JB&JS OPEN ACCESS

An Algorithmic Approach to Understanding Osteoarthritic Knee Pain

Brandon G. Hill, MSc, Travis Byrum, BS, Anthony Zhou, MD, MBA, and Peter L. Schilling, MD, MSc

Investigation performed at Dartmouth Hitchcock Medical Center, Lebanon, New Hampshire

Background: Osteoarthritic knee pain is a complex phenomenon, and multiple factors, both within the knee and external to it, can contribute to how the patient perceives pain. We sought to determine how well a deep neural network could predict osteoarthritic knee pain and other symptoms solely from a single radiograph view.

Methods: We used data from the Osteoarthritis Initiative, a 10-year observational study of patients with knee osteoarthritis. We paired >50,000 weight-bearing, posteroanterior knee radiographs with corresponding Knee Injury and Osteoarthritis Outcome Score (KOOS) pain, symptoms, and activities of daily living subscores and used them to train a series of deep learning models to predict those scores solely from raw radiographic input. We created regression models for specific score predictions and classification models to predict whether the modeled KOOS subscore exceeded a range of thresholds.

Results: The root-mean-square errors were 15.7 for KOOS pain, 13.1 for KOOS symptoms, and 14.2 for KOOS activities of daily living. Modeling was performed to predict whether pain was above or below given pain thresholds, and was able to predict extreme pain (KOOS pain < 40) with an area under the curve (AUC) of 0.78. Notably, the system was also able to correctly predict numerous cases where the Kellgren-Lawrence (KL) grade assigned by the radiologist was 0 but patient pain was high, and cases where the KL grade was 4 but patient pain was low.

Conclusions: A deep neural network can be trained to predict the osteoarthritic knee pain that a patient experienced and other symptoms with reasonable accuracy from a single posteroanterior view of the knee, even using low-resolution images. The system can predict pain and dysfunction that the traditional KL grade does not capture. Deep learning applied to raw imaging inputs holds promise for disentangling sources of pain within the knee from aggravating factors external to the knee.

Level of Evidence: Diagnostic Level III. See Instructions for Authors for a complete description of levels of evidence.

Osteoarthritic knee pain is a complex phenomenon, and multiple factors can contribute to a patient's perception of pain. Clinicians use knee radiographs to help to gauge the extent to which pain and other symptoms may originate from structural factors within the knee. Yet, the process is complicated because a patient's perception of pain can be aggravated by factors external to the knee, including psychosocial stress, coping skills, comorbid conditions, and others[1-7]. Disentangling these "within the knee" and "external to the knee" causes is important, given that they have very different treatment implications. Psychosocial interventions target causes external to the knee, whereas physical therapy, medications, and surgical procedures address causes within the knee[8-10].

This process of assigning causation is a substantial part of an orthopaedic surgeon's thought process when considering whether a patient is a candidate for total knee arthroplasty (TKA). A surgeon must integrate disparate information to judge the extent to which the pain and dysfunction that the patient are experiencing stem from structural factors within the knee, as well as the likelihood of TKA addressing those symptoms to a patient's satisfaction, all while considering the degree to which factors external to the knee may confound matters[6,11-14]. Clinicians may not be as good at this as we think, even when it comes to the first step of determining the degree to which pain originates from structural factors within the knee[15]. Clinicians rely on the Kellgren-Lawrence (KL) grade to gauge the structural severity of osteoarthritis by radiographs[16], yet the KL grade does not explain pain very well[17-20]. It is not uncommon for clinicians to see patients with severe degenerative changes on radiographs (high KL grades) and relatively

Disclosure: The **Disclosure of Potential Conflicts of Interest** forms are provided with the online version of the article (http://links.lww.com/JBJSOA/A563).

Copyright © 2023 The Authors. Published by The Journal of Bone and Joint Surgery, Incorporated. All rights reserved. This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

little pain, as well as those with relatively mild degenerative changes on radiographs (low KL grades) and severe pain. This leaves plenty of room for improvement, starting with: What share of pain can be attributed to the knee itself?

The first step is to determine how well we can predict osteoarthritic knee pain and other symptoms solely from a single radiographic view of the knee. Using deep learning to predict osteoarthritis symptoms may allow for a more direct measure of how much osteoarthritic pain and dysfunction are structurally based. The key innovation is not constraining ourselves to preexisting notions of important structural features in the knee (e.g., joint space narrowing, osteosclerosis) and instead allowing deep learning algorithms to identify the most salient structural features from the images. We hypothesized that a substantial amount of underappreciated information about patient symptoms has gone unrecognized in imaging studies of the knee. In this study, we use a convolutional neural network to predict patients' Knee Injury and Osteoarthritis Outcome Score (KOOS) values from weight-bearing posteroanterior views of the knee using data from the Osteoarthritis Initiative (OAI)[21,22]. We evaluated the accuracy of the models' pain predictions as a minimum prerequisite for adapting and improving the models for future use cases, such as artificial intelligence-based shared decision-making tools. If successful, we anticipated that more sophisticated models and use of advanced imaging will be able to give deeper insights into the discordance between structural and symptomatic severity of knee osteoarthritis.

