Machine-learning predicted and actual 2-year structural progression in the IMI-APPROACH cohort

Jansen, M.P.; Wirth, W.; Bacardit, J.; Helvoort, E.M. van; Marijnissen, A.C.A.; Kloppenburg, M.; ... ; Roemer, F.W.

Brief Report

Machine-learning predicted and actual 2-year structural progression in the IMI-APPROACH cohort

Mylène P. Jansen[1]^, Wolfgang Wirth[2,3,4], Jaume Bacardit[5], Eefje M. van Helvoort[1], Anne C. A. Marijnissen[1], Margreet Kloppenburg[6,7], Francisco J. Blanco[8], Ida K. Haugen[9], Francis Berenbaum[10,11], Cristoph H. Ladel[12], Marieke Loef[6,7], Floris P. J. G. Lafeber[1], Paco M. Welsing[1], Simon C. Mastbergen[1], Frank W. Roemer[13,14]

[1]Department of Rheumatology & Clinical Immunology, University Medical Center Utrecht, Utrecht, The Netherlands; [2]Department of Imaging & Functional Musculoskeletal Research, Institute of Anatomy & Cell Biology, Paracelsus Medical University Salzburg & Nuremberg, Salzburg, Austria; [3]Ludwig Boltzmann Inst. for Arthritis and Rehabilitation, Paracelsus Medical University Salzburg & Nuremberg, Salzburg, Austria; [4]Chondrometrics GmbH, Freilassing, Germany; [5]School of Computing, Newcastle University, Newcastle, UK; [6]Department of Rheumatology, Leiden University Medical Center, Leiden, The Netherlands; [7]Clinical Epidemiology, Leiden University Medical Center, Leiden, The Netherlands; [8]Grupo de Investigación de Reumatología (GIR), INIBIC-Complejo Hospitalario Universitario de A Coruña, SERGAS, Centro de Investigación CICA, Departamento de Fisioterapia y Medicina, Universidad de A Coruña, A Coruña, SpainServicio de Reumatologia, INIBIC-Universidade de A Coruña, A Coruña, Spain; [9]Center for treatment of Rheumatic and Musculoskeletal Diseases (REMEDY), Diakonhjemmet Hospital, Oslo, Norway; [10]Department of Rheumatology, AP-HP Saint-Antoine Hospital, Paris, France; [11]INSERM, Sorbonne University, Paris, France; [12]Independent Consultant, Darmstadt, Germany; [13]Quantitative Imaging Center, Department of Radiology, Boston University School of Medicine, Boston, MA, USA; [14]Department of Radiology, Universitätsklinikum Erlangen and Friedrich-Alexander-University Erlangen-Nürnberg (FAU), Erlangen, Germany

Correspondence to: Mylène P. Jansen, PhD. Department of Rheumatology & Clinical Immunology, UMC Utrecht, HP G02.228, Heidelberglaan 100, 3584CX Utrecht, The Netherlands. Email: m.p.jansen-36@umcutrecht.nl.

Abstract: In the Innovative Medicine's Initiative Applied Public-Private Research enabling OsteoArthritis Clinical Headway (IMI-APPROACH) knee osteoarthritis (OA) study, machine learning models were trained to predict the probability of structural progression (s-score), predefined as >0.3 mm/year joint space width (JSW) decrease and used as inclusion criterion. The current objective was to evaluate predicted and observed structural progression over 2 years according to different radiographic and magnetic resonance imaging (MRI)-based structural parameters. Radiographs and MRI scans were acquired at baseline and 2-year follow-up. Radiographic (JSW, subchondral bone density, osteophytes), MRI quantitative (cartilage thickness), and MRI semiquantitative [SQ; cartilage damage, bone marrow lesions (BMLs), osteophytes] measurements were obtained. The number of progressors was calculated based on a change exceeding the smallest detectable change (SDC) for quantitative measures or a full SQ-score increase in any feature. Prediction of structural progression based on baseline s-scores and Kellgren-Lawrence (KL) grades was analyzed using logistic regression. Among 237 participants, around 1 in 6 participants was a structural progressor based on the predefined JSW-threshold. The highest progression rate was seen for radiographic bone density (39%), MRI cartilage thickness (38%), and radiographic osteophyte size (35%). Baseline s-scores could only predict JSW progression parameters (most P>0.05), while KL grades could predict progression of most MRI-based and radiographic parameters (P<0.05). In conclusion, between 1/6 and 1/3 of participants showed structural progression during 2-year follow-up. KL scores were observed to outperform the machine-learning-based s-scores as progression predictor. The large amount of data collected, and the wide range of disease stage, can be used for further development of more sensitive and successful (whole joint) prediction models. Trial Registration: Clinicaltrials.gov number NCT03883568.

^ ORCID: 0000-0003-1929-6350.

