2014;22(12):2041–2050. 10.1016/j.joca. 2014.09.026. [PubMed: 25305072]

4. Collins JE, Katz JN, Dervan EE, Losina E. Trajectories and risk profiles of pain in persons with radiographic, symptomatic knee osteoarthritis: data from the osteoarthritis initiative. Osteoarthritis Cartilage 2014;22(5):622–630. 10.1016/j.joca.2014.03.009. [PubMed: 24662734]

5. Bastick AN, Verkleij SPJ, Damen J, Wesseling J, Hilberdink WKHA, Bindels PJE, Bierma-Zeinstra SMA. Defining hip pain trajectories in early symptomatic hip osteoarthritis – 5 year results from a nationwide prospective cohort study (CHECK). Osteoarthritis Cartilage 2016;24(5):768–775. 10.1016/j.joca.2015.11.023. [PubMed: 26854794]

6. Wesseling J, Bastick AN, Wolde S ten, Kloppenburg M, Lafeber FPJG, Bierma-Zeinstra SMA, Bijlsma JWJ. Identifying Trajectories of Pain Severity in Early Symptomatic Knee Osteoarthritis: A 5-year Followup of the Cohort Hip and Cohort Knee (CHECK) Study. J Rheumatol 2015;42(8): 1470–1477. 10.3899/jrheum.141036. [PubMed: 26136492]

7. Riddle DL, Dumenci L. Modeling longitudinal osteoarthritis data to identify homogeneous subgroups: opportunities and challenges in a burgeoning literature. Osteoarthritis Cartilage 2015;23(7):1035–1037. 10.1016/j.joca.2015.02.777. [PubMed: 25783482]

8. Nagin DS. Analyzing developmental trajectories: A semiparametric, group-based approach. Psychol Methods 1999;4(2):139–157. 10.1037/1082-989X.4.2.139.

9. Verkleij SPJ, Hoekstra T, Rozendaal RM, Waarsing JH, Koes BW, Luijsterburg PAJ, Bierma-Zeinstra SMA. Defining discriminative pain trajectories in hip osteoarthritis over a 2-year time period. Ann Rheum Dis 2012;71(9):1517–1523. 10.1136/annrheumdis-2011-200687. [PubMed: 22492782]

10. James GM, Sugar CA. Clustering for Sparsely Sampled Functional Data. J Am Stat Assoc 2003;98(462):397–408.

11. James GM, Hastie TJ. Functional linear discriminant analysis for irregularly sampled curves. J R Stat Soc Ser B Stat Methodol 2001;63(3):533–550. 10.1111/1467-9868.00297.

12. James GM, Hastie TJ, Sugar CA. Principal component models for sparse functional data. Biometrika 2000;87(3):587–602. 10.1093/biomet/87.3.587.

13. Bachrach LK, Hastie T, Wang M-C, Narasimhan B, Marcus R. Bone Mineral Acquisition in Healthy Asian, Hispanic, Black, and Caucasian Youth: A Longitudinal Study. J Clin Endocrinol Metab 1999;84(12):4702–4712. 10.1210/jcem.84.12.6182. [PubMed: 10599739]

14. Hastie TJ, Tibshirani R, Friedman J. The Elements of Statistical Learning New York: Springer series in statistics; 2001.

15. Duryea J, Li J, Peterfy CG, Gordon C, Genant HK. Trainable rule-based algorithm for the measurement of joint space width in digital radiographic images of the knee. Med Phys 2000;27(3):580–591. 10.1118/1.598897. [PubMed: 10757609]

16. Neumann G, Hunter D, Nevitt M, Chibnik LB, Kwoh K, Chen H, Harris T, Satterfield S, Duryea J. Location specific radiographic joint space width for osteoarthritis progression. Osteoarthritis Cartilage 2009;17(6):761–765. 10.1016/j.joca.2008.11.001. [PubMed: 19073368]

