

The Diagnostic Performance of the 2010 American College of Rheumatology/European League Against Rheumatism Classification Criteria for Rheumatoid Arthritis: A Systematic Review and Meta-Analysis

Amerikan Romatoloji Derneği/Avrupa Romatizma Birliği 2010 Romatoid Artrit Sınıflama Kriterlerinin Tanısal Değeri: Sistematik Derleme ve Meta-Analiz

Yu-Hui ZHANG,¹ Wei-Min SONG,² Jian-Hong SUN,¹ Jia-Chuan XIONG,³ Guixiu SHI⁴

¹Department of Rheumatology and Clinical Immunology, West China Hospital/West China Medical School, Sichuan University, Chengdu, Sichuan, China

²Department of Rheumatology, Affiliated Hospital of Chengdu University of Traditional Chinese Medicine, Chengdu, Sichuan, China

³Department of Nephrology, West China Hospital/West China Medical School, Sichuan University, Chengdu, Sichuan, China

⁴Department of Rheumatology and Clinical Immunology, The First Affiliated Hospital of Xiamen University, Xiamen, Fujian, China

Objectives: This study aims to assess the diagnostic performance of the 2010 American College of Rheumatology (ACR) European League Against Rheumatism (EULAR) classification criteria for rheumatoid arthritis (RA).

Patients and methods: Between January 2010 and November 2012, an electronic search was conducted using MEDLINE (via PubMed), EMBASE, and Cochrane CENTRAL to find studies related to the diagnostic performance of the 2010 ACR/EULAR classification criteria for RA in patients with inflammatory synovitis. Subgroup analyses were performed according to the adopted gold standards. The sensitivity and specificity of the 2010 criteria were extracted or calculated. Summary receiver operating characteristic (sROC) curves were drawn to evaluate the differences between the 2010 and 1987 criteria.

Results: We found 10 studies which were eligible for inclusion. Based on the 2010 criteria, the pooled information showed a sensitivity of 0.804 (95% confidence interval [CI]; range 0.737 to 0.857) and a specificity of 0.556 (95% CI; range 0.417 to 0.687) in the methotrexate (MTX) group, a sensitivity of 0.706 (95% CI; range 0.585 to 0.803) and a specificity of 0.691 (95% CI; range 0.583 to 0.782) in the disease-modifying antirheumatic drug (DMARD) group, and a sensitivity of 0.901 (95% CI; range 0.856 to 0.933) and a specificity of 0.539 (95% CI; range 0.429 to 0.645) in the expert group. The sensitivity and specificity were pooled under the bivariate binomial mixed model in this review due to the heterogeneous nature of the data. Compared to the 1987 criteria, the 2010 criteria showed a higher sensitivity, lower specificity and similar accuracy rate with more accuracy rate in the MTX group, compared the DMARD and expert groups.

Conclusion: The 2010 ACR/EULAR criteria for RA have a better discriminative and diagnostic ability, higher sensitivity, and lower specificity than the 1987 criteria; therefore, we believe that it cannot be substituted for the 1987 criteria.

Key words: Diagnosis; rheumatoid arthritis; systematic review.

Amaç: Bu çalışmanın amacı romatoid artrit (RA) 2010 Amerikan Romatoloji Derneği (ACR) ve Avrupa Romatizma Birliği (EULAR) sınıflandırma kriterlerinin tanısal değerini değerlendirmektir.

Hastalar ve yöntemler: Ocak 2010 - Kasım 2012 tarihleri arasında inflamatuvar sinoviti olan RA hastalarında 2010 ACR/EULAR sınıflandırma kriterlerinin tanısal değerini inceleyen çalışmalara ulaşmak amacıyla MEDLINE (PubMed aracılığıyla), EMBASE ve Cochrane CENTRAL'da elektronik tarama yapıldı. Kabul gören altın standartlara göre alt grup analizleri yapıldı. 2010 kriterlerinin duyarlılığı ve özgüllüğü elde edildi veya hesaplandı. 2010 ve 1987 kriterleri arasındaki farkı değerlendirmek için özet alıcı işlem karakteristiği (sROC) eğrileri çizildi.

