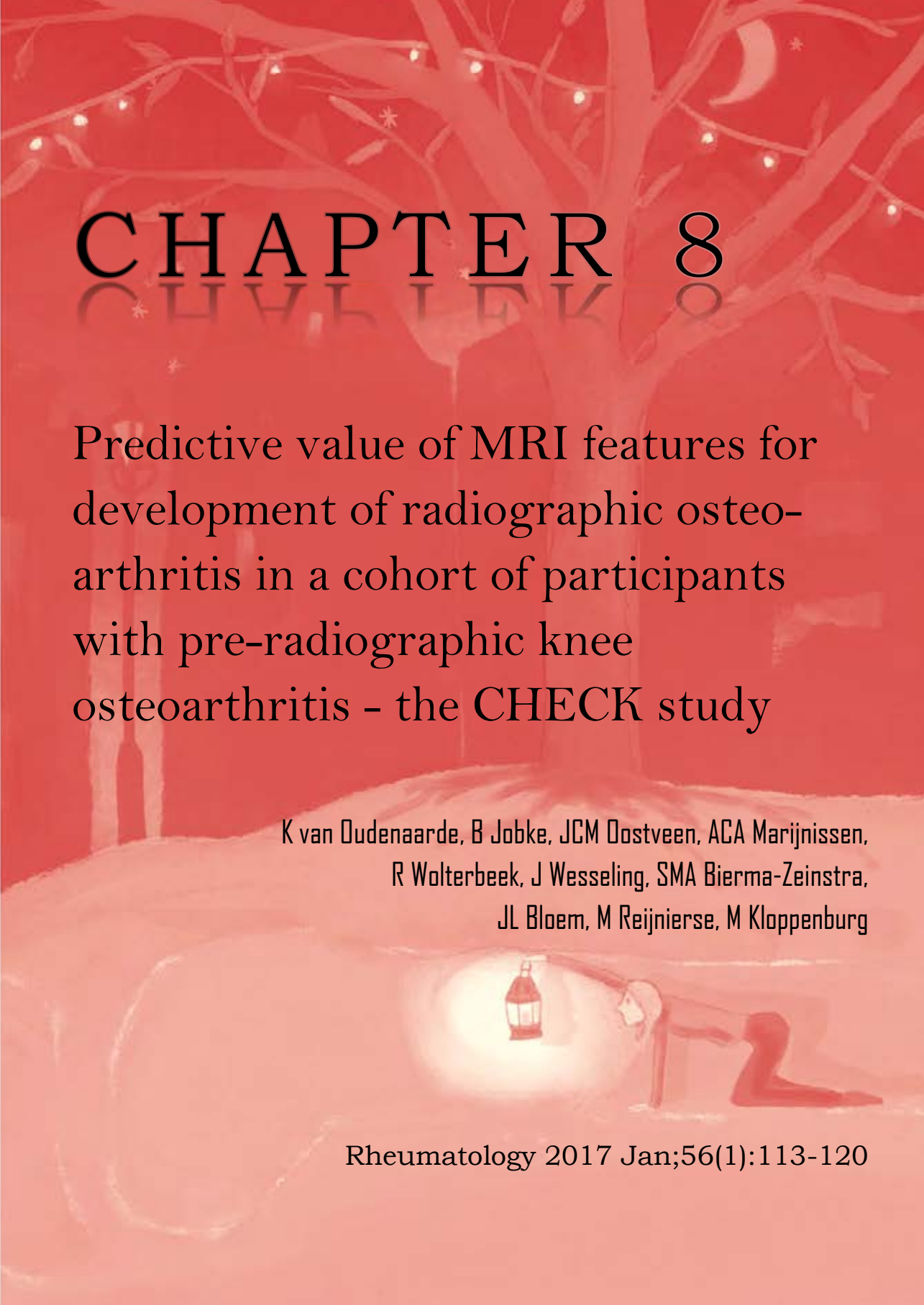
MR imaging of the knee in primary care

Oudenaarde, K. van

CHAPTER 8

Predictive value of MRI features for development of radiographic osteoarthritis in a cohort of participants with pre-radiographic knee osteoarthritis - the CHECK study

K van Oudenaarde, B Jobke, JCM Oostveen, ACA Marijnissen, R Wolterbeek, J Wesseling, SMA Bierma-Zeinstra, JL Bloem, M Reijnierse, M Kloppenburg



Rheumatology 2017 Jan;56(1):113-120

Abstract

Objective

To determine whether MRI features are associated with development of radiographic knee OA and can be used as a predictive tool in early knee OA.

Methods

In 148 participants of the Cohort Hip and Cohort Knee study (mean age 56 years, 78% women), with a Kellgren Lawrence (KL) score≤1, we obtained semi-quantitatively scored knee MRI scans and radiographs at baseline. After 5 years, we determined the development of radiographic knee OA (KL≥2). We calculated odds ratios (ORs), with 95% CIs adjusted for age, sex and BMI, to identify MRI features associated with OA development. With these MRI features, we constructed an internally validated prediction model, for which we measured the area under the receiver operating characteristics curve, sensitivity and specificity.