Keywords: Prediction; structure; osteoarthritis (OA); magnetic resonance imaging (MRI); radiography

Submitted Sep 09, 2022. Accepted for publication Dec 30, 2022. Published online Mar 10, 2023.
doi: 10.21037/qims-22-949
View this article at: https://dx.doi.org/10.21037/qims-22-949

Introduction

Osteoarthritis (OA) is a heterogeneous disease with respect to potential causes, but also in terms of disease progression (1). Many OA patients show little or no progression, which complicates the evaluation of disease-modifying efficacy of treatment candidates in clinical trials (2,3). Predicting structural and/or symptomatic OA progression prior to patient inclusion would be helpful for trials investigating treatments such as disease-modifying OA drugs (DMOADs). Recently, a combination of biomarkers with potential prognostic utility in DMOAD trials based on multivariable modeling has been reported suggesting that once properly qualified, these biomarkers could be used to enrich future trials with participants likely to progress (4). In the Applied Public-Private Research enabling OsteoArthritis Clinical Headway (APPROACH) study, part of the Innovative Medicine's Initiative (IMI), machine learning models were used to select people with knee OA with an increased risk of structural and/or pain progression over 2 years (5). Structural progression was defined as minimum radiographic joint space width (JSW) loss of 0.6 mm over 2 years and expressed as a structure score (s-score). Pain progression was defined as increasing or sustained high self-reported pain and expressed as a pain score (p-score). Both scores ranged from 0–1, reflecting the likelihood of a participant being a progressor at 2 years. Participants with the highest combined s- and p-scores, based on radiographic and questionnaire data from a screening visit, were included in the IMI-APPROACH cohort. The comparison between the 2-year actual and predicted radiographic and pain progression in APPROACH has been reported previously (6). While previous studies have developed and evaluated machine learning models to predict OA progression with varying success (7-9), this is the first cohort that included participants based on a higher likelihood of progression. Further, previous studies generally developed and evaluated models using only one (structural) OA characteristic, usually JSW. A multitude of structural OA parameters were collected in the IMI-APPROACH cohort, including additional radiographic measures and a spectrum of magnetic resonance imaging (MRI) assessments including bone and cartilage measures, which can all be used for evaluation of progression (prediction) (10,11).

Thus, the purpose of the current study was to (I) evaluate the number of progressors based on different radiographic and MRI parameters assessed in this cohort that specifically aimed to include a high number of progressors, and (II) explore whether the predicted progression (s-score) at baseline on radiographs was associated with actual structural 2-year progression in the IMI-APPROACH cohort and how this novel prediction method compares to using baseline OA severity as a more traditional method.

Methods

Study sample

Data from the IMI-APPROACH cohort was used, in which persons with tibiofemoral knee OA were included at five centers throughout Europe, from five completed observational OA cohorts [CHECK (University Medical Center Utrecht) (12), HOSTAS (Leiden University Medical Center) (13), MUST (Diakonhjemmet Hospital, Oslo) (14), PROCOAC (INIBIC-Hospital Universitario, A Coruña) (15), DIGICOD (Sorbonne Université, Paris) (16)] and, when necessary, from outpatient clinics. Machine learning models trained on longitudinal data from the CHECK cohort were used to calculate the s- and p-scores. Details on the design of our machine learning methods for OA progression prediction have been published previously (17). Specifically, a RandomForest algorithm was used to create both the models used to rank potential participants to be invited to the screening visits as well as for the model that created the s- and p-scores used for the final decisions of inclusion into the IMI-APPROACH cohort. All models were trained on historical data from the CHECK cohort filtered to include only participants complying with the inclusion/exclusion criteria of IMI-APPROACH (explained below). The set of features used by the final inclusion model covered all measurements taken at the screening visit of IMI-APPROACH: basic patient information (age, sex, BMI), pain intensity questionnaires (KOOS, NRS), and radiographic features (bone density, eminence height, JSW, femoral-tibial angle, osteophyte

area). The most impactful of these features were minimum JSW and osteophyte size in the medial tibia. The s-scores correspond to the probability estimated by the RandomForest algorithm for the positive class (progression). Stratified ten-fold cross-validation was used to evaluate the predictive capacity of our models and perform hyperparameter tuning (number of trees, tree depth). As the objective of training this model was to perform recruitment decisions, rather than employing a generic predictive capacity metric (e.g., F1) we created a metric tailored to evaluate the effectiveness of the recruitment process. In the IMI-APPROACH screening visits inclusion and exclusion criteria were checked as well, after which the 75% of participants with the highest combined s- and p-scores were included in the IMI-APPROACH cohort. Inclusion criteria were described previously and included: satisfying the American College of Rheumatology (ACR) criteria for knee OA (18), able to walk unassisted, not predominantly patellofemoral OA (using patellar grind test), no contraindications for MRI or CT, and no secondary OA (e.g. due to leg axis deviation >10 degrees or inflammatory joint disease) (5). In total, 297 participants were included and visited the centers at multiple time points, including baseline and 2 years. Among the data collected were radiographs and 1.5T (in 2 centers; n=74) or 3T (in 3 centers; n=223) MRI scans of each participant's index knee, which was determined at screening by the physician. Baseline radiographs were used to determine the most affected compartment (MAC; medial or lateral) and Kellgren-Lawrence (KL) grade of each participant. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the regional ethics committees and Institutional Review Boards (UMC Utrecht, Leiden University Medical Center, Complejo Hospitalario Universitario de A Coruña, AP-HP Saint-Antoine Hospital, and Diakonhjemmet Hospital) and informed consent was taken from all individual participants.