Bulgular: Çalışmaya alınmaya uygun 10 araştırma bulundu. 2010 kriterlerine göre, elde edilen veriler metotreksat (MTX) grubu için 0.804 duyarlılık (%95 güven aralığı [GA]; dağılım 0.737-0.857) ve 0.556 özgüllük (%95 GA; dağılım 0.417-0.687); hastalık modifiye edici antiromatizmal ilaç (DMARD) grubunda 0.706 duyarlılık (%95 GA; dağılım 0.585-0.803) ve 0.691 özgüllük (%95 GA; dağılım 0.583-0.782) ve uzman grubunda 0.901 duyarlılık (%95 GA; dağılım 0.856-0.933) ve 0.539 özgüllük (%95 GA; dağılım 0.429-0.645) gösterdi. Bu incelemede verilerin heterojen olması nedeniyle, duyarlılık ve özgüllük iki değişkenli binominal karma model ile belirlendi. 1987 kriterlerine göre, 2010 kriterlerinin duyarlılığı daha yüksek, özgüllüğü daha düşük ve doğruluk oranı benzer olup, MTX grubunun doğruluk oranı, DMARD ve uzman gruplara kıyasla daha yüksekti.

Sonuç: 2010 ACR/EULAR RA kriterlerinin ayırt edici ve tanı oranı daha iyi, duyarlılığı daha yüksek ve özgüllüğü 1987 kriterlerinden daha düşük olduğundan, 1987 kriterlerinin yerine kullanılmayacağı kanısındayız.

Anahtar sözcükler: Tanı; romatoid artrit; sistematik derleme.

Received: April 07, 2013 Accepted: July 19, 2013

Correspondence: Guixiu Shi, M.D. Department of Rheumatology and Clinical Immunology, The First Affiliated Hospital of Xiamen University, Xiamen, 361003 Fujian, China. Tel: 86-592-2137769 e-mail: guixiu.shi@gmail.com

©2013 Turkish League Against Rheumatism. All rights reserved.

Rheumatoid arthritis (RA) is the most common inflammatory arthritis with a prevalence rate of approximately 1.0% worldwide.^[1] With this disease, progressive joint erosion and deformity cause work disability and mortality; thus, early diagnosis and intervention is crucial to ensure a better prognosis.^[2]

One of the most common criteria, the 1987 American College of Rheumatology (ACR) classification criteria for RA, has been criticized for its lower sensitivity with regard to early arthritis.^[3] Therefore, a joint working group from the ACR and the European League Against Rheumatism (EULAR) developed the 2010 classification criteria to identify and facilitate patients at high risk of persistent disease and erosive damage.^[4] This criteria contains four categories (joint involvement, serology, acute-phase reactants, and duration of symptoms), and a score of $\geq 6/10$ is classified as definite RA since the typical pattern of destructive RA seen on radiographs provides sufficient evidence for this diagnosis, precluding the need for applying additional criteria.

To date, several studies have evaluated the diagnostic performance and discriminative ability of the 2010 criteria. Hence, the aim of this systemic review was to assess the diagnostic values of the 2010 ACR/EULAR criteria for RA and compare them with the 1987 ACR criteria.

PATIENTS AND METHODS

The review was registered in the International Prospective Register of Systematic Reviews (PROSPERO) in November 21, 2012 (No. CRD42012003308), and Checklist 1 for the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) is available as supporting information.

Search strategy

We searched the MEDLINE (through PubMed), EMBASE, and Cochrane CENTRAL databases between January 1, 2010 and November 19, 2012 using the following search terms: RA, prevalence, sensitivity, specificity, accuracy. In addition, we also examined the abstracts of the EULAR annual congresses and the ACR annual meetings and also checked the references of the included studies and reviews to search for clues to identify additional relevant studies. There was no language limitation in this search.

Inclusion criteria

The eligibility criteria were the following: (i) A case-control study or cohort study; (ii) Patients with inflammatory synovitis and no other definitive diagnosis other than RA or undifferentiated arthritis (UA); (iii) The use of the 2010 ACR/EULAR classification criteria for RA as the index test along with the initiation of methotrexate (MTX) or disease-modifying anti-rheumatic drugs (DMARDs) and expert opinion RA (the primary expert's diagnosis of RA) as the reference standard; and (iv) The use of sensitivity and specificity as the primary outcomes so that a 2x2 table (true positive, false positive, true negative, and false negative) could be achieved.

Study selection and data extraction

Two of the authors (Zhang YH and Song WM) of this article independently selected the studies, and disagreements were resolved by discussing or consulting with a third author (Shi GX). The characteristics and outcome data were extracted using the data extraction form for each study. The outcome data, including the sensitivity, specificity, and the 2x2 table, were acquired from the text or calculated using the Review Manager (RevMan) 5.1 software (The Nordic Cochrane Center, the Cochrane Collaboration, Copenhagen, Denmark).

