Arthritis Care & Research
Vol. 76, No. 4, April 2024, pp 511–516
DOI 10.1002/acr.25270

American College of Rheumatology
Empowering Rheumatology Professionals

BRIEF REPORT

Evaluation of the Diagnostic Performance of American College of Rheumatology, EULAR, and National Institute for Health and Clinical Excellence Criteria Against Clinically Relevant Knee Osteoarthritis: Data From the CHECK Cohort

Qiuke Wang,[1] Jos Runhaar,[2] Margreet Kloppenburg,[3] Maarten Boers,[4] Johannes W. J. Bijlsma,[5] and Sita M. A. Bierma-Zeinstra,[2] the CREDO expert group

Objective. Our objective was to evaluate the diagnostic performance of the EULAR, American College of Rheumatology (ACR), and National Institute for Health and Care Excellence (NICE) criteria by using clinical experts' diagnosis of clinically relevant knee osteoarthritis (OA) as the outcome of interest.

Methods. In a previous study, we recruited clinical experts to evaluate longitudinal (5-, 8-, and 10-year follow-up) clinical and radiographic data of symptomatic knees from the Cohort Hip and Cohort Knee (CHECK) study for the presence or absence of clinically relevant OA. In the current study, ACR, EULAR, and NICE criteria were applied to the same 5-, 8-, and 10-year follow-up data; then a knee was diagnosed with OA if fulfilling the criteria at one of the three time points (F1), two of the time points (F2), or at all three time points (F3). Using clinically relevant OA as the reference standard, the sensitivity, specificity, and positive and negative predictive values for the three criteria were assessed.

Results. A total of 539 participants for a total of 833 examined knees were included. Thirty-six percent of knees were diagnosed with clinically relevant OA by experts. Sixty-seven percent to 74% of the knees received the same diagnosis (OA or non-OA) by the three criteria sets for the different definitions (F1 to F3). EULAR consistently (F1 through F3) had the highest specificity, and NICE consistently had the highest sensitivity.

Conclusion. The diagnoses only moderately overlapped among the three criteria sets. The EULAR criteria seemed to be more suitable for study enrollment (when aimed at recruiting clinically relevant OA knees), given the highest specificities. The NICE criteria, given the highest sensitivities, could be more useful for an initial diagnosis in clinical practice.

INTRODUCTION

To help identify knee osteoarthritis (OA) cases in research and/or clinical settings, three sets of diagnostic (or classification) criteria have been proposed[1]: American College of Rheumatology (ACR),[2] EULAR,[3] and National Institute for Health and Care Excellence (NICE) criteria.[4] Specifically, ACR criteria were originally developed as classification criteria to be used in research,[2] EULAR criteria were developed as evidence-based recommendations for diagnosis of knee OA,[3] and NICE's definition of knee OA was proposed as a recommendation in the NICE health care guideline.[4]

Along with the application, the validity of these criteria sets has been questioned. First, the participants used for developing the criteria were mainly limited to patients with OA from secondary care with severe radiographic changes; the diagnostic performance in knees from primary care is insufficiently validated.[5] For this, Skou et al[1] compared the three criteria sets in the individuals treated in primary care, but all the included knees had already been diagnosed with OA, indicating the incapability for assessing the exact diagnostic performance. Second, because of the lack of a gold standard, the criteria were mainly validated by referring to radiographic alterations (eg, Kellgren and Lawrence grade) with

Dr Wang's work was supported by the China Scholarship Council (grant 201906230308).

[1]Qiuke Wang, MD: Erasmus MC University Center Rotterdam, Rotterdam, The Netherlands, and Shanghai Sixth People's Hospital, Shanghai, China; [2]Jos Runhaar, PhD, Sita M. A. Bierma-Zeinstra, PhD: Erasmus MC University Center Rotterdam, Rotterdam, The Netherlands; [3]Margreet Kloppenburg, PhD: Leiden University Medical Center, Leiden, The Netherlands; [4]Maarten Boers MD, PhD: Amsterdam University Medical Center, Amsterdam, The Netherlands; [5]Johannes W. J. Bijlsma, PhD: University Medical Centre Utrecht, Utrecht, The Netherlands.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25270.

Address correspondence via email to Qiuke Wang, MD, at q.wang@erasmusmc.nl.

