CHECKing activity limitations in persons with early osteoarthritis of the knee or hip

Holla, J.F.M.

2014

document version
Publisher's PDF, also known as Version of record

Chapter 10

Diagnostic accuracy of range of motion measurements in early symptomatic hip and/or knee osteoarthritis

Jasmijn F.M. Holla,
Marike van der Leeden,
Leo D. Roorda,
Sita M.A. Bierma-Zeinstra,
Jurgen Damen,
Joost Dekker,
Martijn P.M. Steultjens

Arthritis Care & Research 2012;64:59-65

Abstract

Objective. To examine the diagnostic accuracy of hip internal rotation, hip flexion and knee flexion measurements for the presence of osteophytosis and joint space narrowing (JSN) in early symptomatic osteoarthritis (OA).

Methods. The baseline data of 598 participants of the Cohort Hip and Cohort Knee (CHECK) study were used. Participants underwent a standardized physical and radiographic examination. The active range of motion (ROM) was assessed using a goniometer. The ROM cut-off with the highest discriminative ability for radiographic features of OA was defined by maximizing the sum of the sensitivity and specificity. Several diagnostic measures were calculated to establish the diagnostic accuracy of ROM measurements for the presence of radiographic features.

Results. In patients with hip symptoms hip internal rotation < 24° and flexion < 114° were found to be the cut-offs with the highest discriminative ability to distinguish between patients with and without radiographic features. In patients with knee symptoms knee flexion < 132° was the cut-off with the highest discriminative ability. The American College of Rheumatology (ACR) criterion hip internal rotation < 15° increased the probability of the presence of osteophytosis or JSN from 25% to 58%. The diagnostic accuracy of hip and knee flexion measurements was low.

Conclusions. To reduce the number of patients that is not identified by the ACR criterion hip internal rotation < 15°, it is recommended to change the cut-off to internal rotation < 24° in patients with early symptomatic OA. Individual hip and knee flexion measurements seem to be of little diagnostic value in early symptomatic OA.

Introduction

The American College of Rheumatology (ACR) criteria provide a set of clinical, laboratory, and/or radiographic features to identify patients with osteoarthritis (OA) and to separate them from patients with other diseases.[1] Reduced range of motion (ROM) is one of the possible clinical signs and symptoms (e.g. pain, morning stiffness, crepitus, warmth) of hip and knee osteoarthritis. Reduced hip internal rotation (< 15°) and hip flexion (≤ 115°) are part of the ACR criteria for the classification and reporting of OA of the hip.[1] These values were derived in patients referred to a rheumatology clinic.[2] Measures of diagnostic accuracy vary across different populations.[3,4] Therefore, the validity of these cut-off values for identifying hip OA might be affected in early OA populations.

Only 2 studies were found examining the diagnostic accuracy of ROM measurements for the presence of radiographic features of OA in primary care patients with hip pain. Birrell et al.[5] showed that for reduced hip internal rotation and reduced hip flexion the most discriminatory cut-offs for radiographic features of OA were < 23° and < 94°, respectively. Bierma-Zeinstra et al.[2] used normal values for ROM in adults to define cut-off values for reduced hip internal rotation (< 21°) and reduced hip flexion (< 100°). These findings suggest that the cut-off values used in the ACR criteria may not be optimal for early diagnosis of OA.

Reduced knee motion is not included in the ACR criteria for the classification and reporting of OA of the knee,[6] whereas reduced knee movement is part of 2 of the 10 European League Against Rheumatism (EULAR) recommendations for the diagnosis of knee OA.[7] The EULAR does not provide a cut-off value for reduced knee ROM. Peat et al.[8] found that reduced knee flexion was an independent predictor of radiographic features of OA, in primary care patients with knee pain. They defined reduced knee flexion as a ROM < 120°. No other studies on the diagnostic accuracy of knee flexion in early knee OA were found.