Methodological quality assessment and statistical analysis

The methodological quality of the included studies was assessed using a checklist of 11 items recommended by the Cochrane Collaboration based on Quality Assessment of Diagnostic Accuracy Studies (QUADAS), a tool to determine diagnostic accuracy in systematic reviews.^[5] Since clinical heterogeneity was present in the articles used for our review, when assessing the diagnostic values of the 2010 criteria and the 1987 criteria, three subgroups were utilized based on the main gold standards for RA that were adopted in the studies: the MTX group, the DMARD group, and the expert group. The subgroup analyses were conducted according to the gold standards adopted in the included studies. In addition, the heterogeneity of each subgroup was investigated using the Q statistic and I^2 statistic, with $p \leq 0.10$ and $I^2 \geq 50\%$ signifying substantial heterogeneity. If possible, the sensitivity and specificity were pooled, and hierarchical summary receiver operating characteristic (HSROC) curves were launched. A meta-analysis was also

conducted under the bivariate binomial mixed model if heterogeneity was present, and summary receiver operating characteristic (SROC) curves were performed to evaluate the differences in diagnostic accuracy between the 2010 ACR/EULAR criteria and 1987 ACR criteria. The Stata 12.0 software program (StataCorp LP, College Station, TX, USA) and RevMan 5.1.5 software for Windows (the Nordic Cochrane Center, the Cochrane Collaboration, Copenhagen, Denmark) were used for all data calculations.

RESULTS

Results of the search

Our electronic search identified 2,438 citations, with 1,461 from PubMed, 915 from EMBASE, and 62 from Cochrane CENTRAL. After the titles and abstracts were screened, 25 full-text articles were retrieved. Of these, only 10 studies^[6-15] were eligible for inclusion. A study flow diagram showing these articles is shown in Figure 1.^[16]

Description of the studies

These 10 studies included cohorts from the Netherlands,^[6-8] the United States,^[13] France,^[10] UK,^[9] New Zealand,^[14] Spain,^[15] South Korea,^[11] and

Japan.^[6,12] The three studies^[6-8] from the Netherlands used different cohort databases: the Rotterdam Early Arthritis Cohort[®] (REACH), the Stop Arthritis Very Early (SAVE) trial, and the Leiden early arthritis clinic (EAC) cohort). All of the studies enrolled their participants from 2000 onwards when MTX was already widely used as the first line of treatment. Among these studies, three did not aim to calculate the sensitivity and specificity as the outcomes. One attempted to identify the patients according to whether they required DMARDs or MTX,^[8] another looked for agreement between the 2010 ACR/EULAR criteria and the 1987 ACR criteria,^[9] and one made comparisons between patients with and without RA according to the 2010 ACR/EULAR criteria.^[12] Seven studies accepted two or three gold standards while the others only had one gold standard. This was identified as MTX in five studies,^[6,8-10,15] DMARDs in seven,^[7,9-12,14,15] and the primary expert's diagnosis of RA in five others.^[7,9-12,14,15] The main characteristics of all of the included studies are provided in Table 1.

Methodological quality of the included studies

The methodological quality of the included studies was assessed using a checklist recommended by the Cochrane Collaboration based on the QUADAS.