Submitted for publication February 23, 2023; accepted in revised form November 2, 2023.

SIGNIFICANCE & INNOVATIONS

- We evaluated diagnosis by the three commonly used criteria sets covering longitudinal data (three consecutive follow-up time points) to reflect fluctuations over time.
- Clinical experts' consensus-based diagnosis of clinically relevant knee osteoarthritis was used as the outcome of interest (reference standard).
- Our findings suggest that EULAR criteria seemed to be more suitable for study enrollment, given the highest specificities; the National Institute for Health and Care Excellence criteria, given the highest sensitivities, could be more useful for an initial diagnosis in clinical practice.

unclear clinical relevance.[3,5] Third, the diagnosis made based on the criteria was found unstable over time because the symptoms and physical examination results often fluctuate.[6,7] Fourth, these criteria sets are mostly used in research contexts (no standard diagnostic classification established in the clinic), and their alignment with real clinical scenarios remains unknown. With these in mind, further studies are needed to reevaluate the three criteria sets using clinically relevant references and considering the over-time fluctuations.

In this brief report, we first assessed the overlap of the diagnoses among ACR, EULAR, and NICE criteria in a cohort of patients recruited with knee pain in primary care. To assess the diagnostic performance of these three criteria sets in a clinically relevant circumstance, we adopted clinical experts' consensus-based diagnosis of clinically relevant knee OA as the outcome of interest (reference standard). The reference diagnosis was chosen because it represents the diagnosis in daily clinical practice, and the knees diagnosed with clinically relevant OA are usually the cases needing interventions. The three criteria sets in this study were applied to patient longitudinal data (three consecutive follow-up time points) to reflect fluctuations over time.

PATIENTS AND METHODS

Cohort hip and cohort knee cohort. We used data from the Cohort Hip and Cohort Knee (CHECK) study, a longitudinal cohort study of participants with knee or hip complaints consulted in primary care and observed for 10 years.[8] The protocol of the CHECK study was approved by the Ethical Committee of University Medical Center Utrecht (protocol number 02/017-E). The inclusion criteria of CHECK study were as follows: (1) nontraumatic knee or hip pain or stiffness (knee pain was determined by asking, "Do you feel any knee pain?"), (2) age 45 to 65 years, and (3) at or within six months of first visit to the general practitioner for these complaints. Participants were excluded if the complaints could be explained by diseases other than OA, the participants had comorbidities precluding physical evaluation and/or follow-up of at least 10 years, the participants had malignancies in the past 5 years, or the participants were unable to understand the Dutch language. All the participants were required to fill in questionnaires and to get physical and radiographic examinations at baseline and at 2, 5, 8, and 10 years.

This study included symptomatic knees from the participants who were recruited for knee complaints at baseline and for whom data were available between 5-year (T5) and 10-year (T10) follow-up. Knees with missing values that prevent a diagnosis from being made for any of ACR, EULAR, or NICE criteria were excluded.

Outcome of interest: Clinically relevant knee OA diagnosis. The details of the diagnostic process were presented in our previous study.[9] In brief, we invited both general practitioners and secondary care physicians to evaluate participants' longitudinal (5-, 8-, and 10-year follow-up) clinical and radiographic data. Clinical data consisted of demographics (including sex, age, racial background, marital status, menopausal status, educational level, chronic diseases, occupation, smoking status, and alcohol usage); measurement of body mass index; physical examinations (presence of knee pain, morning stiffness in knee, knee warmth, bony tenderness, crepitus, knee pain on extension and flexion, and range of motion); Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) pain, function, and stiffness scores; numeric rating scale pain scores (pain level in the past week); and incidence of other diseases (quadriceps tendinitis, intraarticular fracture, Baker's cyst, ligament or meniscus damage, osteochondritis dissecans, plica syndrome, and septic arthritis). Radiographic data consisted of scores on tibial attrition, femoral and/or tibial sclerosis, joint space narrowing, femoral and/or tibial osteophytes, and Kellgren and Lawrence (KL) grades.

A final diagnosis of whether clinically relevant knee OA (yes, no, or uncertain) developed during the 5- to 10-year follow-up was made for each knee based on experts' consensus. Knees with a final diagnosis of "uncertain" were excluded from our analysis because these seem to be a mix of OA and non-OA cases (although a higher proportion is attributed to OA knees), according to a previous report.[10] Clinical experts were instructed to use their own clinical expertise to judge the presence of clinically relevant knee OA without being informed of any definition.