In short, little is known about the diagnostic accuracy of ROM measurements in early OA patients. Therefore, the objective of the present study was to examine the diagnostic accuracy of hip internal rotation, hip flexion and knee flexion measurements for the presence of osteophytosis or joint space narrowing (JSN) in patients with early symptomatic hip and/or knee OA. First the ROM cut-offs with the highest discriminative ability for radiographic features of OA were examined. Second the diagnostic accuracy of the ACR definitions of reduced hip internal rotation (< 15°) and hip flexion (≤ 115°),[1] and the cut-off for reduced knee flexion of < 120°[8] for the detection of radiographic features of OA was evaluated.

Materials and methods

Study design and population

The present cross-sectional study presents a secondary analysis of baseline data of a sample of 598 participants of the Cohort Hip & Cohort Knee (CHECK) study.[9] In a previous study,[10] we examined determinants of ROM in the same dataset. We found that several radiographic features, pain, morning stiffness, a higher body-mass index and male sex were associated with lower hip or knee ROM. The associations between radiographic features and ROM measurements encouraged us to examine what these associations mean for the diagnostic accuracy of ROM measurements in early symptomatic OA.

CHECK is a prospective cohort study of 1,002 individuals with early symptomatic OA of the hip or knee. On entry, all participants had pain or stiffness of the hip or knee, and were ages 45-65 years. They had not yet consulted their physician for these symptoms, or the first consultation was within 6 months before entry. Participants with any other pathological condition that could explain the symptoms were excluded (e.g. other rheumatic disease, previous hip or knee joint replacement, congenital dysplasia, osteochondritis dissecans, intra-articular fractures, septic arthritis, Perthes' disease, ligament or meniscus damage, plica syndrome, Baker's cyst). Additional exclusion criteria were: comorbidity that did not allow physical evaluation and/or follow-up of at least 10 years, malignancy in the last 5 years, and inability to understand the Dutch language.

The CHECK cohort was formed from October 2002 till September 2005. Nationwide, 10 general and academic hospitals in the Netherlands are participating, located in urbanized and semi-urbanized regions. General practitioners (GPs) in the surroundings of the participating centres were invited to refer eligible persons. All patients that visited the GP on their own initiative, potentially fulfilling the inclusion criteria, were referred to one of the 10 participating centres. In addition, participants were recruited through advertisements and articles in local newspapers and on the Dutch Arthritis Association website. The physicians in the participating centres checked whether referred patients as well as patients from their outpatient clinics fulfilled the inclusion criteria.

Data for the present study were obtained from 598 participants assessed in 5 CHECK centres (Utrecht, Amsterdam, Nijmegen, Rotterdam and Maastricht). In these centres, the ROM of the lower extremity was assessed in all planes of motion. Two strata were defined based on symptoms: a hip ($n = 344$) stratum and a knee ($n = 497$) stratum. Patients with both hip and knee symptoms were included in both strata. The study was approved by the medical ethics committees of all participating centres, and all participants gave their written informed consent before entering the study.

Index tests: ROM measurements

At baseline, assisted active hip internal rotation, hip flexion and knee flexion were measured bilaterally, using a goniometer. Measurements were taken according to Norkin and White.[11] For each joint action, the protocol provided starting positions for both participant and examiner, reference points for the pivot and distal points of the goniometer. Hip internal rotation was measured in sitting position with the knees flexed to 90° over the edge of the treatment table. The hip was in 0° of abduction and adduction and in 90° of

flexion. The participant carried out maximal internal rotation. The examiner supported the ankle and stabilized the distal end of the femur to prevent adduction or further flexion of the hip. Hip flexion was measured in supine position, with the hip in 0° of abduction, adduction, and rotation. The participant flexed the knee and moved it to the abdomen. The maximal hip flexion was defined as the point at which rotation of the pelvis was observed. Knee flexion was measured in supine position, with the knee in extension. Initially the hip was in 0° of extension, abduction, and adduction, but as the participant maximally flexed the knee, the hip also flexed. The examiner supported the lower leg, and stabilized the femur to prevent rotation, abduction, and adduction of the hip. Supplying a protocol for starting positions and the positioning of the goniometer increases the reliability of ROM measurements.[12] In a previous study the inter-rater reliability of assisted active ROM measurements in hip and knee OA patients was established, and was satisfactory for all joint actions (Pearson's R exceeding 0.75 for all actions).[12] Measurements were carried out by a trained physiatrist, rheumatologist or orthopaedic surgeon, dependent

