Validation of new classification criteria of rheumatoid arthritis in an international multicentre study

R. Li[1], X. Sun[1], H. Ye[1], J. Liu[1], J. Zhao[2], X. Liu[2], Y. Mei[3], Z. Zhang[3], J. Guo[4], L. Bi[4], X. Liu[5], G. Wang[5], J. Liu[6], X. Leng[6], F. Gu[7], L. Sun[7], Q. Zou[8], Y. Fang[8], Q. Jiang[9], K. Chatzidionysiou[10], S. Priya[11], R. Janardana[11], A. Nair[11], A.I. Catrina[10], D. Danda[11], Z.G. Li[1]

[1]Department of Rheumatology and Immunology, Peking University People's Hospital, Beijing Key Laboratory for Rheumatism and Immune Diagnosis (BZ0135) and Peking-Tsinghua Center for Life Sciences, Beijing, China; [2]Department of Rheumatology and Immunology, Peking University Third Hospital, Beijing, China; [3]Department of Rheumatology and Immunology of the First Affiliated Hospital of Harbin Medical University, Harbin, China; [4]Department of Rheumatology and Immunology, China-Japan Union Hospital of Jilin University, Changchun, China; [5]Department of Rheumatology and Immunology, China-Japan Friendship Hospital, Beijing, China; [6]Department of Rheumatology and Immunology, Peking Union Medical College Hospital, Beijing, China; [7]Department of Rheumatology and Immunology of the Affiliated Drum Tower Hospital, Nanjing University, Nanjing, China; [8]Department of Rheumatology and Immunology, Southwest Hospital of Third Military Medical University, Chongqing, China; [9]Guang An Men Hospital of Academy of Traditional Chinese Medicine, Beijing, China; [10]Rheumatology Unit, Department of Medicine Solna, Karolinska Institute, Karolinska University Hospital, Stockholm, Sweden; [11]Department of Clinical Immunology and Rheumatology, Christian Medical College, Vellore, Tamil Nadu, India.

Abstract

Objective

Early identification of patients with rheumatoid arthritis (RA) is essential to allow prompt therapy. In this study, we aimed to evaluate the performance of the newly proposed ERA criteria, compared to the 1987 ACR and 2010 ACR/EULAR criteria in an international multicentre study.

Methods

A total of 606 patients with disease duration ≤2 years and age ≥16 years who were diagnosed as RA or non-RA were enrolled from China, Sweden and India. The clinical and laboratory parameters were recorded. We compared the sensitivity, specificity, predictive value, likelihood ratio (LR), and the area under the ROC curve (AUC) of three criteria in these cohorts. Concordance between the three criteria was calculated with the Kappa coefficient.

Results

Three hundred and twelve RA and 294 non-RA patients were included. The Early Rheumatoid Arthritis (ERA) criteria had significantly higher specificity compared to the 2010 ACR/ EULAR criteria (83.7% vs. 78.2%, p=0.02) and sensitivity were similar (79.2% vs. 78.5%, p=0.883). In comparison with the 1987 ACR criteria, the ERA criteria had higher sensitivity (79.2% vs. 54.5%, p<0.001) but lower specificity (83.7% vs. 89.1%, p<0.001), and the AUC of the ERA criteria (0.878) was comparable to the 2010 ACR/EULAR criteria (0.849) and higher than the 1987 ACR criteria (0.791, p<0.0001). Patients from the three countries, seronegative and very early arthritis cohorts yielded consistent results.

Conclusion

The ERA criteria demonstrates a better performance across ethnics in early RA diagnosis, and is more feasible in daily practice.

Key words

rheumatoid arthritis, early, classification criteria

Ru Li, MD, Xing Sun, MD*, Hua Ye, MD, Jiajia Liu, PhD, Jinxia Zhao, MD, Xiangyuan Liu, MD, Yifang Mei, MD, Zhiyi Zhang, MD, Jialong Guo, MD, Liqi Bi, MD, Xia Liu, MD, Guochun Wang, MD, Jinjing Liu, MD, Xiaomei Leng, MD, Fei Gu, MD, Lingyun Sun, MD, Qinghua Zou, MD, Yongfei Fang, MD, Quan Jiang, MD, Katerina Chatzidionysiou, MD, Sangeetha Priya, MD, Ramya Janardana, MD, Aswin Nair, MD, Anca I. Catrina, MD**, Debashish Danda, MD**, Zhan Guo Li, MD***

**These authors contributed equally to this work.*
***Co-corresponding authors*

Please address correspondence to: Zhan Guo Li, Department of Rheumatology and Immunology, Peking University People's Hospital, 11 Xizhimen South St., Beijing 100044, China.
E-mail: li99@bjmu.edu.cn

Anca Catrina: anca.catrina@ki.se

Debashish Danda: debashisdandacmc@hotmail.com

Reprint requests should be sent to: Sun Xing,
E-mail: 18810650837@163.com

Received on May 30, 2019; accepted in revised form on September 16, 2019.

