

## Original article

## Distinct subtypes of knee osteoarthritis: data from the Osteoarthritis Initiative

Jan H. Waarsing<sup>1</sup>, Sita M. A. Bierma-Zeinstra<sup>1,2</sup> and Harrie Weinans<sup>3,4</sup>

## Abstract

**Objective.** OA is suspected to be a collection of distinct subtypes, each with different aetiology and clinical characteristics. We aimed to explore the existence of different subtypes of knee OA, using cluster analysis of the data of the OA Initiative.

**Methods.** We used latent class cluster analysis (LCA) to cluster baseline data of 518 subjects of the OA Initiative progression cohort. Data included radiographic scores of OA features per compartment, regional quantitative MRI measures of cartilage quantity and denuded bone, and self-reported clinical scores on knee symptoms. To ensure that the clusters were found independently of OA severity, the LCA model was corrected with a measure of OA severity. The resulting clusters were compared with respect to the presence of risk factors and progression.

**Results.** LCA resulted in four clusters containing 47%, 27%, 15% and 12% of the subjects. Clusters 1, 2 and 4 showed OA features at the medial compartment, while cluster 3 only showed lateral OA features. Clusters 3 and 4 showed severe increases in areas of denuded bone, whereas no denuded bone was present in cluster 1. Prevalence of OA progression over 24 months was highest in clusters 3 and 4 and lowest in cluster 1. The clusters also differed significantly in BMI, knee alignment and prevalence of reported trauma.

**Conclusion.** LCA confirmed the existence of distinct subtypes of knee OA with clear differences in structural degradation and symptoms. The fact that subtypes also differed in risk factors suggests that different causes lead to different types of knee OA.

**Key words:** knee osteoarthritis, cluster analysis, phenotype.

## Rheumatology key messages

- Four subtypes of knee OA were identified with distinct patterns of structural degradation.
- The identified subtypes of knee OA differed in risk factors and progression rates, indicating clinical relevance.

## Introduction

One possible explanation of the apparent complex nature and heterogeneity of OA is that the disease is a collection of different subtypes or OA phenotypes [1–3]. A phenotype is defined as a collection of observable traits that arise out of an interaction between environmental and

genetic factors. Distinct subtypes exist when patients can be grouped such that the variation of these phenotypic traits within a subtype is smaller than the variation between subtypes. Distinct phenotypes would suggest distinct underlying causes, whether genotypical or environmental, which could be highly relevant for understanding and treating the disease.

A few studies have been published that have attempted to identify distinct OA phenotypes, all approaching the problem from different angles. At times a distinction has been made between atrophic and hypertrophic OA, distinguishing OA that is mainly characterized by degradation of cartilage from OA that shows predominantly a bony response, leading to osteophytes [4]; however, in the knee, pure atrophic and hypertrophic phenotypes appear to be rather rare [5]. Another approach has been

<sup>1</sup>Department of Orthopaedics, <sup>2</sup>Department of General Practice, Erasmus Medical Center, Rotterdam, <sup>3</sup>Department of Orthopaedics & Rheumatology, UMC Utrecht and <sup>4</sup>Department of Biomedical Engineering, Delft University of Technology, Delft, The Netherlands  
Submitted 2 April 2014; revised version accepted 25 February 2015

Correspondence to: Jan H. Waarsing, Department of Orthopedics, Erasmus Medical Center, Room Ee16-14, Dr Molenwaterplein 50, Rotterdam 3015 GE, The Netherlands.  
E-mail: e.waarsing@erasmusmc.nl

to subdivide OA based on causal factors. One study divides primary OA into OA related to oestrogen deficiency, OA induced genetically and OA due to ageing [6]. Each type appears to have its own aetiological and clinical characteristics.

Most of the above-mentioned studies present hypothesized phenotypes. A more methodological approach to finding OA subtypes would be to use cluster analysis. Knoop *et al.* [7] used *k*-means clustering on clinically relevant patient characteristics and found five distinct phenotypes of knee OA. Another interesting approach is using latent class growth analysis to identify distinct trajectories of OA progression. Bartlett *et al.* [8] identified subtypes with differing rates of progression in joint space narrowing (JSN). Similarly, Verkleij *et al.* [9] identified distinct trajectories in hip pain progression over 2 years. Both studies found that subjects in distinct trajectories also differed in other OA and demographic characteristics.

The approach in this study was to identify distinct phenotypes of knee OA, based on differences in observable OA traits that could be considered to be caused by the progressing disease. Typically, these are measures of structural joint degradation and clinical symptoms. The rationale behind this approach is the idea that if aetiologically distinct subtypes exist, the differences in OA processes connected to these subtypes might lead to distinct observable traits. Thus, these subtypes can then naturally—and perhaps only—be found by cluster analysis of a wide range of these observable traits. We extracted data from the OA Initiative (OAI) and defined subtypes using latent class cluster analysis (LCA), which is a powerful and model-based clustering approach [10] that previously has been used successfully to identify subtypes in other diseases [11, 12].

## Methods

### Study population

The data used in this study are part of the OAI, which is a large multi-centre USA-based prospective observational cohort study of knee OA, for which the data are freely available (<https://oai.epi-ucsf.org>). Since we needed subjects with established knee OA, we used the baseline data of the progression cohort and not the incidence cohort. While the incidence cohort consists of subjects at high risk for knee OA, the progression cohort consists of subjects who, upon inclusion, presented with frequent knee symptoms for at least 1 month in the past year and, in a knee with these symptoms, showed radiographic knee OA (definite osteophytes). More specifically, we used the data for 600 knees in the central reading dataset project 09. If for a subject more than one knee was available, we randomly chose between left or right knee.