Table 1. Characteristics of the study population*

	Hip stratum (*n* = 344)		Knee stratum (*n* = 497)	
	Value	Missing (%)	Value	Missing (%)
Demographics				
Women	269 (78.2%)	0	390 (78.5%)	0
Age, years	55.9 ± 5.4	0	55.8 ± 5.2	0
Body mass index (kg/m²)	25.9 ± 3.9	0	26.2 ± 4.0	0
Clinical signs and symptoms				
Unilateral hip or knee pain	222 (64.5)	0	215 (43.3)	0
Bilateral hip or knee pain	122 (35.5)	0	282 (56.7)	0
Duration of pain (months)	18 (10-36)	47 (13.7%)	18 (9-36)	84 (16.9%)
Clinical hip or knee OA†	97 (28.2)	0	391 (78.7)	0
Range of Motion, degrees				
Hip internal rotation	28.2 ± 9.5	4 (1.2%)		
Hip flexion	115.3 ± 12.2	3 (0.9%)		
Knee flexion			132.2 ± 10.0	5 (1.0)
Radiographic features, moderate/severe‡				
Osteophytosis or JSN	76 (22.1)	38 (11.0%)§	159 (32.0)	60 (12.1%)¶
Osteophytosis and JSN	19 (5.5)	42 (12.2%)#	31 (6.2)	72 (14.5%)**

* Values are the number (percentage), the mean ± SD, or the median (interquartile range). OA = osteoarthritis; JSN = joint space narrowing. † Hip or knee OA according to the American College of Rheumatology clinical classification criteria. ‡ Reference category = normal/mild. § Missing because there was no radiograph available (n = 11) or because the radiograph was of inadequate quality for determination according to Altman and Gold[14] (n = 27). ¶ Missing because there was no radiograph available (n =19) or because the radiograph was of inadequate quality for determination (n = 41). # Missing because there was no radiograph available (n = 12) or because the radiograph was of inadequate quality for determination (n = 30). ** Missing because there was no radiograph available (n =19) or because the radiograph was of inadequate quality for determination (n = 53).

on the CHECK centre in which the participant was assessed. The examiners were blind to the results of the reference standards. The ROM of the "index hip" (most affected hip) or "index knee" (most affected knee) was used as outcome measure.[13]

Reference standards: radiographic features

Radiographs were obtained according to a standardized protocol.[9] Radiographic features were scored in pairs by 4 medical students and 1 GP trainee (JD) independently. The readers of the reference standards were blind to the results of the index tests. Before scoring the features, the examiners were extensively trained by a musculoskeletal radiologist and an experienced reader in 4 separate sessions with training radiographs. At the end of this course the readers' performance was assessed by scoring a new set of radiographs of 12 participants with differing OA severity. The trainers confirmed that all readers had scored the training set adequately.

Table 2. Definitions of diagnostic terms

Diagnostic term	Definition
Sensitivity	The proportion of participants with radiographic features of osteoarthritis (OA) who are correctly identified by a positive test result ("true positive rate")
Specificity	The proportion of participants without radiographic features of OA who are correctly identified by a negative test result ("true negative rate")
Positive predictive value	The proportion of participants with positive test results with radiographic features of OA
Negative predictive value	The proportion of participants with negative test results without radiographic features of OA
Likelihood ratio for a positive test	The ratio of the true positive rate to the false positive rate: sensitivity/(1 – specificity)
Likelihood ratio for a negative test	The ratio of the false negative rate to the true negative rate: (1 – sensitivity)/specificity
Pretest probability (prevalence)	The probability that a participant has radiographic features of OA before the test is carried out
Posttest probability of a positive test	The probability that a participant with a positive test result has radiographic features of OA
Posttest probability of a negative test	The probability that a participant with a negative test result has radiographic features of OA
Area under the receiver operating characteristic (ROC) curve (AUC)	An ROC plot is obtained by plotting sensitivity against 1-specificity. A "curve" that coincided with the left and top sides of the plot indicates that the test perfectly discriminates between the two groups.19 A test that is completely useless would give a straight line from the bottom left corner to the top right corner. The AUC is defined as the area under the ROC curve. This area is equal to the probability that a random participant with radiographic features of OA has a lower value of the ROM measurement than a random participant without radiographic features of OA.