Funding: the study was supported by funds from Beijing Municipal Science and Technology Project (grant no. Z171100000417007), Macau Science and Technology Development Fund (0094/2018/A3) and funding from Foundation for Research in Rheumatology from ERC under EU Horizon 2020 research and innovation programme (CoG 2017-7722209_PREVENT RA and 777387_RT Cure).

Competing interests: none declared.

Introduction

Rheumatoid arthritis (RA) is a chronic destructive disease. Increasing evidence suggests that early intensive treatment can significantly slow the rate of disease progression and increase the probability of achieving remission (1-5). Remission in RA patients receiving prolonged intensive DMARD therapy was achieved, and low disease activity at the start of disease taper leads to less subsequent flares (6). To prevent disease progression, classification criteria for early RA and appropriate treatment are needed.

The 1987 American College of Rheumatology (ACR) criteria (7) have been widely used for the classification of rheumatoid arthritis (RA) for more than two decades, which was developed in patients with established RA for several years, but it has indicated that they have poor sensitivity for the diagnosis of RA in early arthritis cohorts (8). Then, 23 years later, the 2010 ACR/European League Against Rheumatism (2010 ACR/EULAR) classification criteria (9) for RA was published. However, several studies from multiple centres revealed that the 2010 ACR/EULAR criteria had only a "slight improvement" of performance overall compared with the 1987 ACR criteria, others have shown that lower specificity raises the issue of a potential for over-classification and over-treatment in newly onset of inflammatory arthritis (10-15). More reliable classification criteria for early RA should be established. Recently, our multicentre studies established new classification criteria for early rheumatoid arthritis (ERA), which was developed by analysing the clinical and laboratory variables in 803 patients with early inflammatory arthritis with the duration less than 1 year (16). Then a subsequent study in 2016 showed that sensitivity and specificity of the ERA criteria were 72.3% and 87.8% (17).

In this international multicentre study, we aimed to assess the diagnostic value of the ERA criteria, the 1987 criteria, and the 2010 ACR/EULAR criteria in patients with early rheumatoid arthritis in three countries, and analysing the agreement among the three criteria.

Materials and methods

Patients

This study was based on multiple cohorts of patients recruited from three countries including China, Sweden and India. Briefly, participants diagnosed as RA or non-RA by two experienced rheumatologists based on the clinical and laboratory features were consecutively enrolled if the following features were present: 1) with apparent joint swelling at one or more joints; 2) more than 16 years old of age; 3) less than 2 years of duration. Patients with arthritis caused by trauma, suspected septic arthritis were excluded. Chinese cohorts come from the study from H. Ye (17). Swedish and Indian patients were enrolled from October 2015 to October 2018. The gold standard was the clinical diagnosis of RA made by experienced rheumatologists regardless to specific criteria. Ethical Committee approval was obtained at each participating centre and informed consent for participation was signed by all the patients.

Data collection

The demographic and clinical features of individual patients with early arthritis were collected, including age, gender, duration of arthritis, involved joint areas, morning stiffness, symmetric arthritis, arthritis of large joints, arthritis of hand joints, and rheumatoid nodules. Rheumatoid factor (RF), anti-cyclic citrullinated peptide (CCP) antibody, C-reactive protein (CRP) and erythrocyte sedimentation rate (ESR) were measured. Furthermore, all patients were subjected to radiographs of their hands such as x-ray, ultrasonography (US) or magnetic resonance imaging (MRI), and their radiological images were interpreted by two radiologists. The typical RA-related radiological characters included joint erosions or unequivocal bony decalcification localised in or most marked adjacent to the involved joints in x-ray (7). The ERA includes 3 out of the following 5 factors: 1) morning stiffness ≥30 min; (2) affecting at least three joints; (3) affecting at least one joint in the wrist, metacarpophalangeal, or proximal interphalangeal joints; (4) RF positive; and (5) positive for anti-CCP antibodies (16). The ERA

Table I. The demographic and clinical characteristics of overall.

	ERA (n=312)	non-RA (n=294)	*p*-value
Female, n (%)	234 (75.5)	189 (64.5)	0.003
Age (years), median (P_{25}, P_{75})	49 (39, 60)	45 (31, 55)	<0.001
Duration (months), median (P_{25}, P_{75})	4.5 (2.2, 8.0)	3.5 (1.7, 7.9)	0.044
Morning Stiffness (minutes), median (P_{25}, P_{75})	45 (15, 90)	20 (15, 60)	<0.001
RF positive, n(%)	198 (63.5)	42 (14.3)	<0.001
Anti-CCP positive, n(%)	209 (67.0)	10 (3.4)	<0.001
CRP (mg/dL), median (P_{25}, P_{75})	8.0 (2.0, 22.1)	4.5 (1.0, 20.5)	0.025
ESR (mm/1h), median (P_{25}, P_{75})	38.0 (18.0, 60.0)	30.0 (12, 52.8)	0.001
Erosion, n (%)	77 (24.8)	19 (6.5)	<0.001

RF: rheumatoid factor; CCP: cyclic citrullinated peptide; CRP: C-reactive protein; ESR: erythrocyte sedimentation rate.

criteria and the 1987 ACR criteria (7) were evaluated, and the score of the 2010 ACR/EULAR criteria (9) was calculated respectively in all the patients.