Materials and Methods

Population

The study participants were selected from the National Institutes of Health (NIH)-sponsored OAI. The OAI was a multicenter, 10-year longitudinal, observational study of 4,796 men and women who were 45 to 79 years of age at enrollment, of whom 29% had and 68% were at risk for symptomatic femorotibial knee osteoarthritis. Detailed descriptions of the eligibility criteria and study protocol have been published, and the study data can be found on the NIH website[21]. Briefly, individuals who had inflammatory arthritides, severe joint space narrowing, or total knee replacement in both knees, or required ambulatory aids (other than a cane) for most of their walking were excluded. Annual evaluations of the OAI participants began in 2004 at 4 study sites. Missing clinic visit data for the entire OAI sample ranged from 10% at the 1-year clinic follow-up visit to 35% at the 8-year follow-up.

In this study, 4,794 participants had ≥1 bilateral, weight-bearing, posteroanterior knee radiographs. Across all visits, this provided 26,520 radiographs. The mean number of radiographs per patient was 5.5, with as few as 1 and as many as 7. Each radiograph was split into left and right knees and was cropped and normalized. Dropping 10 radiographs for which the cropping algorithm failed left 53,030 radiographs.

Participants in the OAI were asked KOOS pain, symptoms, and activities of daily living questions for each knee[22-24]. These questions were filtered to ensure that only scores acquired within 180 days of their closest radiograph were included. This resulted in the radiograph and KOOS score pair counts in Table I.

The data were split into training, validation, and test groups (70/10/20 splits) by patients, not by radiograph, to prevent the models from leveraging individual patient characteristics. The final counts are in Table I.

Model Creation

Each set of radiograph and KOOS subscore pairs was used to build and test a series of convolutional neural network (CNN) models to predict patients' self-reported measures from their posteroanterior radiographs. All CNN models were built using a ResNet18 model pretrained on the ImageNet data set. Regression models were constructed to predict exact patient scores, and binary classification models were built to predict whether the patient score was beyond a set threshold. For regression models, the final layer was replaced with a fully connected layer to

TABLE I Data Set Creation and Final Sizes*

	KOOS					
	Pain		Symptoms		Activities of Daily Living	
	Scores	Patients	Scores	Patients	Scores	Patients
KOOS subscore count	59,039	4,796	59,088	4,796	58,637	4,796
With matching radiograph	52,983	4,794	53,013	4,794	52,753	4,792
With recent radiograph	52,927	4,794	52,993	4,794	52,733	4,792
After removal of flawed radiographs	52,917	4,794	52,983	4,794	52,723	4,792
Training	36,940	3,356	36,902	3,356	36,903	3,355
Validation	5,278	479	5,406	479	5,241	479
Test	10,699	959	10,675	959	10,579	958
Test with KL grade	8,566	877	8,596	883	8,475	868

*Each data set was broken into training, validation, and test sets using a 70/10/20 split by patients.

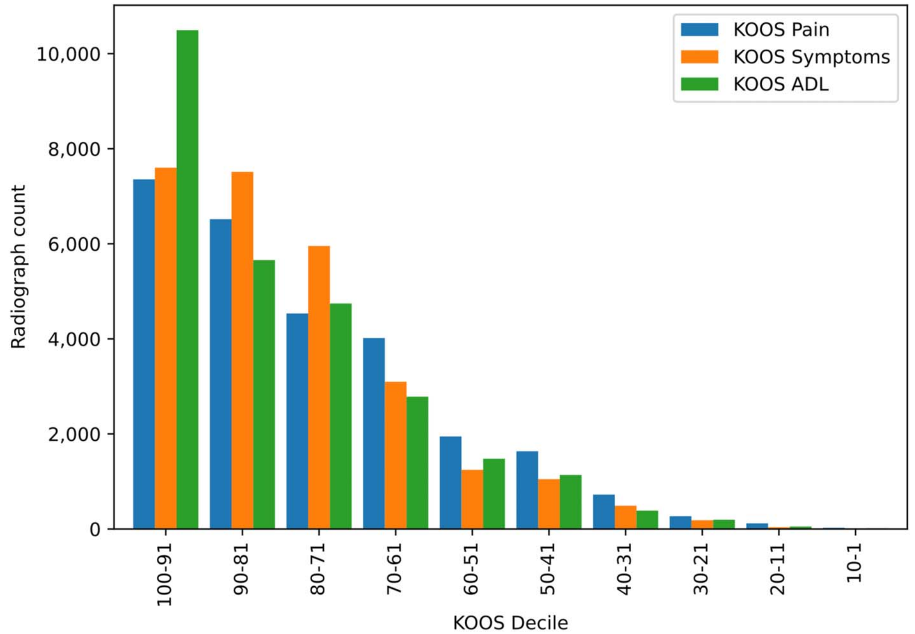

Fig. 1
Frequency of KOOS subscores by decile. ADL = activities of daily living.

predict a single output. For binary classification models, an additional sigmoid layer was added. To counter the distribution of the KOOS seen in Figure 1, imbalance weights were used in the binary classification models' loss functions.

Inputs

Radiographs were first split into left and right knees to create single-knee inputs for the models. Knee centers were located using a support vector machine trained to locate the tibiofemoral joint space[25]. This was cropped to a 140-mm square (creating 600 to 800-pixel squares, depending on the radiograph). Cropping preserved more resolution from the original image and prevented the models from learning bias resulting from radiograph artifacts unique to individual collection sites. To fit the input needs of ResNet18, the resulting images were rescaled to 224 × 224, and the z-score was normalized.

Outputs and Labels

The labels used for training and evaluating the models were the patients' self-reported KOOS knee pain and symptoms (both per knee) and activities of daily living. The KOOS ranges from 100 (perfect health) to 0 (extreme impairment). Figure 1 shows the distribution of scores for each subscore by decile.

