

Identifying Phenotypes of Knee Osteoarthritis by Separate Quantitative Radiographic Features May Improve Patient Selection for More Targeted Treatment

Margot B. Kinds, Anne C.A. Marijnissen, Max A. Viergever, Pieter J. Emans, Floris P.J.G. Lafeber, and Paco M.J. Welsing

ABSTRACT. Objective. Expression of osteoarthritis (OA) varies significantly between individuals, and over time, suggesting the existence of different phenotypes, possibly with specific etiology and targets for treatment. Our objective was to identify phenotypes of progression of radiographic knee OA using separate quantitative features.

Methods. Separate radiographic features of OA were measured by Knee Images Digital Analysis (KIDA) in individuals with early knee OA (the CHECK cohort: Cohort Hip & Cohort Knee), at baseline and at 2-year and 5-year followup. Hierarchical clustering was performed to identify phenotypes of radiographic knee OA progression. The phenotypes identified were compared for changes in joint space width (JSW), varus angle, osteophyte area, eminence height, bone density, for Kellgren-Lawrence (K-L) grade, and for clinical characteristics. Logistic regression analysis evaluated whether baseline radiographic features and demographic/clinical characteristics were associated with each of the specific phenotypes.

Results. The 5 clusters identified were interpreted as “Severe” or “No,” “Early” or “Late” progression of the radiographic features, or specific involvement of “Bone density.” Medial JSW, varus angle, osteophyte area, eminence height, and bone density at baseline were associated with the Severe and Bone density phenotypes. Lesser eminence height and bone density were associated with Early and Late progression. Larger varus angle and smaller osteophyte area were associated with No progression.

Conclusion. Five phenotypes of radiographic progression of early knee OA were identified using separate quantitative features, which were associated with baseline radiographic features. Such phenotypes might require specific treatment and represent relevant subgroups for clinical trials. (First Release May 1 2013; J Rheumatol 2013;40:891–902; doi:10.3899/jrheum.121004)

Key Indexing Terms:

KNEE

OSTEOARTHRITIS

RADIOGRAPHY

From the Department of Rheumatology and Clinical Immunology, University Medical Center Utrecht, Utrecht; Image Sciences Institute, University Medical Center Utrecht, Utrecht; Department of Orthopaedic Surgery, Maastricht University Medical Center, Maastricht; and the Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands.

Supported by the Dutch Arthritis Association.

M.B. Kinds, PhD, Department of Rheumatology and Clinical Immunology, Image Sciences Institute, University Medical Center Utrecht; A.C.A. Marijnissen, PhD, Department of Rheumatology and Clinical Immunology, University Medical Center Utrecht; M.A. Viergever, PhD, Professor, Image Sciences Institute, University Medical Center Utrecht; P.J. Emans, MD, PhD, Department of Orthopaedic Surgery, Maastricht University Medical Center; F.P.J.G. Lafeber, PhD, Professor, Department of Rheumatology and Clinical Immunology, University Medical Center Utrecht; P.M.J. Welsing, PhD, Department of Rheumatology and Clinical Immunology, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht.

Address correspondence to F. Lafeber, Department of Rheumatology and Clinical Immunology, University Medical Center Utrecht, PO Box 85500, Utrecht, 3508 GA Utrecht, The Netherlands.

E-mail: F.Lafeber@umcutrecht.nl

Accepted for publication January 21, 2013.

Osteoarthritis (OA) is a degenerative joint disease characterized by pain and functional disability. In particular, knee OA has a high and increasing prevalence and is considered a major health and economic problem¹. Structural changes affect the whole joint and include cartilage, bone, and soft tissues². Definition of the disease and of diagnostic criteria remains difficult despite research on OA over many years^{3,4}. This is mainly due to the generally slow progression of the degenerative process early in the disease⁵, and the (apparent) inconsistent relation between clinical symptoms and radiographic characteristics of OA representing structural damage (which are directly or indirectly assessed)^{6,7,8,9}. Consequently, magnetic resonance imaging (MRI) techniques were developed, which show promise in directly visualizing morphologic and premorphologic changes of cartilage and other joint tissues using both conventional and complex MRI techniques¹⁰. However, radiography is still the primary method to prove the disease-modifying efficacy

(tissue structure modification) of treatment^{2,11}, because image acquisition is noninvasive, inexpensive, fast, and generally available^{1,12}.

In clinical practice, expression of disease varies significantly between patients and over time, and therefore it is appreciated that different phenotypes (subpopulations) of OA exist^{2,11,13}. For instance, in patients with prominent inflammation, a more destructive type of OA is found¹⁴. It is hypothesized also that radiographic phenotypes of OA exist. For example, some patients may suffer mainly from bone changes, while others predominantly have damage of cartilage. These radiographic phenotypes may also have their specific clinical characteristics. For example, patients with predominantly bone changes may sense more pain¹⁵, and patients with osteophyte growth may have more joint inflammation, as these phenomena have been linked¹⁶. The rate of progression and the sequence of occurrence of different radiographic characteristics may vary for different phenotypes of radiographic knee OA. Such subtle differences will be overlooked when progression of radiographic joint damage is evaluated by the commonly used Kellgren-Lawrence (K-L) grading, which is a rough score (0 to IV) summarizing multiple characteristics¹⁷. These limitations hamper selection of subgroups of individuals for whom specific treatment strategies might be helpful. In cases of bone involvement, bisphosphonates might be effective, but this benefit will be leveled out and will not be detected in the average OA population. Intensive treatment of inflammation might do more harm than good in the overall OA population¹⁸, but might be very helpful for subgroups of patients with evident inflammation. Identification of radiographic phenotypes is expected to improve through quantitative evaluation of separate features on radiographs.

The objectives of our study were to identify radiographic phenotypes of early knee OA and to describe their radiological and clinical characteristics.

MATERIALS AND METHODS

Study population; Cohort Hip & Cohort Knee (CHECK). Development of knee OA was evaluated from baseline to 5-year followup in CHECK (Cohort Hip & Cohort Knee). In this cohort, 1002 participants with pain and/or stiffness of hip and/or knee, age 45–65 years, and without a previous visit or with a first visit not more than 6 months previously to the general practitioner for these complaints were included¹⁹. At baseline, 82% of the participants had knee complaints (18% had hip complaints only), and the radiographic knee damage of the entire cohort was limited, with K-L grade in the knee of 0 in 81%, I in 16%, II in 3%, and III in 0.4%.

The study procedures were in accord with the standards of the medical ethics committees of all 10 participating hospitals and with the Helsinki Declaration of 1975 (revised 2000), and all participants gave their written informed consent.