Statistical methods

The sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive LR (LR+), negative LR (LR-), and the Receiver Operating Characteristic (ROC) curve of three criteria were analysed in the total, three countries, seronegative (both ACPA and RF negative) and the very early arthritis (duration ≤3 months) patients, respectively. Data analysis was performed with the standard software packages, SPSS 24.0. Quantitative numerical variables were expressed by the median (P_{25}, P_{75}) and category data as the real numbers or the percentages. The differences between two groups were analysed by Student's t-test and Mann-Whitney U-test, respectively. The differences in the sensitivity and specificity of the criteria were determined by the McNemar. A 2-tailed *p*-value <0.05 was taken to indicate statistical significance.

The sensitivity, specificity, PPV, NPV at the proposed cut-off value was calculated. Diagnostic value of three classification criteria was analysed by the LR+, LR- and the AUC. The sensitivity was plotted against 1-specificity to obtain the ROC curve, and the corresponding AUC was calculated. The agreement between the two criteria was based on the κ statistic.

Results

The demographic and clinical characteristics of patients

Six hundred and six patients who were diagnosed as definite inflammatory arthritis by two experienced rheumatologists were recruited, including 312 RA and 294 non-RA patients. Non-RA patients included 61 osteoarthritis (OA), 44 systemic lupus erythematosus (SLE), 28 psoriatic arthritis (PsA), 26 Gout, 19 primary Sjögren's syndrome (pSS), 14 ankylosing spondylitis (AS), 12 polymyositis(PM), 5 reactive arthritis (ReA), 4 systemic sclerosis (SSc), 3 pseudo-gout, 2 polymyalgia rheumatica (PMA), 1 Adult Onset Still Disease (AOSD), and 75 of other diseases. Table I depicts the characteristics of all patients, which showed a direct comparison between RA and non-RA patients. Overall, ERA patients had higher rates of female, RF positivity, anti-CCP positivity, and bone erosion than non-RA patients (p<0.01, respectively). Longer morning stiffness, higher ESR and CRP levels were also shown in ERA patients (p<0.05, respectively). Moreover, the differences between the ERA and non-RA group in three countries was consistent with the total. (Table II).

The diagnostic value of different classification criteria in the diagnosis of ERA

The diagnostic value of the 1987 ACR criteria, the 2010 ACR/ EULAR criteria, and the ERA criteria were evaluated by comparing sensitivities, specificities, PPV, NPV, LR+, LR- and the AUC. The sensitivity of the ERA criteria was much higher than that of the 1987 ACR criteria in total cohorts (79.2% *vs*. 54.5%, p<0.001), and similar to 2010 ACR/ EULAR criteria (79.2% *vs*. 78.5%, p=0.883). The specificity of ERA criteria was slightly lower than that of 1987 ACR criteria but higher than 2010 ACR/EULAR criteria (83.7% *vs*.

Table II. The demographic and clinical characteristics of three countries.

	China			Sweden			India		
	ERA (n=202)	non-RA (n=197)	*p*-value	ERA (n=50)	non-RA (n=41)	*p*-value	ERA (n=60)	non-RA (n=56)	*p*-value
Female, n (%)	146 (73.0)	1 33 (67.9)	0.262	35 (70.0)	21 (51.2)	0.067	53 (88.3)	35 (62.5)	0.001
Age (years), median (P_{25}, P_{75})	50.0 (40.0, 61.0)	46.0 (31.5, 55.0)	<0.001	58.5 (39.5, 71.0)	57.5(42.3, 67.0)	0.826	44.5(26.0, 51.8)	35.0(25.0, 49.5)	0.03
Duration (months), median (P_{25}, P_{75})	4.4 (2.2, 8.1)	3.2 (1.6, 7.8)	0.062	3.8 (1.7, 7.9)	2.7 (1.4, 5.4)	0.227	5.6 (3.3, 11.1)	5.7(3.2, 9.4)	0.543
erosion (n; %)	21 (10.5)	1 (0.5)	<0.001	24 (48.0)	9 (22.0)	0.010	32 (53.3)	9 (16.1)	<0.001
Morning Stiffness (minutes), median (P_{25}, P_{75})	45(15, 90)	15 (15, 45)	<0.001	60 (30, 105)	0 (0, 30)	<0.001	30 (15, 120)	30 (22.5, 60)	0.355
RF positive (n, %)	130 (64.4)	34 (17.3)	<0.001	26 (52.0)	3 (7.3)	<0.001	42 (70.0)	5 (8.9)	<0.001
Anti-CCP positive (n, %)	151 (74.8)	10 (5.1)	<0.001	23 (46.0)	0 (0.0)	<0.001	35 (58.3)	0 (0.0)	<0.001
ESR (mm/1h), median (P_{25}, P_{75})	34.0(17.0, 64.5)	25.0(10.0, 48.0)	0.002	28.0(15.0, 53.5)	24.0(15.0, 50.5)	0.534	48.0(39.0, 56.25)	47.5(28.8, 59.0)	0.577
CRP (mg/dL). median (P_{25}, P_{75})	5.6 (1.2, 17.3)	3.1 (0.7, 14.2)	0.052	10.5(3.0, 21.0)	8 .0(2.0, 27.5)	0.481	15.4(6.9, 34.6)	14.4(3.6, 33.6)	0.366