17. Rousseeuw PJ. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math 1987;20:53–65. 10.1016/0377-0427(87)90125-7.

18. Tibshirani R Regression Shrinkage and Selection via the Lasso. J R Stat Soc Ser B Methodol 1996;58(1):267–288.

19. Blagojevic M, Jinks C, Jeffery A, Jordan KP. Risk factors for onset of osteoarthritis of the knee in older adults: a systematic review and meta-analysis. Osteoarthritis Cartilage 2010;18(1):24–33. 10.1016/j.joca.2009.08.010. [PubMed: 19751691]

20. Creamer P, Hunt M, Dieppe P. Pain mechanisms in osteoarthritis of the knee: Effect of intraarticular anesthetic. J Rheumatol 1996;23(6):1031–1036. [PubMed: 8782136]

21. Felson DT, McLaughlin S, Goggins J, LaValley MP, Gale ME, Totterman S, Li W, Hill C, Gale D. Bone marrow edema and its relation to progression of knee osteoarthritis. Ann Intern Med 2003;139(5):330–336. [PubMed: 12965941]

22. Conaghan PG, Felson DT. Structural associations of osteoarthritis pain: lessons from magnetic resonance imaging. Novartis Found Symp 2004;260:191–201; discussion 201–205, 277–279. [PubMed: 15283451]

23. Szebenyi B, Hollander AP, Dieppe P, Quilty B, Duddy J, Clarke S, Kirwan JR. Associations between pain, function, and radiographic features in osteoarthritis of the knee. Arthritis Rheum 2006;54(1):230–235. 10.1002/art.21534. [PubMed: 16385522]

24. Dieppe PA, Lohmander LS. Pathogenesis and management of pain in osteoarthritis. The Lancet 2005;365(9463):965–973. 10.1016/S0140-6736(05)71086-2.

25. Neogi T, Felson D, Niu J, Nevitt M, Lewis CE, Aliabadi P, Sack B, Torner J, Bradley L, Zhang Y. Association between radiographic features of knee osteoarthritis and pain: results from two cohort studies. BMJ 2009;339 10.1136/bmj.b2844.

26. Lethbridge-Çejku M, Scott WW, Reichle R, Ettinger WH, Zonderman A, Costa P, Plato CC, Tobin JD, Hochberg MC. Association of radiographic features of osteoarthritis of the knee with knee pain: Data from the baltimore longitudinal study of aging. Arthritis Rheum 1995;8(3):182–188. 10.1002/art.1790080311.

27. Lawrence JS, Bremner JM, Bier F. Osteo-arthrosis. Prevalence in the population and relationship between symptoms and x-ray changes. Ann Rheum Dis 1966;25(1):1–24. [PubMed: 5905334]

28. Dieppe PA, Cushnaghan J, Shepstone L. The Bristol "OA500" study: progression of osteoarthritis (OA) over 3 years and the relationship between clinical and radiographic changes at the knee joint. Osteoarthritis Cartilage 1997;5(2):87–97. [PubMed: 9135820]

29. Cicuttini FM, Baker J, Hart DJ, Spector TD. Association of pain with radiological changes in different compartments and views of the knee joint. Osteoarthritis Cartilage 1996;4(2):143–147. [PubMed: 8806116]

30. Dieppe PA, Lohmander LS. Pathogenesis and management of pain in osteoarthritis. The Lancet 2005;365(9463):965–973. 10.1016/S0140-6736(05)71086-2.