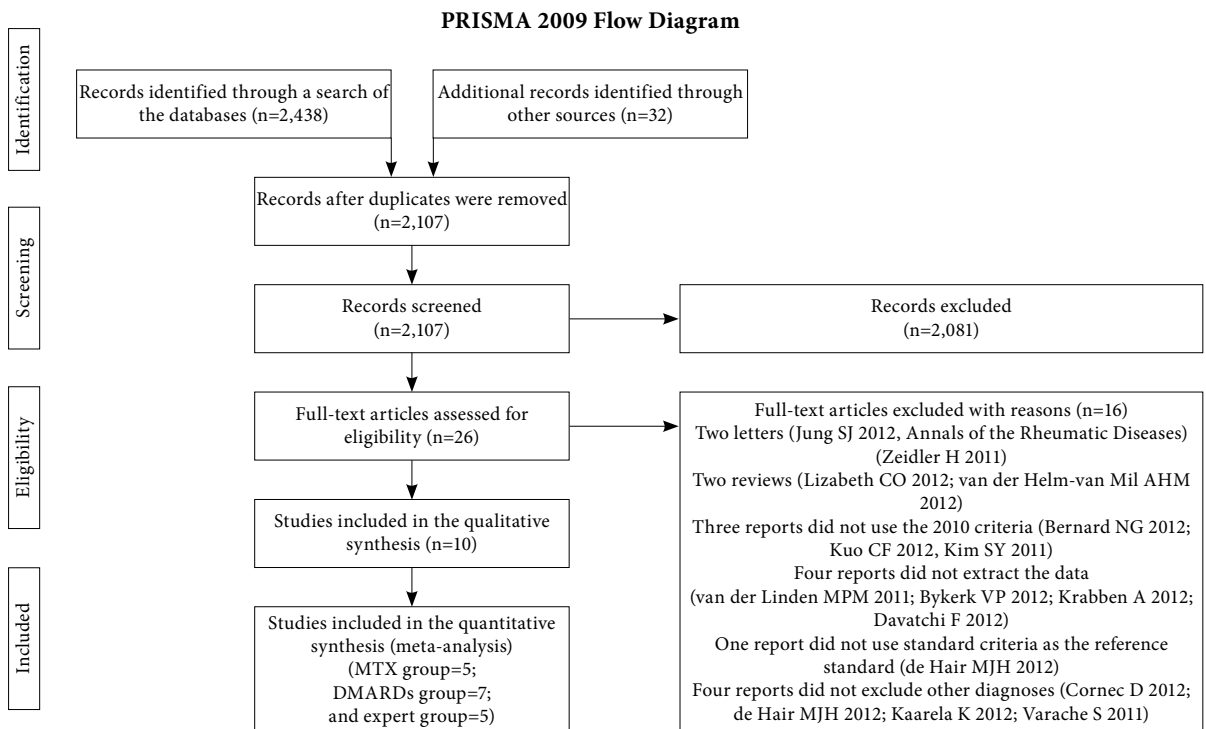


Figure 1. PRISMA 2009 flow diagram.^[16]

Table 1. Characteristics of all included studies

Study ID	Gold standard	Women (%)	Age (years)	Sample size	Duration	Swollen joint	Follow-up	Withdrawal	Country
Alves et al. ^[6]	MTX, persistent	68	53±14	207	Less than 12 months	4 (mean)	12 months	3.54% (31/875)	The Netherlands
Britsemmer et al. ^[8]	MTX, expert, erosion	69	52±13	455	Less than 24 months	5	12 months	NA	The Netherlands
Cader et al. ^[9]	MTX, DMARDs	52	49 (median)	205	42 days (median)	3	18 months	2.21% (6/271)	UK
Fautrel et al. ^[10]	MTX, DMARDs, expert	76.8	48.1±12.5	811	6 weeks-6 months	7.2±5.4	At baseline, 6 months	14.67% (119/811)	France
Kaneko et al. ^[12]	DMARDs	79	54 (median)	82	18 weeks (median)	At least 1	NA	0.32% (1/314)	Japan
Kennish et al. ^[13]	Expert	77.7	48.2	112	5.3 years	NA	NA	11.11% (14/126)	USA
Reneses et al. ^[15]	MTX, DMARDs, expert	71.6	51.4±17.2	201	6.3±3.8 months	12.0±6.9	12 months	7.04% (33/469)	Spain
Jung et al. ^[11]	DMARDs	71.6	46.8±14.3	102	10.1±13.7 months	1.6±1.0	19.7±12.8 months	NA	South Korea
Biliavska et al. ^[7]	DMARDs	75.2	47.72±16.47	303	57 days (median)	5 (median)	12 months	20.89% (80/389)	The Netherlands
Raja et al. ^[14]	DMARDs, erosion	67	57.6 (mean)	79	Less than 12 months	3 (median)	24 months	31.62% (37/117)	New Zealand

NA: No Available.

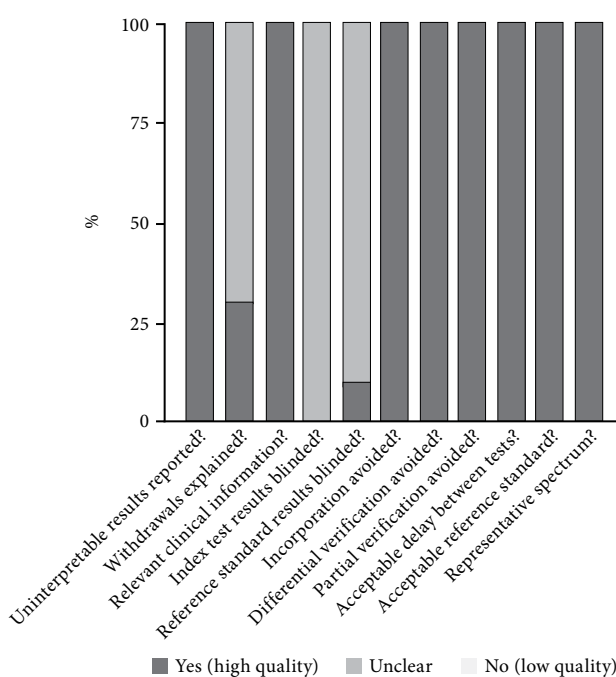
Blindness was an important in the evaluation of the methodological quality of the diagnostic accuracy test; however, there was no information related to this item in most of the included studies. Only one article^[7] included blindness, and it stated that all patients received a single intramuscular injection