Model Evaluation

For the regression models, the coefficient of determination (R^2) (i.e., how close the model's predictions are to a fitted line) and the root-mean-square error (RMSE) (i.e., the difference between the model's predicted values and those observed) were the primary measures of model performance.

For the binary classification models, we created a receiver operating characteristic (ROC) curve and captured the area under the curve (AUC) for each prediction threshold for each KOOS measure. To give a finer-grained look at performance, we also reported the accuracy, sensitivity, and specificity when each model's decision threshold was set to 0.5.

Finally, systems such as the KL grade fail to capture the range of reported pain and function at various levels of knee joint degeneration[17-20]. Figure 2 demonstrates this in a violin plot of patient pain compared with KL grade. A key question is whether the CNN models can find evidence in the radiograph to correctly predict anomalies regularly encountered by clinicians, such as patients with a high KL grade but low self-reported pain or a low KL grade but high self-reported pain. For this, we grouped patient data with equal KL grades and examined how well the classifier models could predict dysfunction above or below thresholds for pain (86.1), symptoms (85.7), and impairment of activities of daily living (86.8)[26].

Source of Funding

There was no source of external funding for this study.

Results

Descriptive Statistics

Most patients reported relatively low levels of pain; as shown in Figure 1, 74.1% had KOOS pain of 81 to 100). However, very high and very low levels of pain were reported across all KL grades, including severe pain (e.g., KOOS < 20) in KL grade-0 knees and no pain whatsoever (e.g., KOOS = 100) in KL grade-4 knees. The distribution of reported pain increased gradually from KL grade 0 (minimal symptoms) to KL grade 3 (mild to moderate symptoms), whereas patients with KL grade 4 had a much wider distribution of pain levels spanning the full range of possible KOOS pain (Fig. 2).

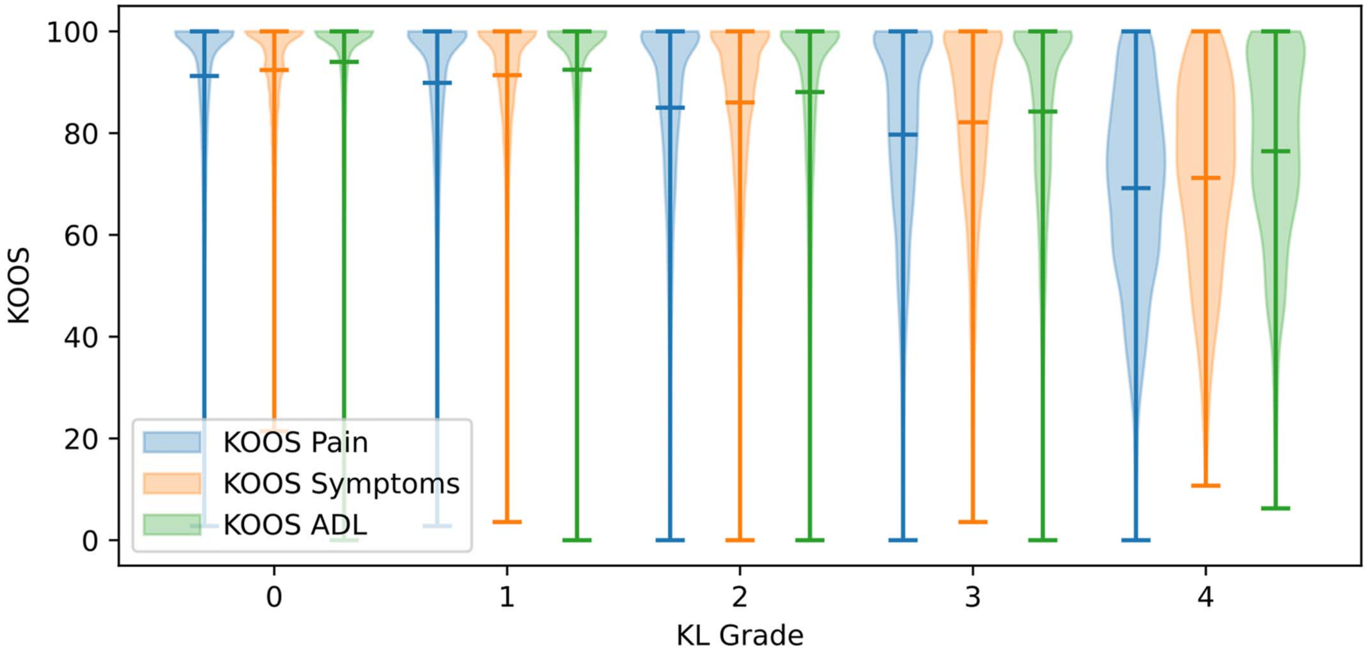

Fig. 2

KOOS subscores compared with the KL grade for the most recent radiograph across the entire cohort. Radiographs are grouped by KL grade. Darker horizontal bars indicate the group's maximum, mean, and minimum KOOS subscores. The width indicates the percentage of the group receiving that specific KOOS subscore. This plot only shows the subset of the entire data set with a KL grade available for the radiograph (pain = 42,463, symptoms = 42,483, activities of daily living [ADL] = 42,302).

Similar patterns can be seen in KOOS symptoms and activities of daily living subscores.