### Study measures

We used baseline data, first of semi-quantitative radiographic readings (Kellgren–Lawrence (KL), osteophytes, JSN, cysts, sclerosis, chondrocalcinosis and attrition, per compartment for the tibia and femur). Radiographs

were obtained according to a fixed-flexion protocol. Scores were read according to the OA Research Society International atlas [13]. The second set of data included region-specific quantitative MRI measures of cartilage thickness, and relative volume and relative areas of denuded bone. The scans were analysed by Chondrometrics GmbH (Ainring, Germany) and Paracelcus University (Salzburg, Austria), following the same protocol [14]. The third set of data included scores of OA symptoms for the knee obtained through questionnaires (WOMAC pain, function and disability, Visual Analogue Scale (VAS) pain during the past month, VAS pain during the past week, knee baseline symptom status).

After cluster analysis, the resulting clusters were compared using baseline demographic data and data on specific knee OA risk factors: age, gender, BMI, knee alignment (measured by goniometer while standing), isometric muscle strength (maximum force during isometric contraction and mean of flexion and extension, measured by the Good Strength Chair), self-reported knee trauma (knee ever injured badly enough to limit walking for at least 1 week) and presence of Heberden's nodes (visual inspection, defined as more than one bony enlargement). Also, we tested cluster differences with respect to OA progression in JSN and KL scores separately. Using JSN scores and KL scores at the 24-month follow-up, a case was defined as being progressive for each score when it increased by at least one grade during the 24 months.

### Pre-processing of the data

The OAI data offer quantitative MRI measures for many distinct areas in the tibiofemoral knee compartment [15], resulting in a large set of strongly correlated measures. We used principal component analysis (PCA) to reduce the number of measures, and this was performed separately on the 54 measures of cartilage volume and thickness and on the 20 measures of area of denuded bone. PCA was performed using IBM SPSS v20.0 (SPSS Inc., Chicago, IL, USA).

### Severity score

A population of OA patients will include patients at various stages of disease severity. Using cluster analysis would likely result in clusters of patients with similar OA severity, which would not represent true OA subtypes. To prevent this, cluster analysis should be performed independently of OA severity. This can be accomplished by correcting for OA severity in the clustering approach.

We considered the KL score not suitable for this purpose in this cohort because 99% of the subjects had a KL score of 2 or 3, making the score not very discriminative. Furthermore, the KL score only reflects features on radiographs and might not reflect other measures well. Therefore, we constructed an alternative measure for OA severity using all the measures that we used to perform the cluster analysis. The severity measure was obtained using latent class factor analysis (LFA), which is similar to PCA, but suitable for a combination of continuous and

categorical data [16]. The result was an ordinal variable, of which the levels represent increasing OA severity. The optimal number of levels was decided upon using an optimum in the Bayesian Information Criterion (BIC) that informs on the goodness of fit of the LFA model.

### LCA

LCA is a state of the art model-based clustering approach that is assumed to be more powerful than traditional clustering approaches [10, 17]. The total multivariate distribution of the data is modelled as a combination of multivariate distributions that are distinct for each cluster. LCA was performed using Latent Gold v4.5 (Statistical Innovations Inc., Belmont, MA, USA).

All radiographic scores, MRI quantitative measures (after PCA) and the various clinical scores were incorporated into the model. We corrected for OA severity by including a direct effect of the OA severity score on each variable.

The optimal number of clusters was determined based on (1) the BIC which informs on the goodness of fit of the model by balancing the log-likelihood of the model fit with the increasing complexity when adding clusters [17] and (2) the approximate weight of evidence, which also informs on the ability to classify subjects into the different clusters [18]. To investigate the stability of the solution, we randomly divided the data into two and repeated the cluster analysis on each half. We then compared the two resulting cluster definitions (and the classification of subjects into the clusters) with the results of the previous analysis on the complete set.

### Statistical analysis

Differences between the clusters for a specific measure were tested using regression models corrected for OA severity. In cases of continuous or ordinal measures we used a general linear model, while logistic regression was used for dichotomous measures. *Post hoc* analysis with a Bonferroni adjustment was used to test for which clusters the measures were different. All analyses were performed using IBM SPSS v20.0 (SPSS Inc., Chicago, IL, USA).

## Results

### Study population

We included 518 subjects of the progression cohort, for whom both the detailed radiographic scores and the quantitative MRI measures were available at the time of this study. From the original 600 subjects, 26 were excluded due to missing pain scores, 35 due to missing radiographic data and 11 due to single missing values. Of the remaining 518 subjects, 9 subjects had both knees in the cohort and we randomly excluded one of their knees. Subjects were on average 61 years old (s.d. 9), range 45–79; 58% were female. KL scores of the included knees ranged from 1 to 4, although the majority had a KL-score of 2 or 3 (KL=1: 4; KL=2: 236; KL=3: 276; KL=4: 2).

### PCA of the MRI data

PCA resulted in five principle components that explained 80% of the variation in cartilage thickness/volume and four components that explained 70% of the variation in measures of denuded bone (Table 1). The descriptions of the different principal components (PCs) were based on the factor loadings (supplementary Tables S1 and S2, available at *Rheumatology* Online). For instance, PC1 showed strong factor loadings for all included MRI measures of cartilage volume and thickness, thus indicating overall cartilage quantity. PC2 gave strong negative factor loadings for all medial measures, while giving strong positive loadings for the lateral measures, meaning that this PC contrasts medial vs lateral cartilage quantity, etc.