Results

Radiographic OA developed in 28% of the participants after 5 years. Statistically significant associations were: cartilage defects OR= 1.7 (95% CI: 1.1, 2.6), osteophytes OR= 3.1 (1.7, 5.7), bone marrow lesions OR= 2.0 (1.2, 3.4), effusion OR= 2.1 (1.2, 3.5) and meniscal pathology OR= 2.8 (1.3, 6.3). With the combined MRI features in a prediction model, the sensitivity was 66%, the specificity 67% and the optimism-corrected area under the receiver operating characteristics curve 0.685.

Conclusion

In early knee OA, MRI depicts significantly associated pathology in cartilage, bone and menisci, whereas the radiograph fails to detect these changes. Although MRI has potential for identifying patients at risk for developing radiographic knee OA, it cannot be used as an absolute diagnostic tool in early knee OA due to its low discriminative ability.

Introduction

Knee OA is a complex and slowly developing disease involving the entire joint.[1] During the development of OA, several stages can be distinguished: a preclinical/molecular stage, a pre-radiographic stage, a radiographic stage and an end stage.[2] The Kellgren and Lawrence (KL) scoring system has been the classic method for diagnosing and categorizing OA on radiographs. It comprises the presence of osteophytes (OSTs), joint space narrowing sclerosis and bony deformities.[3] It is widely accepted that radiographic OA is present with a KL score of≥2.[4]

MRI has become popular as a more comprehensive method that is more sensitive than conventional imaging to OA-related changes, including abnormalities in cartilage, subchondral bone, menisci, ligaments and synovia.[5–7] These abnormalities can already be seen on MRI of patients with knee pain, but a KL score of ≤1.[8] It is still unknown which of the patients in this early, pre-radiographic OA stage are at increased risk of progressing to definite OA. Earlier identification of patients at risk of developing radiographic OA (KL≥2) might provide a window of opportunity for modifying the course of this disease. OA-related features on MRI are therefore potentially interesting as biomarkers in clinical trials aimed at modifying disease. Earlier studies investigating these imaging biomarkers often included patients in more advanced stages of OA, or investigated only specific MRI features and their association with development of OA.[2,9]

Therefore, the objective of this study was to determine which structural abnormalities depicted on MRI in pre-radiographic knee OA (KL score≤1) are associated with development of definite radiographic OA, and whether these MRI features can be used as predictors (single or in combination) of radiographic knee OA. Related is the question of whether MRI is potentially useful for diagnosing early knee OA.

Methods

Study design and participants

We included participants of the Cohort Hip and Cohort Knee study (CHECK), who were suspected of having early symptomatic knee OA. The CHECK study is a longitudinal

prospective observational cohort of 1002 participants with pain and/or stiffness of a knee and/or hip, recruited in 10 centres in the Netherlands in 2002-2005. The study population and selection have been described previously in detail.[10] In short, inclusion criteria were patients with pain and/or stiffness of the knee and/or hip, age between 45 and 65 years, who had never, or not >6 months ago, visited their general practitioner for these symptoms for the first time. Exclusion criteria were conditions other than OA explaining their existing complaints, like rheumatic diseases and previous hip or knee joint replacement. For the current study, participants with knee complaints and a KL score of ≤1 at baseline were asked to enter this substudy. When the participant had knee pain on both sides, the knee causing the patient most difficulty was designated the signal knee and used in this study. MRI scans of each participant's most affected knee was acquired. These CHECK participants were selected in three centres (Leiden University Medical Center, University Medical Center Utrecht and Medical Spectrum Twente). The Medical Ethics Committees of all participating centres approved this study, and all participants gave written informed consent before entering the study. The study is in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Declaration of Helsinki of 1975, as revised in 2000.

Radiographic assessment

Standardized weight-bearing, posterior-anterior (PA) view, semi-flexed (7-10°) radiographic views were acquired of the femorotibial joint (FTJ).[11] The baseline and 5-year follow-up knee radiographs were scored pairwise by trained observers using the 5-point KL score (grade 0-4), with the observers being blinded for MRI information.[3,12] The interobserver reliability was assessed with prevalence-adjusted bias-adjusted kappa (PABAK) scores and the percentage agreement.[12] The PABAK score is calculated as 2ρ - 1 (with ρ = observed proportion of agreement) and takes into account the effects of bias of the low prevalence of radiographic features in this CHECK study. The PABAK score for reliability on progression of OA (KL score) in the knee from 0 to 5 years was 0.82, with a 90% average agreement.[12] We defined development of radiographic OA of the signal knee as a KL grade of ≥2 on the 5-year radiograph or when the participant received a total knee arthroplasty.