Model Performance

Table II shows the regression models' performance in predicting each KOOS subscore. In the test groups for pain (10,699 radiographs), symptoms (10,675 radiographs), and activities of daily living (10,579 radiographs), the models achieved RMSEs of 15.7 for pain, 13.1 for symptoms, and 14.2 for activities of daily living. The coefficient of determination (R^2) was 0.140 for pain, 0.141 for symptoms, and 0.143 for activities of daily living. We used violin plots to overlay the regression models' predicted KOOS values on top of the patient-reported KOOS subscores for each of the 5 KL grades (Fig. 3). Table III shows the accuracy of the binary classifier models when used to predict pain worse than specific KOOS thresholds. Figures 4, 5, and 6 present the corresponding ROC curves and AUCs for these binary predictions. Together, Table III and Figures 4, 5, and 6 illustrate that the models' predictive performance improved as the threshold was changed to more extreme pain levels. For example, the model was only 73% accurate when predicting KOOS pain worse (less) than 80 but 96% accurate when predicting KOOS pain worse than 50 (severe pain). Naturally, there was a trade-off between sensitivity and specificity as the threshold was altered, and parity between the measures was achieved when predicting KOOS pain worse than 90, at which point both sensitivity and specificity were approximately 60%.

As Figure 2 showed, patient pain did not tightly follow KL grades. The violin plot in Figure 3 overlays the regression model predictions over the actual KOOS values. It shows the portion of the real score range for which the regression models could find evidence in the set of radiographs for each KL grade. In general, models tended to underpredict KOOS values at lower KL grades and underpredict the variation in KOOS values seen at higher KL grades. Comparatively, the KOOS predictions based on KL grade alone would be the means seen in Figure 2, which do not allow for variations within each KL grade.

Finally, the models' ability to predict KOOS results that would be unexpected based on the KL grade was investigated. For the test cohort, 26.9% of patients with KL grade 4 (88 of 327) reported unexpectedly low KOOS pain of >86.1, and the model captured 14.8% of these 88 cases. On the other end of the

TABLE II Primary Measures of Models' Predictive Performance*

KOOS Measure	RMSE	R^2
Pain	15.7	0.140
Symptoms	13.1	0.141
Activities of daily living	14.2	0.143

*RMSE = root-mean-square error, and R^2 = coefficient of determination.

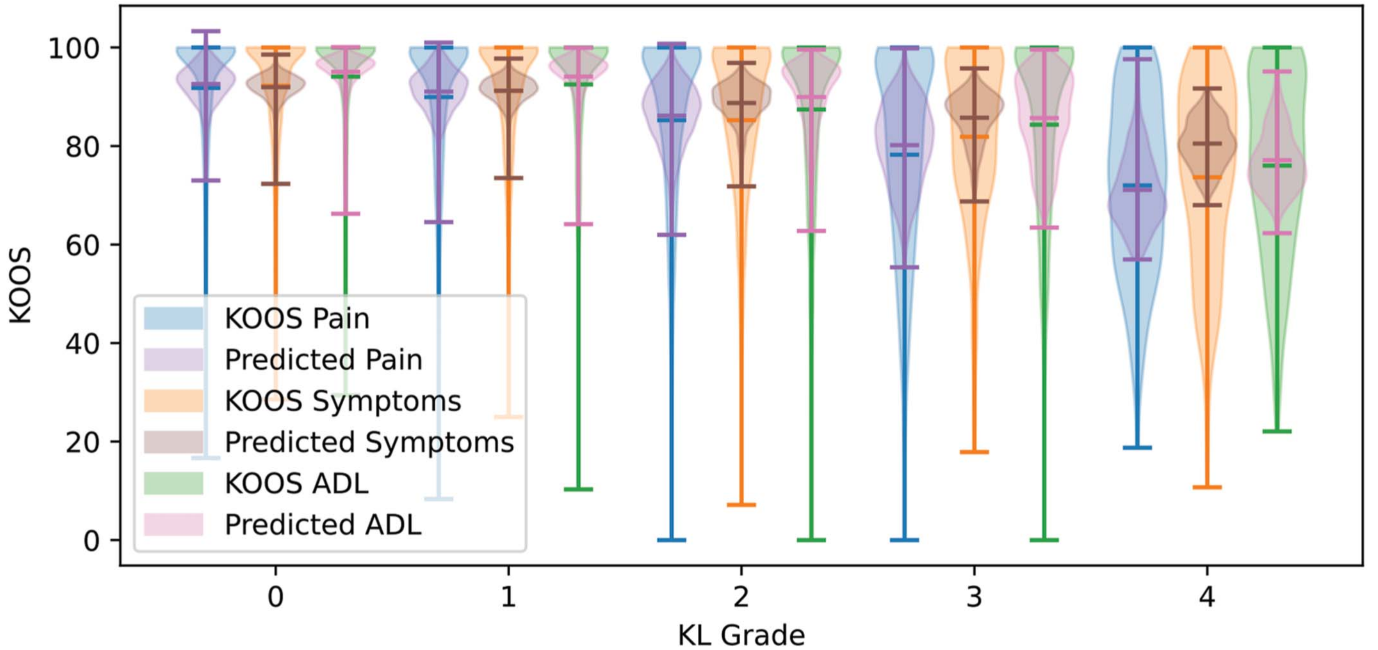

Fig. 3

Range of predicted KOOS subscore values compared with actual KOOS subscore ranges. Radiographs are grouped by KL grade. Darker horizontal bars indicate the group's maximum, mean, and minimum KOOS subscores. The width indicates the percentage of the group receiving that specific KOOS subscore. This plot only shows the subset of the test data set with a KL grade available for the radiograph (pain = 8,566, symptoms = 8,596, and activities of daily living [ADL] = 8,475).