ACR, EULAR, and NICE criteria. For ACR criteria, we used the commonly used clinical version of "three of six items."[2] Knees were diagnosed with OA if individuals reported pain and fulfilled three of the following six items: (1) aged 50 years or older, (2) with no early morning stiffness or morning stiffness less than 30 minutes, (3) with crepitus, (4) with bony tenderness, (5) with bony enlargement, and (6) with no palpable warmth.

We used proposition number five of the EULAR criteria for EULAR diagnosis, the same as a previous report.[1] Knees were diagnosed with OA if individuals were older than 40 years, had

movement-related knee pain, had morning stiffness less than 30 minutes, had functional limitation, and had one or more typical examination findings (crepitus, restricted movement, and bony enlargement). We deemed "functional limitation" as reporting any limitation in the WOMAC subscale of function, and we deemed "restricted movement" as knee extension deficit ≥1° or flexion ≤115°.[10] The NICE criteria recommend a diagnosis of OA in the knees for individuals aged 45 years or older with movement-related knee pain and without or with less than 30 minutes of morning stiffness.[4]

We applied the previously mentioned ACR, EULAR, and NICE criteria to the 5-, 8-, and 10-year follow-up data and obtained the diagnoses for the three criteria sets at each time point. Next, to make it comparable to experts' consensus-based diagnosis, we made a final criteria-based diagnosis (OA or non-OA) for each criteria set in each knee if it fulfilled the criteria at one or more of the three time points (F1), at two or more of the time points (F2), or at all three time points (F3). For example, "ACR F1 OA" means a knee fulfills the ACR criteria at least one of the three time points and is diagnosed with OA; "ACR F1 non-OA" means a knee does not fulfill the ACR criteria at any of the three time points and is diagnosed with non-OA.

Statistical analysis. We described 5-, 8-, and 10-year follow-up characteristics of the included knees by means ± SD or numbers (percentages), when appropriate. We performed the analysis on the knee level and treated knees as independent observations. We calculated the prevalence of clinically relevant knee OA as well as the final criteria-based OA diagnosis. We assessed the "agreements" among ACR, EULAR, and NICE criteria for each situation of fulfillment separately (ie, for F1, F2, and F3, respectively) via a Venn diagram. Using clinically relevant knee OA as the reference standard, we calculated sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV), as well as their 95% confidence intervals, for each final criteria-based diagnosis. No statistical test was performed; the comparisons among the three criteria regarding diagnostic performance were interpreted based on observed values.

RESULTS

A total of 539 participants for a total of 833 examined knees were included in the analysis; 80% were women, the mean ± SD age was 60.8 ± 5.0 years, and the mean ± SD body mass index was 26.4 ± 4.3 at the 5-year follow-up. Table 1 presents clinical and radiographic characteristics for the included participants at 5-, 8-, and 10-year follow-ups. From the 5- to 10-year follow-up, the prevalence of radiographic OA (KL grade ≥2) had increased from 38% to 60%, whereas prevalence of fulfillment of the three

Table 1. Characteristics of knees included in the analysis (N = 833)*

Characteristic	5-y follow-up	8-y follow-up	10-y follow-up
Age, mean (SD), y	60.8 (5.0)	63.8 (5.0)	65.8 (5.0)
Female, n (%)	666 (80)	666 (80)	666 (80)
BMI, mean (SD)	26.4 (4.3)	26.5 (4.4)	26.6 (4.2)
Any knee pain at present, n (%)	488 (59)	451 (55)	405 (49)
With morning stiffness, duration <30 min, n (%)	525 (63)	486 (58)	475 (57)
Joint line tenderness, n (%)	281 (34)	272 (33)	312 (38)
Knee crepitus, n (%)	370 (44)	368 (44)	388 (47)
Warmth, n (%)	20 (2)	13 (2)	16 (2)
Active knee extension range, mean (SD), degree	2.0 (2.3)	1.8 (2.3)	1.9 (2.6)
Active knee flexion range, mean (SD), degree	134.9 (8.4)	132.9 (8.4)	132.8 (9.1)
KL grade, n (%)			
0	172 (21)	96 (12)	41 (5)
1	343 (41)	369 (44)	283 (34)
2	271 (33)	320 (38)	434 (52)
3	21 (2)	33 (4)	48 (6)
TKR	0	0	6 (1)
Missing	26 (3)	15 (2)	21 (2)
ACR OA, n (%)[a]	430 (48)	406 (49)	382 (46)
EULAR OA, n (%)[b]	346 (41)	322 (39)	305 (37)
NICE OA, n (%)[c]	488 (59)	451 (54)	405 (49)