Radiographs

Standardized posterior-anterior weight-bearing semi-flexed knee radiographs were performed according to the Buckland-Wright protocol (19). In the IMI-APPROACH cohort, radiographs were analyzed semi-automatically with Knee Images Digital Analysis (KIDA) software to determine mean medial, mean lateral, and minimum JSW, subchondral bone density of the medial and lateral tibia and femur, and osteophyte area of the medial and lateral tibia and femur (20,21). The pre-determined definition of progression was a decrease in minimum JSW of at least 0.3 mm per year (i.e., at least 0.6 mm over 2 years). Additionally, for all parameters, progression was determined as a deterioration of at least the smallest detectable change (SDC) in the MAC, determined previously for all parameters using KIDA software on similar knee radiographs and the same observer (20).

MRI

The IMI-APPROACH MRI protocol included 3D spoiled gradient recalled echo (SPGR) scans for quantitative, manual, quality-controlled cartilage segmentation (qMRI) to obtain the mean medial and lateral cartilage thickness (Chondrometrics GmbH, Freilassing, Germany). Progression was defined as a decrease exceeding the SDC in the MAC. The qMRI SDC was determined previously in the IMI-APPROACH cohort (10).

Triplanar proton density weighted and coronal T1 weighted scans were used for semi-quantitative (SQ) MRI Osteoarthritis Knee Score (MOAKS) scoring of cartilage damage (size of cartilage loss as a % of surface area and % of area that is full-thickness loss), bone marrow lesions (BMLs; number and size) and osteophytes (size) scores in IMI-APPROACH (22). Readings were performed by one reader with 18 years of experience of standardized MRI OA assessment at the time of assessment (FWR). MOAKS scores (0–3) of the five medial or lateral tibiofemoral subregions were summarized to one score for each feature and included only if all subregions in the compartment could be scored; progression was defined as an increase of at least one full score in the MAC. Progressors for the patellofemoral compartment were analyzed as well, where the same MOAKS scores were assigned and summarized.

Statistical analysis

Statistical evaluation of 2-year changes in JSW, changes and test-retest precision of qMRI, and changes and reliability of MOAKS scoring have been performed and published previously (6,10,11). Logistic regression was used to evaluate whether the s-score could predict actual structural progression; the s-score was first rescaled from 0–1 to 0–10, so that the odds ratios correspond with a 0.1 increase in s-score. To compare results with a more traditional parameter, logistic regression was used to analyze whether baseline KL grade could predict structural progression as well. Regression models were not adjusted for confounders (such as age or sex), since they were already included in

Table 1 Baseline data of included participants

Parameters	Included participants (n=237)
Age (years)	66.4±7.1
BMI (kg/m^2)	27.9±5.1
Sex	
Male	56 (23.6)
Female	181 (76.4)
Index knee	
Right	134 (56.5)
Left	103 (43.5)
Kellgren-Lawrence grade	
Grade 0	45 (19.0)
Grade 1	66 (27.8)
Grade 2	51 (21.5)
Grade 3	65 (27.4)
Grade 4	8 (3.4)
Center	
Utrecht	128 (54.0)
Leiden	43 (18.1)
A Coruna	32 (13.5)
Oslo	22 (9.3)
Paris	12 (5.1)
Most affected compartment	
Medial	201 (84.8)
Lateral	36 (15.2)

Data are expressed as mean ± SD or n (%). BMI, body mass index; SD, standard deviation.

the machine learning model development. Baseline values were compared between progressors and non-progressors for all structural parameters using independent *t*-tests for continuous parameters and chi square tests for categorical parameters (MOAKS). The agreement of being a progressor on similar parameters (of JSW/cartilage thickness, subchondral bone, or osteophytes) was analyzed with Cohen's κ. Only participants with at least one of KIDA, qMRI and MOAKS results at both time points were included. P values <0.05 were considered statistically significant.

Results

The required data was available of 237 participants, and baseline data can be found in *Table 1*. Descriptive statistics by sex can be found in Table S1.

The number of progressors for each parameter is shown in *Table 2* (based on SDC or one score as mentioned in Methods, see *Table 2* for exact cutoffs). Of the 221 participants that could be evaluated on minimum JSW, 40 (16.9%) was a structural progressor according to the predefined criterion of minimum JSW decrease of at least 0.6 mm over 2 years and 51 (23.1%) based on the minimum JSW SDC. The highest rates of progression were seen for radiographic subchondral bone density (85 of 221; 38.5%), quantitative MRI (qMRI) cartilage thickness (86 of 226; 38.1%), and radiographic osteophyte size (78 of 221; 35.3%). In the patellofemoral compartment, progression was low (<15%), except for MOAKS full thickness cartilage loss progression (38 of 207; 18.4%; *Table 2*). Baseline values for progressors and non-progressors are shown in Table S2 for all parameters. Comparing progressors based on parameters evaluating similar characteristics showed only slight agreement in most cases (κ≤0.20; Tables S3-S5). Only progressors based on the number and size of MOAKS BMLs showed moderate agreement (κ=0.59), and radiographic and MOAKS osteophytes showed fair agreement (κ=0.22).