89.1%, $p<0.001$; 83.7% *vs*. 78.2 %, $p=0.02$, respectively). In addition, the NPV of the ERA criteria were similar to the 2010 ACR/EULAR and higher than the 1987 ACR criteria (79.1% *vs*. 77.4% and 64.9%). The PPV of ERA criteria was similar to 1987ACR criteria, and both were higher than 2010 ACR/EULAR (83.7%, 84.2%, and 79.3%, respectively) (Table III).

The LR+ of the ERA criteria was higher than the 2010 ACR/EULAR criteria and lower than the 1987 ACR criteria, while the LR- of the ERA criteria was similar to the 2010 ACR/EULAR criteria and both were lower than the 1987 ACR criteria (Table III). The ROC curves of three criteria in overall population were plotted in Figure 1, which showed that the AUC of the ERA criteria was similar to the 2010 ACR/EULAR criteria (0.878 *vs*. 0.849, $p=0.071$), and much higher than that of the 1987 ACR criteria (0.878 *vs*. 0.792, $p<0.0001$), respectively.

Three countries were involved in our study, and the diagnostic values of the ERA criteria were compared among them. The sensitivity, specificity, PPV, NPV, LR+ and LR- of the three criteria sets at the proposed cut-off points were shown in Table III. The ERA criteria had higher sensitivity and NPV comparing to the 1987 ACR, which had similar sensitivity and specificity comparing to the 2010 ACR/EULAR criteria. And the results of LR in three countries were similar to the total cohort.

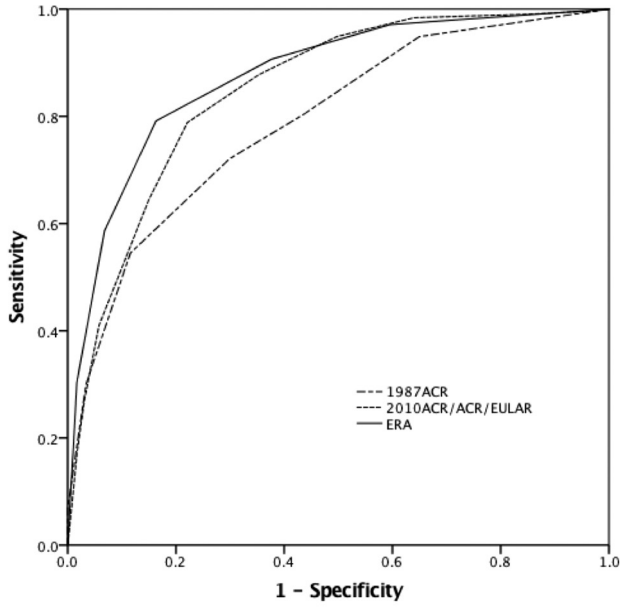
In the Chinese cohort, the AUC of the ERA criteria was 0.877, comparable to that of the 2010 ACR/EULAR criteria (0.855, $p=0.055$), and higher than that of the 1987 ACR criteria (0.792, $p<0.0001$). The Swedish and Indian cohorts yielded consistent results with Chinese cohorts (not shown).

Table III. Sensitivity, specificity, PPV, NPV of the three countries.

	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	LR+	LR-
All patients (n=589)						
ERA	79.2	83.7	83.7	79.1	4.86	0.25
2010 ACR/EULAR	78.3	78.2	79.3	77.4	3.59	0.28
1987 ACR	54.5	89.1	84.2	64.9	5.00	0.51
China (n=399)						
ERA	72.3	87.8	85.9	75.5	5.93	0.32
2010 ACR/EULAR	72.3	83.2	81.6	74.5	4.30	0.33
1987 ACR	39.1	92.4	84.0	59.7	5.14	0.66
Sweden (n=91)						
ERA	88.0	90.2	91.7	86.0	8.98	0.13
2010 ACR/EULAR	82.0	90.2	91.1	80.4	8.37	0.20
1987 ACR	80.0	92.7	93.0	79.2	10.96	0.22
India (n=116)						
ERA	95.0	64.3	74.0	92.3	2.66	0.08
2010 ACR/EULAR	96.7	51.8	68.2	93.5	2.01	0.06
1987 ACR	85.0	75.0	78.5	82.4	3.40	0.20

PPV: positive predictive value; NPV: negative predictive value; LR+: positive likelihood ratio; LR-: negative likelihood ratio.