Knee images digital analysis. Standardized weight-bearing semiflexed views [metatarsophalangeal (MTP), according to Buckland-Wright, *et al*^{20,21}] of both knees were acquired at baseline and 2-year and 5-year followup (T0, T2y, and T5y). Radiographs were analyzed for 14 separate OA variables by use of Knee Images Digital Analysis (KIDA)²²: minimum

joint space width (JSW, in mm), mean medial and mean lateral JSW, femur-tibia varus angle (in degrees), eminence height (to represent spiking of the tibial eminence; in mm), osteophyte area (in mm²) in lateral and medial femur and lateral and medial tibia, and bone density in these 4 compartments. The varus angle between the femur and tibia was determined in the frontal plane using the intersection points that determine the bone and cartilage interface; a positive value represents (more) varus alignment. Bone density was expressed in mmAl equivalents by comparison with an aluminum reference wedge that was added in each radiograph, which was found to be a reliable method to measure bone density²³. Intra- and interobserver variation of KIDA was described²². In the CHECK study¹⁹ the KIDA measurements were performed by 1 experienced observer (M. Lafeber) in random order blinded to information on timepoint, severity, and characteristics of individual patients. The numbers of analyzed knees are indicated and vary slightly for the different radiographic variables, because poor radiographic quality can hamper KIDA measurement. The intraobserver variation, tested by random reanalysis of 108 radiographs several months later, revealed good intra-observer variability (intraclass correlation coefficient 0.73–0.99) for the different features²⁴.

Statistical analysis. Using principal component analysis, the measurements of 14 separate KIDA variables were reduced into 5 components to represent the following radiographic features: (1) medial JSW (mean of 4 predefined locations); (2) lateral JSW (mean of 4 predefined locations); (3) osteophyte area (sum of lateral and medial femur and lateral and medial tibia); (4) eminence height (sum of lateral and medial); and (5) bone density (mean of lateral and medial femur and tibia)^{25,26}. By multiplying the factor loadings from the principal component analysis of the individual KIDA measurements, 5 component scores were calculated. These 5 component scores (standardized using z-scores) were used in a hierarchical cluster analysis (Ward's method) to identify possible phenotypes of progression of radiographic knee OA. Per individual, the component scores of the left and right knee at T0, T2y, T5y and the change scores (T5y – T2y and T2y – T0) were all used as separate variables in this analysis. The number of selected clusters of individuals was based on inspection of dendrograms (MBK, PMJW).

To interpret the clusters (phenotypes), the following straightforward and/or well known prespecified features were evaluated over time and compared between clusters: minimum JSW, medial JSW, lateral JSW, varus angle, osteophyte area (log-transformed sum of 4 compartments + 1; because normal distribution is preferred for statistical analysis), eminence height (sum of both), and bone density (mean of 4 compartments)²⁵. Further, the presence of knee and/or hip pain as assessed by the physician during physical examination per joint, and the Western Ontario and McMaster Universities Osteoarthritis index pain and function scores (WOMAC; 0–100 scale, 100 = worst condition), assessed as an overall measure for the individual, were compared between clusters.

Subsequently, logistic regression analyses were performed to evaluate whether the radiographic features measured at T0, in addition to demographic and clinical characteristics [age, sex, body mass index (BMI), erythrocyte sedimentation rate, and pain intensity measured by visual analog scale (VAS; 0–10 scale, 10 = worst possible pain)] at T0, could be used to predict to which specific phenotype an individual belonged. These analyses were performed in participants that were included in CHECK with at least knee complaints at T0, because these individuals visit a physician with early complaints and are suspected for development of radiographic knee OA. Univariate and multivariate analyses were performed. In the multivariate analyses, all variables were initially included, and were removed manually using a backward selection strategy to generate a model including only variables that are significantly related (based on p value < 0.05 and size of the OR) to the outcome. Models including only demographic and clinical variables were compared to models where radiographic features were added and to models where conventional K-L grading was added. To represent the total burden of radiographic damage, for each participant the sum of the left and right knee was used in the models. Since radiographic features might be very characteristic of an

individual specifically early in the disease²⁴, the difference between a knee and the contralateral knee for the radiographic features was also studied as an independent variable. These difference scores might detect small changes by using the contralateral knee as a reference in this early OA population with only subtle damage in 1 joint.

Prognostic ability of the final models was summarized and compared using the area under the curve (AUC) of the receiver-operating characteristic (ROC) curve. The AUC-ROC provides a measure for the ability to discriminate between a specific phenotype and the other phenotypes²⁷. Additionally, per phenotype the regression coefficients of the final models were corrected for overfitting using the van Houwelingen and Le Cessie method²⁸, and were converted into a simple score. Three cutoff points were determined: optimal sensitivity, optimal tradeoff between sensitivity and specificity, and optimal specificity. For these cutoff values, positive predictive values (PPV) were calculated as an estimate of predictive ability.

Analyses were performed using SPSS version 15.0 and SAS version 9.1.3; p value < 0.05 was considered statistically significant.

RESULTS

Identification of radiographic phenotypes. Based on development over time of component scores of the knees, 5 clusters could be identified. Participants were classified only when complete data of KIDA measurements were available for all 3 timepoints, leading to evaluation of 417 of 1002 participants. The 5 clusters were interpreted as follows: (1) “Severe”: severe progression; (2) “Bone density”: prominent involvement of the bone density feature; (3) “Early”: progression mainly in an early phase (T0 to T2y); (4) “Late”: progression mainly in a later phase (T2y to T5y); and (5) “No”: no progression.

Figure 1 depicts the development of the separate radiographic features over time per cluster, representing the level of progression, the prominent involvement of bone density, and the different phases of progression, as follows.

In general, the radiographic features showed OA progression during 5 years’ followup: minimum and medial JSW decreased, and lateral JSW, varus angle, osteophyte area, eminence height, and bone density increased. Participants in the Severe cluster ($n = 17$; 4% of 417 available participants; Figure 1) progressed more evidently than participants in the other clusters on all radiographic features. Notably, at T0 features of these participants were already more affected than those of participants in the other clusters. The Bone density cluster ($n = 113$; 27% of participants) represented prominent involvement of bone density at all 3 timepoints compared with the other phenotypes. In this cluster the other features were only mildly affected. Participants in the Early cluster ($n = 110$; 26% of participants) progressed mainly between T0 and T2y, most evidently for lateral JSW, varus angle, and bone density. Participants in the Late cluster ($n = 69$; 17% of participants) progressed mainly between T2y and T5y on lateral JSW, varus angle, and eminence height. In the No progression cluster ($n = 108$; 26% of participants) the radiographic features did not progress during followup; small changes in radiographic features might be due to random error.

Characterization of radiographic phenotypes. Baseline characteristics are depicted per phenotype in Table 1.

Radiographic and clinical development. For further interpretation of the phenotypes, Table 2 depicts K-L grades at T0, T2y, and T5y of both knees of participants within the 5 clusters. The frequency of K-L grades was statistically significantly different between the clusters (chi-square test, $p < 0.0001$ at all timepoints). Notably, in the Severe cluster a substantial percentage of knees already had evident radiographic OA (K-L grade \geq II) at T0, which increased over time, to 27% ($n = 9$ of 34 knees; 17 participants) at T0; 35% ($n = 12$ knees) at T2y; and 39% ($n = 13$ knees) at T5y. In the Bone density cluster the portion of knees with K-L grade \geq II was substantial, specifically at T5y. In the Early and Late cluster the percentage of knees with K-L grade = 0 at T0 was highest (compared to other clusters). In the No cluster only 2%, 4%, and 6% of knees ($n = 4$, 9, and 13 knees) had radiographic damage (K-L grade \geq II) at T0, T2y, and T5y, respectively.