In general, s-scores could not significantly predict a participant being a progressor (*Table 2*), except for minimum JSW progression based on the predefined criterion or on the SDC [both P≤0.03 and odds ratio (OR) >1.6]. However, baseline minimum JSW was used for calculation of the s-score. Correcting for baseline minimum JSW by including it in the regression model, to evaluate whether the s-score has predictive value additional to minimum JSW alone, resulted in the s-score no longer being significantly predictive of progression (both progression definitions P>0.38), confirming that baseline minimum JSW was the main driver of the s-score. Moreover, baseline minimum JSW seemed a stronger predictor than s-score based on P value (P=0.05 and P=0.10 for predefined and SDC-based progression). The s-score significantly predicted patellofemoral progression based on the MOAKS number of BMLs, but the odds ratio of 0.48 indicates that a higher s-score was actually associated with less progression. KL grade, on the other hand, could in most cases significantly

Table 2 Structural progressors and associations of s-score and KL grades with progression for all tibiofemoral and patellofemoral parameters

Parameters	Progressors most affected compartment			Association s-score		Association KL grade	
	Total No. [237]	Progression cut-off	Progressors, n (%)	P value	OR (95% CI)	P value	OR (95% CI)
Predefined progression (minimum JSW decrease ≥0.3 mm/y)							
KIDA minimum JSW	221	−0.6 mm	40 (16.9)	0.030*	1.63 (1.05–2.53)	0.084	1.30 (0.97–1.75)
TF JSW and cartilage thickness measures (change ≥ SDC or 1 full MOAKS score)							
KIDA minimum JSW	221	−0.49 mm	51 (23.1)	0.007*	1.76 (1.17–2.66)	0.051	1.31 (1.00–1.72)
KIDA mean JSW	221	−0.67/−1.53 mm**	16 (7.2)	0.669	1.15 (0.60–2.21)	0.015	1.81 (1.12–2.93)
MRI quantitative cartilage thickness	226	−0.132/−0.120 mm**	86 (38.1)	0.446	1.14 (0.81–1.61)	<0.001	1.69 (1.34–2.17)
MOAKS % area cartilage loss	187	1 score	14 (7.5)	0.485	0.77 (0.38–1.59)	0.056	1.60 (0.99–2.58)
MOAKS % full thickness loss	187	1 score	31 (16.6)	0.061	0.60 (0.35–1.02)	0.001	1.77 (1.26–2.49)
TF subchondral bone measures (change ≥ SDC or 1 full MOAKS score)							
KIDA bone density	221	0.84–1.08 mm Al Eq***	85 (38.5)	0.384	1.17 (0.82–1.67)	0.916	0.99 (0.78–1.25)
MOAKS BML number	231	1 score	28 (12.1)	0.373	1.25 (0.76–2.05)	<0.001	2.07 (1.39–3.08)
MOAKS BML size	200	1 score	25 (12.5)	0.514	0.84 (0.49–1.42)	<0.001	3.66 (2.13–6.29)
TF osteophyte measures (change ≥ SDC or 1 full MOAKS score)							
KIDA osteophyte size	221	3.2–8.1 mm^2***	78 (35.3)	0.214	0.79 (0.55–1.14)	<0.001	2.39 (1.79–3.19)
MOAKS osteophyte size	229	1 score	30 (13.1)	0.853	0.95 (0.58–1.56)	0.001	1.90 (1.30–2.77)
PF scores (change ≥1 full MOAKS score)							
MOAKS % area cartilage loss	207	1 score	27 (13.0)	0.071	1.59 (0.96–2.61)	0.596	1.10 (0.77–1.56)
MOAKS % full thickness loss	207	1 score	38 (18.4)	0.065	1.51 (0.97–2.35)	0.606	1.09 (0.80–1.48)
MOAKS BML number	231	1 score	32 (13.9)	0.009	0.48 (0.28–0.84)	0.048	1.40 (1.00-1.94)
MOAKS BML size	179	1 score	21 (11.7)	0.173	0.65 (0.35–1.21)	0.120	1.37 0.92–2.02)
MOAKS osteophyte size	230	1 score	14 (6.1)	0.956	0.98 (0.49–1.96)	0.010	2.04 (1.19–3.49)

*, S-scores are partly based on baseline minimum JSW. After adjusting these models for baseline minimum JSW, s-scores no longer show statistically significant association with progression (both models P>0.38); **, cut off depended on whether the most affected compartment was the medial side (first number) or lateral side (second number) of the joint; ***, range for different regions (medial and lateral femur and tibia). Participants were progressors if at least one of two areas in the most affected compartment surpassed the progression cut off. KIDA radiographic bone density is measured in mm Aluminum Equivalent (mm Al Eq) using an aluminum step wedge. KL, Kellgren-Lawrence; KIDA, Knee Images Digital Analysis; JSW, joint space width; TF, tibiofemoral; SDC, smallest detectable change; MOAKS, MRI Osteoarthritis Knee Score; MRI, magnetic resonance imaging; BML, bone marrow lesion; PF, patellofemoral.

predict progression, as a higher KL grade frequently resulted in increased odds of being a progressor, especially in the tibiofemoral compartment (*Table 2*).