The anterioposterior radiographs of the hip were scored according to Altman and Gold.[14] Superior and medial JSN, superior and inferior acetabular osteophytes, and superior and inferior femoral osteophytes were scored on a 0-3 scale (where 0 = normal, 1 = mild or 1-33% abnormal, 2 = moderate or 34-66% abnormal, and 3 = severe or 67-100% abnormal). The faux profil radiographs of the hip were taken according to Lequesne and Loredo.[15] The faux profil view provides a lateral projection of the femoral head and neck, and an oblique view of the acetabulum tangential to its superoanteromedial edge.[15] On the faux profile radiographs, superior JSN was scored on a 0-3 scale according to Altman and Gold.[14]

The posterioanterior radiographs of the knee were scored according to Altman and Gold.[14] Medial and lateral JSN, femoral medial and lateral osteophytes, and tibial medial and lateral osteophytes were scored on a 0-3 scale. Radiographs of the patellofemoral joints were made by a single standing mediolateral view in 30° flexion and a non-weight-bearing skyline (inferior superior) view in 30° flexion.[16,17] The mediolateral and skyline radiographs of the knee were scored using the radiographic atlas of Burnett et al.[18] On these radiographs patellofemoral JSN and osteophytes were scored on a 0-3 scale.[14]

For the analyses features were combined with each other (**Table 1**), and all features were dichotomized into normal/mild vs. moderate/severe. A feature was scored as present if a 2 (moderate) or 3 (severe) was scored in ≥ 1 radiograph.

Statistical analysis

Assessment of diagnostic test accuracy for all tests was calculated using 2 x 2 tables. As index tests we used the ROM measurements which we dichotomized using hip internal rotation < 15°, hip flexion ≤ 115° and knee flexion < 120° as cut-offs for reduced hip and knee ROM. In addition we used the cut-offs for hip internal rotation, hip flexion and knee flexion with the highest discriminative ability for radiographic features of OA as index tests. The cut-off with the highest discriminative ability for radiographic features of OA (optimal cut-off) was calculated by maximizing the sum of the sensitivity and specificity.[19] The presence of osteophytosis or JSN and the presence of osteophytosis and JSN were used as reference standards. Test sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), likelihood ratio for a positive test (LR+), likelihood ratio for a negative test (LR-), pretest probability, posttest probability of a positive test, posttest probability of a negative test, and area under the receiver operating characteristic curve (AUC) were calculated (**Table 2**).[19,20] The AUC was nonparametrically estimated.[21] For all effect measures, a 95% confidence interval was calculated.[22,23]

Results

Diagnostic accuracy of hip internal rotation and hip flexion

The baseline characteristics of the cohort and information on missing values are shown in **Table 1**. Hip internal rotation < 24° was found to be the cut-off with the highest discriminative ability for radiographic features based on maximizing the sum of the sensitivity and specificity. Hip internal rotation < 24° increased the probability of osteo-

phytosis or JSN from 25% (pretest probability = prevalence) to 46% (posttest probability = PPV), and hip internal rotation ≥ 24° decreased the probability of osteophytosis or JSN from 25% to 16% (1 – NPV). Additional measures of diagnostic accuracy (sensitivity, specificity, PPV, NPV, LR+, LR-, and AUC) are shown in **Table 3**. The ACR criterion hip internal rotation < 15° increased the probability of the presence of osteophytosis or JSN from 25% to 58%, and hip internal rotation ≥ 15° decreased the probability of osteophytosis or JSN from 25% to 22%.

Hip flexion < 114° was found to be the cut-off with the highest discriminative ability for radiographic features. Hip flexion < 114° increased the probability of osteophytosis or JSN from 25% to 36% and hip flexion ≥ 114° decreased the probability of osteophytosis or JSN from 25% to 18% (**Table 3**). The ACR criterion hip flexion ≤ 115° increased the probability of the presence of osteophytosis or JSN from 25% to 31%, and hip flexion > 115° decreased the probability of osteophytosis or JSN from 25% to 19%.