31. Wesseling J, Dekker J, Berg WB van den, Bierma-Zeinstra SMA, Boers M, Cats HA, Deckers P, Gorter KJ, Heuts PHTG, Hilberdink WKHA, Kloppenburg M, Nelissen RGHH, Oosterveld FGJ, Oostveen JCM, Roorda LD, Viergever MA, Wolde S ten, Lafeber FPJG, Bijlsma JWJ. CHECK (Cohort Hip and Cohort Knee): similarities and differences with the Osteoarthritis Initiative. Ann Rheum Dis 2009;68(9):1413–1419. 10.1136/ard.2008.096164. [PubMed: 18772189]
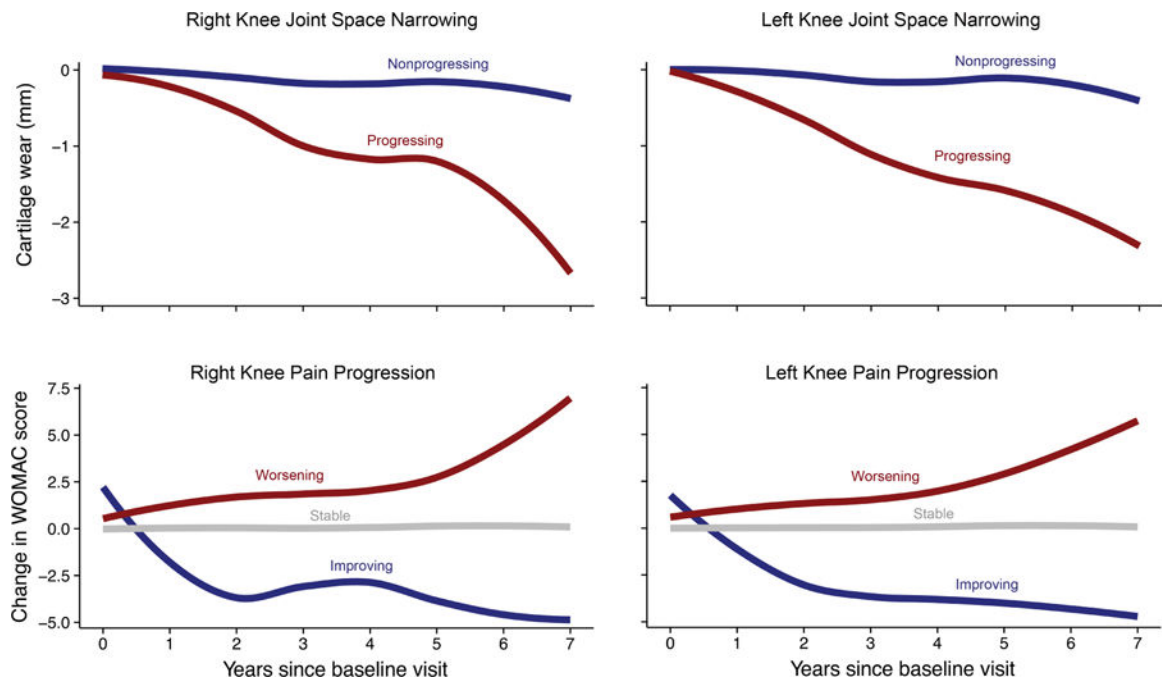
Figure 1. Functional Clustering of Osteoarthritis Progression.
Top: X-ray data revealed two major profiles of joint space narrowing—non-progressing and progressing. Bottom: WOMAC pain scores revealed three profiles of pain progression: stable levels of pain, worsening, and improving pain.
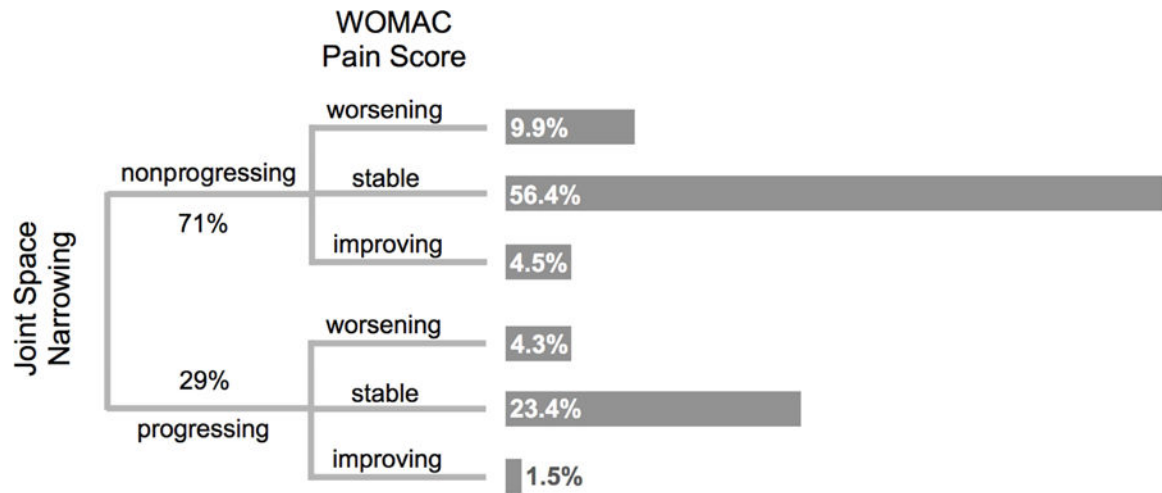
WOMAC Pain Score