MRI assessment

Signal knees were imaged in a dedicated knee coil in a 1.5 T magnet using a standardized protocol. We focused on the femorotibial joint because of our correlation with the KL scores as assessed on the conventional AP images of the knee. We used a validated and semi-quantitative knee OA scoring system to assess OA defects in the femorotibial joint in which cartilage defects, OSTs, bone marrow lesions (BMLs), subchondral cysts, Baker's cysts, effusion and meniscal extrusion were scored graded from 0 (absence) to 3 (severe), except for a meniscal tear, which is either absent or presen.[5] In addition to this, we defined meniscal pathology as presence of a meniscal tear and/or presence of meniscal extrusion (grade≥1). Furthermore, we calculated the number of knees with structural FTJ OA on MRI at baseline, according to the proposed criteria by Hunter et al.[7], modified for our available data. We determined FTJ OA on MRI to be present when a definite OST (grade 2 or higher) and a full-thickness cartilage loss (grade 3) was present, or when one of these features was present, combined with two or more other OA features: a subchondral BML or cyst, extrusion or tearing of a meniscus, or partial cartilage loss, where full-thickness cartilage loss was absent. The MRI images were scored by two different teams, scoring in consensus within that team. The first team consisted of a musculoskeletal radiologist (IW, with >20 years of experience) and two trained research fellows (PK and RS). The second team consisted of a musculoskeletal radiologist (JLB., also with >20 years of experience), a fellow in musculoskeletal radiology (BJ) and a trained research fellow (KvO). During the assessment, the teams were blinded to radiographic information, the patient's symptoms, the patient's age and other clinical data. We measured the interobserver reliability for the two teams, using 104 MRI scans that were scored by both teams. The weighted kappa for cartilage defects was 0.42, for OSTs 0.69, for BMLs 0.43, for Baker's cysts 0.61, for effusion 0.67 and for meniscal pathology 0.65.

Statistical analyses

We used descriptive statistics for the baseline patient characteristics. We calculated the prevalence frequencies of MRI features and tested the difference in prevalence of pathology between the medial and the lateral menisci using the Chi-square test.

Explanatory analyses

We determined the associations of various MRI features with development of radiographic OA using binary logistic regression analyses and calculated odds ratios (ORs) with 95% CIs. The dependent variable was development of radiographic OA, and the independent variables were the various MRI features. We calculated crude ORs as well as the adjusted ORs corrected for age, sex and BMI. Since the highest grades of most MRI features in this early knee OA cohort were rare, we pooled grades 2 and 3 into one category. This accounted for OSTs, BMLs, Baker's cysts and effusion. Consequently, the provided ORs are for 1 U increase in score.

Prediction analyses

First, we calculated the predictive values of the MRI features found to have statistically significant association with development of radiographic OA, encompassing sensitivity, specificity, the positive and negative predictive values and the area under the curve (AUC) of the receiver operating characteristics curve. Second, we developed a prediction model that incorporated all single MRI features found to have statistically significant association with development of radiographic OA. We chose this construction, since stepwise methods are prone to unstable predictor selection in small sample sizes, especially when the variables are correlated.[13] We used rounded scores of 1 or 2[14] to value each predictor in the model instead of estimating the regression coefficients, since the risk of overestimating these coefficients is relatively high in a small sample size.[13] We assigned a score of 2 for independent predictors identified in a multivariable logistic regression analysis. We dichotomized predictors that were only statistically significant in the univariable analyses and graded these as 1 point per presence of this MRI feature.

Performance of the prediction model

Finally, we tested the performance of the prediction model for accuracy, discriminative power, calibration and validity. Accuracy was measured with sensitivity and specificity. For assessment of the discriminative ability we calculated the AUC. To test for the calibration of the score we used the Hosmer-Lemeshow goodness-of-fit statistic. A $P < 0.05$ was considered to indicate that the model was not well calibrated. We tested the validity of the

prediction model using a comprehensive bootstrap procedure with 100 iterations, for optimism-correction calculations.[14,15] The statistical analyses were performed using IBM SPSS version 23.0.

Results

Study population

We selected 169 participants with knee pain and a KL score of≤1 at baseline. Of these, 162 MRI scans were of sufficient quality and available for analyses. On follow-up, 148 participants attended the hospital for their 5-year visit. Two participants had received a total knee arthroplasty in the past 5 years and were thus graded as 'development of radiographic OA'. In total, we included the data for 148 participants in the analyses. For details on the study population, see Table 1.

Table 1 Baseline characteristics of participants with preradiographic knee OA

Baseline characteristics	Study population (n=148)
Age, mean (S.D.), years	56.0 (5.0)
Female, sex, n (%)	115 (78)
Right knee side, n (%)	93 (63)
BMI, median (range), kg/m2	25.7 (17.543.2)
KL score of imaged knee (KL 0/1/2/3/4)	59/89/0/0/0

KL: Kellgren and Lawrence.

