ASSOCIATIONS BETWEEN PREDICTED AND ACTUAL STRUCTURAL PROGRESSION IN THE APPROACH COHORT

M.P. Jansen[1], W. Wirth[2,3], F.W. Roemer[4,5], J. Bacardit[6], E.M. van Helvoort[1], A.C. Marijnissen[1], M. Kloppenburg[7], F.J. Blanco[8], I.K. Haugen[9], F. Berenbaum[10], F.P. Lafeber[1], P.M. Welsing[1], S.C. Mastbergen[1]

[1]University Medical Center Utrecht, Utrecht, The Netherlands; [2]Paracelcus Medical University, Salzburg, Austria; [3]Chondrometrics GmbH, Freilassing, Germany; [4]Universitätsklinikum Erlangen, Erlangen, Germany; [5]Boston University School of Medicine, Boston, US; [6]Newcastle University, Newcastle upon Tyne, UK [7]Leiden University Medical Center, Leiden, The Netherlands; [8]INIBIC-Hospital Universitario A Coruña, A Coruña, Spain; [9]Diakonhjemmet Hospital, Oslo, Norway; [10]Sorbonne University, INSERM, AP-HP Saint-Antoine hospital, Paris, France

Purpose

Since osteoarthritis (OA) is a heterogeneous disease in terms of progression, many trials include patients showing barely or no progression, particularly in early stages of the disease.
In the APPROACH study, machine learning models were trained to predict structural and pain progression of individual participants before inclusion. Predicted progression was differentiated into a structural (S) and pain (P) progression score, where the score (range 0-1) reflected the likelihood of a participant being a progressor. Structural progression was defined as a decrease in minimum joint space width (JSW) of at least 0.3mm/year; pain progression as sustained or increasing KOOS-based pain. Only participants with the estimated highest 75% combined (P+S) score after screening were included.

The purpose of the current study was to evaluate the associations between the predicted and actual structural progression according to different parameters over two years.

Methods

- 297 Participants were included at five European centers. Radiographs and 1.5T or 3T MRIs were acquired at baseline and two years.
- KIDA was used to evaluate radiographic minimum, medial and lateral tibiofemoral (TF) JSW, and subchondral bone and osteophyte size in medial and lateral femur and tibia.
- Medial and lateral compartment cartilage thickness was measured from 3D SPGR MRI scans using manual, quality-controlled segmentations.
- Semi-quantitative MOAKS cartilage damage (size of cartilage loss as a % of surface area and % of area that is full-thickness loss), bone marrow lesions (BML; number and size) and osteophytes (size) scores were assessed from triplanar PDW and coronal T1W sequences. MOAKS scores of the five medial or lateral tibiofemoral subregions were summarized and included only if all subregions could be scored in a compartment.
- Progression was defined as two-year change greater than the smallest detectable change (SDC; Table 1) for continuous measures or ≥1 full (summarized) MOAKS score change, in the most affected compartment. For minimum JSW, progressors as predefined in the study protocol (≤-0.6mm over two years) were analyzed as well.
- The five MOAKS parameters were also summarized for all patellofemoral (PF) subregions, and PF progressors were defined as a change in ≥1 full MOAKS score.
- The number of progressors was compared between centers and Kellgren-Lawrence (KL) grades with chi-square tests. Association of S scores with progression was analyzed with logistic regression. The agreement between progressors on different parameters that measure similar characteristics (of cartilage thickness, subchondral bone, or osteophytes) was analyzed using Cohen's κ.
- Only participants with ≥1 of KIDA, cartilage thickness and MOAKS results at both time points were included.

Acknowledgements

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.