* ACR, American College of Rheumatology; BMI, body mass index; KL, Kellgren and Lawrence; NICE, National Institute for Health and Care Excellence; OA, osteoarthritis; TKR, total knee replacement.
[a] According to the ACR criteria, knees were diagnosed with OA if patients reported pain and fulfilled three of the following six items: (1) aged 50 years or older, (2) with no early morning stiffness or morning stiffness less than 30 minutes, (3) with crepitus, (4) with bony tenderness, (5) with bony enlargement, and (6) with no palpable warmth.
[b] According to the EULAR criteria, knees were diagnosed with OA if individuals were older than 40 years, had movement-related knee pain, had morning stiffness less than 30 minutes, had functional limitation, and had one or more typical examination findings (crepitus, restricted movement, and bony enlargement).
[c] According to the NICE criteria, knees were diagnosed with OA for individuals aged 45 years or older with movement-related knee pain and without or with less than 30 minutes of morning stiffness.

criteria sets had decreased (ACR 48%–46%, EULAR 41%–37%, NICE 59%–49%).

The final criteria-based knee OA was diagnosed in 72%, 48%, and 27% of the knees, according to the ACR F1, F2, and F3 criteria, respectively; in 61%, 37%, and 19% according to the EULAR F1, F2, and F3 criteria, respectively; and in 79%, 54%, and 31% according to the NICE F1, F2, and F3 criteria, respectively. In general, the overlaps (for OA or non-OA) among the three criteria were moderate (nearly 70%); the overlaps were similar between F1 and F2 criteria and slightly improved (from 67% to 74%) among F3 criteria (Figure 1).

A total of 368 (44%) knees were diagnosed with clinically relevant OA by the clinical experts. For all the three criteria sets, the sensitivity and NPV decreased from F1 to F3, whereas the specificity and PPV increased. For either F1, F2, or F3, the differences in diagnostic performance among the three criteria sets were small to moderate. The EULAR criteria consistently had the highest specificity and PPV; the NICE criteria consistently had the highest sensitivity and NPV among the three criteria, but the relevance of the differences among criteria is debatable (Table 2).

DISCUSSION

Among the knees of individuals from the CHECK study, there was only a moderate overlap among the ACR, EULAR, and NICE criteria. By referring to the experts' clinically relevant knee OA diagnosis, we found none of the criteria sets remarkably outperformed the other two. Specifically, the EULAR criteria presented the highest specificity and PPVs, and the NICE criteria presented the highest sensitivities and NPVs among the three criteria sets.

To properly interpret the results of diagnostic tests, the reference standard should be considered concurrently. The outcome of interest of this study was the expert consensus-based diagnosis of clinically relevant knee OA, which might significantly diverge from standardized OA definitions (eg, radiographic OA). It is