Table 2 also depicts whether pain was present in the knee and/or the hip at T0, T2y, and T5y (bottom panel). The location of pain was significantly different between the phenotypes (chi-square test: $p = 0.002$ at T0, $p = 0.001$ at T2y, and $p < 0.0001$ at T5y). Participants with Severe radiographic progression specifically presented with knee pain. Participants in the Late cluster reported pain in “hip only” more commonly than participants in the other clusters, which might suggest (early) hip involvement, followed by knee involvement. Interestingly, a substantial portion of the participants reported “neither knee nor hip” pain at T2y and T5y, specifically in the No progression cluster. This may indicate that this phenotype is characterized by acute transient joint pain that does not lead to progressive radiographic joint damage.

Figure 2 depicts the development of the average WOMAC pain and function score over time per progression phenotype. The WOMAC scores were moderate at all timepoints and did not increase notably during followup. Although the average WOMAC pain score over time was statistically significantly lower in the No progression phenotype than in the Severe ($p = 0.003$), Bone density ($p = 0.02$), and Late ($p = 0.02$) progression phenotypes, the development over time was not significantly different between the phenotypes (tested using longitudinal regression analysis including an interaction term for time \times phenotype). Also, the average WOMAC function score was significantly lower in the No progression phenotype than in the Severe ($p = 0.004$), Bone density ($p = 0.03$), Early ($p = 0.04$), and Late ($p = 0.01$) phenotypes. Further, progression over time was significantly different between the No and Late phenotypes.

Membership in a phenotype. Which baseline variables are associated with belonging to a specific phenotype (compared to belonging to any of the other phenotypes) was

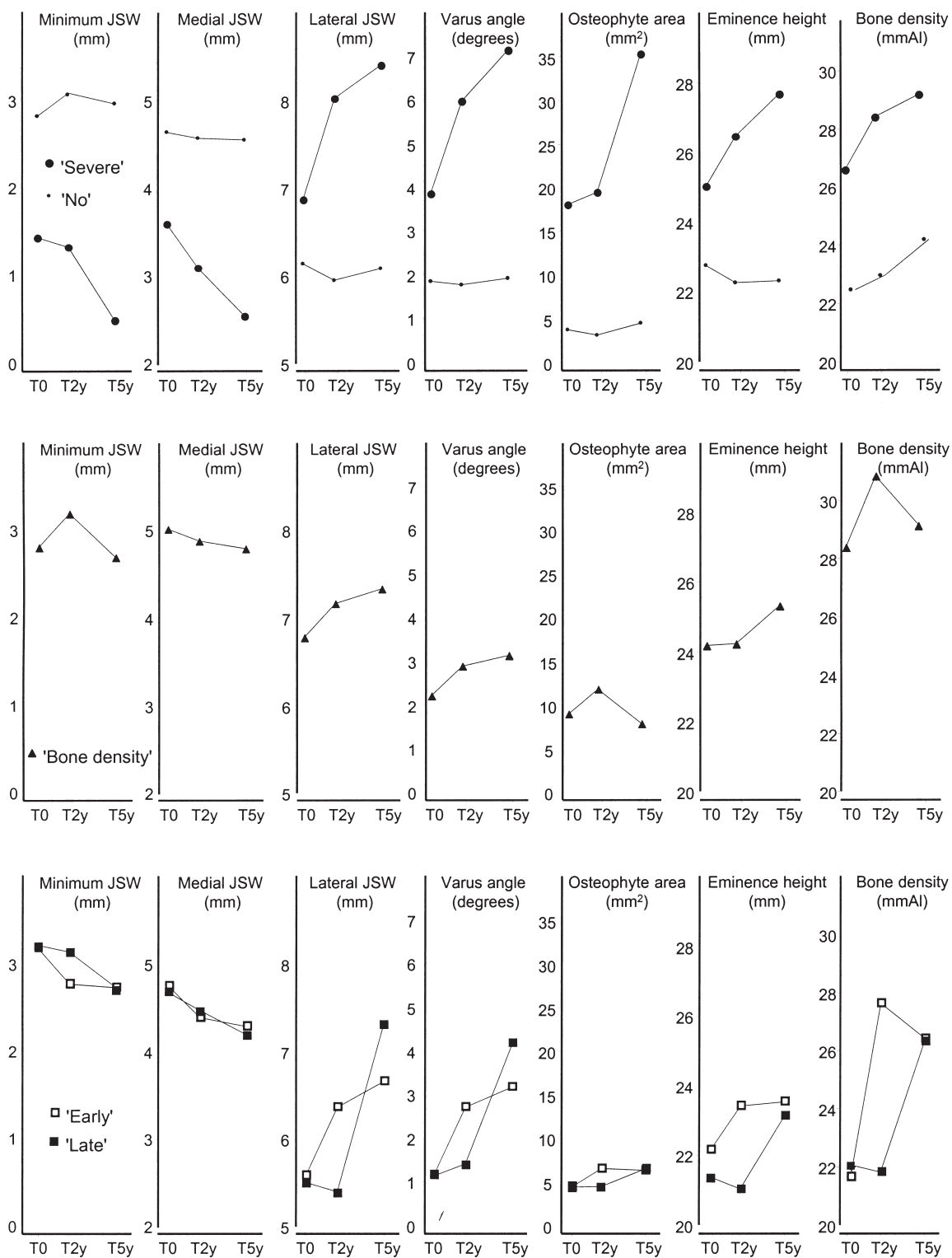


Figure 1. Development of the separate radiographic features per phenotype of knee osteoarthritis (right knee) over time per cluster. Results were similar for left knee (data not shown). Top row: clusters representing the level of progression (Severe progression group, n = 17; 4% of 417 available participants. No progression group, n = 108; 26% of participants). Middle row: prominent involvement of bone density (Bone density cluster, n = 113; 27% of participants). Bottom row: different phases of progression (Early cluster, n = 110; 26% of participants. Late cluster, n = 69; 17% of participants). JSW: joint space width.

Table 1. Baseline characteristics per cluster of radiographic knee osteoarthritis (OA).

Characteristic	Severe, n = 17	Bone Density, n = 113	Early, n = 110	Late, n = 69	No, n = 108	p Overall
Age, yrs	58 (4)	56 (5)	56 (5)	56 (5)	57 (5)	0.16
Female, %	82	55	81	88	92	< 0.0001
Body mass index, kg/m ²	27 (25–31)	27 (24–30)	24 (23–27)	24 (23–27)	24 (22–27)	< 0.0001
Erythrocyte sedimentation rate, mm/h	9 (5–15)	6 (4–12)	8 (5–15)	8 (5–13)	9 (5–15)	0.07

Data are mean (SD) or median (25–75th percentile). P values are based on ANOVA (age), chi-square (sex), and Kruskal-Wallis test (body mass index and erythrocyte sedimentation rate).