Joint Space Narrowing

nonprogressing 71%
- worsening 9.9%
- stable 56.4%
- improving 4.5%

progressing 29%
- worsening 4.3%
- stable 23.4%
- improving 1.5%

Figure 2. Overlap Between Joint Space Narrowing and Pain Progression.
Joint space narrowing over the course of eight years was not accompanied by worsening of pain—the majority of radiographic progressors had stable levels of pain (4.3% < 23.4%). Similarly, worsening pain not associated with radiographic progression (4.3% < 9.9%).
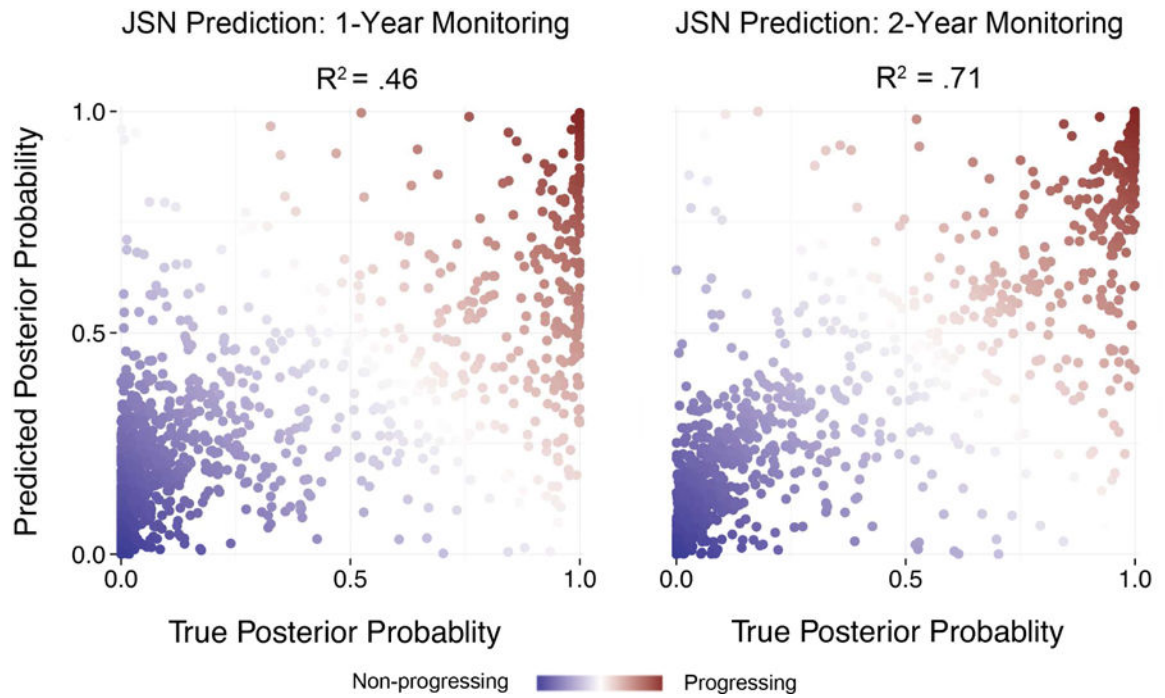
Figure 3. Predicting Longitudinal Joint Space Narrowing Using One- and Two-Year Data.
True posterior probabilities of being a progressor compared against probabilities computed using the LASSO predictive model (test data). One-year follow-up data explain 46% of the variance, whereas two-year follow-up data explain 71% of the variance.