of 120 mg of methylprednisolone or a placebo. This achieved the blindness between the interpretation of the index test results and knowledge of the reference standard results. The results of each methodological quality item are presented in Figure 2, but no funnel plot of publication bias was done because of the small amount of studies included in each group.

Findings

Five studies from Europe that used the 2010 criteria in the MTX group^[6,8-10,15] were included in our review. The sensitivity of this group ranged from 0.67 to 0.88, and the specificity was between 0.30 and 0.72. The data was pooled using the bivariate binomial mixed model because of the identified heterogeneity ($p < 0.001$, $I^2 = 97.30$). The 2010 ACR/EULAR classification criteria for RA showed a pooled sensitivity of 0.804 [95% confidence interval (CI); range 0.737-0.857], a pooled specificity of 0.556 (95% CI; range 0.417-0.687), a positive likelihood ratio (LR+) of 1.811 (range 1.408-2.330), a negative likelihood ratio (LR-) of 0.352 (range 0.298-0.417), a diagnostic odds ratio (DOR) of 5.141 (range 3.740-7.069), and an area under the receiver operating characteristic (AUROC) curve of 0.78 (range 0.74-0.81).

Seven studies^[7,9-12,14,15] were included in the DMARD group with a sensitivity between 0.34 and 0.86, and a specificity between 0.40 and 0.78.

**Figure 2.** Methodological quality graph.

The heterogeneity was also identified ($p < 0.001$, $I^2 = 97.07$). The pooled sensitivity was 0.706 (95% CI; range 0.585-0.803), and the pooled specificity was 0.691 (range 0.583-0.782). In addition, the LR+ was calculated as 2.286 (range 1.837-2.845) and the LR- as 0.426- (range 0.325-0.558). Furthermore, there was a DOR of 5.370 (range 3.922-7.352) and an AUROC of 0.75 (range 0.71-0.79).

The expert opinion group (expert group for short) was composed of five studies,^[7,8,10,13,15] and

heterogeneity was also present ($p < 0.001$, $I^2 = 95.87$). The sensitivity of this group ranged from 0.85 to 0.97 while the specificity varied from 0.35 to 0.73. The 2010 criteria had a pooled sensitivity of 0.901 (95% CI; range 0.856-0.933), a pooled specificity of 0.539 (range 0.429-0.645), an LR+ of 1.945 (range 1.589-2.403), an LR- of 0.183 (range 0.142-0.236), a DOR of 10.674 (range 8.056-14.144), and an AUROC of 0.84 (range 0.81-0.87). The forest plots and HSROCs of the three groups are shown in Figures 3 and 4. Compared to the SROC, the HSROC allows for different parameters to