Fig. 1. Receiver operating characteristic curve of three criteria in the overall population (n=606).



Comparison of different classification criteria in seronegative ERA

We divided the patients into seropositive and seronegative groups, and the values of the three classification criteria areshown in Table IV. The ERA criteria had lower sensitivity but higher specificity in seronegative patients than in seropositive patients (86.3% *vs*. 54.9% and 52.2% *vs*. 89.5%, respectively). For seronegative patients, the sensitivity of ERA and 2010 ACR/EULAR criteria were similar to each other but higher than that of 1987 ACR criteria. The specificity and PPV of the ERA criteria were similar to the 2010 ACR/EULAR criteria. The LR+ of the ERA criteria was between the 2010 ACR/EULAR and 1987 ACR criteria (5.23, 3.80 and 6.22) while the LR- of three criteria were comparable (Table IV).

The AUC of ERA and 2010 ACR/EULAR were comparable 0.785, 0.772 and 0.769 (Fig. 2).

Performance of three classification criteria in very early RA

To evaluate the diagnostic performance of ERA criteria in early stages, we analysed the three criteria for sensitivity, specificity, PPV, NPV according to the disease course, as shown in Figure 3. For the very early patients with symptoms less than 3 months, the ERA criteria had higher sensitivity, specificity, PPV, NPV than 2010 ACR/EULAR criteria (78.4% *vs*. 74.5% for sensitivity, $p=0.454$; 86.0% *vs*. 79.5% for specificity, $p<0.05$). Meanwhile the sensitivity and NPV were higher than 1987 ACR

Table IV. Sensitivity, specificity, PPV, NPV of seropositive and seronegative patients.

	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	LR+	LR-
Seropositive patients (n=288)						
ERA	86.3	52.2	90.4	42.1	1.81	0.26
2010 ACR/EULAR	86.3	34.8	87.4	32.7	1.32	0.39
1987 ACR	56.4	71.7	91.3	23.9	1.99	0.61
Seronegative patients (n=318)						
ERA	54.9	89.5	60.0	87.4	5.23	0.50
2010 ACR/EULAR	52.1	86.3	52.1	86.3	3.80	0.56
1987 ACR	47.9	92.3	64.2	86.1	6.22	0.56

Seropositive: RF and/or anti-CCP were/was positive; Seronegative: both RF and anti-CCP were negative; LR+: positive likelihood ratio; LR-: negative likelihood ratio.

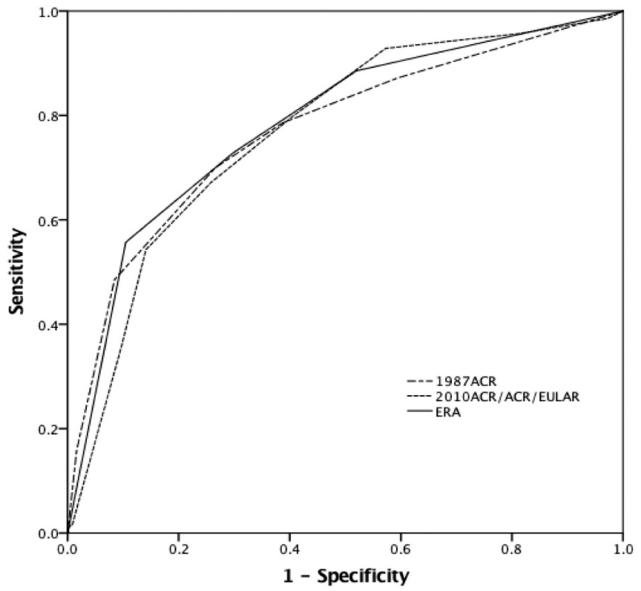


Fig. 2. Receiver operating characteristic curve of three criteria in the seronegative patients (n=318).

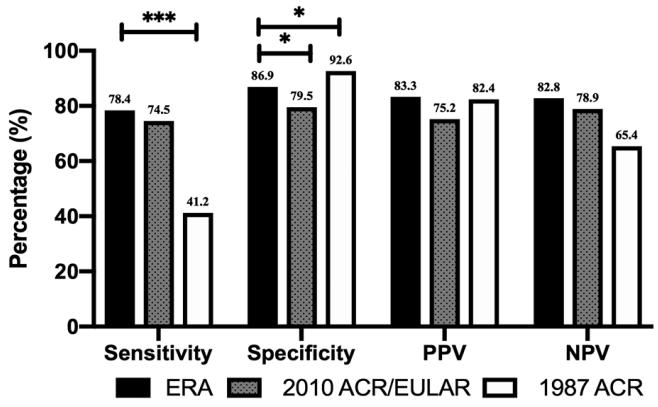


Fig. 3. Sensitivity, specificity, PPV, NPV of three criteria in the very early arthritis patients with duration less than 3 months (significance indicated as *p<0.05, and ***p<0.001).

criteria (78.4% *vs*. 41.2% for sensitivity, p<0.001), respectively.