Table 2. Frequency of Kellgren-Lawrence (K-L) grades (0–IV) and location of pain (knee only, hip only, knee and hip, neither knee nor hip) over time per cluster of radiographic knee osteoarthritis.

	Severe, n = 17			Bone Density, n = 113			Early, n = 110			Late, n = 69			No, n = 108		
	T0	T2y	T5y	T0	T2y	T5y	T0	T2y	T5y	T0	T2y	T5y	T0	T2y	T5y
K-L [†] , %															
0	38	29	26	74	69	35	86	78	57	88	77	40	83	79	64
I	35	35	35	22	24	43	10	16	27	10	20	47	15	17	30
II	24	29	18	4	6	20	3	4	12	2	3	10	2	3	6
III	3	6	15	0	1	2	0	1	4	0	0	2	0	1	0
IV	0	0	6	0	0	1	0	0	0	0	0	0	0	0	0
Pain ^{††} , %															
Knee only	71	70	53	44	32	29	40	32	30	30	31	27	34	28	21
Knee and hip	24	12	35	42	38	40	41	46	39	33	37	41	49	33	29
Hip only	5	0	0	14	13	13	19	13	13	36	23	17	17	14	6
Neither knee nor hip	—*	18	12	—*	17	18	—*	8	18	—*	9	15	—*	25	44

[†] K-L: percentage of knees, taking into account the left and right knee of participants within cluster (e.g., 17 participants in Severe cluster for a total of 34 knees). ^{††} Pain: reflects whether an individual experiences pain in, e.g., any knee (1 or both). * Participants with neither knee nor hip pain were not included in the CHECK study. T0: baseline; T2y: 2-year followup; T5y: 5-year followup.

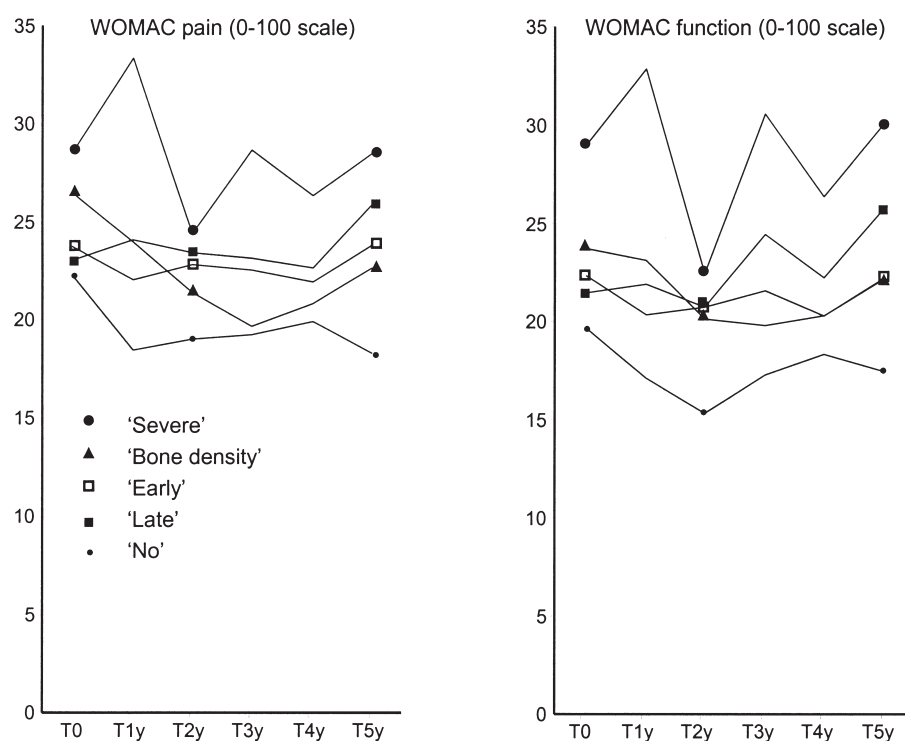


Figure 2. Progression of WOMAC scores (Western Ontario and McMaster Universities Osteoarthritis Index: 0–100 scale, 100 = worst) per phenotype of knee osteoarthritis.

evaluated in 336 participants with knee pain at T0, because these are the individuals suspected of developing radiographic knee OA. Table 3 presents a summary of these logistic regression analyses by depicting per phenotype the AUC of the multivariate model with radiographic features and demographic and clinical variables, and the direction of the effect for the significant dependent variables (+: OR > 1 and -: OR < 1). Details of the regression analyses are given in the Appendix.

Because the Severe phenotype consisted of only 16 participants with knee pain, multivariate analyses were not performed for this outcome. In the univariate evaluation, almost all radiographic features were significantly associated with the outcome (severe phenotype vs other phenotypes), as were K-L grade and BMI.

In general, the multivariate analyses showed that the discriminative ability (AUC-ROC) of the models improved when radiographic features were added to the demographic and clinical variables. The K-L grade was not significantly associated with any of the phenotypes. The predictors for Early, Late, and No progression phenotypes generally had an effect opposite to that of the predictors for the Severe and Bone density phenotypes.

Female sex reduced and higher BMI increased the risk of belonging to the Bone density phenotype together with multiple radiographic features (Table 3), resulting in a model with AUC-ROC = 0.91 (95% CI 0.88–0.94), and with AUC-ROC = 0.87 (95% CI 0.83–0.91) after correction for overfitting and rounding of coefficients. The PPV, the chance of belonging to the Bone density phenotype, was 83% in individuals with a score above the cutoff for optimal sensitivity (Table 4). The final predictive score was calculated as

$$0.3*\text{medial JSW} + 0.1*\text{varus angle} + 1*(\log[\text{osteophyte area} + 1]) + 0.05*\text{eminence height} + 0.1*\text{bone density} - 0.5*\text{gender (male is 1; female is 2)}$$

The Early progression phenotype was associated only with radiographic features. AUC-ROC of this model was 0.79 (95% CI 0.74–0.84) and decreased to 0.70 (95% CI 0.64–0.76) after correction for overfitting and rounding of coefficients. Final score was calculated as

$$-0.1*\text{varus angle} - 0.05*\text{eminence height} - 0.05*\text{bone density} - 0.3*\text{absolute difference in varus angle} - 0.2*\text{absolute difference in bone density}$$

Female sex and several radiographic features were associated with the Late progression phenotype; AUC-ROC was 0.76 (95% CI 0.69–0.83) and remained unchanged (predictive score $-0.2*\text{lateral JSW} - 0.1*\text{eminence height} - 0.05*\text{bone density}$).