Prevalence of MRI features at baseline

Table 2 shows the distribution of MRI features at baseline for 148 knees. Sixteen patients (11%) had a completely normal MRI scan. We determined that 89 knees (60%) had some form of cartilage loss, 81 knees (55%) had at least one OST, and 25 knees (17%) had a BML. Only seven knees (5%) had a subchondral cyst in the FTJ, and because of this low prevalence we excluded this MRI feature from further statistical analyses. A Baker's cyst was present in

38 knees (26%) and effusion was seen in 54 knees (36%). In total, 80 knees (54%) showed meniscal pathology. Prevalence of meniscal pathology was higher in the medial meniscus, as compared with the lateral meniscus, with 70 medial menisci (47%) and 22 lateral menisci (15%) with a tear and/or extrusion, P<0.001. Structural FTJ OA on MRI was seen in 14 knees (9%) at baseline.

Table 2 Associations between baseline MRI abnormalities and development of radiographic knee OA after 5 years (n=148)

	Cartilage defects		**Osteophytes**		**Bone marrow lesions**		**Baker's cyst 1.5 months**		**Effusion**		**Meniscal pathology[1]**	
Grade	No.	ROA (%)	No.	ROA (%)	No.	ROA (%)	No.	ROA (%)	No.	ROA (%)	No.	ROA (%)
0	59	13 (22)	67	12 (18)	123	30 (24)	110	28 (26)	94	20 (21)	68	12 (18)
1	61	16 (26)	66	18 (27)	10	3 (30)	24	6 (25)	39	12 (31)	80	29 (36)
2	15	5 (33)	15	11 (73)	15	8 (53)	14	7 (50)	15	9 (60)		
3	13	7 (54)										
Crude OR (95%CI)	1.5 (1.0, 2.2)		2.9 (1.6, 5.2)		1.8 (1.1, 3.1)		1.5 (0.9, 2.6)		2.1 (1.3, 3.6)		2.7 (1.2, 5.7)	
P-values	0.030		<0.001		0.026		0.116		0.004		0.013	
Adjusted OR (95%CI)[2]	1.7 (1.1, 2.6)		3.1 (1.7, 5.7)		2.0 (1.2, 3.4)		1.5 (0.9, 2.6)		2.1 (1.2, 3.5)		2.8 (1.3, 6.3)	
P-values	0.013		<0.001		<0.014		0.116		0.007		0.011	

MR features in bold indicate statistically significant association with development of radiographic OA. [1]Meniscal pathology (tear and/or extrusion): 0 for absence of and 1 for presence of meniscal pathology. [2]Adjusted for age, sex and BMI. No.: number; ROA: absolute number of participants developing radiographic knee OA, with percentage in parentheses; OR: odds ratio.

Association of baseline MRI features with development of radiographic knee OA

Radiographic knee OA developed in 41 participants (28%). Cartilage lesions, OSTs, BMLs, effusion and meniscal pathology seen on MRI all had statistically significant association with development of radiographic OA (Table 2). A Baker's cyst was not associated with development of radiographic OA. Furthermore, the composite score FTJ OA on MRI was

found to have statistically significant association with development of radiographic OA, with an adjusted OR of 9.7 (95%CI 2.6, 35.6). In the multivariable logistic regression analysis, only OSTs were a statistically independent MRI predictor [OR = 2.4 (95%CI 1.2, 4.7), P = 0.009].

Predictive values of MRI features at baseline for development of radiographic knee OA

Table 3 shows the predictive values of the MRI features found to have statistically significant association with development of radiographic knee OA. Sensitivity was highest for cartilage defects, OSTs and meniscal pathology; however, it was ≤71%. Full-thickness cartilage defects (grade 3), OSTs grade≥2, BMLs grade≥2 and moderate-to-large effusion had a specificity above 90%. Furthermore, FTJ OA on MRI showed a high specificity of 96% (95%CI 90, 99), but had a low sensitivity of 24% (95%CI 13, 41). In a few of these MRI features with high specificity (OSTs grade≥2, effusion grade≥2 and FTJ OA on MRI), the pre-test probability of radiographic OA after 5 years at least doubled, from 28% to positive predictive values above 60%.

Prediction model

Age, BMI and sex were not associated with the outcome in our sample and therefore excluded from the prediction model. The final model was built up from the five (single) MRI features with statistically significant association with development of radiographic knee OA. Presence of OSTs was the only MRI feature incorporated in the final model as a categorical feature, since this was the strongest predictor in the multivariate analysis.