TABLE III Results After Training, Presented as a Series of Binary Classification Models*

Measure	Measure Threshold	Accuracy	P	TP	FP	N	TN	FN	Sensitivity	Specificity
Pain	<90	0.616	4,485	2,797	2,416	6,214	3,798	1,688	0.623	0.611
	<80	0.728	2,646	1,152	1,420	8,053	6,633	1,494	0.435	0.824
	<70	0.826	1,754	367	474	8,945	8,471	1,387	0.209	0.947
	<60	0.900	960	126	234	9,739	9,505	834	0.131	0.976
	<50	0.955	472	4	18	10,227	10,209	468	0.008	0.998
	<40	0.977	246	5	14	10,453	10,439	241	0.020	0.999
Symptoms	<90	0.650	4,798	2,970	1,912	5,877	3,965	1,828	0.619	0.675
	<80	0.756	2,602	806	811	8,073	7,262	1,796	0.310	0.900
	<70	0.870	1,192	260	459	9,483	9,024	932	0.218	0.952
	<60	0.936	557	44	169	10,118	9,949	513	0.079	0.983
	<50	0.977	227	9	30	10,448	10,418	218	0.040	0.997
	<40	0.986	108	1	39	10,567	10,528	107	0.009	0.996
Activities of daily living	<90	0.601	3,466	1,859	2,613	7,113	4,500	1,607	0.536	0.633
	<80	0.795	2,202	597	559	8,377	7,818	1,605	0.271	0.933
	<70	0.881	1,260	128	126	9,319	9,193	1,132	0.102	0.986
	<60	0.931	687	38	83	9,892	9,809	649	0.055	0.992
	<50	0.964	327	23	79	10,252	10,173	304	0.070	0.992
	<40	0.988	129	0	0	10,450	10,450	129	0.000	1.000

*Models were trained to predict whether the patient measure would be worse than a given threshold. The values reflect a decision threshold of 0.5 (i.e., the model had ≥0.5 confidence in a prediction). P = positive, TP = true positive, FP = false positive, N = negative, TN = true negative, and FN = false negative.

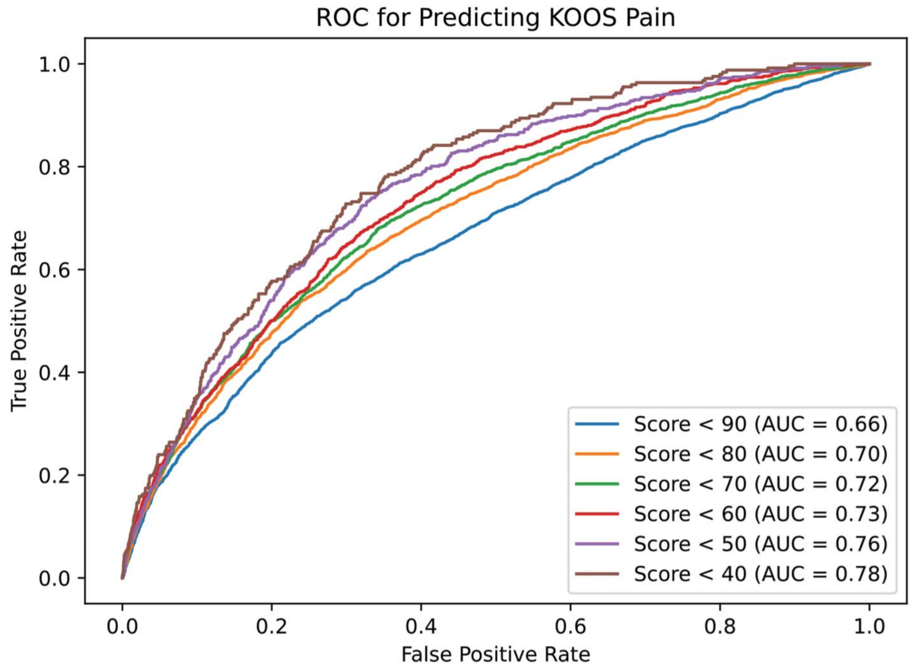

Fig. 4

ROC curves after training as a series of binary classification models. Models were trained to predict whether the patient's KOOS pain would be worse than a given threshold.

spectrum, 19.4% of patients with KL grade 0 (710 of 3,668) reported symptomatic KOOS pain of <86.1. The model captured 56.0% of these cases. These results and similar measures for KOOS symptoms and activities of daily living can be seen in Table IV.

Discussion

We found that deep learning could predict osteoarthritis knee pain and dysfunction with reasonable accuracy from a single, low-resolution radiograph. More interestingly, the