Women and participants with lower BMI were more likely to belong to the No progression phenotype, and several radiographic features were also associated with this phenotype. Unexpectedly, individuals with a larger varus angle were more likely to belong to the No radiographic progression phenotype. The discriminative ability of the model was fair, with AUC-ROC = 0.72 (95% CI 0.66–0.78) decreasing to 0.68 (95% CI 0.62–0.74) for the predictive score $(-0.1*\text{varus angle} - 0.5*(\log[\text{osteophyte area} + 1]) + 0.2*\text{absolute difference in eminence height} + 0.05*\text{age} + 1*\text{gender} - 0.1*\text{BMI})$.

Table 4 shows PPV for the different cutoffs for the predictive scores per phenotype.

DISCUSSION

Our study describes a first attempt to identify specific

Table 3. Summary of regression analyses with phenotype as dependent variable.

AUC	Severe, NA	Bone Density, 0.91	Early, 0.79	Late, 0.76	No, 0.72
Radiographic feature					
Minimum JSW					
Medial JSW	+				
Lateral JSW			–		
Varus angle	+	–		+	
Osteophyte area	+			–	
Eminence height	+	–	–	+	
Bone density	+	–	–		
Demographic and clinical					
Age				+	
Female sex	–			+	
Body mass index				–	

Area under receiver-operator characteristic curve (AUC-ROC) depicted for multivariate model with radiographic features, and demographic and clinical variables as dependent variables (measured at baseline). For the significant dependent variables + represents OR > 1 (and 95% CI > 1), and – represents OR < 1 (and 95% CI < 1). NA: multivariate analyses not performed because of limited n-value; JSW: joint space width.

Table 4. Ability to predict radiographic phenotypes for 3 cutoff points of predictive score.

	% (n) Present	Predictive Score		% Identified (> cutoff)
		Cutoff	PPV, %	
Severe	5 (16)	NA	NA	NA
Bone density	29 (97)	Sn 9.00	53	54
		S&S 10.30	56	44
		Sp 12.00	83	14
Early	26 (89)	Sn -5.60	33	77
		S&S -5.00	40	50
		Sp -4.10	48	10
Late	13 (44)	Sn -9.80	20	63
		S&S -9.00	28	34
		Sp -8.00	38	8
No	27 (90)	Sn 0.50	34	77
		S&S 1.30	35	50
		Sp 2.40	51	11

NA: not applicable; Sn: optimal sensitivity; S&S: optimal tradeoff between sensitivity and specificity; Sp: optimal specificity; PPV: positive predictive value.

phenotypes of progression of radiographic knee OA, specifically in participants with complaints of early OA. Phenotypes were found to represent the level of disease progression (Severe and No progression), the phase of progression (Early and Late), and the prominent involvement of Bone density. Although the definition of the phenotypes should be validated in other datasets, these phenotypes might represent a (partly) different etiology. Such phenotypes may benefit from different treatment strategies, e.g., an intense regimen that combines pain medication with cartilage-safe nonsteroidal antiinflammatory drugs in cases of the Severe phenotype, and treatment aimed at bone quality (e.g., bisphosphonates) in cases of the Bone density phenotype. The percentage of participants with K-L grade \geq II was significantly higher in the Severe cluster than in the other clusters. Clinical characteristics were not evidently different between the clusters, and the WOMAC scores were only slightly lower in the No cluster than in the other clusters. This is in agreement with the limited relation between clinical and radiographic OA in earlier studies^{6,8,29}.

Previous work investigating possible subtypes of radiographic joint damage was performed in more established OA^{13,30,31}. We found that it is of high value to evaluate phenotypes in an early phase of the disease, because this might enable early intervention before structural damage is established. That we were able to identify specific phenotypes with different progression of radiographic features of OA using detailed KIDA measurement justifies continuing development of more precise evaluation of plain radiographs³² in the early phase of OA. For example, the finding by Oka, *et al* that varus alignment is a predictor for progression of OA³² emphasizes that this radiographic feature should be measured separately. Adding specific separate radiographic features to demographic and

clinical characteristics also substantially improved ability to discriminate between the progression phenotypes, contrary to K-L grading of overall damage. Applying measurements of specific separate radiographic features in clinical trials is therefore recommended, and this will also enable our results to be confirmed and extended.

Female sex³⁴ and BMI^{35,36} are known risk factors for onset and progression of OA, and were also identified as predictors for most (but not all) phenotypes of radiographic progression in this study. Interestingly, being female was protective of belonging to the Bone density phenotype, and was significantly (OR 3.87) associated with belonging to the No progression phenotype. This might be related to the fact that women have lower bone density than men³⁷. Osteophyte area was identified as the most important predictor for Severe progression and Bone density involvement, and was protective for the No progression phenotype, which might support the notion that osteophyte formation is assumed to occur early in the disease¹⁷. Unexpectedly, however, osteophyte area was not associated with belonging to the Early phenotype, and this requires further evaluation. The radiographic features that were identified to be associated with the Early and Late progression phenotypes (e.g., eminence height, bone density, varus angle, and JSW) actually had a protective effect, which also calls for further evaluation.

Generally, the PPV based on the predictive scores using demographic and clinical characteristics combined with specific radiographic features were not high enough for prediction at the individual level. However, defining subgroups for inclusion in clinical trials might be significantly improved (e.g., smaller groups needed; less time-consuming and more cost-efficient studies) based on these scores and hence enable the development of a more personalized treatment approach. For instance, 54% of our

population could be classified as belonging to the Bone density phenotype with a certainty of 53% (PPV) when the predictive score was > 9.0. In our study, overall, 24% of participants (113 of 417) were classified as belonging to the Bone density phenotype, so by using the predictive score a substantially different (sub)population can potentially be identified that might react differently to treatment. Clearly, however, these phenotypes and prediction models should be validated before they are used in this way.

Cluster analysis is a technique to group individuals who are “similar” regarding the variables that are included in the analysis. To derive a set of phenotypes, “subjective” choices also have to be made. The value of clustering individuals is determined by the relevance and characteristics of the clusters, in our case underlying etiology, disease severity, need for treatment, and longterm outcome. Performing a cluster analysis with a different set of variables, for instance including clinical characteristics, might result in different clusters, e.g., phenotypes in which radiographic and clinical characteristics are strongly related to each other. Also, when such evaluations can be verified in an even larger population, this can limit overfitting of the model by evaluating a large number of variables in a relatively small population.

In our study, cluster analysis aimed at identifying radiographic progression phenotypes by exploring radiographic features at and between different timepoints. We also deliberately chose to cluster participants and not knees. When performing cluster analysis with radiographic features at T0, T2y, and T5y separately, a Severe cluster with involvement of all feature scores, a cluster with Bone density involvement, and a cluster with No progression of all feature scores were identified, which adds to the validity of the defined progression clusters. Of note, no clusters were identified with specific progression of, for example, 1 knee (and not the other knee). Radiographic features within an individual might explain this; they are quite similar, and small differences are overlooked because of much larger differences between individuals or knees²⁴. Also, this finding might be a reflection of the systemic character of OA³⁸ affecting the whole joint and also more joints within an individual³⁹. This might also be the reason that the scores of differences of the radiographic features were not related to membership in a specific phenotype.