Discussion

The IMI-APPROACH cohort used machine learning models to predict pain and/or structure progression in people with knee OA and included those with the highest predicted progression likelihood. The resulting number of structural progressors, especially those based on the predefined progressor criterion of at least 0.6 mm minimum JSW decrease over 2 years, was lower than expected. Still, looking at the number of progressors in the Osteoarthritis

Initiative (OAI), which may be considered a somewhat comparable cohort, the IMI-APPROACH cohort seems to have included a higher percentage of participants showing structural progression. Compared to the 23% of participants in the current study showing minimum JSW progression based on the SDC, only 8% of the OAI showed early/short-term JSW progression, with another 6% of patients showing late JSW progression that would likely not be picked up during the 2-year follow-up of the IMI-APPROACH cohort (3). Also, 18–24% of patients in the OAI showed SDC-based 2-year MRI cartilage thickness progression in the OAI, compared to 38% in the current study (23). MOAKS scoring progression was higher in the OAI though, although they counted within-grade changes as well, which were not considered progression in the current study (11,24). Also, MOAKS readings in the OAI are only available for highly selected subsamples based on different outcomes (24). Still, the OAI did not aim to specifically include participants showing progression, while IMI-APPROACH did, with some success.

In the IMI-APPROACH cohort, participants were included from five previous observational cohorts. While this provided the advantage of utilizing available data of these participants for the initial progression prediction, it also meant that participants had OA for many years without undergoing (joint replacement) surgery. Previous research has shown that knee OA radiographic progression follows a pattern of inertia, where knees that have shown stable OA usually remain stable and do not show significant structural progression (25). Following this reasoning, participants included in the IMI-APPROACH cohort would be expected to remain relatively stable, at least structurally. It has also been shown previously that KL grade 2 and 3 knees show more cartilage thickness loss than those with KL grade <2, which was confirmed in the current study as well, as participants with a higher KL grade showed more progression (26). Looking only at participants with radiographic OA (KL grade ≥2) resulted in higher progression rates (e.g., 50% for MRI cartilage thickness and 30% for minimum JSW based on SDC; data not shown) but also in this subgroup, the s-score was not significantly associated with progression.

The traditional and rather crude KL grade seemed to predict structural progression better than the s-score did, especially progression of BMLs and osteophytes based on OR (*Table 2*). While the machine learning model predicting progression with the s-score was not developed for progression prediction of most of the parameters evaluated in the current study, as it aimed to predict only the likelihood of a minimum JSW loss exceeding 0.3 mm per year, the machine learning model did include other structural parameters such as osteophyte size and was hypothesized to also have predictive value for OA progression in other structural parameters. Given that minimum JSW can be predicted by the s-score and not by the KL grade, perhaps the machine learning model was too strongly influenced or constricted by minimum JSW to be of value for other parameters as well. Also, the current study revealed that the different progressor definitions do not show a high agreement (Tables S3-S5), which may explain why the s-score cannot be reused to predict other progression definitions as well. In the comparison between JSW and MRI cartilage thickness this could potentially be the result of differences in acquisition (such the difference in weight-bearing, or the fact that JSW progression is a composite result of cartilage loss and meniscal damage and extrusion) (27), but even progression on qMRI cartilage thickness and MOAKS cartilage scores showed only low agreement in this study. This means that, even if the progression in one parameter (in this case minimum JSW) could have been predicted perfectly, this would not necessarily have resulted in a similarly high number of progressors in the other structural parameters. However, since a previous FNIH study demonstrated that qMRI cartilage thickness and MOAKS cartilage scores did show good agreement in terms of progression, the low agreement here could have been the result of dichotomization of the outcomes in the current study, as knees could barely exceed the progression threshold for one outcome but barely fail for the others.

In conclusion, despite the fact that the machine-learning determined s-scores did not significantly predict progression and the number of progressors was somewhat lower than expected, the IMI-APPROACH cohort seems to have included an adequate number of people with knee OA showing structural progression. Though the s-score could significantly predict minimum JSW progression, it could not predict structural progression in any other OA characteristic, while the KL grade could for most measures. When aiming to predict multiple whole joint OA changes in future studies, broader machine learning models should be developed, which are trained for multiple outcome parameters. The large amount of data collected in the cohort, and the inclusion of participants at different stages of the disease, can be used for further development of models that can predict (whole joint) structural OA

progression.

Acknowledgments

Funding: The research leading to these results have received support from the Innovative Medicines Initiative Joint Undertaking under Grant Agreement no 115770, resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in-kind contribution. See www.imi.europa.eu and www.approachproject.eu.

Footnote

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://qims.amegroups.com/article/view/10.21037/qims-22-949/coif). WW reports serving as an employee and shareholder of Chondrometrics GmbH and receiving consulting fees from Galapagos NV; MK reports consulting fees from Abbvie, Pfizer, Kiniksa, Flexion, Galapagos, Jansen, CHDR, Novartis, UCB, all paid to institution; FJB reports Funding from Gedeon Richter Plc., Bristol-Myers Squibb International Corporation (BMSIC), Sun Pharma Global FZE, Celgene Corporation, Janssen Cilag International N.V, Janssen Research & Development, Viela Bio, Inc., Astrazeneca AB, UCB BIOSCIENCES GMBH, UCB BIOPHARMA SPRL, AbbVie Deutschland GmbH & Co.KG, Merck KGaA, Amgen, Inc., Novartis Farmacéutica, S.A., Boehringer Ingelheim España, S.A, CSL Behring, LLC, Glaxosmithkline Research & Development Limited, Pfizer Inc, Lilly S.A., Corbus Pharmaceuticals Inc., Biohope Scientific Solutions for Human Health S.L., Centrexion Therapeutics Corp., Sanofi, TEDEC-MEIJI FARMA S.A., Kiniksa Pharmaceuticals, Ltd; IKH reports Research grant (ADVANCE) from Pfizer (payment to institution) and consulting fees from Novartis, outside of the submitted work; FB reports Institutional grants from TRB Chemedica and Pfizer. Consulting fees from AstraZeneca, Boehringer Ingelheim, Bone Therapeutics, Cellprothera, Galapagos, Gilead, Grunenthal, GSK, Eli Lilly, MerckSerono, MSD, Nordic Bioscience, Novartis, Pfizer, Roche, Sandoz, Sanofi, Servier, UCB, Peptinov, 4P Pharma. Honoraria for lectures from Expanscience, Pfizer, Viatris. Payment for expert testimony from Pfizer and Eli Lilly. Travel support from Nordic Pharma, Pfizer, Eli Lilly, Novartis. Stock owner of 4Moving Biotech and Peptinov; CHL reports employee of Merck KGaA at start of the study; FWR reports serving as a shareholder of Boston Imaging Core Lab (BICL), LLC and consultant to Calibr and Grünenthal. The other authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the regional ethical committees and Institutional Review Boards (UMC Utrecht, Leiden University Medical Center, Complejo Hospitalario Universitario de A Coruña, AP-HP Saint-Antoine Hospital, and Diakonhjemmet Hospital) and informed consent was taken from all individual participants.