The ERA criteria had the highest LR+ (10.59, 3.63 and 3.15) and lowest LR- (0.23, 0.32 and 0.68) in comparison with the 2010 and 1987 criteria. The ERA criteria had an AUC of 0.885 in this cohort. In comparison, the AUC of the 1987 ACR criteria and 2010 ACR/EULAR criteria were 0.829 (p=0.007) and 0.743 (p<0.0001) (Fig. 4).

Concordance between the three criteria

In 312 RA patients, 158 patients met three criteria simultaneously. 65 patients met both ERA and 2010 ACR/EULAR criteria, while 8 patients met both ERA and 1987 ACR criteria. Kappa coefficient of the ERA criteria with 2010 ACR/EULAR and 1987 ACR criteria were 0.71 and 0.65, respectively. In RA patients, forty patients were considered discordant: 3 patients only met the 1987 ACR criteria and 21 patients fulfilled only 2010 ACR/EULAR criteria, while 16 patients fulfilled only the ERA criteria respectively. Meanwhile, in non-RA patients, seventy-nine patients met at least one of the three criteria. Among these, 64 met the 2010 ACR/EULAR, more than the patients who met the ERA and 1987 ACR criteria (64 *vs*. 48 and 32, respectively) (Fig. 5).

Discussion

Diagnosing RA is a highly individualised process led by the rheumatologist. The 1987 ACR classification criteria are well accepted as providing the benchmark for disease classification, but have been criticised for their lack of sensitivity in early disease because they were derived by trying to discriminate patients with established RA from other definite rheumatic diseases. Moreover, it does not help achieve the goal of identifying patients who would benefit from early effective intervention (9). Until 2010, ACR and EULAR developed a new classification criteria for RA to facilitate early recognition of RA, to guide therapeutic intervention and also to form homogeneous early RA patient groups for clinical trials (14). Several studies have assessed the diagnostic accuracy of the 2010 ACR/EULAR classification criteria in comparison with the 1987 criteria in several cohorts of patients, which indicates a substantially higher sensitivity but lower specificity, and raises a potential for overclassification and overtreatment (10-12, 18). In the past ten years, we have addressed and classified a variety of RA-related issues, including epidemiology, mechanisms of pathogenesis, early diagnosis and prolonged intention therapy (6, 19, 20). One of them was the establishment of the ERA classification criteria. Then JX Zhao and H Ye *et al*. reported that the ERA criteria showed a sensitivity of 72.3–78.4 % and specificity of 86.3–87.8 % in Chinese cohorts. These prospective studies demonstrate that the ERA criteria may be more sensitive

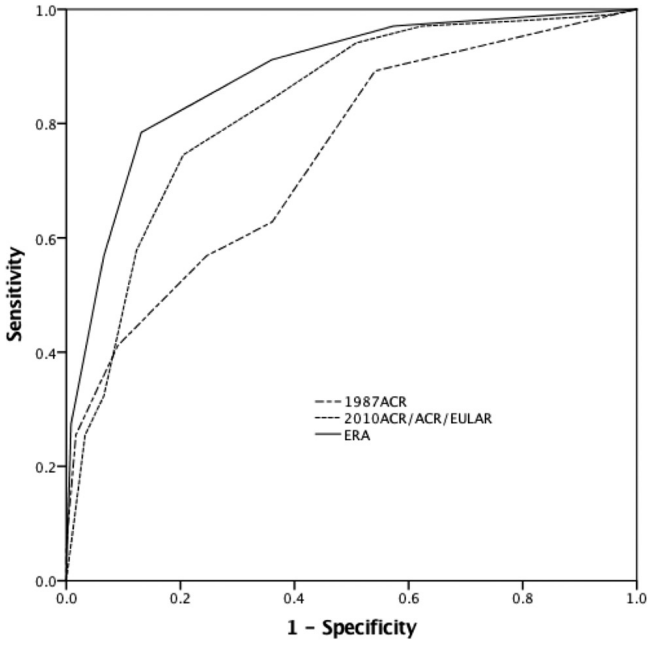


Fig. 4. Receiver operating characteristic curve of three criteria in the very early arthritis patients (n=224).