Limitations of our study are that the number of participants was evidently decreased by the requirement for complete data for both knees at all 3 timepoints. However, this was not considered to be systematic bias because the reason for missing data was only radiographic quality. Age, sex, pain, and K-L grading were comparable between the participants who were and those who were not included in our analyses. Importantly, we did not select radiographs that had perfect tibial alignment. Although this might have influenced outcome regarding, for example, JSW⁴⁰, this approach most closely represents clinical trial practice.

Further, although it seems intuitive that the different radiographic features at baseline are associated with the phenotypes, this was not the case. It was found that radiographic features at baseline were associated with the development over 5-year followup, because radiographs at 3 timepoints were assessed to define the phenotypes. It was the detailed evaluation of the separate radiographic features that enabled identification of phenotypes, which could not have been done in this early phase of the disease by K-L grading (because only a small portion of participants had radiographic OA based on K-L grading).

Because our results represent a first attempt to define different phenotypes of OA based on radiographic features in early OA, results should be replicated and further validated. Future investigation might also include clinical OA characteristics and other measurements regarding structural joint damage, e.g., MRI¹⁰, to further define subgroups of OA.

Based on separate radiographic features, phenotypes with different levels and phases of progression and prominent involvement of “bone density” were detected in our cohort of participants with early complaints related to OA. These phenotypes might represent potential subgroups for the evaluation of preventive therapies in clinical trials and the discovery of better-targeted treatment strategies.

ACKNOWLEDGMENT

We acknowledge Marja Lafeber for KIDA analyses. The CHECK steering committee comprises 16 members with expertise in different fields of OA and is chaired by Prof. J.W.J. Bijlsma and coordinated by J. Wesseling. We acknowledge the participation of the following institutions: Academic Hospital Maastricht; Erasmus Medical Center Rotterdam; Jan van Breemen Institute – VU Medical Center Amsterdam; Kennemer Gasthuis Haarlem; Martini Hospital Groningen – Allied Health Care Center for Rheumatology and Rehabilitation Groningen; Medical Spectrum Twente Enschede – Twenteborg Hospital Almelo; St. Maartenskliniek Nijmegen; Leiden University Medical Center; and University Medical Center Utrecht and Wilhelmina Hospital Assen.

REFERENCES

1. Nevitt MC. Definition of hip osteoarthritis for epidemiological studies. *Ann Rheum Dis* 1996;55:652-5.
2. Bijlsma JW, Berenbaum F, Lafeber FP. Osteoarthritis: An update with relevance for clinical practice. *Lancet* 2011;377:2115-26.
3. McAlindon T, Dieppe P. Osteoarthritis: definitions and criteria. *Ann Rheum Dis* 1989;48:531-2.
4. Peat G, Greig J, Wood L, Wilkie R, Thomas E, Croft P. Diagnostic discordance: We cannot agree when to call knee pain ‘osteoarthritis’. *Fam Pract* 2005;22:96-102.
5. Boegard TL, Rudling O, Petersson IF, Jonsson K. Joint space width of the tibiofemoral and of the patellofemoral joint in chronic knee pain with or without radiographic osteoarthritis: A 2-year follow-up. *Osteoarthritis Cartilage* 2003;11:370-6.
6. Dieppe PA, Cushnaghan J, Shephstone L. The Bristol ‘OA500’ study: Progression of osteoarthritis (OA) over 3 years and the relationship between clinical and radiographic changes at the knee joint. *Osteoarthritis Cartilage* 1997;5:87-97.
7. Hochberg MC, Lawrence RC, Everett DF, Cornoni-Huntley J. Epidemiologic associations of pain in osteoarthritis of the knee:

APPENDIX. Details of regression analyses with phenotype as dependent variable.

'Severe' progression							
	Univariate		p	Multivariate (NA)			
	OR	(95% CI)					
Demographic and clinical							
Age	1.08	(0.97-1.19)	0.15				
Female sex	1.28	(0.36-4.62)	0.71				
BMI	1.18	(1.07-1.30)	0.001				
ESR	0.99	(0.93-1.06)	0.84				
Pain intensity	1.20	(0.96-1.51)	0.12				
Radiographic							
Feature sum							
Minimum JSW	0.58	(0.46-0.74)	<0.001				
Medial JSW	0.49	(0.35-0.68)	<0.001				
Lateral JSW	1.15	(0.95-1.40)	0.15				
Varus angle	1.37	(1.13-1.60)	<0.001				
Osteophyte	15.48	(4.53-52.9)	<0.001				
Eminence	1.13	(1.05-1.23)	0.002				
Bone density	1.04	(1.00-1.09)	0.06				
Feature abs d							
Minimum JSW	2.85	(1.69-4.79)	<0.001				
Medial JSW	4.15	(2.18-7.90)	<0.001				
Lateral JSW	1.68	(1.18-2.38)	0.004				
Varus angle	0.57	(0.45-0.74)	<0.001				
Osteophyte	31.0	(6.1-158.1)	<0.001				
Eminence	1.72	(1.29-2.28)	<0.001				
Bone density	0.90	(0.64-1.26)	0.54				
K-L sum	2.62	(1.79-3.84)	<0.001				
K-L abs d	2.17	(0.76-6.20)	0.15				
'Bone density' involvement							
	Univariate		p	Multivariate		p	AUC
	OR	(95% CI)			OR		
Demographic and clinical				Demographic and clinical		0.66	
Age	0.97	(0.93-1.02)	0.26				
Female sex	0.19	(0.11-0.32)	<0.001	Female sex	0.17	(0.10-0.31)	<0.001
BMI	1.13	(1.06-1.20)	<0.001	BMI	1.14	(1.07-1.21)	<0.001
ESR	0.96	(0.93-0.99)	0.03				
Pain intensity	1.06	(0.95-1.18)	0.33				
Radiographic				Radiographic (demographic and clinical)			
Feature sum				Feature sum			0.91
Minimum JSW	0.98	(0.88-1.09)	0.67				
Medial JSW	1.25	(1.09-1.41)	0.01	Medial JSW	1.37	(1.13-1.65)	0.001
Lateral JSW	1.40	(1.26-1.57)	<0.001				
Varus angle	1.15	(1.06-1.24)	<0.001	Varus angle	1.14	(1.02-1.28)	0.02
Osteophyte	3.38	(2.23-5.14)	<0.001	Osteophyte	2.82	(1.64-4.82)	<0.001
Eminence	1.11	(1.06-1.16)	<0.001	Eminence	1.09	(1.02-1.16)	0.007
Bone density	1.15	(1.11-1.19)	<0.001	Bone density	1.16	(1.12-1.21)	<0.001
Feature abs d							
Minimum JSW	1.20	(0.84-1.71)	0.31				
Medial JSW	0.95	(0.60-1.51)	0.84				
Lateral JSW	1.26	(0.98-1.61)	0.07				
Varus angle	0.90	(0.76-1.07)	0.22				
Osteophyte	2.71	(1.40-5.26)	0.003				
Eminence	0.87	(0.72-1.05)	0.15				
Bone density	1.22	(1.07-1.39)	0.003				
				Female sex	0.45	(0.20-0.98)	0.04
K-L sum	1.12	(0.88-1.43)	0.36				
K-L abs d	1.11	(0.63-1.95)	0.71				