### Severity score

LFA resulted in an OA severity score with four levels. Nearly all measures associated significantly with our severity score, showing that the various levels in our score represented increasing severity of these measures. Typically, measures of self-reported clinical symptoms showed the lowest association with our severity score. Measured scores of cysts, attrition and chondrocalcinosis did not associate, mainly due to a lack of variation in these measures within the included subjects. Also VAS pain scores and the fourth PCA component of cartilage quantity, which represented differences in cartilage quantity between the femur and the tibia in the lateral compartment, did not associate significantly with the OA severity score.

### Clusters

We performed the LCA for one to seven clusters. The approximate weight of evidence showed an optimum at four clusters, but the BIC did not show an optimum. Since the BIC showed only a small improvement after LCA with

**TABLE 1** Description of principle components of quantitative MRI measures

	Variance, %	Description
Cartilage quantity		
PC1	41	General cartilage thickness
PC2	22	Medial vs lateral
PC3	9	Local variation in thickness
PC4	6	Femur vs tibia at lateral compartment
PC5	4	Femur vs tibia at medial compartment
Denuded bone		
PC1	29	General denuded bone
PC2	25	Lateral vs medial
PC3	8	Lateral femur and medial tibia vs medial femur
PC4	7	Femur vs tibia

four clusters, we chose to continue with the simpler model containing four clusters.

The first cluster consisted of 47% of the subjects, the second of 27%, the third of 15% and the fourth of 12% of the subjects. All clusters contained subjects of the entire range of the OA severity score, though the third and fourth clusters contained notably more subjects with higher severity (Table 2). Repeating the cluster analysis on each half of the dataset after randomly splitting it in two, resulted in the same cluster definitions, with only 3% ( $n = 17$ ) of the subjects classified into a different cluster.

#### OA characteristics of the various clusters

After correction for severity, most radiographic scores showed significant differences between the clusters. Exceptions were scores on cysts, attrition and chondrocalcinosis, because these measures showed hardly any variation within the cohort. *Post hoc* analysis showed that it was mainly cluster 3 that differed from the other clusters. Whereas clusters 1, 2 and 4 showed structural OA changes in the medial compartment, cluster 3 showed OA characteristics in the lateral compartment only. Further, cluster 4 showed slightly more osteophytes and medial sclerosis, but less JSN, than cluster 1, after correction for severity. These differences were significant for lateral femoral osteophytes, medial JSN and medial sclerosis.

The quantitative MRI measures, after PCA, also differed significantly between the clusters, with the exception of the first and fifth components of the cartilage quantity measures. Quantitative MRI measures showed the same compartment-specific patterns between the clusters as

the radiographic scores (Fig. 1A–C). Cluster 1 did not show any denuded bone, independent of OA severity. The amount of denuded bone in cluster 2 was only mild, whereas clusters 3 and 4 exhibited the most dramatic increase in area of denuded bone with increasing OA severity (Fig. 1C). Interestingly, for cluster 3, denuded bone was mainly limited to the tibia, while for clusters 2 and 4, denuded bone was present equally on the articular surface of both tibia and femur, as indicated by the third and fourth components of the denuded bone measures (Table 3).

The measures of clinical OA symptoms were not significantly different between the clusters after correction for severity. Details on the differences between clusters for the various OA characteristics can be found in Table 3.

#### Demographic and risk factor differences between the clusters

Cluster 3 had the highest proportion of women (65%), while the proportion of women in clusters 2 and 4 was lowest (52%). However, these differences were not significant. There were no differences in age between the clusters.