Table 3 Predictive values of MR features for development of radiographic knee OA after 5 years

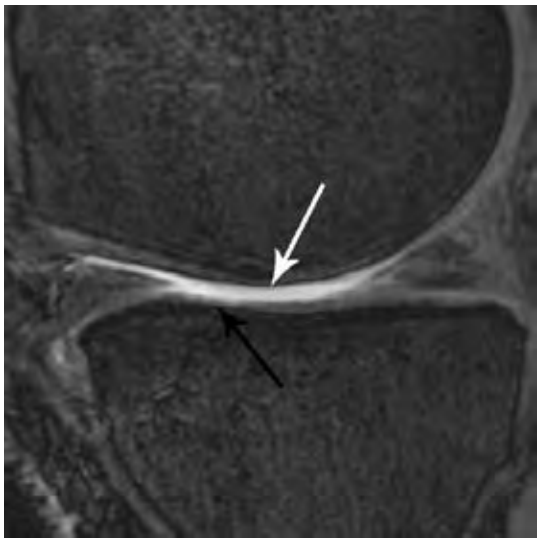
MR features	Sensitivity (95%CI)	Specificity (95%CI)	PPV (95%CI)	NPV (95%CI)	AUC (s.e.)
Cartilage defects, any (n=89)	68 (52, 81)	43 (34, 53)	31 (22, 42)	78 (65, 87)	0.602
Cartilage defects grade 3 (n=13)	17 (9, 31)	94 (88, 97)	54 (26, 80)	75 (66, 82)	0.536
Osteophytes, any (n=81)	71 (54, 83)	51 (42, 61)	36 (26, 47)	82 (70, 90)	0.664
Osteophytes, grade ≥2 (n=15)	27 (15, 43)	96 (90, 99)	73 (45, 91)	77 (69, 84)	0.626
BMLs, any (n=25)	27 (15, 43)	87 (79, 92)	44 (25, 65)	76 (67, 83)	0.592
BMLs, grade ≥2 (n=15)	20 (9, 35)	93 (87, 97)	53 (27, 78)	75 (67, 82)	0.583
Effusion, any (n=54)	51 (35, 67)	69 (59, 78)	39 (26, 53)	79 (69, 86)	0.637
Effusion, grade ≥2 (n=15)	22 (11, 38)	94 (88, 98)	60 (33, 83)	76 (68, 83)	0.598
Meniscal pathology (n=80)	71 (54, 83)	52 (43, 62)	36 (26, 48)	82 (71, 90)	0.609
FTJ OA on MR (n=14)	24 (13, 41)	96 (90, 99)	71 (42, 90)	77 (68, 84)	0.620

Sensitivity, specificity and the predictive values are all depicted as percentages. PPV: positive predictive value; NPV: negative predictive value; AUC: area under the curve, of the receiver operating characteristics curve; n: number of subjects; BML: bone marrow lesion; FTJ OA; femorotibial joint OA.

We assigned one point for detection of small OSTs at most (OST = 1), or two points for presence of moderate-to-large OSTs (OST = 2). We dichotomized the remaining MRI features to absence (0) or presence (1). For cartilage defects, we scored presence of a (grade 3) full-thickness cartilage defect [Cart_Gr3] (Fig. 1); for a BML the cut-off grade was≥2 [BML], and for effusion the cut-off grade was also≥2 [Eff]. Meniscal pathology was already dichotomized as absent or present [Men_path]. The individual score per participant was then calculated with the algorithm, score = OST + Cart_Gr3 +BML + Eff + Men_Path. This score ranged, consequently, from 0 to 6.

The AUC of the final prediction model was 0.722. After internal validation, the optimism-corrected AUC was 0.685, the sensitivity was 65.9% and the specificity was 67.0%. The optimism-corrected Hosmer-Lemeshow test had a P value of 0.645, indicating good calibration of scores.

Figure 1 MRI image of full-thickness cartilage loss (grade 3)



Sagittal gradient echo sequence showing full-thickness cartilage loss (grade 3) of the medial femoral condyle (white arrow) and grade 2 cartilage loss of the medial tibia plateau (black arrow).

Discussion

The participants in this study had knee pain and were suspected of having early knee OA. We found that the MRI features OST, cartilage defects, BMLs, effusion and meniscal pathology were associated with development of radiographic knee OA after 5 years. These MRI features had insufficient discriminative power to be useful as single predictors. We combined the five MRI features into one prediction model and reached fair discriminative power. However, after internal validation of the prediction model, we again observed poor discriminative power. Our MRI-based prediction model improved the risk assessment for the development of radiographic OA after 5 years, but cannot be used as an absolute diagnostic tool, due to the poor discriminative power.