Table 3. Diagnostic accuracy of hip internal rotation and hip flexion measurements for the presence of radiographic features of hip OA*

	Hip internal rotation vs. radiographic features		Hip flexion vs. radiographic features	
Reference standard	Osteophytosis or JSN	Osteophytosis and JSN	Osteophytosis or JSN	Osteophytosis and JSN
Prevalence	0.25	0.06	0.25	0.06
AUC	0.69 (0.62 - 0.77)	0.84 (0.73 - 0.94)	0.66 (0.58 - 0.73)	0.73 (0.59 - 0.86)
ACR cut-off	< 15°	<15°	≤ 115°	≤ 115°
Results, no. tp/fp/fn/tn	11/8/64/220	7/11/11/270	45/100/31/128	13/129/6/152
Sensitivity	0.15 (0.08 - 0.24)	0.39 (0.20 - 0.61)	0.59 (0.48 - 0.70)	0.68 (0.46 - 0.85)
Specificity	0.97 (0.93 - 0.98)	0.96 (0.93 - 0.98)	0.56 (0.50 - 0.62)	0.54 (0.48 - 0.60)
PPV	0.58 (0.36 - 0.77)	0.39 (0.20 - 0.61)	0.31 (0.24 - 0.39)	0.09 (0.05 - 0.15)
NPV	0.78 (0.72 - 0.82)	0.96 (0.93 - 0.98)	0.81 (0.74 - 0.86)	0.96 (0.92 - 0.98)
LR+	4.18 (1.75 - 10.00)	9.93(4.38 - 22.54)	1.35 (1.07 - 1.71)	1.49 (1.07 - 2.08)
LR-	0.88 (0.80 - 0.97)	0.64 (0.44 - 0.92)	0.73 (0.54 - 0.98)	0.58 (0.30 - 1.14)
Optimal cut-off†	< 24°	< 26°	< 114°	< 110°
Results, no. tp/fp/fn/tn	42/50/33/178	17/108/1/173	44/78/32/150	11/47/8/234
Sensitivity	0.56 (0.45 - 0.67)	0.94 (0.74 - 0.99)	0.58 (0.47 - 0.68)	0.58 (0.36 - 0.77)
Specificity	0.78 (0.72 - 0.83)	0.62 (0.56 - 0.67)	0.66 (0.59 - 0.72)	0.83 (0.79 - 0.87)
PPV	0.46 (0.36 - 0.56)	0.14 (0.09 - 0.21)	0.36 (0.28 - 0.45)	0.19 (0.11 - 0.31)
NPV	0.84 (0.79 - 0.89)	0.99 (0.97 - 1.00)	0.82 (0.76 - 0.87)	0.97 (0.94 - 0.98)
LR+	2.55 (1.86 - 3.51)	2.46 (2.04 - 2.96)	1.69 (1.30 - 2.20)	3.46 (2.18 - 5.50)
LR-	0.56 (0.43 - 0.73)	0.09 (0.01 - 0.61)	0.64 (0.48 - 0.85)	0.51 (0.30 - 0.86)

* Values are the measure (95% confidence interval) unless otherwise indicated. OA = osteoarthritis; JSN = joint space narrowing; AUC = area under the receiver operating characteristic curve; ACR = American College of Rheumatology; tp = true positives; fp = false positives; fn = false negatives; tn = true negatives; PPV = positive predictive value; NPV = negative predictive value; LR+ = likelihood ratio for a positive test; LR- = likelihood ratio for a negative test. † Cut-off for which the sum of the sensitivity and specificity was maximal.

Diagnostic accuracy of knee flexion

Knee flexion < 132° was found to be the cut-off with the highest discriminative ability for radiographic features. Knee flexion < 132° increased the probability of osteophytosis or JSN from 37% to 44% and knee flexion ≥ 132° decreased the probability of osteophytosis or JSN from 37% to 31%. Additional measures of diagnostic accuracy (sensitivity, specificity, PPV, NPV, LR+, LR-, and AUC) are shown in **Table 4**. Knee flexion < 120°[8] increased the probability of the presence of osteophytosis or JSN from 37% to 42%. Knee flexion ≥ 120° did not decrease the probability of osteophytosis or JSN. The posttest probability of a negative test was equal to the pretest probability of osteophytosis or JSN.