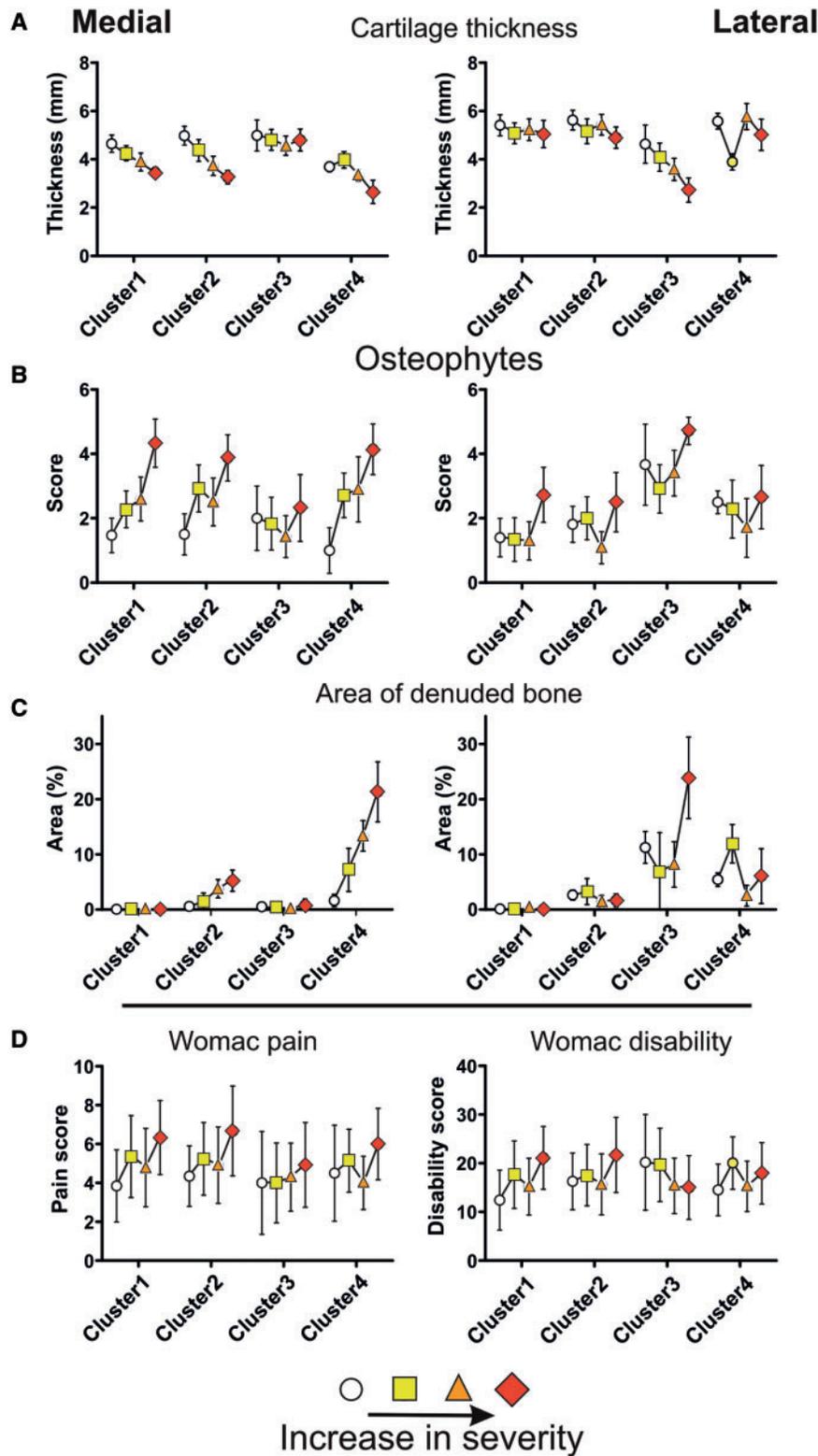
BMI differed significantly between the clusters, although the differences were small. In cluster 3 the BMI was lowest (28.5), while BMI in clusters 1 and 2 was highest (30.5) (Fig. 2A). Knee alignment was significantly the lowest in cluster 3 ( $-2.9^\circ$ ), representing a valgus alignment. In cluster 1 the alignment was highest ( $0.1^\circ$ ), representing neutral alignment. With increasing OA severity, alignment increased towards more varus-like alignment in all clusters (Fig. 2B). The subjects in cluster 3 had the

**TABLE 2** Cluster differences for knee OA risk factors and prevalence of progression

	Severity <i>P</i> -value	Cluster <i>P</i> -value	Clusters <sup>a</sup>			
			1	2	3	4
Number of subjects (%)			238 (46)	140 (27)	80 (15)	60 (12)
OA severity						
1			85 (36)	32 (23)	3 (4)	2 (3)
2			68 (29)	29 (21)	12 (15)	7 (12)
3			70 (30)	47 (33)	37 (46)	10 (17)
4			15 (6)	32 (23)	28 (35)	41 (68)
Risk factors						
Gender: female, %	0.44	0.21	60	53	65	52
Age, years	<b>0.01</b>	0.43	60	62	63	62
BMI	<b>0.049</b>	<b>0.02</b>	<b>30.6<sup>3</sup></b>	30.4	<b>28.6<sup>1</sup></b>	29.4
Alignment, <sup>b</sup> angle	0.11	<b>&lt;0.001</b>	<b>0.1<sup>3</sup></b>	<b>-0.7<sup>3</sup></b>	<b>-2.9<sup>1,2,4</sup></b>	<b>-0.4<sup>3</sup></b>
Muscle strength, <i>n</i>	0.49	0.14	242	234	214	241
Trauma, %	0.58	<b>0.035</b>	<b>33<sup>4</sup></b>	44	48	<b>55<sup>1</sup></b>
Heberden's nodes, %	0.38	0.27	65	63	76	65
Progression, %						
KL	0.81	<b>&lt;0.001</b>	<b>7<sup>3,4</sup></b>	<b>15<sup>4</sup></b>	<b>25<sup>1</sup></b>	<b>35<sup>1,2</sup></b>
JSN	<b>&lt;0.001</b>	<b>&lt;0.001</b>	<b>21<sup>3</sup></b>	<b>34<sup>3</sup></b>	<b>54<sup>1,2</sup></b>	37

<sup>a</sup>Indices (indicated in superscript next to value) indicate that the specific cluster is significantly different from the clusters indicated by the index numbers ( $P < 0.05$ ). <sup>b</sup>Negative values represent valgus alignment. Bold values indicate  $P < 0.05$ . JSN: joint space narrowing; KL: Kellgren-Lawrence.

**Fig. 1** Differences between the clusters for a selection of measures of OA characteristics that have been used to define the clusters



Data are shown for radiological markers of OA (A–C) and symptomatic markers of OA (D). For each cluster, the data are shown as a function of increasing OA severity. The data are given as mean  $\pm$  half the s.d. The radiological markers are shown for medial and lateral knee compartments separately, while clinical markers represent per patient scores.