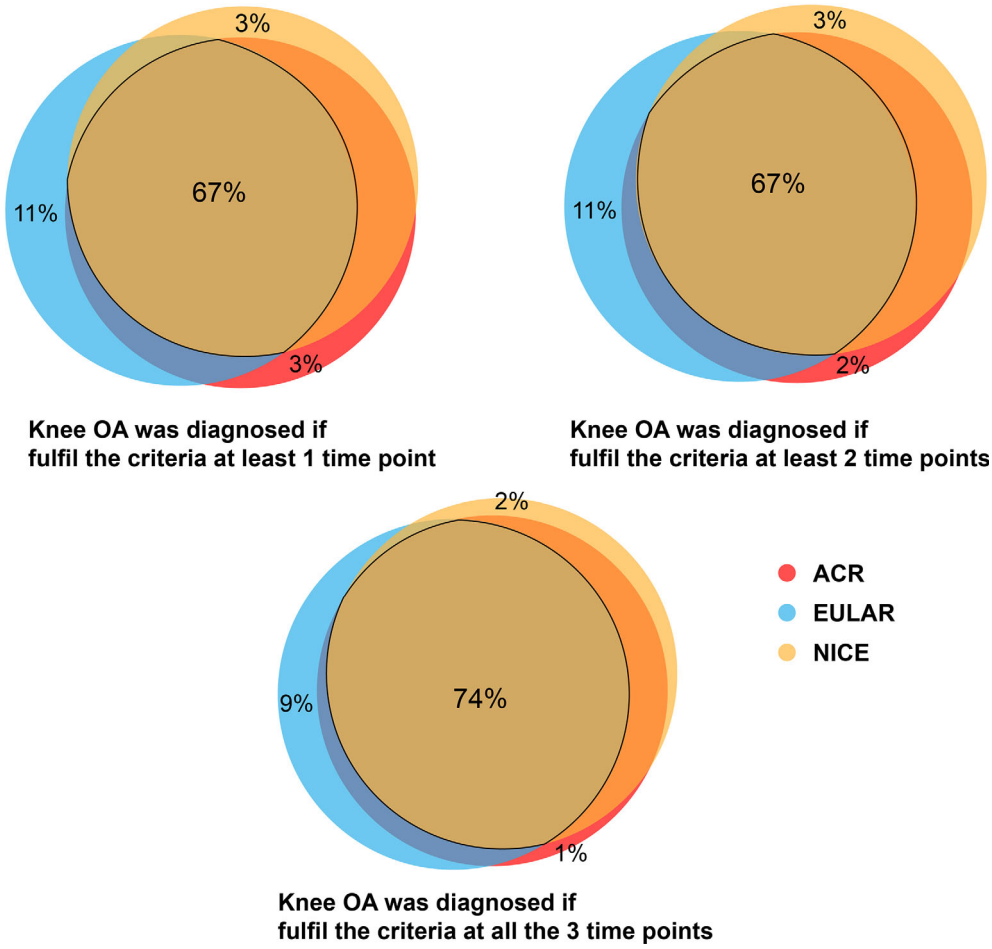

Figure 1. Venn diagrams illustrating the agreement between criteria-based diagnosis and fulfilling at least one set of criteria, at one, two, or all three time points. Each circle represents the diagnoses (OA and non-OA knees, N = 833) by one of the criteria sets. The percentage at the center indicates the proportion of clinical diagnoses agreed by all three criteria sets. The percentages on the periphery indicate the proportion of diagnoses made by one of the criteria that are "disagreed" by the other two. ACR, American College of Rheumatology; NICE, National Institute for Health and Care Excellence; OA, osteoarthritis.

Table 2. Diagnostic performance of ACR, EULAR, and NICE criteria against clinical experts' diagnosis (N = 833)*

	Experts' diagnosis					
Criteria set	OA, n (%)	Non-OA, n (%)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
ACR F1						
OA	335 (40)	265 (32)	0.91 (0.88–0.94)	0.43 (0.38–0.48)	0.56 (0.54–0.58)	0.86 (0.81–0.89)
Non-OA	33 (4)	200 (24)				
ACR F2						
OA	263 (32)	131 (16)	0.71 (0.67–0.76)	0.72 (0.67–0.76)	0.67 (0.63–0.70)	0.76 (0.73–0.79)
Non-OA	105 (12)	334 (40)				
ACR F3						
OA	155 (19)	69 (8)	0.42 (0.37–0.47)	0.85 (0.82–0.88)	0.69 (0.64–0.74)	0.65 (0.63–0.67)
Non-OA	213 (25)	396 (47)				
EULAR F1						
OA	295 (35)	214 (26)	0.80 (0.76–0.84)	0.54 (0.49–0.59)	0.58 (0.55–0.61)	0.77 (0.73–0.81)
Non-OA	73 (9)	251 (30)				
EULAR F2						
OA	203 (24)	104 (13)	0.55 (0.50–0.60)	0.78 (0.74–0.81)	0.66 (0.62–0.70)	0.69 (0.66–0.71)
Non-OA	165 (20)	361 (43)				
EULAR F3						
OA	108 (13)	49 (6)	0.29 (0.25–0.34)	0.89 (0.86–0.92)	0.69 (0.62–0.75)	0.61 (0.60–0.63)
Non-OA	260 (31)	416 (50)				
NICE F1						
OA	345 (41)	312 (38)	0.94 (0.91–0.96)	0.33 (0.29–0.37)	0.52 (0.51–0.54)	0.87 (0.81–0.91)
Non-OA	23 (3)	153 (18)				
NICE F2						
OA	283 (34)	168 (20)	0.77 (0.72–0.81)	0.64 (0.59–0.68)	0.63 (0.60–0.66)	0.78 (0.74–0.81)
Non-OA	85 (10)	297 (36)				
NICE F3						
OA	169 (20)	93 (11)	0.46 (0.41–0.51)	0.80 (0.76–0.84)	0.64 (0.59–0.69)	0.65 (0.63–0.67)
Non-OA	199 (24)	372 (45)				