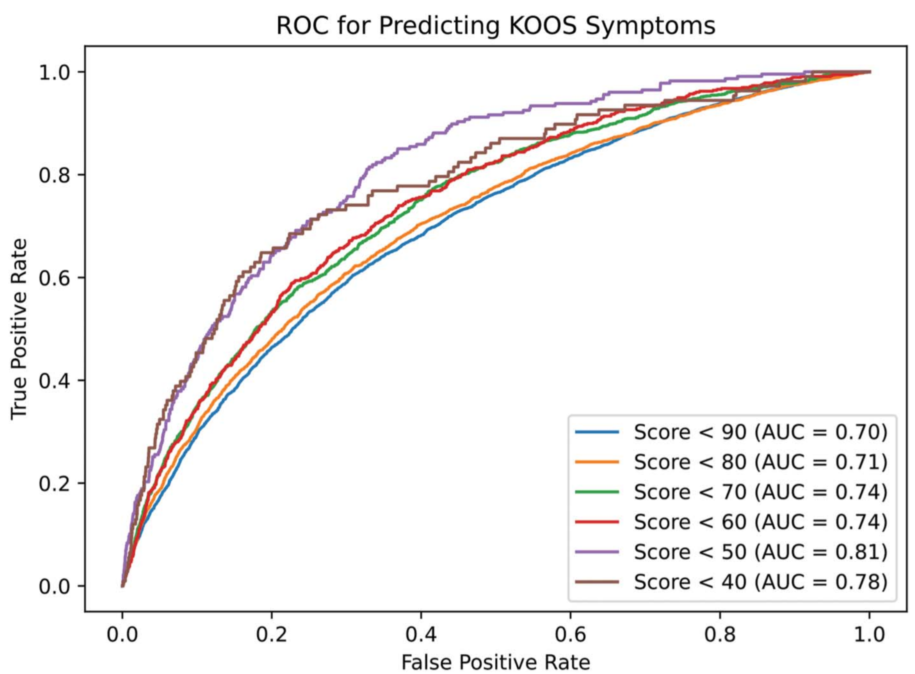

Fig. 5

ROC curves after training as a series of binary classification models. Models were trained to predict whether the patient's KOOS symptoms score would be worse than a given threshold.

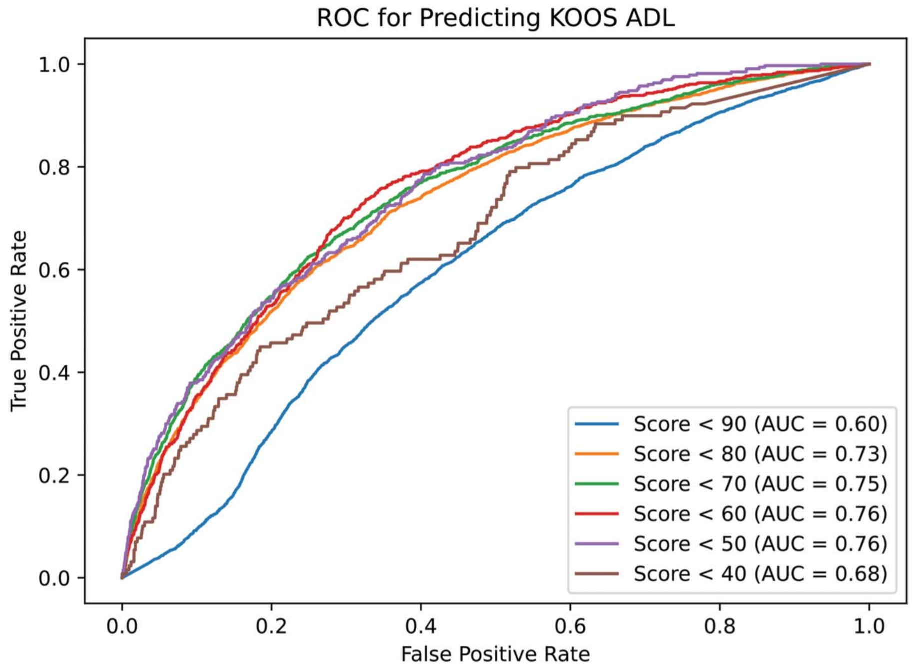

Fig. 6

ROC curves after training as a series of binary classification models. Models were trained to predict whether the patient's KOOS activities of daily living (ADL) score would be worse than a given threshold.

algorithm captured a small but notable share of cases where the radiologist assigned a KL grade of 0 but patient pain was high, and cases where the KL grade was 4 but patient pain was low. In other words, the algorithm demonstrated the ability to understand something about the relationship between radiographic imaging features and pain that a human-derived algorithmic approach (KL grade) does not capture. What the algorithm is detecting and what it means for our understanding of knee pain and osteoarthritis will be the topic of future research.

Interpreting these results benefits from context. The relationship between radiographic severity of osteoarthritis and pain has long been debated[27,28]. The KL grade has been the standard radiographic measure of osteoarthritis severity for >75 years[16]. Nevertheless, KL grades do not predict pain or other symptoms very well[17-20]. Many patients with mild or no disease as measured by radiographs experience pain, and many patients with structural damage on radiographs or magnetic resonance imaging (MRI) experience no or very little pain. This discordance between structural and symptomatic severity has represented a fundamental gap in our understanding of osteoarthritis. In general, it has been explained by indirectly implicating factors "external to the knee"[29]. For example, patients with the same radiographic severity of disease ("within the knee" causes of pain) may have experienced very different pain due to psychosocial stress, coping skills, comorbid conditions, or other factors external to the knee. Parsing "within the knee" and "external to the knee" causes is important, given that they have very different treatment implications[8-10].

TABLE IV Prediction of KOOS Levels Not Suggested by the KL Grades*

Measure	KL Grade	Count	Anomaly Threshold	No. of Anomalies	Percentage of KL Grade	Predicted	Percentage Predicted
Pain	0	3,668	≤86.1	710	19.4%	404	56.0%
	4	327	>86.1	88	26.9%	13	14.8%
Symptoms	0	3,382	≤85.7	575	17.0%	58	10.1%
	4	331	>85.7	116	35.0%	33	28.4%
Activities of daily living	0	3,383	≤86.8	538	15.9%	182	33.8%
	4	321	>86.8	106	33.0%	4	3.8%

*For radiographs with the lowest KL grades, this was how many patients reported substantial dysfunction, and for radiographs with the highest KL grades, this was how many patients reported less than substantial dysfunction. Both include how many of those cases were caught by the models.