TABLE 3 Differences between the clusters for all measures used for cluster definitions

	OA severity <i>P</i> -value	Cluster <i>P</i> -value	Cluster <sup>a</sup>			
			1	2	3	4
Osteophytes						
Medial tibia	<0.001	<0.001	1.3 <sup>3</sup>	1.3 <sup>3</sup>	0.6 <sup>1,2,4</sup>	1.4 <sup>3</sup>
Medial femur	<0.001	<0.001	1.3 <sup>3</sup>	1.3 <sup>3</sup>	0.8 <sup>1,2,4</sup>	1.6 <sup>3</sup>
Lateral tibia	0.07	<0.001	0.7 <sup>3</sup>	0.8 <sup>3</sup>	1.8 <sup>1,2,4</sup>	1 <sup>3</sup>
Lateral femur	0.003	<0.001	0.8 <sup>3,4</sup>	1 <sup>3</sup>	1.9 <sup>1,2,4</sup>	1.3 <sup>1,3</sup>
JSN						
Medial	<0.001	<0.001	1.4 <sup>3,4</sup>	1.4 <sup>3,4</sup>	-0.3 <sup>1,2,4</sup>	1.2 <sup>1,2</sup>
Lateral	0.04	<0.001	0.1 <sup>3</sup>	0.1 <sup>3</sup>	1.8 <sup>1,2,4</sup>	0.1 <sup>3</sup>
Sclerosis						
Medial tibia	<0.001	<0.001	0.9 <sup>2,3,4</sup>	1.1 <sup>1,3</sup>	-0.2 <sup>1,2,4</sup>	1.1 <sup>1,3</sup>
Medial femur	<0.001	<0.001	0.9 <sup>3</sup>	1 <sup>3</sup>	-0.3 <sup>1,2,4</sup>	1 <sup>3</sup>
Lateral tibia	<0.001	<0.001	0.04 <sup>3</sup>	0.02 <sup>3</sup>	1.2 <sup>1,2,4</sup>	0 <sup>3</sup>
Lateral femur	<0.001	<0.001	0.06 <sup>3</sup>	0.02 <sup>3</sup>	1.3 <sup>1,2,4</sup>	0.03 <sup>3</sup>
Cysts						
Medial tibia	0.2	0.2	0.02	0.03	0	0.05
Medial femur	0.6	4	0	0.01	0	0
Lateral tibia	0.1	0.4	0.01	0.04	0.03	0.02
Lateral femur	0.1	0.02	0	0	0	0.02
Attrition						
Medial	0.1	0.1	0	0.01	0	0.03
Lateral	0.4	0.2	0	0	0.01	0
Chondrocalcinosis						
Medial	0.5	0.5	0.03	0.01	0.02	0.05
Lateral	0.6	0.2	0.03	0.01	0.02	0.07
KL	<0.001	0.04	2.55 <sup>4</sup>	2.53	2.57 <sup>4</sup>	2.45 <sup>1,3</sup>
WOMAC						
Pain	0.001	0.24	5	5.2	4.1	4.9
Stiffness	0.01	0.76	2.6	2.7	2.8	2.9
Disability	0.05	0.55	15.8	17.5	15.4	16.5
Total	0.02	0.6	23.4	25.5	22.3	24.3
VAS						
7 days	0.1	0.2	4.4	4.7	4.1	5
30 days	0.5	0.06	4.9	5.2	4.5 <sup>4</sup>	5.5 <sup>3</sup>
SymptOA	0.4	0.7	0.95	0.95	0.99	0.98
SymptOAstat	0.02	0.05	2.2	2.1	1.9	2.1
Cartilage quantity (MRI)						
PC1	<0.001	0.19	0.01	0.12	-0.09	-0.18
PC2	<0.001	<0.001	0.2 <sup>3,4</sup>	0.3 <sup>3,4</sup>	-1.6 <sup>1,2,4</sup>	0.8 <sup>1,2,3</sup>
PC3	<0.001	<0.001	-0.5 <sup>2,3,4</sup>	0.02 <sup>1,3,4</sup>	0.4 <sup>1,2,4</sup>	1.6 <sup>1,2,3</sup>
PC4	0.56	0.001	-0.01 <sup>3</sup>	0.09 <sup>3</sup>	-0.37 <sup>1,2,4</sup>	0.29 <sup>3</sup>
PC5	0.01	0.22	-0.04	-0.02	-0.11	0.23
Denuded bone (MRI)						
PC1	<0.001	<0.001	-0.5 <sup>2,3,4</sup>	-0.2 <sup>1,3,4</sup>	0.8 <sup>1,2</sup>	1.2 <sup>1,2</sup>
PC2	0.004	<0.001	-0.11 <sup>2,3,4</sup>	0.08 <sup>1,3,4</sup>	-1 <sup>1,2,4</sup>	1.6 <sup>1,2,3</sup>
PC3	<0.001	<0.001	-0.06 <sup>2,3</sup>	0.1 <sup>1,3</sup>	-0.4 <sup>1,2,4</sup>	0.5 <sup>3</sup>
PC4	0.008	<0.001	0.06 <sup>2,3</sup>	-0.2 <sup>1,3</sup>	0.5 <sup>1,2,4</sup>	-0.5 <sup>3</sup>

<sup>a</sup>Indices (indicated in superscript next to value) indicate that the specific cluster is significantly different from the clusters indicated by the index numbers ( $P < 0.05$ ). Bold values indicate  $P < 0.05$ . JSN: joint space narrowing; KL: Kellgren-Lawrence; VAS: visual analogue scale.