In the proposed MRI definition of FTJ OA, developed with a Delphi consensus method, the presence of a definite OST next to a full-thickness cartilage defect would enable an OA diagnosis. Presence of OSTs is a key feature for scoring and classifying radiographic knee OA, according to the KL scale and the ACR classification criteria.[3,16] Our results are in line with these standards, with the most significant predictor in our model being OSTs. However,

the importance of MRI-depicted OSTs has received less attention in previous publications. In our sample of knees with a KL score of 0 or 1 on the radiograph, OSTs were frequently present on the MRI images, although their size was usually small. We observed most of the OSTs in the intercondylar regions of the femur, and these OSTs were often not visible on standard PA radiographs (Fig. 2). Our findings are in line with the general consensus, that MRI has a higher sensitivity than radiography for the detection of OSTs.[3,17] Other predictors from our model, BMLs and effusion, have been identified in earlier studies as important biomarkers for OA development.[18–20] A systematic review provided evidence for their correlation, not only with radiographic OA development, but also with clinical findings such as pain and stiffness of the knee, underscoring the robustness of our findings.[9] The last predictor of the prediction model, meniscal pathology (Fig. 3), has been studied thoroughly in OA. Although a fair proportion of meniscal abnormalities are regarded as incidental findings in the elderly without knee complaints[21], additional evidence is provided that meniscal pathology is associated with early-stage knee OA.[19,22]

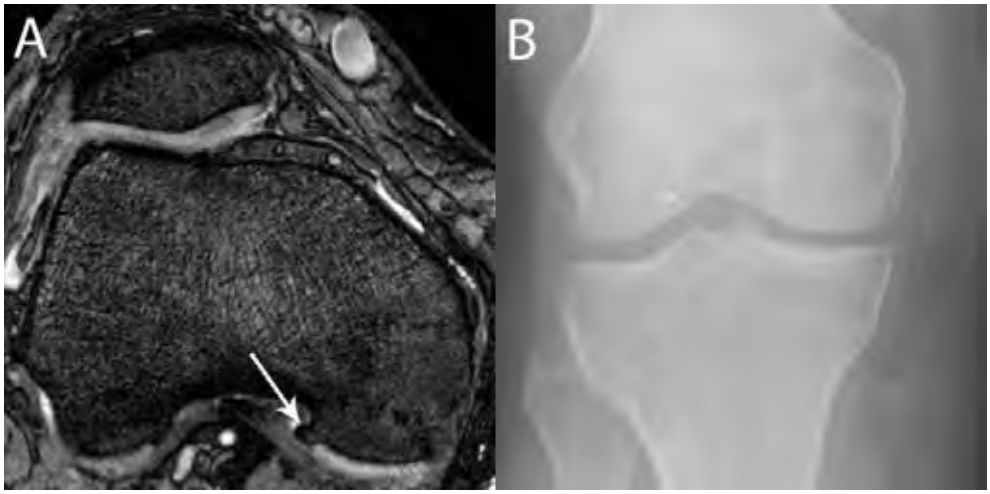
The strength of this study is its unique study population of patients with knee pain suspected of having early knee OA. To the best of our knowledge, we are the first to propose an MRI-based prediction model for the identification of these early knee OA patients. Other prediction models have been proposed for the development of radiographic knee OA, but these models have mainly focused on the established risk factors of OA, including age, sex, BMI, genetic predisposition and occupational risk factors.[23–25]

In the Rotterdam study, an externally validated prediction model was presented using patient factors, baseline KL score and genetic and biochemical markers.[23] Without the baseline KL score in the model, an AUC of 0.67 with an explained variance of 8% was calculated in that study. Adding the baseline KL score to the model raised the AUC to 0.79, with an explained variance of 34%, indicating the additive value of the baseline KL score. Their AUC of 0.79 is slightly higher than our AUC of 0.72. In our model, we chose not to incorporate the KL score at baseline, because our aim was to construct an MRI-based prediction model. Furthermore, the Rotterdam study was of a population-based cohort, with likely greater contrast between those considered and those not considered to be at risk of developing OA, which could explain the higher AUC. In the Nottingham study, another

externally validated prediction model, based on risk factors, showed a slightly lower AUC of 0.69.[24] The predictors in this model for incident radiographic knee OA after a maximum of 12 years were age, sex, BMI, occupational risk, family history and knee injury. MRI was, however, not used in that study. In an earlier publication from the CHECK study, a prediction model was presented (based on quantified radiographic features) with an AUC of 0.73, which is comparable to our AUC.[25] Unfortunately, this model was not validated. Joint space width and OST area were the main predictors in this model, and these are also the key features of the KL score. Again, no MRI imaging was used. These studies indicate the potency of prediction modelling in early knee OA; with our study, the additive value of MRI imaging is comparatively defined.