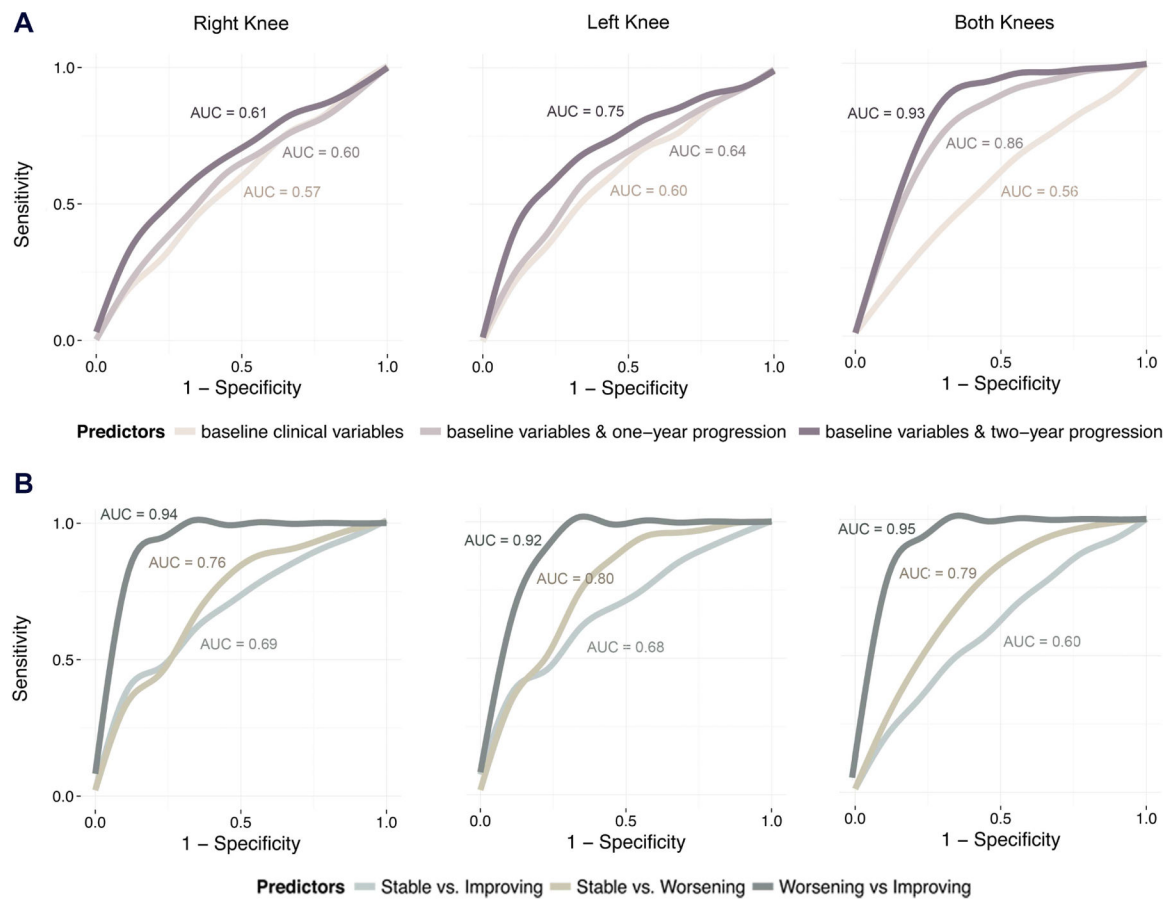
Figure 4. Model Performance Evaluation. A.
Joint space narrowing could not be predicted with high accuracy using baseline clinical variables. Data from follow-up visits improved predictive performance. Models built to predict progression in both knees simultaneously were more accurate than those built for the right and left knees separately. **B.** Pain progression could be predicted with high accuracy using data from a single visit. Models to distinguish worsening or improving subjects from stable ones were moderately accurate, whereas those distinguishing worsening from improving ones were highly accurate.