References

1. Deveza LA, Loeser RF. Is osteoarthritis one disease or a collection of many? Rheumatology (Oxford) 2018;57:iv34-42.
2. Karsdal MA, Michaelis M, Ladel C, Siebuhr AS, Bihlet AR, Andersen JR, Guehring H, Christiansen C, Bay-Jensen AC, Kraus VB. Disease-modifying treatments for osteoarthritis (DMOADs) of the knee and hip: lessons learned from failures and opportunities for the future. Osteoarthritis Cartilage 2016;24:2013-21.
3. Collins JE, Neogi T, Losina E. Trajectories of Structural Disease Progression in Knee Osteoarthritis. Arthritis Care Res (Hoboken) 2021;73:1354-62.
4. Hunter DJ, Deveza LA, Collins JE, Losina E, Katz JN, Nevitt MC, Lynch JA, Roemer FW, Guermazi A, Bowes MA, Dam EB, Eckstein F, Kwoh CK, Hoffmann S, Kraus VB. Multivariable Modeling of Biomarker Data From the Phase I Foundation for the National Institutes of Health Osteoarthritis Biomarkers Consortium. Arthritis Care Res (Hoboken) 2022;74:1142-53.

5. van Helvoort EM, van Spil WE, Jansen MP, Welsing PMJ, Kloppenburg M, Loef M, Blanco FJ, Haugen IK, Berenbaum F, Bacardit J, Ladel CH, Loughlin J, Bay-Jensen AC, Mobasheri A, Larkin J, Boere J, Weinans HH, Lalande A, Marijnissen ACA, Lafeber FPJG. Cohort profile: The Applied Public-Private Research enabling OsteoArthritis Clinical Headway (IMI-APPROACH) study: a 2-year, European, cohort study to describe, validate and predict phenotypes of osteoarthritis using clinical, imaging and biochemical markers. BMJ Open 2020;10:e035101.
6. van Helvoort EM, Jansen MP, Marijnissen ACA, Kloppenburg M, Blanco FJ, Haugen IK, Berenbaum F, Bay-Jensen AC, Ladel C, Lalande A, Larkin J, Loughlin J, Mobasheri A, Weinans HH, Widera P, Bacardit J, Welsing PMJ, Lafeber FPJG. Predicted and actual 2-year structural and pain progression in the IMI-APPROACH knee osteoarthritis cohort. Rheumatology (Oxford) 2022;62:147-57.
7. Tiulpin A, Klein S, Bierma-Zeinstra SMA, Thevenot J, Rahtu E, Meurs JV, Oei EHG, Saarakkala S. Multimodal Machine Learning-based Knee Osteoarthritis Progression Prediction from Plain Radiographs and Clinical Data. Sci Rep 2019;9:20038.
8. Almhdie-Imjabbar A, Nguyen KL, Toumi H, Jennane R, Lespessailles E. Prediction of knee osteoarthritis progression using radiological descriptors obtained from bone texture analysis and Siamese neural networks: data from OAI and MOST cohorts. Arthritis Res Ther 2022;24:66.
9. Chan LC, Li HHT, Chan PK, Wen C. A machine learning-based approach to decipher multi-etiology of knee osteoarthritis onset and deterioration. Osteoarthr Cartil Open 2021;3:100135.
10. Wirth W, Maschek S, Marijnissen ACA, Lalande A, Blanco FJ, Berenbaum F, van de Stadt LA, Kloppenburg M, Haugen IK, Ladel CH, Bacardit J, Wisser A, Eckstein F, Roemer FW, Lafeber FPJG, Weinans HH, Jansen M. Test-retest precision and longitudinal cartilage thickness loss in the IMI-APPROACH cohort. Osteoarthritis Cartilage 2023;31:238-48.
11. Roemer FW, Jansen M, Marijnissen ACA, Guermazi A, Heiss R, Maschek S, Lalande A, Blanco FJ, Berenbaum F, van de Stadt LA, Kloppenburg M, Haugen IK, Ladel CH, Bacardit J, Wisser A, Eckstein F, Lafeber FPJG, Weinans HH, Wirth W. Structural tissue damage and 24-month progression of semi-quantitative MRI biomarkers of knee osteoarthritis in the IMI-APPROACH cohort. BMC Musculoskelet Disord 2022;23:988.
12. Wesseling J, Boers M, Viergever MA, Hilberdink WK, Lafeber FP, Dekker J, Bijlsma JW. Cohort Profile: Cohort Hip and Cohort Knee (CHECK) study. Int J Epidemiol 2016;45:36-44.
13. Damman W, Liu R, Kroon FPB, Reijnierse M, Huizinga TWJ, Rosendaal FR, Kloppenburg M. Do Comorbidities Play a Role in Hand Osteoarthritis Disease Burden? Data from the Hand Osteoarthritis in Secondary Care Cohort. J Rheumatol 2017;44:1659-66.