Within this context, the implications of our findings are broad and foundational. This study demonstrates a new way for researchers to use machine learning to discriminate between the "within the knee" causes of pain and the "external to the knee" causes. Our approach may begin to illuminate the apparent discordance between structural and symptomatic disease.

Determining the maximum amount of predictive signal from an imaging study of the knee (e.g., a radiograph or MRI) is a beneficial exercise because it establishes a lower bound for the amount of structural information within the knee that can explain symptomatic knee pain. The lower bound will increase as we incorporate more granular information about the knee, including multiple radiographic views (e.g., standard views as well as multiple flexion angles) and more advanced imaging studies such as MRI. As this lower bound converges to its true upper limit, we can be more confident that the residual may be caused by or aggravated by factors external to the knee. Variables that characterize these external factors can be added to the models with increasingly sophisticated modeling techniques.

It is important to note that prior studies have accurately demonstrated the ability to predict KL grades from radiographs using deep learning[30]. Nevertheless, automation of KL grading does not advance our scientific understanding of the relationship between radiographic and symptomatic findings. To be clear, our approach described herein does. By predicting pain (not the KL grade) from radiographs, we are pivoting from human mimicry to an approach for knowledge creation and hypothesis generation.

The study had limitations that also inform the next steps. First, we do not know what the algorithm "sees" in the knee radiograph that explains pain better than KL grading. We cannot yet explain what it is about a radiograph with a KL grade of 0 or 1 that allows the algorithm to correctly identify it as a painful knee. This is the black-box problem manifested. We know what goes into the models and what comes out, but we do not understand their inner workings; the explanations for the models' predictions remain opaque. The answer to this question is beyond the scope of this article, but it is already the topic of future research efforts using explainable artificial intelligence tools. Second, the information in a single, low-resolution knee radiograph is quite limited. The models' coefficient of determination was generally <0.15. As such, the lower bound for within-the knee information will likely be substantially higher once advanced imaging like MRI can be incorporated into future models. Finally, this data set had limitations. Figure 1 shows that not all KOOS values are equally represented within the OAI data set. The models' predictive performances would likely benefit appreciably from training on more radiographs associated with poor (i.e., low) KOOS values. Also, in the parlance of machine learning, KOOS values are noisy labels, which further hampered the models' predictive performances. The noise level within the KOOS can best be characterized by estimates of the minimal detectable change in the KOOS, which is ≥20 for the age range of OAI participants[23,24]. This highlights the largest challenge of this technique. Any measure of dysfunction will have some measurement error, so the limit on the models' accuracy will always be twofold: sources of external pain and measurement errors in the dysfunction measures.

The approaches presented here are likely to someday be useful in the clinical setting for artificial intelligence-assisted shared decision-making tools. As described above, parsing "within the knee" and "external to the knee" causes is essential, given that they have very different treatment implications.

In conclusion, we present a machine learning-based approach to understanding the degree to which pain can be explained by structural factors within the knee. Even a single, low-resolution radiographic image of the knee has more information about pain than our current benchmark approaches enable us to discern. The techniques presented here do not give us all of the answers; however, they are excellent tools for asking interesting questions and generating novel hypotheses about knee osteoarthritis and its treatments. ■

Brandon G. Hill, MSc[1]
Travis Byrum, BS[2]
Anthony Zhou, MD, MBA[1]
Peter L. Schilling, MD, MSc[1,2]

[1]Dartmouth Hitchcock Medical Center, Lebanon, New Hampshire

[2]The Geisel School of Medicine at Dartmouth, Hanover, New Hampshire

Email for corresponding author: pschillin@gmail.com

References

1. Eberly L, Richter D, Comerci G, Ocksrider J, Mercer D, Mlady G, Wascher D, Schenck R. Psychosocial and demographic factors influencing pain scores of patients with knee osteoarthritis. PLoS One. 2018 4 9;13(4):e0195075.

2. Lingard EA, Riddle DL. Impact of psychological distress on pain and function following knee arthroplasty. J Bone Joint Surg Am. 2007 Jun;89(6):1161-9.

3. Allen KD, Helmick CG, Schwartz TA, DeVellis RF, Renner JB, Jordan JM. Racial differences in self-reported pain and function among individuals with radiographic hip and knee osteoarthritis: the Johnston County Osteoarthritis Project. Osteoarthritis Cartilage. 2009 Sep;17(9):1132-6.

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 May;22(5):622-30.

5. Lichtenberg PA, Swensen CH, Skehan MW. Further investigation of the role of personality, lifestyle and arthritic severity in predicting pain. J Psychosom Res. 1986;30(3):327-37.

6. Summers MN, Haley WE, Reveille JD, Alarcón GS. Radiographic assessment and psychologic variables as predictors of pain and functional impairment in osteoarthritis of the knee or hip. Arthritis Rheum. 1988 Feb;31(2):204-9.

7. Creamer P, Lethbridge-Cejku M, Hochberg MC. Determinants of pain severity in knee osteoarthritis: effect of demographic and psychosocial variables using 3 pain measures. J Rheumatol. 1999 Aug;26(8):1785-92.