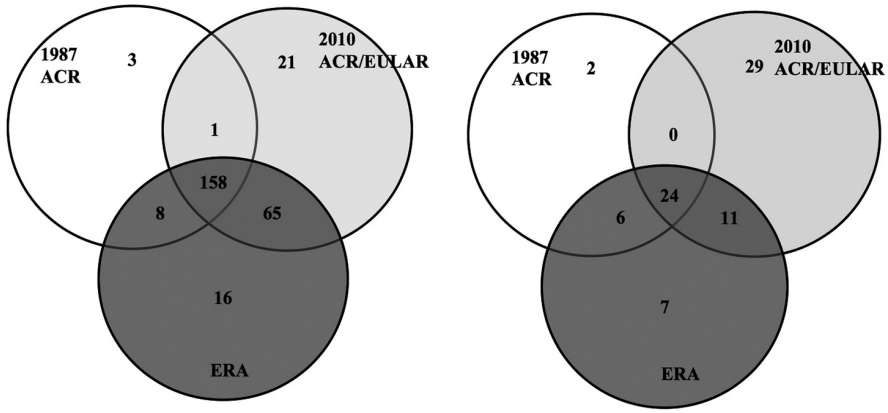


Fig. 5. Venn diagram of the number of patients fulfilling the three different classification criteria in ERA (n=312) and non-RA patients (n=294): (**A**) In ERA patients, 158 patients met the all three criteria, and in patients who did not meet the 1987 ACR and 2010 ACR/EULAR criteria, there were still 16 patients who met the ERA criteria. (**B**) In non-RA patients, 79 fulfilled one of the three criteria. Among them there were 64, 32 and 48 patients who met the 2010 ACR/EULAR, 1987 ACR and the ERA criteria, respectively.

or feasible than the 1987 ACR and 2010 ACR/EULAR criteria for the detection of RA in early stages (16, 17).

Our aim in this study was to validate the real-world performance of the new criteria in different rheumatic diseases in an international multicentre cross-sectional study. In the current study we compared three different classification criteria for early RA and investigated the performance of the ERA classification criteria in the international multicentre cohorts. Comparing to the 1987 criteria, the sensitivity, NPV, and the AUC were higher under the ERA criteria, while the specificity was slightly lower. The higher detection ability of the newly proposed criteria was desirable because several studies reviewed that the 1987 criteria would 'miss' RA at an early stage (8, 9). The sensitivity, specificity, LR+ and AUC of the ERA criteria were comparable to or slightly higher than the 2010 ACR/EULAR criteria in our study, while the LR- was comparable between the ERA and 2010 ACR/EULAR criteria, indicating that the performance of the ERA criteria might be similar to or better than 2010 criteria in early arthritis patients. In addition, the ERA criteria seems to provide an easier way than the 2010 ACR/EULAR criteria which assesses all of the joints and uses a scoring system complicated in clinical practice lacking of the application convenience. However, international multicentre prospective studies are still needed to determine causal relationships and to avoid recall bias.

Three countries were involved in our study. The diagnostic value of the ERA criteria was evaluated among them. The results of three cohorts were consistent with the overall. However, the sensitivity and specificity of Sweden was higher while the Indian cohort had higher sensitivity but lower specificity than Chinese cohort. The reason why the performance of the new criteria differs in these cohorts is not clear, which may be due to the higher rate of bone erosion in Swedish and Indian cohorts, the influence of geographic location and ethnicity for the diagnosis of RA.

Individuals who have seronegative arthritis may be missed for classification as RA due to the relative weight of the autoantibodies among the criteria before (14). Our study showed that the sensitivity of the ERA criteria in seronegative cohort is 54.9%, similar to the 2010 ACR/EULAR and higher than the 1987 ACR criteria, which means that both the ERA and 2010 ACR/EULAR criteria are useful to detect RA patients, even for those seronegative patients. However, the specificity of three criteria in seropositive patients were decreased, especially of the 2010 ACR/EULAR criteria, consistent with the study of Kaneko *et al*. (13), which might result from the misclassification as RA caused by a small number of swollen joints, non-specific high-titre RF positivity and mildly elevated ESR. In very early arthritis (≤3 months), the sensitivity, LR+, and the AUC of the ERA criteria exceeded that of the 1987 ACR and 2010 ACR/EULAR criteria, and LR- of the ERA criteria was the lowest, indicating that the ERA criteria might have the highest diagnostic value in the very early inflammatory arthritis patients. However, further studies with larger populations, combined with objective measures is needed to confirm it.

In addition, the commonly used gold standards, such as objective diagnosis by

physicians, experimental treatment with methotrexate (MTX), or disease-modifying anti-rheumatic drugs (DMARDs), may cause moderate levels of variations (12, 14, 15, 21). The diagnosis by physicians may be affected by long-time influence of the existed criteria, so the sensitivity might be overestimated, and its accuracy also depends on the physicians' understanding of RA. However, all rheumatologists in this study were experienced and most of the diagnoses were believed to be correct. Since the centres involved were major academic medical institutes, there are possibilities that the enrolled subjects were more likely to have RA, which lead to the fact that the PPV might be highly estimated. However, the strength of the study lies in its multicentre and international nature so that the results can be seen as a representative of early arthritis cohorts anywhere.

In conclusion, the ERA criteria showed a better sensitivity and NPV than 1987 ACR criteria, and compared to 2010 ACR/EULAR criteria, the ERA criteria is more specific and feasible in daily practice for early RA diagnosis. There is a considerable concordance of the ERA criteria with 2010 ACR/EULAR and 1987 ACR criteria.