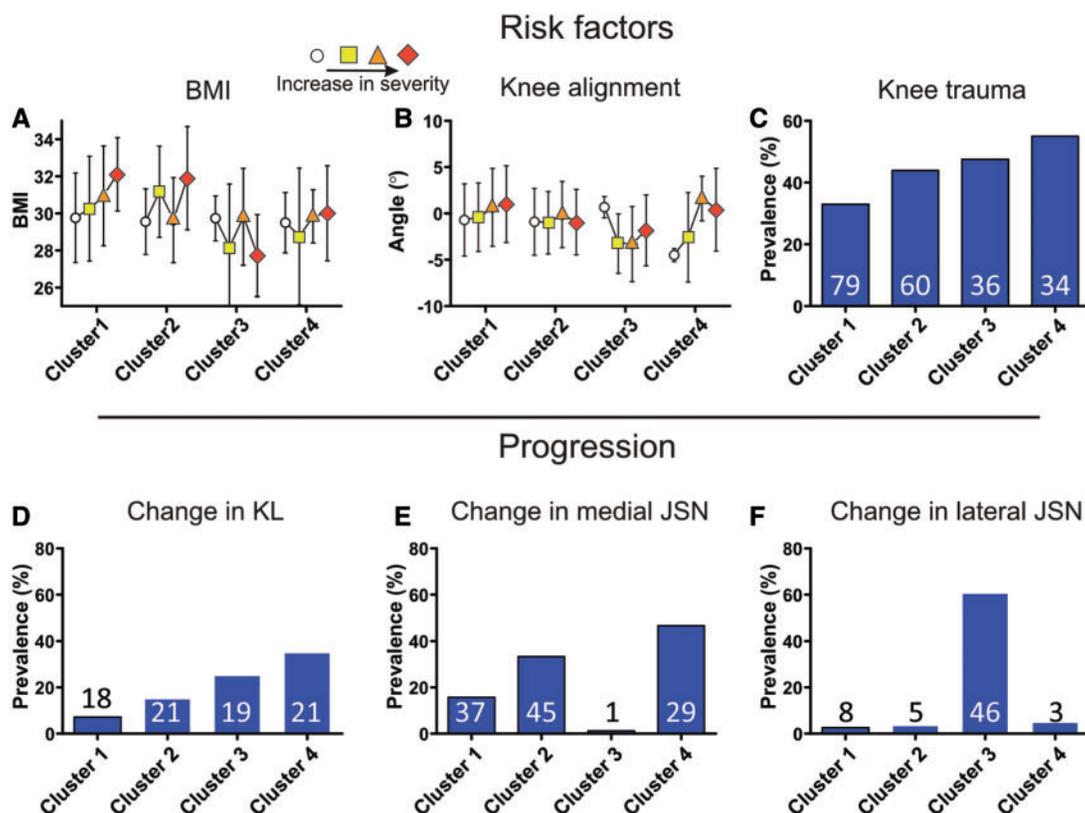
weakest muscles (212 N), while the subjects in cluster 1 had the strongest muscles (243 N), although these differences were not significant.

Subjects in cluster 4 most often reported to have suffered from knee trauma for which they had visited a physician (55%), which was significantly more often than in cluster 1 (33%) (Fig. 2C). The prevalence of reported

trauma was independent of OA severity. The presence of Heberden's nodes did not differ between the clusters.

### Progression

Progression was significantly different between the clusters, when defined using either the KL score or the score

**Fig. 2** Differences between the clusters for measures of OA risk factors (A–C) and measures of 2-year progression of OA (D–F)

These data were not used in performing the cluster analysis. Continuous measures (A and B) are presented as mean and half s.d. as a function of OA severity for each cluster. Prevalences (C–F) are presented as bars for each cluster, the number inside the bars indicating the actual number of cases.

for JSN, after correction for OA severity. Progression was lowest in cluster 1 (7% for KL score, 21% for JSN). Subjects in cluster 4 progressed most often with respect to the KL score (35%) and most often in cluster 3 with respect to JSN (54%) (Fig. 2D–F). In cluster 3, progression of JSN was purely lateral, while in clusters 1, 2 and 4, progression only occurred medially.

## Discussion

Using LCA, we were able to distinguish four distinct phenotypes of knee OA, based on differences in OA characteristics such as structural degradation and clinical symptoms. The first cluster, the most common, could be described as a mild type of OA, with no areas of denuded bone and limited progression. Cluster 2 was very similar to cluster 1, but with small areas of denuded bone. The third and fourth clusters were more aggressive types of OA, showing larger areas of denuded bone with increasing OA severity and a higher prevalence of progression. Cluster 3 was purely a lateral type of OA, while the other clusters were medial, in all aspects (Table 4).

The presence of areas of denuded bone was a strong deterministic factor in separating the four clusters. This

specific measure can either indicate areas where cartilage has been eroded completely so the underlying bone becomes visible, or it can indicate the presence of central osteophytes [19]. It has been shown that the amount of denuded bone is associated with progression [20]. We also observed that the amount of denuded bone is significantly associated with progression, defined either by KL score or by JSN ( $P < 0.001$  for both definitions), which underscores the fact that progression was strongest in clusters with more severe presence of denuded bone. However, adding cluster membership to the models reduced the association of denuded bone with progression to non-significant levels (KL:  $P = 0.06$ ; JSN:  $P = 0.50$ ). This suggests that the amount of denuded bone is only associated with progression indirectly and that a factor that leads to a specific cluster influences both progression and the increase in denuded bone.