8. Richmond J, Hunter D, Irrgang J, Jones MH, Levy B, Marx R, Snyder-Mackler L, Watters WC 3rd, Haralson RH 3rd, Turkelson CM, Wies JL, Boyer KM, Anderson S, St Andre J, Sluka P, McGowan R; American Academy of Orthopaedic Surgeons. Treatment of osteoarthritis of the knee (nonarthroplasty). J Am Acad Orthop Surg. 2009 Sep;17(9):591-600.

9. Jevsevar DS. Treatment of osteoarthritis of the knee: evidence-based guideline, 2nd edition. J Am Acad Orthop Surg. 2013 Sep;21(9):571-6.

10 Deveza LA, Bennell K. Management of knee osteoarthritis. 2023 5 1. Accessed 2023 Jul 28. https://www.medilib.ir/uptodate/show/111177.

11. Gademan MGJ, Hofstede SN, Vliet Vlieland TPM, Nelissen RGHH, Marang-van de Mheen PJ. Indication criteria for total hip or knee arthroplasty in osteoarthritis: a state-of-the-science overview. BMC Musculoskelet Disord. 2016 11 9;17(1):463.

12. Franklin PD, Li W, Ayers DC. The Chitranjan Ranawat Award: functional outcome after total knee replacement varies with patient attributes. Clin Orthop Relat Res. 2008 Nov;466(11):2597-604.

13. Lingard EA, Katz JN, Wright EA, Sledge CB; Kinemax Outcomes Group. Predicting the outcome of total knee arthroplasty. J Bone Joint Surg Am. 2004 Oct;86(10):2179-86.

14. Kane RL, Saleh KJ, Wilt TJ, Bershadsky B. The functional outcomes of total knee arthroplasty. J Bone Joint Surg Am. 2005 Aug;87(8):1719-24.

15. Bedson J, Croft PR. The discordance between clinical and radiographic knee osteoarthritis: a systematic search and summary of the literature. BMC Musculoskelet Disord. 2008 Sep 2;9:116.

16. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis. 1957 Dec;16(4):494-502.

17. Kim SH, Park KN. Evaluation of the relationships between Kellgren-Lawrence radiographic score and knee osteoarthritis-related pain, function, and muscle strength. Phys Ther Korea. 2019;26(2):69-75.

18. Özden F, Nadiye Karaman Ö, Tuğay N, Yalın Kilinç C, Mihriban Kilinç R, Umut Tuğay B. The relationship of radiographic findings with pain, function, and quality of life in patients with knee osteoarthritis. J Clin Orthop Trauma. 2020 Jul;11(Suppl 4):S512-7.

19. Steenkamp W, Rachuene PA, Dey R, Mzayiya NL, Ramasuvha BE. The correlation between clinical and radiological severity of osteoarthritis of the knee. SICOT J. 2022;8:14.

20. Cubukcu D, Sarsan A, Alkan H. Relationships between pain, function and radiographic findings in osteoarthritis of the knee: a cross-sectional study. Arthritis. 2012;2012:984060.

21 Nevitt MC, Felson DT, Lester G. The Osteoarthritis Initiative protocol for the cohort study. 2006 Jun 21. Accessed 2023 Jul 28. https://nda.nih.gov/static/docs/StudyDesignProtocolAndAppendices.pdf.

22. Roos EM, Roos HP, Lohmander LS, Ekdahl C, Beynnon BD. Knee Injury and Osteoarthritis Outcome Score (KOOS)—development of a self-administered outcome measure. J Orthop Sports Phys Ther. 1998 Aug;28(2):88-96.

23. Roos EM, Toksvig-Larsen S. Knee injury and Osteoarthritis Outcome Score (KOOS) - validation and comparison to the WOMAC in total knee replacement. Health Qual Life Outcomes. 2003 May 25;1:17.

24. Collins NJ, Prinsen CAC, Christensen R, Bartels EM, Terwee CB, Roos EM. Knee Injury and Osteoarthritis Outcome Score (KOOS): systematic review and meta-analysis of measurement properties. Osteoarthritis Cartilage. 2016 Aug;24(8):1317-29.

25. Tiulpin A, Thevenot J, Rahtu E, Saarakkala S. A novel method for automatic localization of joint area on knee plain radiographs. In: Image Analysis. New York: Springer International; 2017. p 290-301.

26. Englund M, Roos EM, Lohmander LS. Impact of type of meniscal tear on radiographic and symptomatic knee osteoarthritis: a sixteen-year followup of meniscectomy with matched controls. Arthritis Rheum. 2003 Aug;48(8):2178-87.

27. Hannan MT, Felson DT, Pincus T. Analysis of the discordance between radiographic changes and knee pain in osteoarthritis of the knee. J Rheumatol. 2000 Jun;27(6):1513-7.

28. Davis MA, Ettinger WH, Neuhaus JM, Barclay JD, Segal MR. Correlates of knee pain among US adults with and without radiographic knee osteoarthritis. J Rheumatol. 1992 Dec;19(12):1943-9.

29. Pierson E, Cutler DM, Leskovec J, Mullainathan S, Obermeyer Z. An algorithmic approach to reducing unexplained pain disparities in underserved populations. Nat Med. 2021 Jan;27(1):136-40.

30. Thomas KA, Kidziński Ł, Halilaj E, Fleming SL, Venkataraman GR, Oei EHG, Gold GE, Delp SL. Automated classification of radiographic knee osteoarthritis severity using deep neural networks. Radiol Artif Intell. 2020 Mar 18;2(2):e190065.