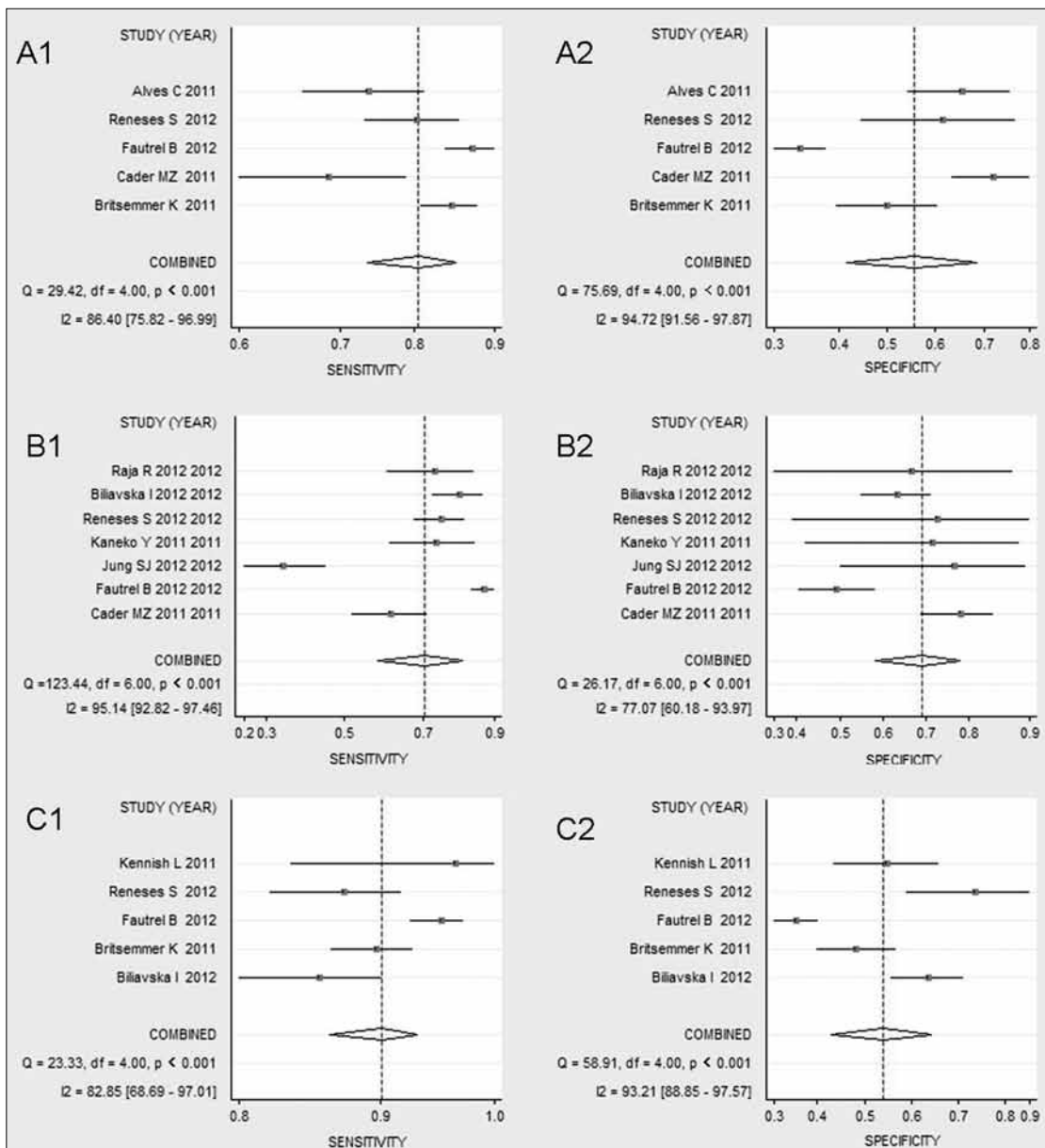


Figure 3. Forest plot of the 2010 ACR/EULAR criteria. A1: Sensitivity of the MTX group; A2: Specificity of the MTX group; B1: Sensitivity of the DMARD group; B2: Specificity of the DMARD group; C1: Sensitivity of the expert group; C2: Specificity of the expert group.