However, the true value of MRI imaging might not have been assessed completely with our study. Two reasons could be mentioned. First, we validated our MRI-based prediction model against conventional imaging. Radiography is known to have several disadvantages, including the inability to detect small changes, and poor correlation with knee function and pain.[2]

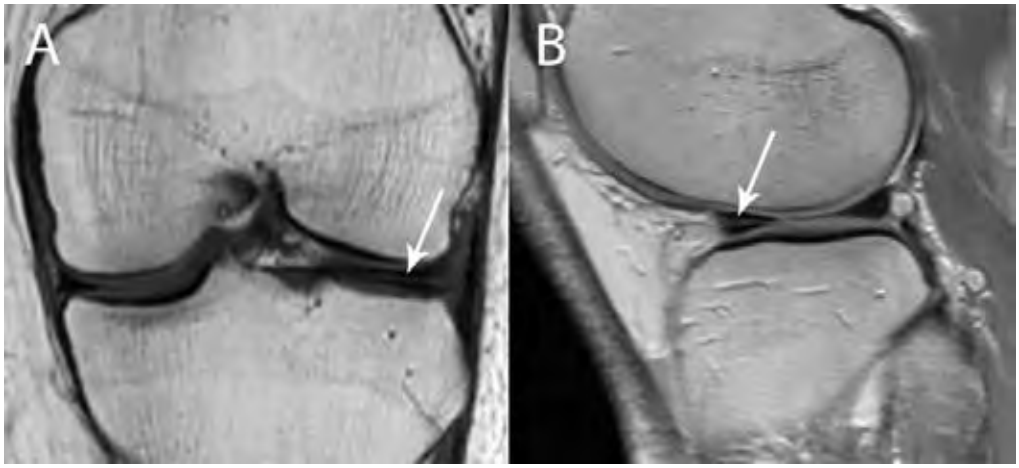
Figure 2 MRI image and corresponding conventional image of a grade 3 osteophyte



(A) Axial gradient echo sequence showing a large osteophyte (grade 3) intercondyllar of the medial femoral condyle (white arrow). (B) On the corresponding radiograph, scored as a KL grade 1, the large osteophyte is not visible.

Despite this, a KL score of≥2 on the radiograph is often used as an outcome in explanatory and prediction analyses[2,9,26], presumably since radiography is still the most widely used imaging technique to accomplish the diagnosis of knee OA.[4,16] Whether incident radiographic knee OA, determined on repeated PA views of the FTJ, is the ideal outcome is therefore questionable. The lack of a true gold standard has been discussed before, but to date no other generally accepted imaging outcome exists.[26] Second, more promising MRI techniques are appearing in the form of quantifiable MRI and higher spatial resolutions. We chose a semi-quantitative scoring method, because in routine clinical MRI reporting, quantified MRI scoring is neither a standard, nor a nimble procedure. We aimed to construct a straightforward diagnostic tool, useful in clinical practice. A limitation to this procedure, at least in our study, was the moderate-to-fair interobserver reliability for the MRI features cartilage defects and BMLs. Semi-quantitative MRI reporting of these features is known to be prone to observer variability.[27] According to Landis and Koch, our lowest-weighted Kappa of 0.43 for cartilage defects is moderate; our highest-weighted Kappa of 0.83 for effusion is almost perfect.[28] Using Fleiss' interpretation, these values are fair to good and excellent, respectively.[29]

Figure 3 MRI images of a horizontal meniscus tear



A) Coronal T1-weighted sequence and (B) a sagittal proton density sequence of a left knee showing a horizontal tear (white arrow) of the anterior horn of the lateral meniscus.

In conclusion, an internally validated prediction model with five combined MRI features resulted in an optimism-corrected AUC of 0.676 for incident radiographic knee OA after 5 years, for participants with knee pain suspected of having early knee OA. Future research is needed for external validation of our findings, with the possible addition of other predictors identified in earlier research, to further investigate the role of MRI imaging in early knee OA.