Discussion

In the present study the diagnostic accuracy of hip internal rotation, hip flexion and knee flexion measurements for the presence of radiographic features was examined in patients with early symptomatic hip and/or knee OA.

Hip internal rotation < 24° was found to be the cut-off with the highest sum of sensitivity and specificity. For this cut-off the probability of the presence of osteophytosis or JSN increased from 25% to 46% with a positive test result and decreased to 16% with a negative test result. Hip internal rotation < 24° lies close to the most discriminatory cut-off for radiographic features of OA of hip internal rotation < 23° as reported by Birrell et al.[5] in a primary care population with hip pain. For the ACR cut-off hip internal rotation < 15° the probability of the presence of osteophytosis or JSN increased from 25% to 58% with a positive test result, and decreased to 22% with a negative test result. The ACR criterion hip internal rotation < 15° contributes substantially to the classification of OA, however at this cut-off many patients with radiographic features are not identified (the percentage false negatives was 21%).

In general, OA is not diagnosed by 1 examination or by a limited number of tests performed at the same time. The diagnostic process usually starts with some simple tests at a GP visit and ends with laboratory, radiographic and/or even magnetic resonance imaging examinations for final diagnosis. ROM measurements are performed relatively early in the diagnostic process. In this stage, it is not desirable to exclude potential patients from further diagnostics. Therefore, in early stage OA highly sensitive tests (few false negatives) are preferred over highly specific tests (few false positives). Later on in the diagnostic process tests must be highly specific to confirm the results of the first sensitive but less specific tests. For the cut-off hip internal rotation < 24° the percentage false negatives was 11%. Therefore, to reduce the percentage false negatives and because Birrell et al.[5] found almost the same cut-off, we recommend to use hip internal rotation < 24° as classification criteria for radiographic features in early symptomatic OA patients.

The most discriminatory cut-off for hip flexion of hip flexion < 114°, and the ACR classification criterion hip flexion ≤ 115° were of little diagnostic value for the presence of radiographic features. The tests changed the probability of radiographic features with less than 12% and 7%, respectively.

In the EULAR recommendations for the diagnosis of knee OA, reduced knee movement is not defined. Therefore, we examined knee flexion < 120° as reported by Peat et al.[8]

as measure for reduced knee flexion. Results showed that the most discriminatory cut-off (knee flexion < 132°) and knee flexion < 120° were both of little diagnostic value. The tests increased the probability of radiographic features with 7% and 5%, respectively. However, when we changed the reference standard to osteophytosis and JSN instead of osteophytosis or JSN knee flexion < 120° increased the probability of a positive test result with 15% (**Table 4**). Therefore, knee flexion < 120° may be useful for detecting more severe radiographic damage. Knee flexion < 120° did not change the probability of a negative test result (no osteophytosis or JSN): the test seems not useful for ruling out radiographic OA in patients with a negative test result.

Some issues need to be addressed concerning the methodology of this study and its impact on the presented results. First, the study population consisted of new presenters with symptoms of pain and/or stiffness of the hip/and or knee. Therefore, the results of the present study can only be extrapolated to comparable populations with early symptomatic hip and/or knee OA.[4,20] The ACR classification criteria for hip OA[1] and EULAR recommendations for the diagnosis of knee OA[7] are intended to identify patients with OA and

Table 4. Diagnostic accuracy of knee flexion measurements for the presence of radiographic features of knee OA*