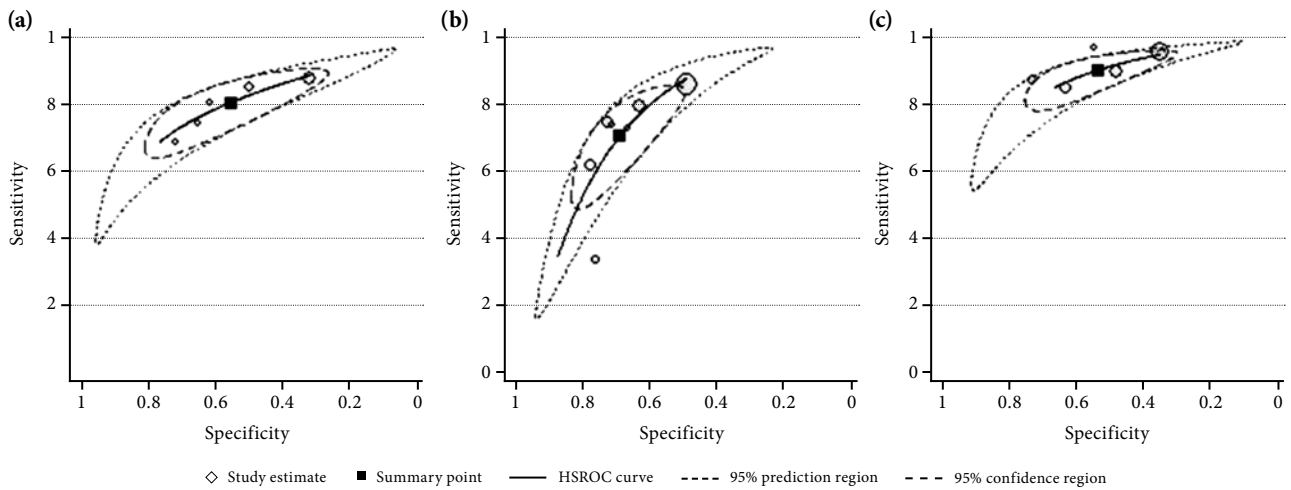


Figure 4. Hierarchical summary receiver operating characteristic (HSROC) curves of the 2010 ACR/EULAR criteria. The horizontal axis shows the specificity, the vertical coordinate shows the sensitivity, and the solid line shows the HSROC curve. The dot shows the sensitivity and the specificity of each study while the square frame shows the summary point (pooled specificity and pooled sensitivity). A: HSROC of the MTX group; B: HSROC of the DMARD group; C: HSROC of the expert group.

be defined within the same model which provides a general framework for the meta-analysis of diagnostic test studies.

Even though most of the included studies did not consider radiographic information when using the 2010 criteria, one study^[9] judged a participant with erosion and a score of less than six as being positive under this criteria.

We also evaluated the sensitivity and specificity of the 1987 criteria and found measurements of 0.42-0.82 and 0.40-0.88 in MTX group, 0.38-0.79 and 0.50-0.93 in DMARD group, 0.69-0.94 and 0.49-0.94 in the expert group, respectively. Our data showed that the 2010 criteria had more sensitivity and less specificity

in each study. Furthermore, the SROC curves of the 2010 were more accurate in the MTX group but less accurate in the DMARD and expert groups. The SROC curve results for the 1987 and 2010 criteria are given in Figure 5.

DISCUSSION

Methodological quality

Blindness is an important item that can be utilized to evaluate the methodological quality of diagnostic accuracy tests. Additionally, the interpretation of the results of the index test may be influenced by the knowledge of the results of the reference standard and vice versa. Blindness was the 11th item in the

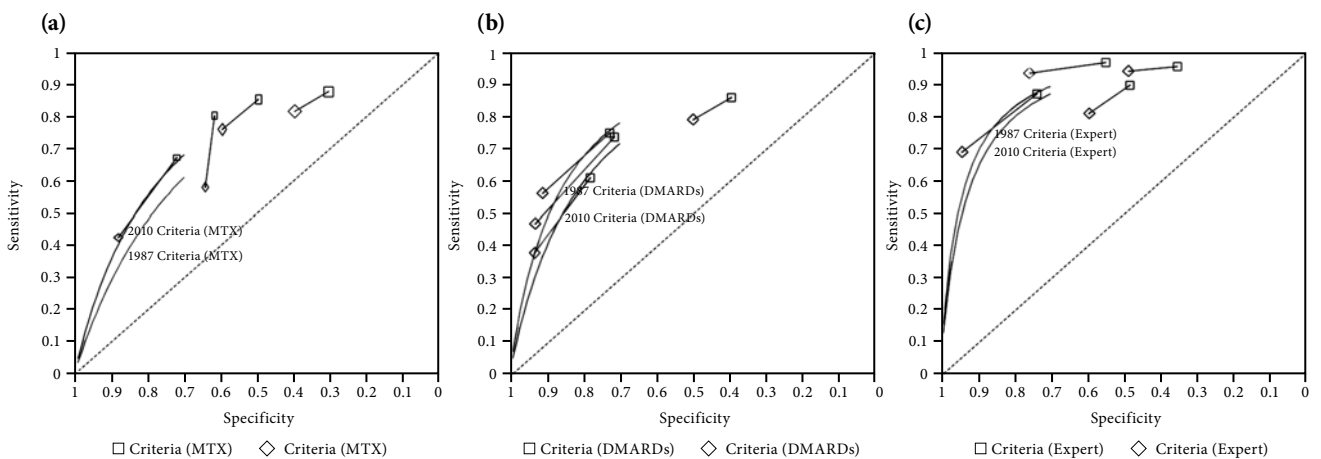


Figure 5. Summary receiver operating characteristic (SROC) curves of the 2010 criteria versus the 1987 criteria. A: SROC of the MTX group; B: SROC of the DMARD group; C: SROC of the expert group.

reporting checklist developed by the Standards for the Reporting of Diagnostic Accuracy (STARD) steering committee.^[17] However, it was difficult to assess the blindness item in this review because there was no explicit information for this in most of the included studies.