* ACR, American College of Rheumatology; CI, confidence interval; F1, diagnosed with OA if fulfilling the criteria at one or more time points; F2, diagnosed with OA if fulfilling the criteria at two or more time points; F3, diagnosed with OA if fulfilling the criteria at all three time points; NICE, National Institute for Health and Care Excellence; NPV, negative predictive value; OA, osteoarthritis; PPV, positive predictive value.

important to note that these are usually applied in research, and there is no single gold standard definition for use in the clinic. The diagnostic test results presented here demonstrated the performance of the three criteria sets in scenarios in which the population of interest pertains to consensus-based clinically relevant OA knees.

To our knowledge, this is the first study to evaluate diagnosis by the three commonly used criteria sets covering multiple time points. A cross-sectional design was adopted in nearly all the previous studies, including the one comparing the same three criteria sets as in this study.[1] One interesting finding was that the prevalence of fulfillment of the three criteria sets had decreased. The decreased prevalence of OA should not be interpreted as disease remission in a subset of persons but rather as the intermittent fulfillment of the criteria. It has been well known that patients with OA often present with intermittent and fluctuating symptoms,[11,12] which could result in inconsistent fulfillment of the criteria among different time points. Schiphof et al[7] evaluated ACR criteria fulfillment in the CHECK study and found that 42% of the knees fulfilled the criteria at baseline, but only 17% fulfilled the criteria throughout the 10 years. Additionally, those findings imply the limitation of the three criteria sets that none of them can provide stable diagnosis over time. Our findings showed that the more times the knee fulfilled the criteria (from F1 to F3), the more likely the knee had clinically relevant OA, which was applicable for all the three criteria sets. For diagnostic purposes, in daily practice, the scenario can be different from what we applied in this study; in daily practice, patients usually consult because of recent onset or worsening of symptoms. The current study assessed diagnostic performance for patients with a history of knee complaints, applied during periodical inspections in primary care.

Despite originally developed as diagnostic recommendations, the EULAR criteria showed the highest specificity among the three, suggesting that they seem to be able to identify a more homogeneous group of knees and should be more suitable for participant recruitment in research (eg, clinical trials aimed at recruiting clinically relevant OA knees).[13] Similar to a previous report,[1] we found nearly all clinically relevant OA knees fulfilled the NICE criteria at least at one of the three time points. As a price, the low specificity indicated there could be a certain amount of non-OA knees misdiagnosed as OA when using NICE criteria. On the other hand, a knee that never fulfilled the NICE criteria had a very low probability of having clinically relevant OA. Therefore, the NICE criteria seem to be suitable for initial

clinical diagnosis in primary care, which identifies nearly all clinically relevant OA knees. It was unexpected that the ACR criteria, developed as classification criteria, performed just slightly worse than the NICE criteria and moderately better than the EULAR regarding its sensitivity.

This study was limited by the fact that the CHECK study recruited participants at 45 to 65 years at baseline, meaning that all the study participants had already fulfilled the age requirement of the three criteria sets from the 5-year follow-up. This may cause an overestimation of criteria-based OA prevalence. In addition to that, the generalizability of our findings could also be limited by the cohort design. For example, the CHECK study excluded patients with differential diseases (eg, other rheumatic disease, previous hip or knee joint replacement, congenital dysplasia, osteochondritis dissecans, intraarticular fractures, septic arthritis, Perthes disease, ligament or meniscus damage, plica syndrome, and Baker's cyst), so the diagnostic performance may not be generalizable to the patients with these conditions, whereas it should be applicable to situations such as trials with exclusion criteria for other diseases or after clinicians have done examinations for excluding other causes of knee symptoms. Finally, because radiographic OA (KL grade ≥2) was not used as the reference standard, we were unable to identify how our consensus-based clinically relevant knee OA differed from conventional radiographic knee OA. This choice was deliberate because employing radiographic OA alone would result in missing symptomatic patients with mild structural changes.