'Early' progression							
	Univariate			Multivariate			
	OR	(95% CI)	p		OR	(95% CI)	p
Demographic and clinical				Demographic and clinical			
Age	0.99	(0.94-1.04)	0.68				
Female sex	1.11	(0.61-1.99)	0.74				
BMI	0.96	(0.91-1.03)	0.26				
ESR	1.01	(0.98-1.04)	0.44				
Pain intensity	0.92	(0.82-1.04)	0.17				
Radiographic				Radiographic (demographic and clinical)			
Feature sum				Feature sum			0.79
Minimum JSW	1.13	(1.01-1.28)	0.04				
Medial JSW	0.98	(0.86-1.12)	0.80				
Lateral JSW	0.76	(0.68-0.85)	<0.001				
Varus angle	0.80	(0.73-0.88)	<0.001	Varus angle	0.85	(0.77-0.95)	0.003
Osteophyte	0.54	(0.38-0.76)	<0.001				
Eminence	0.92	(0.88-0.97)	0.001	Eminence	0.93	(0.89-0.98)	0.006
Bone density	0.92	(0.90-0.95)	<0.001	Bone density	0.94	(0.91-0.97)	<0.001
Feature abs d							
Minimum JSW	0.67	(0.43-1.04)	0.07				
Medial JSW	0.82	(0.50-1.35)	0.43				
Lateral JSW	0.56	(0.37-0.84)	0.006				
Varus angle	0.64	(0.49-0.84)	0.001	Varus angle	0.75	(0.57-0.98)	0.04
Osteophyte	0.58	(0.29-1.17)	0.13				
Eminence	0.81	(0.66-0.99)	0.05				
Bone density	0.75	(0.62-0.91)	0.004	Bone density	0.78	(0.63-0.98)	0.03
K-L sum	0.73	(0.54-1.00)	0.05				
K-L abs d	0.54	(0.28-1.05)	0.07				
'Late' progression							
	Univariate			Multivariate			
	OR	(95% CI)	p		OR	(95% CI)	p
Demographic and clinical				Demographic and clinical			
Age	0.98	(0.92-1.04)	0.45				
Female sex	7.13	(1.08-30.17)	0.008	Female sex	7.13	(1.08-30.17)	0.008
BMI	0.94	(0.86-1.02)	0.15				
ESR	1.00	(0.96-1.04)	0.96				
Pain intensity	1.02	(0.88-1.19)	0.79				
Radiographic				Radiographic (demographic and clinical)			
Feature sum				Feature sum			0.76
Minimum JSW	1.08	(0.93-1.26)	0.30				
Medial JSW	0.92	(0.77-1.09)	0.33				
Lateral JSW	0.77	(0.67-0.89)	<0.001	Lateral JSW	0.83	(0.71-0.98)	0.02
Varus angle	0.85	(0.77-0.95)	0.005				
Osteophyte	0.58	(0.38-0.89)	0.01				
Eminence	0.88	(0.83-0.94)	<0.001	Eminence	0.91	(0.85-0.97)	0.004
Bone density	0.94	(0.91-0.97)	0.001	Bone density	0.95	(0.92-0.99)	0.006
Feature abs d							
Minimum JSW	0.69	(0.38-1.25)	0.22				
Medial JSW	0.96	(0.51-1.79)	0.89				
Lateral JSW	0.78	(0.50-1.22)	0.27				
Varus angle	0.92	(0.71-1.20)	0.55				
Osteophyte	0.46	(0.18-1.17)	0.10				
Eminence	0.88	(0.68-1.14)	0.34				
Bone density	1.01	(0.85-1.20)	0.89				
K-L sum	0.77	(0.51-1.18)	0.23				
K-L abs d	0.82	(0.36-1.86)	0.63				

'No' progression							
	Univariate			Multivariate			
	OR	(95% CI)	p	OR	(95% CI)	p	AUC
Demographic and clinical				Demographic and clinical	0.68		
Age	1.04	(0.99-1.09)	0.15				
Female sex	3.92	(1.80-8.52)	0.001	Female sex	3.87	(1.75-8.54)	0.001
BMI	0.87	(0.81-0.94)	<0.001	BMI	0.87	(0.81-0.94)	<0.001
ESR	1.03	(1.00-1.06)	0.10				
Pain intensity	0.97	(0.86-1.08)	0.55				
Radiographic Feature sum				Radiographic (demographic and clinical) Feature sum			0.72
Minimum JSW	1.00	(0.89-1.11)	0.96				
Medial JSW	0.98	(0.86-1.11)	0.74				
Lateral JSW	1.03	(0.93-1.13)	0.59				
Varus angle	1.05	(0.97-1.13)	0.24	Varus angle	1.09	(1.00-1.19)	0.04
Osteophyte	0.60	(0.43-0.84)	0.003	Osteophyte	0.60	(0.41-0.87)	0.007
Eminence	0.99	(0.96-1.04)	0.80				
Bone density	0.97	(0.95-0.99)	0.009				
Feature abs d				Feature abs d			
Minimum JSW	0.90	(0.61-1.33)	0.59				
Medial JSW	0.63	(0.36-1.10)	0.10				
Lateral JSW	1.03	(0.19-1.35)	0.82				
Varus angle	1.00	(0.83-1.19)	0.97				
Osteophyte	0.41	(0.20-0.83)	0.01				
Eminence	1.22	(1.03-1.45)	0.03	Eminence	1.25	(1.03-1.51)	0.02
Bone density	0.99	(0.86-1.13)	0.84				
				Age	1.06	(1.00-1.11)	0.04
				Female sex	3.81	(1.69-8.62)	0.001
				BMI	0.88	(0.82-0.95)	0.001
K-L sum	0.86	(0.65-1.14)	0.29				
K-L abs d	1.39	(0.78-2.47)	0.27				

Independent variables: measures at baseline (T0); Feature and K-L abs d: absolute difference between the radiographic measurement of the right and left knee; NA: multivariate analyses not performed due to limited n-value; abs d: absolute difference; osteophyte: osteophyte area; eminence: eminence height; AUC: area under receiver operator characteristic curve (AUC-ROC).