Table 1:

Subject characteristics at the enrollment visit for the included subjects, left out ones, and whole OAI incidence subcohort

Characteristic	Included (n = 1243)		Left Out (n = 2041)		Incidence Cohort (n = 3284)	
Age (years)	62.1 ± 8.8		60.8 ± 9.4		61.3 ± 9.2	
Sex (% women)	62.8%		56.6%		59.0%	
BMI (kg/m^2)	29.0 ± 4.8		27.6 ± 4.5		28.1 ± 4.6	
	Right	**Left Knee**	**Right Knee**	**Left Knee**	**Right Knee**	**Left Knee**
WOMAC Pain (0 – 20)	1.7 ± 2.5	1.6 ± 2.7	1.9 ± 2.7	1.7 ± 2.8	1.8 ± 2.6	1.7 ± 2.9
WOMAC Function (0 – 68)	5.4 ± 7.9	5.7 ± 8.8	5.9 ± 8.8	6.0 ± 9.7	5.7 ± 8.5	5.9 ± 9.3
Joint Space Width (mm)	4.3 ± 1.1	4.3 ± 1.1	4.1 ± 1.3	4.2 ± 1.2	4.3 ± 1.2	4.2 ± 1.1

Table 2:

Subject characteristics at the enrollment visit for each joint space narrowing profile

Characteristic	Nonprogressors (n = 880)		Progressors (n = 363)	
Age (years)	62.1 ± 9.0		62.3 ± 8.2	
Sex (% women)	66.7 %		53.2 %	
BMI (kg/m²)	29.0 ± 4.8		29.1 ± 4.6	
	Right Knee	**Left Knee**	**Right Knee**	**Left Knee**
WOMAC Pain (0 – 20)	1.6 ± 2.4	1.5 ± 2.6	1.8 ± 2.7	1.7 ± 2.9
WOMAC Function (0 – 68)	5.3 ± 7.7	5.5 ± 8.6	5.8 ± 8.3	6.3 ± 9.4
Joint Space Width (mm)	4.3 ± 1.1	4.2 ± 1.0	4.4 ± 1.3	4.2 ± 1.0

Table 3:

Subject characteristics at the enrollment visit for each pain progression profile

Characteristic	Worsening (n = 176)		Stable (n = 992)		Improving (n = 75)	
Age (years)	63.1 ± 8.7		62.1 ± 8.8		61.2 ± 9.0	
Sex (% women)	69.3 %		61.1 %		69.3 %	
BMI (kg/m^2)	29.8 ± 5.0		28.7 ± 4.6		30.9 ± 5.5	
	Right Knee	**Left Knee**	**Right Knee**	**Left Knee**	**Right Knee**	**Left Knee**
WOMAC Pain (0 – 20)	2.0 ± 2.6	1.9 ± 2.7	1.3 ± 2.0	1.2 ± 2.1	6.4 ± 3.6	6.4 ± 4.6
WOMAC Function (0 – 68)	7.4 ± 9.6	7.5 ± 9.5	4.3 ± 6.5	4.5 ± 7.4	16.3 ± 10.4	18.2 ± 13.4
Joint Space Width (mm)	4.0 ± 1.3	4.3 ± 1.1	4.3 ± 1.1	4.3 ± 1.1	4.4 ± 1.0	4.3 ± 1.1

Table 4.

Variables included in at least 9 of the 10 cross-validated models for predicting pain progression in both knees

Stable vs. Improving	Worsening vs. Stable	Improving vs. Worsening
KOOS: Left knee pain while bending knee fully	20-meter Walk Test: Number of steps (longer time)	KOOS: Left knee pain while walking
KOOS: Either knee difficulty while kneeling	Right Knee X-ray: Evidence of knee medial joint space narrowing	WOMAC: Right knee aggregate pain score
Block Brief 2000: frequency of dairy (cheese)	Block Brief 2000: Frequency of beef steaks/roasts/pot roast/frozen dinners/sandwiches	KOOS: Left knee pain while in bed, last 7 days
	CES-D: How often sleep was restless	