In conclusion, among the knees of middle-aged individuals observed in primary care, about one of four to one of three knees would receive inconsistent diagnosis by the ACR, EULAR, and NICE criteria. With the clinical experts' diagnosis of clinically relevant knee OA as the reference, no remarkable differences were observed for the diagnostic performances of the three criteria. Based on small to moderate differences, the EULAR criteria seemed to be more suitable for participant enrollment in research (when aimed at recruiting clinically relevant OA knees), given the highest specificity and PPV of the three criteria sets. The NICE criteria could be more appropriate for initial clinical diagnosis because of the highest sensitivity and NPV of the three criteria sets.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr Wang had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Wang, Runhaar.
Acquisition of data. Wang, Runhaar, Kloppenburg, Boers, Bijlsma, Bierma-Zeinstra.
Analysis and interpretation of data. Wang, Runhaar, Kloppenburg, Boers, Bijlsma, Bierma-Zeinstra.

REFERENCES

1. Skou ST, Koes BW, Gronne DT, et al. Comparison of three sets of clinical classification criteria for knee osteoarthritis: a cross-sectional study of 13,459 patients treated in primary care. Osteoarthritis Cartilage 2020;28(2):167–172.
2. Altman R, Asch E, Bloch D, et al; Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association. Development of criteria for the classification and reporting of osteoarthritis. Classification of osteoarthritis of the knee. Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association. Arthritis Rheum 1986;29(8):1039–1049.
3. Zhang W, Doherty M, Peat G, et al. EULAR evidence-based recommendations for the diagnosis of knee osteoarthritis. Ann Rheum Dis 2010;69(3):483–489.
4. National Clinical Guideline Centre (UK). Osteoarthritis: Care and Management in Adults. National Institute for Health and Care Exellence; 2014.
5. Peat G, Thomas E, Duncan R, et al. Clinical classification criteria for knee osteoarthritis: performance in the general population and primary care. Ann Rheum Dis 2006;65(10):1363–1367.
6. Damen J, van Rijn RM, Emans PJ, et al. Prevalence and development of hip and knee osteoarthritis according to American College of Rheumatology criteria in the CHECK cohort. Arthritis Res Ther 2019; 21(1):4.
7. Schiphof D, Runhaar J, Waarsing JH, et al. The clinical and radiographic course of early knee and hip osteoarthritis over 10 years in CHECK (Cohort Hip and Cohort Knee). Osteoarthritis Cartilage 2019;27(10):1491–1500.
8. Wesseling J, Boers M, Viergever MA, et al. Cohort profile: Cohort Hip and Cohort Knee (CHECK) study. Int J Epidemiol 2016;45(1):36–44.
9. Wang Q, Runhaar J, Kloppenburg M, et al. The added value of radiographs in diagnosing knee osteoarthritis is similar for general practitioners and secondary care physicians; data from the CHECK early osteoarthritis cohort. J Clin Med 2020;9(10):3374.
10. Runhaar J, Kloppenburg M, Boers M, et al. Towards developing diagnostic criteria for early knee osteoarthritis: data from the CHECK study. Rheumatology (Oxford) 2021;60(5):2448–2455.
11. Parry E, Ogollah R, Peat G. Significant pain variability in persons with, or at high risk of, knee osteoarthritis: preliminary investigation based on secondary analysis of cohort data. BMC Musculoskelet Disord 2017;18(1):80.
12. Soni A, Kiran A, Hart DJ, et al. Prevalence of reported knee pain over twelve years in a community-based cohort. Arthritis Rheum 2012; 64(4):1145–1152.
13. Aggarwal R, Ringold S, Khanna D, et al. Distinctions between diagnostic and classification criteria? Arthritis Care Res (Hoboken) 2015; 67(7):891–897.