- Data from the National Health and Nutrition Examination Survey and the National Health and Nutrition Examination-I Epidemiologic Follow-up Survey. *Semin Arthritis Rheum* 1989;18 Suppl 2:4-9.
- Kinds MB, Welsing PM, Vignon EP, Bijlsma JW, Viergever MA, Marijnissen AC, et al. A systematic review of the association between radiographic and clinical osteoarthritis of hip and knee. *Osteoarthritis Cartilage* 2011;19:768-78.
 - Lethbridge-Cejku M, Scott WW Jr, Reichle R, Ettinger WH, Zonderman A, Costa P, et al. Association of radiographic features of osteoarthritis of the knee with knee pain: Data from the Baltimore Longitudinal Study of Aging. *Arthritis Care Res* 1995;8:182-8.
 - Guerhazi A, Roemer FW, Burstein D, Hayashi D. Why radiography should no longer be considered a surrogate outcome measure for longitudinal assessment of cartilage in knee osteoarthritis. *Arthritis Res Ther* 2011;13:247.
 - Duryea J, Neumann G, Niu J, Totterman S, Tamez J, Dabrowski C, et al. Comparison of radiographic joint space width with magnetic resonance imaging cartilage morphometry: Analysis of longitudinal data from the Osteoarthritis Initiative. *Arthritis Care Res* 2010;62:932-7.
 - Buckland-Wright JC. Quantitative radiography of osteoarthritis. *Ann Rheum Dis* 1994;53:268-75.
 - Kerkhof HJ, Meulenbelt I, Akune T, Arden NK, Aromaa A, Bierma-Zeinstra SM, et al. Recommendations for standardization and phenotype definitions in genetic studies of osteoarthritis: The TREAT-OA Consortium. *Osteoarthritis Cartilage* 2011;19:254-64.
 - Banks SE. Erosive osteoarthritis: A current review of a clinical challenge. *Clin Rheumatol* 2011;29:697-706.
 - Hunter DJ, Gerstenfeld L, Bishop G, Davis AD, Mason ZD, Einhorn TA, et al. Bone marrow lesions from osteoarthritis knees are characterized by sclerotic bone that is less well mineralized. *Arthritis Res Ther* 2009;11:R11.
 - van der Kraan PM, van den Berg WB. Osteophytes: relevance and biology. *Osteoarthritis Cartilage* 2007;15:237-44.
 - Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthritis. *Ann Rheum Dis* 1957;16:494-502.
 - Ding C, Cicuttini F, Jones G. Do NSAIDs affect longitudinal changes in knee cartilage volume and knee cartilage defects in older adults? *Am J Med* 2009;122:836-42.
 - Wesseling J, Dekker J, van den Berg WB, Bierma-Zeinstra SM, Boers M, Cats HA, et al. CHECK (Cohort Hip and Cohort Knee): Similarities and differences with the Osteoarthritis Initiative. *Ann Rheum Dis* 2009;68:1413-9.
 - Buckland-Wright C. Protocols for precise radio-anatomical

- positioning of the tibiofemoral and patellofemoral compartments of the knee. *Osteoarthritis Cartilage* 1995;3 Suppl A:71-80.
21. 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.
 22. Marijnissen AC, Vincken KL, Vos PA, Saris DB, Viergever MA, Bijlsma JW, et al. Knee Images Digital Analysis (KIDA): A novel method to quantify individual radiographic features of knee osteoarthritis in detail. *Osteoarthritis Cartilage* 2008;16:234-43.
 23. Kinds MB, Bartels LW, Marijnissen AC, Vincken KL, Viergever MA, Lafeber FP, et al. Feasibility of bone density evaluation using plain digital radiography. *Osteoarthritis Cartilage* 2011;19:1343-8.
 24. Kinds MB, Vincken KL, Vignon E, ten Wolde S, Bijlsma JW, Welsing PM, et al. Radiographic features of knee and hip osteoarthritis represent characteristics of an individual, in addition to severity of osteoarthritis. *Scand J Rheumatol* 2012;41:141-9.
 25. Kinds MB, Marijnissen AC, Bijlsma JW, Boers M, Lafeber FP, Welsing PM. Quantitative radiographic features of early knee osteoarthritis: Development over 5 years and relationship with symptoms in the CHECK Cohort. *J Rheumatol* 2013;40:58-65.
 26. Kinds MB, Marijnissen AC, Vincken KL, Viergever MA, Drossaers-Bakker KW, Bijlsma JW, 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. *Osteoarthritis Cartilage* 2012;20:548-56.
 27. Muller MP, Tomlinson G, Marrie TJ, Tang P, McGeer A, Low DE, et al. Can routine laboratory tests discriminate between severe acute respiratory syndrome and other causes of community-acquired pneumonia? *Clin Infect Dis* 2005;40:1079-86.
 28. Van Houwelingen JC, Le Cessie S. Predictive value of statistical models. *Stat Med* 1990;9:1303-25.
 29. Bedson J, Croft PR. The discordance between clinical and radiographic knee osteoarthritis: A systematic search and summary of the literature. *BMC Musculoskelet Disord* 2008;9:116.
 30. Kumm J, Tamm A, Lintrop M, Tamm A. The prevalence and progression of radiographic knee osteoarthritis over 6 years in a population-based cohort of middle-aged subjects. *Rheumatol Int* 2012;32:3545-50.
 31. Felson DT. Identifying different osteoarthritis phenotypes through epidemiology. *Osteoarthritis Cartilage* 2010;18:601-4.
 32. Oka H, Muraki S, Akune T, Mabuchi A, Suzuki T, Yoshida H, et al. Fully automatic quantification of knee osteoarthritis severity on plain radiographs. *Osteoarthritis Cartilage* 2008;16:1300-6.
 33. Brouwer GM, van Tol AW, Bergink AP, Belo JN, Bernsen RM, Reijman M, et al. Association between valgus and varus alignment and the development and progression of radiographic osteoarthritis of the knee. *Arthritis Rheum* 2007;56:1204-11.
 34. Felson DT, Zhang Y. An update on the epidemiology of knee and hip osteoarthritis with a view to prevention. *Arthritis Rheum* 1998;41:1343-55.
 35. Grotle M, Hagen KB, Natvig B, Dahl FA, Kvien TK. Obesity and osteoarthritis in knee, hip and/or hand: An epidemiological study in the general population with 10 years follow-up. *BMC Musculoskelet Disord* 2008;9:132.
 36. Heliövaara M, Makela M, Impivaara O, Knekt P, Aromaa A, Sievers K. Association of overweight, trauma and workload with coxarthrosis. A health survey of 7,217 persons. *Acta Orthop Scand* 1993;64:513-8.
 37. Breijawi N, Eckardt A, Pitton MB, Hoelzl AJ, Giesa M, von Stechow D, et al. Bone mineral density and vitamin D status in female and male patients with osteoarthritis of the knee or hip. *Eur Surg Res* 2008;42:1-10.
 38. Kraus VB, Kepler TB, Stabler T, Renner J, Jordan J. First qualification study of serum biomarkers as indicators of total body burden of osteoarthritis. *PLoS One* 2010;5:e9739.
 39. Hunter DJ, Felson DT. Osteoarthritis. *BMJ* 2006;332:639-42.
 40. Kinds MB, Vincken KL, Hoppinga TN, Bleys RL, Viergever MA, Marijnissen AC, et al. Influence of variation in semiflexed knee positioning during image acquisition on separate quantitative radiographic parameters of osteoarthritis, measured by Knee Images Digital Analysis. *Osteoarthritis Cartilage* 2012; 20:997-1003.