14. Magnusson K, Hagen KB, Østerås N, Nordsletten L, Natvig B, Haugen IK. Diabetes is associated with increased hand pain in erosive hand osteoarthritis: data from a population-based study. Arthritis Care Res (Hoboken) 2015;67:187-95.
15. Oreiro-Villar N, Raga AC, Rego-Pérez I, Pértega S, Silva-Diaz M, Freire M, Fernández-López C, Blanco FJ. PROCOAC (PROspective COhort of A Coruña) description: Spanish prospective cohort to study osteoarthritis. Reumatol Clin (Engl Ed) 2022;18:100-4.
16. Sellam J, Maheu E, Crema MD, Touati A, Courties A, Tuffet S, Rousseau A, Chevalier X, Combe B, Dougados M, Fautrel B, Kloppenburg M, Laredo JD, Loeuille D, Miquel A, Rannou F, Richette P, Simon T, Berenbaum F. The DIGICOD cohort: A hospital-based observational prospective cohort of patients with hand osteoarthritis - methodology and baseline characteristics of the population. Joint Bone Spine 2021;88:105171.
17. Widera P, Welsing PMJ, Ladel C, Loughlin J, Lafeber FPFJ, Petit Dop F, Larkin J, Weinans H, Mobasheri A, Bacardit J. Multi-classifier prediction of knee osteoarthritis progression from incomplete imbalanced longitudinal data. Sci Rep 2020;10:8427.
18. Altman R, Asch E, Bloch D, Bole G, Borenstein D, Brandt K, Christy W, Cooke TD, Greenwald R, Hochberg M. Development of criteria for the classification and reporting of osteoarthritis. Classification of osteoarthritis of the knee. Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association. Arthritis Rheum 1986;29:1039-49.
19. Buckland-Wright JC, Wolfe F, Ward RJ, Flowers N, Hayne C. Substantial superiority of semiflexed (MTP) views in knee osteoarthritis: a comparative radiographic study, without fluoroscopy, of standing extended, semiflexed (MTP), and schuss views. J Rheumatol 1999;26:2664-74.
20. Marijnissen AC, Vincken KL, Vos PA, Saris DB, Viergever MA, Bijlsma JW, Bartels LW, Lafeber FP. Knee Images

Digital Analysis (KIDA): a novel method to quantify individual radiographic features of knee osteoarthritis in detail. Osteoarthritis Cartilage 2008;16:234-43.
21. Jansen MP, Welsing PMJ, Vincken KL, Mastbergen SC. Performance of knee image digital analysis of radiographs of patients with end-stage knee osteoarthritis. Osteoarthritis Cartilage 2021;29:1530-9.
22. Hunter DJ, Guermazi A, Lo GH, Grainger AJ, Conaghan PG, Boudreau RM, Roemer FW. Evolution of semi-quantitative whole joint assessment of knee OA: MOAKS (MRI Osteoarthritis Knee Score). Osteoarthritis Cartilage 2011;19:990-1002.
23. Wirth W, Larroque S, Davies RY, Nevitt M, Gimona A, Baribaud F, Lee JH, Benichou O, Wyman BT, Hudelmaier M, Maschek S, Eckstein F; OA Initiative Investigators Group. Comparison of 1-year vs 2-year change in regional cartilage thickness in osteoarthritis results from 346 participants from the Osteoarthritis Initiative. Osteoarthritis Cartilage 2011;19:74-83.
24. Roemer FW, Guermazi A, Collins JE, Losina E, Nevitt MC, Lynch JA, Katz JN, Kwoh CK, Kraus VB, Hunter DJ. Semi-quantitative MRI biomarkers of knee osteoarthritis progression in the FNIH biomarkers consortium cohort - Methodologic aspects and definition of change. BMC Musculoskelet Disord 2016;17:466.
25. Felson D, Niu J, Sack B, Aliabadi P, McCullough C, Nevitt MC. Progression of osteoarthritis as a state of inertia. Ann Rheum Dis 2013;72:924-9.
26. Maschek S, Wirth W, Ladel C, Hellio Le Graverand MP, Eckstein F. Rates and sensitivity of knee cartilage thickness loss in specific central reading radiographic strata from the osteoarthritis initiative. Osteoarthritis Cartilage 2014;22:1550-3.
27. Jansen MP, Roemer FW, Marijnissen AKCA, Kloppenburg M, Blanco FJ, Haugen IK, Berenbaum F, Lafeber FPJG, Welsing PMJ, Mastbergen SC, Wirth W. Exploring the differences between radiographic joint space width and MRI cartilage thickness changes using data from the IMI-APPROACH cohort. Skeletal Radiol 2023. [Epub ahead of print]. doi: 10.1007/s00256-022-04259-3.