Results

- Data was available for 237 participants. TF progression (**Table 1**) was seen most often in KIDA bone density (39%), MRI cartilage thickness (38%), and KIDA osteophyte size (35%), and least often in KIDA mean JSW (7%) and MOAKS % of area with cartilage loss (8%).
- There were no significant differences in TF or PF progressors between centers, except for TF MOAKS number of BML (p=0.04).
- Higher KL grades showed significantly more progressors than lower KL grades for most TF parameters.
- S scores were significantly associated only with minimum JSW, both based on the predefined progression criterion of ≤-0.6mm and on the SDC of <-0.49mm over two years. However, baseline minimum JSW was used for calculation of S scores, and after adjustment for baseline minimum JSW, the associations disappeared (both p>0.38).
- S scores were significantly associated with PF progressors based on BML number, but the odds ratio was <1 (i.e. higher S scores were associated with less progression).
- Selecting only participants with radiographic OA (KL grade ≥2, 53%) resulted in a higher percentage of progressors for all parameters (e.g. 50% MRI cartilage thickness, 52% KIDA osteophyte size) but this did not change results for S score association.
- Cohen's κ only showed little agreement in progressors based on similar TF parameters (all κ<0.2), except BML number and size (κ=0.59) and KIDA and MOAKS osteophyte size (κ=0.22).

Parameter	Number of participants and progressors most affected compartment			Association S scores (OR for 0.1 S score change)	
	Total n (237)	**Progression cut-off**	**Progressors n (%)**	**P-value**	**OR (95%CI)**
Predefined progression (minimum JSW decrease >0.3mm/y)					
KIDA minimum JSW	221	-0.6 mm	40 (16.9)	**0.030***	1.63 (1.05-2.53)
TF JSW and cartilage 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)
KIDA mean JSW	221	-0.67/-1.53 mm**	16 (7.2)	0.669	1.15 (0.60-2.21)
MRI quantitative cartilage thickness	226	-0.132/-0.120 mm**	86 (38.1)	0.446	1.14 (0.81-1.61)
MOAKS % area cartilage loss	187	1 score	14 (7.5)	0.485	0.77 (0.38-1.59)
MOAKS % full thickness loss	187	1 score	31 (16.6)	0.061	0.60 (0.35-1.02)
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)
MOAKS BML #	231	1 score	28 (12.1)	0.373	1.25 (0.76-2.05)
MOAKS BML size	200	1 score	25 (12.5)	0.514	0.84 (0.49-1.42)
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)
MOAKS osteophyte size	229	1 score	30 (13.1)	0.853	0.95 (0.58-1.56)
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)
MOAKS % full thickness loss	207	1 score	38 (18.4)	0.065	1.51 (0.97-2.35)
MOAKS BML #	231	1 score	32 (13.9)	**0.009**	0.48 (0.28-0.84)
MOAKS BML size	179	1 score	21 (11.7)	0.173	0.65 (0.35-1.21)
MOAKS osteophyte size	230	1 score	14 (6.1)	0.956	0.98 (0.49-1.96)

Table 1: Progressors for structural parameters and association of S scores with progression.
*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 bone density is measured in mm Aluminum Equivalent (mm Al Eq) using an aluminum step wedge.

Conclusions

Despite the use of machine learning models in the APPROACH cohort to only include participants that were predicted to show fast pain and/or structural progression, the number of participants that showed actual structural progression was not high. Around 1 in 6 participants was a structural progressor based on the predefined criterion, but it is possible this number would have been lower without the model-based inclusion, or progressors were more pain-based. Still, the S scores were generally not associated with progression, while KL grade often was. The fact that S scores could no longer predict minimum JSW progression when correcting for baseline, indicates that the S scores were likely influenced heavily by minimum JSW in the machine learning model. Progressors in similar characteristics but measured with different scoring techniques showed poor agreement, indicating that even successfully selecting progressors on one aspect (e.g. JSW) may not result in progressors on different aspect (e.g. cartilage thickness). The machine learning model was trained using a different progression criterion than most criteria used in this validation; machine learning models trained on the actual criteria used for this analysis may have been more effective and is now possible with the data from the APPROACH cohort.