	Knee flexion vs. radiographic features			
Reference standard	**Osteophytosis or JSN**		**Osteophytosis and JSN**	
Prevalence	0.37		0.07	
AUC	0.57	(0.51 - 0.63)	0.63	(0.53- 0.73)
Cut-off reported by Peat et al.[8]	< 120°		< 120°	
Results, no. tp/fp/fn/tn	8/11/151/262		4/14/27/375	
Sensitivity	0.05	(0.03 - 0.10)	0.13	(0.05- 0.29)
Specificity	0.96	(0.93 - 0.98)	0.96	(0.94 - 0.98)
PPV	0.42	(0.23 - 0.64)	0.22	(0.09 - 0.45)
NPV	0.63	(0.59 - 0.68)	0.93	(0.90 - 0.95)
LR+	1.25	(0.51 - 3.04)	3.60	(1.26 - 10.24)
LR-	0.99	(0.95 - 1.03)	0.90	(0.79 - 1.04)
Optimal cut-off†	< 132°		< 127°	
Results, no. tp/fp/fn/tn	90/117/69/156		14/90/17/299	
Sensitivity	0.57	(0.49 - 0.64)	0.45	(0.29 - 0.62)
Specificity	0.57	(0.51 - 0.63)	0.77	(0.72 - 0.81)
PPV	0.44	(0.37 - 0.50)	0.14	(0.08 - 0.21)
NPV	0.69	(0.63 - 0.75)	0.95	(0.92 - 0.97)
LR+	1.32	(1.09 - 1.60)	1.95	(1.27 - 3.00)
LR-	0.76	(0.62 - 0.93)	0.71	(0.52 - 0.99)

* Values are the measure (95% confidence interval) unless otherwise indicated. OA = osteoarthritis; JSN = joint space narrowing; AUC = area under the receiver operating characteristic curve; ACR = American College of Rheumatology; tp = true positives; fp = false positives; fn = false negatives; tn = true negatives; PPV = positive predictive value; NPV = negative predictive value; LR+ = likelihood ratio for a positive test; LR- = likelihood ratio for a negative test. † Cut-off for which the sum of the sensitivity and specificity was maximal.

to separate them from patients with other diseases. Especially in early stage OA in which symptoms commence these criteria have to function well to facilitate early diagnosis. In populations with established OA there is less need for a first sensitive test to diagnose OA, because in this stage the suspicion of OA is already so high that further diagnostics are already indicated.

Second, the intra- and inter-rater reliability of the ROM measurements was not tested. Therefore we have no indication if and to what extent measurements were biased. However, all measurements were taken according to a strict protocol. A previous study has shown that ROM measurements according to a similar protocol yielded reliable measurements.[12]

Third, the reliability of the x-ray readings was not tested. Although we have no exact figure on the reading reliability, readings with similar training show in general, intra-class correlation coefficients of 0.5-0.9, depending on the type of feature.[24]

Fourth, in clinical practice, besides reduced ROM, other criteria are tested to identify patients with OA (e.g. age > 50 years, pain, morning stiffness, crepitus on active motion of the knee, bony tenderness of the knee, no palpable warmth of the knee).[1,6] In the present study the diagnostic accuracy of ROM measurements was analysed for each joint action separately. The study does not provide quantitative knowledge about the probability of radiographic OA for combinations of test results.[4]

In conclusion, to reduce the number of patients that is not identified by the ACR cut-off hip internal rotation < 15°, we recommend to change this cut-off to hip internal rotation < 24° in patients with early symptomatic OA. Individual hip flexion and knee flexion measurements seem to be of little diagnostic value for the presence of radiographic features. Additional research is needed to study the diagnostic accuracy of combinations of test results (e.g. reduced hip internal rotation and reduced hip flexion) for the presence of radiographic OA in early symptomatic hip and/or knee OA.

References

1. Altman R, Alarcon G, Appelrouth D, Bloch D, Borenstein D, Brandt K et al. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hip. *Arthritis Rheum* 1991;**34**:505-14.
2. Bierma-Zeinstra SM, Oster JD, Bernsen RM, Verhaar JA, Ginai AZ, Bohnen AM. Joint space narrowing and relationship with symptoms and signs in adults consulting for hip pain in primary care. *J Rheumatol* 2002;**29**:1713-8.
3. Deeks JJ, Altman DG. Diagnostic tests 4: likelihood ratios. *BMJ* 2004;**329**:168-9.
4. Moons KG, Harrell FE. Sensitivity and specificity should be de-emphasized in diagnostic accuracy studies. *Acad Radiol* 2003;**10**:670-2.
5. Birrell F, Croft P, Cooper C, Hosie G, Macfarlane G, Silman A. Predicting radiographic hip osteoarthritis from range of movement. *Rheumatology (Oxford)* 2001;**40**:506-12.
6. Altman R, Asch E, Bloch D, Bole G, Borenstein D, Brandt K et al. Development of criteria for the classification and reporting of osteoarthritis. Classification of osteoarthritis of the knee. *Arthritis Rheum* 1986;**29**:1039-49.
7. Zhang W, Doherty M, Peat G, Bierma-Zeinstra SM, Arden NK, Bresnihan B et al. EULAR evidence based recommendations for the diagnosis of knee osteoarthritis. *Ann Rheum Dis* 2010;**69**:483-9.
8. Peat G, Thomas E, Duncan R, Wood L, Wilkie R, Hill J et al. Estimating the probability of radiographic osteoarthritis in the older patient with knee pain. *Arthritis Rheum* 2007;**57**:794-802.