Cite this article as: Jansen MP, Wirth W, Bacardit J, van Helvoort EM, Marijnissen ACA, Kloppenburg M, Blanco FJ, Haugen IK, Berenbaum F, Ladel CH, Loef M, Lafeber FPJG, Welsing PM, Mastbergen SC, Roemer FW. Machine-learning predicted and actual 2-year structural progression in the IMI-APPROACH cohort. Quant Imaging Med Surg 2023;13(5):3298-3306. doi: 10.21037/qims-22-949

Supplementary

Table S1 Baseline data of included participants by sex

Parameter	Male participants (n=56)	Female participants (n=181)
Age (years)	66.5±6.9	66.0±7.8
BMI (kg/m^2)	27.8±5.2	28.2±4.6
Index knee		
Right	26 (46.4)	108 (59.7)
Left	30 (53.6)	73 (40.3)
Kellgren-Lawrence grade		
Grade 0	10 (17.9)	35 (19.6)
Grade 1	20 (35.7)	46 (25.7)
Grade 2	7 (12.5)	44 (24.6)
Grade 3	15 (26.8)	50 (27.9)
Grade 4	4 (7.1)	4 (2.2)
Center		
Utrecht	38 (67.9)	90 (49.7)
Leiden	10 (17.9)	33 (18.2)
A Coruna	2 (3.6)	30 (16.6)
Oslo	6 (10.7)	16 (8.8)
Paris	0 (0.0)	12 (6.6)
Most affected compartment		
Medial	49 (87.5)	152 (84.0)
Lateral	7 (12.5)	29 (16.0)

Data are expressed as mean ± SD or n (%).

https://dx.doi.org/10.21037/qims-22-949

Table S2 Baseline statistics for progressors and non-progressors

Parameter	Progressors		Non-progressors		P value
	Total No.	Baseline	Total No.	Baseline	
Predefined progression (minimum JSW decrease ≥0.3mm/y)					
KIDA minimum JSW	40	3.1±1.0	181	2.5±1.2	0.003
TF JSW and cartilage thickness measures (change ≥ SDC or 1 full MOAKS score)					
KIDA minimum JSW	51	3.0±1.1	170	2.4±1.2	0.002
KIDA mean JSW	16	4.1±1.7	205	4.1±1.2	0.927
MRI quantitative cartilage thickness	86	2.7±0.8	140	3.0±0.6	0.022
MOAKS % area cartilage loss	14	5.0 (6.0)	173	4.0 (7.0)	0.531
MOAKS % full thickness loss	31	0.0 (3.0)	156	0.0 (2.0)	0.641
TF subchondral bone measures (change ≥ SDC or 1 full MOAKS score)					
KIDA bone density	85	29.5±5.5	136	33.2±5.2	<0.001
MOAKS BML number	28	1.0 (3.0)	203	0.0 (2.0)	0.204
MOAKS BML size	25	1.0 (4.0)	175	0.0 (1.0)	0.032
TF osteophyte measures (change ≥ SDC or 1 full MOAKS score)					
KIDA osteophyte size	78	7.9±8.0	143	3.6±4.8	<0.001
MOAKS osteophyte size	30	3.0 (2.3)	199	1.0 (3.0)	<0.001
PF scores (change ≥1 full MOAKS score)					
MOAKS % area cartilage loss	27	4.0 (4.0)	180	4.0 (4.0)	0.310
MOAKS % full thickness loss	38	1.0 (3.3)	169	1.0 (3.0)	0.254
MOAKS BML number	32	1.0 (2.0)	199	1.0 (2.0)	0.124
MOAKS BML size	21	1.0 (1.0)	158	0.0 (2.0)	0.011
MOAKS osteophyte size	14	5.0 (5.0)	216	1.0 (3.0)	<0.001

P values were calculated with independent t-tests for continuous parameters and chi square tests for categorical parameters (MOAKS). Mean ± standard deviation is shown for continuous parameters; n (%) is shown for categorical parameters.

Table S3 Agreement of being a progression for different cartilage-related parameters

	KIDA minimum JSW	KIDA mean JSW	MRI cartilage thickness	MOAKS % area cartilage loss	MOAKS % full thickness loss
KIDA minimum JSW	x				
KIDA mean JSW	0.178	x			
MRI cartilage thickness	0.101	0.123	x		
MOAKS % area cartilage loss	0.022	0.011	0.115	x	
MOAKS % full thickness loss	0.060	0.060	0.081	0.182	x

Cohen's κ values of agreement are shown.

https://dx.doi.org/10.21037/qims-22-949

Table S4 Agreement of being a progression for different subchondral bone parameters

	KIDA bone density	MOAKS BML number	MOAKS BML size
KIDA bone density	x		
MOAKS BML number	−0.050	x	
MOAKS BML size	−0.016	0.592	x

Cohen's κ values of agreement are shown.

Table S5 Agreement of being a progression for different osteophyte parameters

	KIDA osteophyte size	MOAKS osteophyte size
KIDA osteophyte size	x	
MOAKS osteophyte size	0.219	x

Cohen's κ values of agreement are shown.

https://dx.doi.org/10.21037/qims-22-949