References

1. RAZA K: The Michael Mason prize: early rheumatoid arthritis--the window narrows. *Rheumatology* (Oxford) 2010; 49: 406-10.
2. NELL VP, MACHOLD KP, EBERL G, STAMM TA, UFFMANN M, SMOLEN JS: Benefit of very early referral and very early therapy with disease-modifying anti-rheumatic drugs in patients with early rheumatoid arthritis. *Rheumatology* (Oxford) 2004; 43: 906-14.
3. QUINN MA, CONAGHAN PG, O'CONNOR PJ *et al*.: Very early treatment with infliximab in addition to methotrexate in early, poor-prognosis rheumatoid arthritis reduces magnetic resonance imaging evidence of synovitis and damage, with sustained benefit after infliximab withdrawal: results from a twelve-month randomized, double-blind, placebo-controlled trial. *Arthritis Rheum* 2005; 52: 27-35.
4. FINCKH A, LIANG MH, VAN HERCKENRODE CM, DE PABLO P: Long-term impact of early treatment on radiographic progression in rheumatoid arthritis: A meta-analysis. *Arthritis Rheum* 2006; 55: 864-72.
5. VAN DER LINDEN MP, LE CESSIE S, RAZA K *et al*.: Long-term impact of delay in assessment of patients with early arthritis. *Arthritis Rheum* 2010; 62: 3537-46.
6. LI R, ZHAO JX, SU Y *et al*.: High remission and low relapse with prolonged intensive DMARD therapy in rheumatoid arthritis (PRINT): A multicenter randomized clinical trial. *Medicine* (Baltimore) 2016; 95: e3968.
7. ARNETT FC, EDWORTHY SM, BLOCH DA *et al*.: The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988; 31: 315-24.
8. SARAUX A, BERTHELOT JM, CHALES G *et al*.: Ability of the American College of Rheumatology 1987 criteria to predict rheumatoid arthritis in patients with early arthritis and classification of these patients two years later. *Arthritis Rheum* 2001; 44: 2485-91.
9. ALETAHA D, NEOGI T, SILMAN AJ *et al*.: 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Ann Rheum Dis* 2010; 69: 1580-8.
10. ALVES C, LUIME JJ, VAN ZEBEN D *et al*.: Diagnostic performance of the ACR/EULAR 2010 criteria for rheumatoid arthritis and two diagnostic algorithms in an early arthritis clinic (REACH). *Ann Rheum Dis* 2011; 70: 1645-7.
11. CADER MZ, FILER A, HAZLEHURST J, DE PABLO P, BUCKLEY CD, RAZA K: Performance of the 2010 ACR/EULAR criteria for rheumatoid arthritis: comparison with 1987 ACR criteria in a very early synovitis cohort. *Ann Rheum Dis* 2011; 70: 949-55.
12. VAN DER LINDEN MP, KNEVEL R, HUIZINGA TW, VAN DER HELM-VAN MIL AH: Classification of rheumatoid arthritis: comparison of the 1987 American College of Rheumatology criteria and the 2010 American College of Rheumatology/European League Against Rheumatism criteria. A*rthritis Rheum* 2011; 63: 37-42.
13. KANEKO Y, KUWANA M, KAMEDA H, TAKEUCHI T: Sensitivity and specificity of 2010 rheumatoid arthritis classification criteria. *Rheumatology* (Oxford) 2011; 50: 1268-74.
14. BILIAVSKA I, STAMM TA, MARTINEZ-AVILA J *et al*.: Application of the 2010 ACR/EULAR classification criteria in patients with very early inflammatory arthritis: analysis of sensitivity, specificity and predictive values in the SAVE study cohort. *Ann Rheum Dis* 2013; 72: 1335-41.
15. RAJA R, CHAPMAN PT, O'DONNELL JL *et al*.: Comparison of the 2010 American College of Rheumatology/European League Against Rheumatism and the 1987 American Rheumatism Association classification criteria for rheumatoid arthritis in an early arthritis cohort in New Zealand. *J Rheumatol* 2012; 39: 2098-103.
16. ZHAO J, SU Y, LI R *et al*.: Classification criteria of early rheumatoid arthritis and validation of its performance in a multi-centre cohort. *Clin Exp Rheumatol* 2014; 32: 667-73.
17. YE H, SU Y, LI R *et al*.: Comparison of three classification criteria of rheumatoid arthritis in an inception early arthritis cohort. *Clin Rheumatol* 2016; 35: 2397-401.
18. VARACHE S, CORNEC D, MORVAN J *et al*.: Diagnostic accuracy of ACR/EULAR 2010 criteria for rheumatoid arthritis in a 2-year cohort. *J Rheumatol* 2011; 38: 1250-7.
19. LI ZG: A new look at rheumatology in China--opportunities and challenges. *Nat Rev Rheumatol* 2015; 11: 313-7.
20. LI R, SUN J, REN LM *et al*.: Epidemiology of eight common rheumatic diseases in China: a large-scale cross-sectional survey in Beijing. *Rheumatology* (Oxford) 2012; 51: 721-9.
21. KENNISH L, LABITIGAN M, BUDOFF S *et al*.: Utility of the new rheumatoid arthritis 2010 ACR/EULAR classification criteria in routine clinical care. *BMJ Open* 2012; 2.