References

1. Glyn-Jones S, Palmer AJR, Agricola R, et al. Osteoarthritis. *Lancet*. 2015;386(9991):376-387.
2. Kraus VB, Burnett B, Coindreau J, et al. Application of biomarkers in the development of drugs intended for the treatment of osteoarthritis. *Osteoarthr Cartil*. 2011;19(5):515-542.
3. Kellgren J, Lawrence J. Radiological assessment of osteo-arthrosis. *Ann Rheum Dis*. 1957;16(4):494-502.
4. Roemer FW, Eckstein F, Hayashi D, Guermazi A. The role of imaging in osteoarthritis. *Best Pract Res Clin Rheumatol*. 2014;28(1):31-60.
5. Kornaat PR, Ceulemans RYT, Kroon HM, et al. MRI assessment of knee osteoarthritis: Knee Osteoarthritis Scoring System (KOSS)?inter-observer and intra-observer reproducibility of a compartment-based scoring system. *Skeletal Radiol*. 2005;34(2):95-102.
6. Hunter DJ, Guermazi A, Lo GH, et al. Evolution of semi-quantitative whole joint assessment of knee OA: MOAKS (MRI Osteoarthritis Knee Score). *Osteoarthr Cartil*. 2011;19(8):990-1002.
7. Hunter DJ, Arden N, Conaghan PG, et al. Definition of osteoarthritis on MRI: results of a Delphi exercise. *Osteoarthr Cartil*. 2011;19(8):963-969.
8. Schiphof D, Oei EHG, Hofman A, Waarsing JH, Weinans H, Bierma-Zeinstra SMA. Sensitivity and associations with pain and body weight of an MRI definition of knee osteoarthritis compared with radiographic Kellgren and Lawrence criteria: a population-based study in middle-aged females. *Osteoarthr Cartil*. 2014;22(3):440-446.
9. Hunter DJ, Zhang W, Conaghan PG, et al. Systematic review of the concurrent and predictive validity of MRI biomarkers in OA. *Osteoarthr Cartil*. 2011;19(5):557-588.
10. Wesseling J, Dekker J, van den Berg WB, et al. CHECK (Cohort Hip and Cohort Knee): similarities and differences with the Osteoarthritis Initiative. *Ann Rheum Dis*. 2009;68(9):1413-1419.
11. Buckland-Wright C. Protocols for precise radio-anatomical positioning of the tibiofemoral and patellofemoral compartments of the knee. *Osteoarthr Cartil*. 1995;3 Suppl A:71-80.
12. Damen J, Schiphof D, Wolde S Ten, Cats HA, Bierma-Zeinstra SMA, Oei EHG. Inter-observer reliability for radiographic assessment of early osteoarthritis features: the CHECK (cohort hip and cohort knee) study. *Osteoarthr Cartil*. 2014;22(7):969-974.
13. Steyerberg EW, Eijkemans MJC, Harrell FE, Habbema JD. Prognostic modelling with logistic regression analysis: a comparison of selection and estimation methods in small data sets. *Stat Med*. 2000;19(8):1059-1079.
14. Steyerberg EW. Chapter 5: Overfitting and optimism in prediction models. In: Clinical prediction models: a practical approach to development, validation, and updating, 1st edn. New York: Springer Science + Business Media, LLC, 2008: 94-96.
15. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the Performance of Prediction Models. *Epidemiology*. 2010;21(1):128-138.
16. Altman R, Asch E, Bloch D, et al. 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(8):1039-1049.
17. Chan WP, Lang P, Stevens MP, et al. Osteoarthritis of the knee: comparison of radiography, CT, and MR imaging to assess extent and severity. *Am J Roentgenol*. 1991;157(4):799-806.
18. Felson DT, McLaughlin S, Goggins J, et al. Bone marrow edema and its relation to progression of knee osteoarthritis. *Ann Intern Med*. 2003;139(5 Pt 1):330-336.
19. Sharma L, Chmiel JS, Almagor O, et al. Significance of Pre-Radiographic MRI Lesions in Persons at Higher Risk for Knee Osteoarthritis. *Arthritis Rheum*. 2014;66(7):1-23.

20. Roemer FW, Guermazi A, Felson DT, et al. Presence of MRI-detected joint effusion and synovitis increases the risk of cartilage loss in knees without osteoarthritis at 30-month follow-up: the MOST study. *Ann Rheum Dis*. 2011;70(10):1804-1809.
21. Englund M, Guermazi A, Gale D, et al. Incidental meniscal findings on knee MRI in middle-aged and elderly persons. *N Engl J Med*. 2008;359(11):1108-1115.
22. Englund M, Niu J, Guermazi A, et al. Effect of meniscal damage on the development of frequent knee pain, aching, or stiffness. *Arthritis Rheum*. 2007;56(12):4048-4054.
23. Kerkhof HJM, Bierma-Zeinstra SMA, Arden NK, et al. Prediction model for knee osteoarthritis incidence, including clinical, genetic and biochemical risk factors. *Ann Rheum Dis*. 2014;73(12):2116-2121.
24. Zhang W, McWilliams DF, Ingham SL, et al. Nottingham knee osteoarthritis risk prediction models. *Ann Rheum Dis*. 2011;70(9):1599-1604.
25. Kinds MB, Marijnissen ACA, Vincken KL, et al. Evaluation of separate quantitative radiographic features adds to the prediction of incident radiographic osteoarthritis in individuals with recent onset of knee pain: 5-year follow-up in the CHECK cohort. *Osteoarthr Cartil*. 2012;20(6):548-556.
26. Kraus VB, Nevitt M, Sandell LJ. Summary of the OA biomarkers workshop 2009 – biochemical biomarkers: biology, validation, and clinical studies. *Osteoarthr Cartil*. 2010;18(6):742-745.
27. Hunter DJ, Zhang W, Conaghan PG, et al. Responsiveness and reliability of MRI in knee osteoarthritis: a meta-analysis of published evidence. *Osteoarthr Cartil*. 2011;19(5):589-605.
28. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159-174.
29. Fleiss JL. Measuring nominal scale agreement among many raters. *Psychol Bull*. 1971;5(5):378-382.