Although differences between the clusters were clear with respect to measures of structural degradation, cluster differences were only marginal and not significant in measures that represent clinical symptoms. However, some patterns were observed. Generally, subjects in cluster 3 with purely lateral OA experienced the fewest symptoms, and this was nearly significant for VAS pain during

TABLE 4 Short description of the four clusters

Type	Population, %	Affected side	Description
Cluster 1 Mild OA	47	Medial	No denuded bone and slow progression
Cluster 2 Classical OA	27	Medial	Minor areas of denuded bone and moderate progression
Cluster 3 Aggressive OA	15	Lateral	Strong cartilage destruction and progression
Cluster 4 Aggressive OA	12	Medial	Strong cartilage destruction and progression

the past month. Most symptomatic measures showed a significant association between the intensity of the symptoms and OA severity. Interestingly, this relationship was dependent on cluster type. The relationship was strongest in the subjects in cluster 1 and weakest in subjects in cluster 3 (Fig. 1D). This difference was significant for WOMAC stiffness and disability scores ( $P < 0.05$ ). Apparently, the relationship between structural degradation and clinical symptoms depends on OA subtype, which might partly explain the weak associations in a general OA population.

It is interesting to consider what the causes might be that created these different clusters. Are the causes intrinsic (i.e. physiological or genetic) or do the different clusters arise as a result of environmental factors? Since the data used are cross sectional, the answers remain speculative. The subjects in cluster 3—the lateral OA cluster—have more valgus-aligned knees than the subjects in the medial clusters, which would suggest a biomechanical explanation. Possibly, the loads experienced in the knee determine in which compartment OA develops [21, 22]. More curious is that OA in clusters 3 and 4 is more aggressive, with large areas of denuded bone, while denuded bone in clusters 1 and 2 is absent or near absent. Since subjects in clusters 3 and 4 have a lower BMI and a higher prevalence of reported trauma, these subjects could represent a more active population in which OA has developed due to mechanical overload or trauma-related cartilage lesions. On the other hand, the subjects in cluster 1, which appears to represent a mild form of OA, are generally more obese and more often female (compared with the other medial clusters) and have the lowest prevalence of reported trauma, which might suggest that metabolic syndrome plays a role. We checked this suggestion by comparing measures that are indicators of metabolic syndrome between the clusters. Indeed, abdominal circumference was larger and hypertension ( $\geq 140/90$  mmHg) was more common in clusters 1 and 2. This was significant only for hypertension, after correction for BMI and OA severity. The idea that different causes lead to clear differences in OA characteristics and progression is intriguing and worth further investigation.

The result of cluster analysis should be stable and meaningful [23]. We tested the stability of our results by randomly dividing our dataset into two, and repeating the cluster analysis on both sets separately. Not only did this result in similar clusters, but 97% of the subjects were classified into the same cluster as before. Differences between clusters in measures that were not part of the

cluster definition are considered an indication that the clusters are meaningful. The four clusters we found differed in several OA risk factors and in the rate of progression, which thus indicates that the clusters represent true OA subtypes and are not just a creation of the clustering method.

We intended to perform cluster analysis independent of OA severity, which is crucial for a progressive disease. Indeed, each of the clusters contained knees from all levels of our severity measure. Clusters 3 and 4, however, contained significantly more knees with a high severity score than clusters 1 and 2. On first thought this might lead one to think that the correction for severity was only partial. However, randomly sampling OA knees from a population that consists of subpopulations that vary in speed of progression, which our data suggest is the case, would naturally lead to more severe knees from the rapidly progressing subpopulations, clusters 3 and 4 in our case. Further, to confirm our assumptions, we performed the cluster analysis without correcting for OA severity. As expected, this led to one lateral cluster and three medial clusters that only differed in general OA severity.

The result of any clustering approach depends on the choice of measures on which the clustering is performed. While we used a broad set of measures representing various aspects of structural degradation and clinical symptoms, it would have been interesting to include e.g. MRI scores of bone marrow lesions, synovial effusion and meniscal damage. However, these data were not present in the OAI database at the time of this study. Also, our analysis was limited to the femorotibial compartment and did not include measures of femoropatellar OA.

Finding distinct OA subtypes might have a fundamental impact on the way we look at the disease and could eventually influence how the disease is treated [1]. Though our work only presents a first step, our results are interesting, both from a fundamental and a clinical perspective. The fact that subtypes differ with respect to the presence of denuded bone suggests that certain pathological processes are present in some subtypes, but not in others, which could affect how the disease should be treated. Also, differences in progression between subtypes are highly relevant when conducting clinical trials.

Our results confirmed the suspected existence of distinct subtypes of knee OA. The various clusters showed different patterns of structural degradation and clinical symptoms. Moreover, the clusters were different with respect to rates of progression and risk factors, which

suggests that classification of knee OA into different subtypes is clinically relevant.

## Acknowledgements

The OAI is a public-private partnership comprised of five contracts (N01-AR-2-2258, N01-AR-2-2259, N01-AR-2-2260, N01-AR-2-2261, N01-AR-2-2262) funded by the National Institutes of Health, a branch of the Department of Health and Human Services, and conducted by the OAI Study Investigators. Private funding partners include Merck Research Laboratories, Novartis Pharmaceuticals Corporation, GlaxoSmithKline and Pfizer, Inc. Private sector funding for the OAI is managed by the Foundation for the National Institutes of Health. This manuscript was prepared using an OAI public use data set and does not necessarily reflect the opinions or views of the OAI investigators, the National Institutes of Health or the private funding partners.