9. Wesseling J, Dekker J, van den Berg WB, Bierma-Zeinstra SM, Boers M, Cats HA et al. CHECK: Cohort Hip & Cohort Knee; similarities and differences with the Osteoarthritis Initiative. *Ann Rheum Dis* 2009;**68**:1413-9.
10. Holla JF, Steultjens MP, van der Leeden M, Roorda LD, Bierma-Zeinstra SM, den Broeder AA et al. Determinants of range of joint motion in patients with early symptomatic osteoarthritis of the hip and/or knee: an exploratory study in the CHECK cohort. *Osteoarthritis Cartilage* 2011;**19**:411-9.
11. Norkin CC, White DJ. Measurement of joint motion: a guide to goniometry. 2nd ed. Philadelphia: F.A. Davis Company; 1995.
12. Steultjens MP, Dekker J, van Baar ME, Oostendorp RA, Bijlsma JW. Range of joint motion and disability in patients with osteoarthritis of the knee or hip. *Rheumatology (Oxford)* 2000;**39**:955-61.
13. Holla JF, Steultjens MP, Roorda LD, Heymans MW, Ten Wolde S, Dekker J. Prognostic factors for the two-year course of activity limitations in early osteoarthritis of the hip and/or knee. *Arthritis Care Res (Hoboken)* 2010;**62**:1415-25.
14. Altman RD, Gold GE. Atlas of individual radiographic features in osteoarthritis, revised. *Osteoarthritis Cartilage* 2007;**15 Suppl A**:A1-56.
15. Lequesne MG, Laredo JD. The faux profil (oblique view) of the hip in the standing position. Contribution to the evaluation of osteoarthritis of the adult hip. *Ann Rheum Dis* 1998;**57**:676-81.
16. Chaisson CE, Gale DR, Gale E, Kazis L, Skinner K, Felson DT. Detecting radiographic knee osteoarthritis: what combination of views is optimal? *Rheumatology (Oxford)* 2000;**39**:1218-21.
17. Laurin CA, Dussault R, Levesque HP. The tangential x-ray investigation of the patellofemoral joint: x-ray technique, diagnostic criteria and their interpretation. *Clin Orthop Relat Res* 1979;16-26.
18. Burnett DJ, Hart DJ, Cooper C, Spector A. A radiographic atlas of osteoarthritis. London: Springer; 1994.
19. Altman DG, Bland JM. Diagnostic tests 3: receiver operating characteristic plots. *BMJ* 1994;**309**:188.
20. Pewsner D, Battaglia M, Minder C, Marx A, Bucher HC, Egger M. Ruling a diagnosis in or out with "SpPIn" and "SnNOut": a note of caution. *BMJ* 2004;**329**:209-13.
21. Park SH, Goo JM, Jo CH. Receiver operating characteristic (ROC) curve: practical review for radiologists. *Korean J Radiol* 2004;**5**:11-8.
22. Simel DL, Samsa GP, Matchar DB. Likelihood ratios with confidence: sample size estimation for diagnostic test studies. *J Clin Epidemiol* 1991;**44**:763-70.
23. Newcombe RG. Two-sided confidence intervals for the single proportion: comparison of seven methods. *Stat Med* 1998;**17**:857-72.
24. Sun Y, Gunther KP, Brenner H. Reliability of radiographic grading of osteoarthritis of the hip and knee. *Scand J Rheumatol* 1997;**26**:155-65.