**Funding:** The work presented in this article was funded by the Dutch Arthritis Association under grant NR 10-1-102.

**Disclosure statement:** The authors have declared no conflicts of interest.

## Supplementary data

Supplementary data are available at *Rheumatology* Online.

## References

- Felson DT. Identifying different osteoarthritis phenotypes through epidemiology. *Osteoarthritis Cartilage* 2010;18:601-4.
- Driban JB, Sittler MR, Barbe MF, Balasubramanian E. Is osteoarthritis a heterogeneous disease that can be stratified into subsets? *Clin Rheumatol* 2010;29:123-31.
- Bierma-Zeinstra SM, Verhagen AP. Osteoarthritis subpopulations and implications for clinical trial design. *Arthritis Res Ther* 2011;13:213.
- Conrozier T, Ferrand F, Poole AR *et al.* Differences in biomarkers of type II collagen in atrophic and hypertrophic osteoarthritis of the hip: implications for the differing pathobiologies. *Osteoarthritis Cartilage* 2007;15:462-7.
- Roemer FW, Guermazi A, Niu J *et al.* Prevalence of magnetic resonance imaging-defined atrophic and hypertrophic phenotypes of knee osteoarthritis in a population-based cohort. *Arthritis Rheum* 2012;64:429-37.
- Herrero-Beaumont G, Roman-Blas JA, Castaneda S, Jimenez SA. Primary osteoarthritis no longer primary: three subsets with distinct etiological, clinical, and therapeutic characteristics. *Semin Arthritis Rheum* 2009;39:71-80.
- Knoop J, van der Leeden M, Thorstensson CA *et al.* Identification of phenotypes with different clinical outcomes in knee osteoarthritis: data from the Osteoarthritis Initiative. *Arthritis Care Res* 2011;63:1535-42.
- Bartlett SJ, Ling SM, Mayo NE, Scott SC, Bingham CO III. Identifying common trajectories of joint space narrowing over two years in knee osteoarthritis. *Arthritis Care Res* 2011;63:1722-8.
- Verkleij SP, Hoekstra T, Rozendaal RM *et al.* Defining discriminative pain trajectories in hip osteoarthritis over a 2-year time period. *Ann Rheum Dis* 2012;71:1517-23.
- Vermunt JK, Magidson J. Latent class cluster analysis. In: Hagenaars JA, McCutcheon AL, eds. *Applied Latent Class Analysis*. New York: Cambridge University Press, 2002:89-106.
- Wessman J, Paunio T, Tuulio-Henriksson A *et al.* Mixture model clustering of phenotype features reveals evidence for association of DTNBP1 to a specific subtype of schizophrenia. *Biol Psychiatry* 2009;66:990-6.
- Ganesalingam J, Stahl D, Wijesekera L *et al.* Latent cluster analysis of ALS phenotypes identifies prognostically differing groups. *PLoS One* 2009;4:e7107.
- Altman RD, Hochberg M, Murphy WA Jr, Wolfe F, Lequesne M. Atlas of individual radiographic features in osteoarthritis. *Osteoarthritis Cartilage* 1995;3 (Suppl A): 3-70.
- Wirth W, Hellio Le Graverand MP, Wyman BT *et al.* Regional analysis of femorotibial cartilage loss in a subsample from the Osteoarthritis Initiative progression sub-cohort. *Osteoarthritis Cartilage* 2009;17:291-7.
- Eckstein F, Ateshian G, Burgkart R *et al.* Proposal for a nomenclature for magnetic resonance imaging based measures of articular cartilage in osteoarthritis. *Osteoarthritis Cartilage* 2006;14:974-83.
- Magidson J, Vermunt JK. Latent class factor and cluster models, bi-plots and related graphical displays. *Sociol Methodol* 2001;31:223-64.
- Fraley C, Raftery AE. How many clusters? Which clustering method? Answers via model-based cluster analysis. *Comput J* 1998;41:578-88.
- Celex G, Biernacki C, Govaert G. Choosing models in model-based clustering and discriminant analysis. Technical Report. Rhone-Alpes: INRIA, 1997.
- Frobell RB, Wirth W, Nevitt M *et al.* Presence, location, type and size of denuded areas of subchondral bone in the knee as a function of radiographic stage of OA - data from the OA initiative. *Osteoarthritis Cartilage* 2010;18:668-76.
- Eckstein F, Wirth W, Hudelmaier MI *et al.* Relationship of compartment-specific structural knee status at baseline with change in cartilage morphology: a prospective observational study using data from the Osteoarthritis Initiative. *Arthritis Res Ther* 2009;11:R90.
- Englund M. The role of biomechanics in the initiation and progression of OA of the knee. *Best Pract Res* 2010;24:39-46.
- Segal NA, Kern AM, Anderson DD *et al.* Elevated tibiofemoral articular contact stress predicts risk for bone marrow lesions and cartilage damage at 30 months. *Osteoarthritis Cartilage* 2012;20:1120-6.
- Clatworthy J, Buick D, Hankins M, Weinman J, Horne R. The use and reporting of cluster analysis in health psychology: a review. *Br J Health Psychol* 2005;10(Pt 3): 329-58.