Diagnostic values of the 2010 criteria

To evaluate the diagnostic value of the 2010 criteria, we did not include the studies which enrolled participants who were not eligible for assessment by the 2010 criteria. The participants with at least one swollen joint and with no diagnosis other than RA or UA represented the spectrum of those who received the test in practice. Under the 2010 criteria, the results of the pooled sensitivity and specificity for the three groups corroborated the findings of another reviewer^[18] who presented only the diagnostic values of the 2010 ACR/EULAR criteria for RA but did not compare the 2010 and 1987 criteria.

Difference between the 2010 criteria and the 1987 criteria

Though the data of both the 2010 criteria and the 1987 criteria were not pooled due to the heterogeneity, the data in each study showed that the 2010 criteria had more sensitivity and less specificity than the 1987 criteria. With regard to the SROC, both criteria performed similarly in terms of diagnostic accuracy, but the 2010 criteria did slightly better in the MTX group. Moreover, the lower specificity of the 2010 criteria as it related to the diagnostic criteria could lead to a higher false positive rate and unnecessary treatment, particularly with the potentially toxic DMARDs. In addition, the 2010 criteria had a better discriminative and diagnostic ability which allowed it to identify the risks of symptom persistence and structural damage at an earlier stage. However, the cost-effectiveness of the early initiation of DMARDs is difficult to estimate.

Limitation of the review

There were only 10 studies included, and heterogeneity was significantly identified in this review. After checking the data again, we tried to explore the source of the heterogeneity via a meta-regression method but failed because of a lack of valid information. For example, some studies reported the symptom duration using the median data while others used the mean. In actuality, inflammatory arthritis, especially UA, spans a wide spectrum of heterogeneous conditions with a variety of natural

courses, and patients from different countries and regions may experience different courses depending on the proportions of benign, self-limiting forms and the development of overt RA. Hence, we believe that more trials are needed to verify the diagnostic values of the 2010 criteria.

The 2010 ACR/EULAR classification criteria did not include erosion in the scoring system in order to focus on the earlier course of the disease. In the later stages of RA, patients with typical erosions were deemed to have prima facie evidence of RA. Only one of the included studies^[12] enrolled patients with early arthritis as well as those with a symptom duration of more than three years. Furthermore, another study^[9] considered the radiological evidence of erosion by using only RA or UA as a definitive RA under the 2010 ACR/ EULAR criteria. As a classification and diagnosis criteria system, the aim of the 2010 criteria was that it could be applied to all patients.

Conclusion

Although the 2010 ACR/EULAR criteria for RA had a better discriminative and diagnostic ability as well as a higher sensitivity and lower specificity, it cannot be substituted for the 1987 criteria. Further studies are needed that focus on the development and improvement of the diagnostic performance of the 2010 classification criteria for the RA.

Acknowledgements

We would like to thank the authors of the included studies who replied to our e-mail correspondence in our efforts to gather the necessary relative information and data. However, there were no competing interests between their studies and ours.

Declaration of conflicting interests

The authors declared no conflicts of interest with respect to the authorship and/or publication of this article.

Funding

This article is supported from National Natural Science Foundation of China (NSFC) Grant NO. 81273285 to Dr. Guixiu Shi.

REFERENCES

1. Vogt T. Rheumatoid arthritis--clinical picture and important differential diagnoses. *Ther Umsch* 2005;62:265-8. [Abstract]
2. 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.

3. Banal F, Dougados M, Combes C, Gossec L. Sensitivity and specificity of the American College of Rheumatology 1987 criteria for the diagnosis of rheumatoid arthritis according to disease duration: a systematic literature review and meta-analysis. *Ann Rheum Dis* 2009;68:1184-91.
4. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO 3rd, et al. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Arthritis Rheum* 2010;62:2569-81.
5. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003;3:25.
6. Alves C, Luime JJ, van Zeben D, Huisman AM, Weel AE, Barendregt PJ, 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.
7. Biliavska I, Stamm TA, Martinez-Avila J, Huizinga TW, Landewé RB, Steiner G, 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.
8. Britsemmer K, Ursun J, Gerritsen M, van Tuyl LH, van Schaardenburg D. Validation of the 2010 ACR/EULAR classification criteria for rheumatoid arthritis: slight improvement over the 1987 ACR criteria. *Ann Rheum Dis* 2011;70:1468-70.
9. 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.
10. Fautrel B, Combe B, Rincheval N, Dougados M; ESPOIR Scientific Committee. Level of agreement of the 1987 ACR and 2010 ACR/EULAR rheumatoid arthritis classification criteria: an analysis based on ESPOIR cohort data. *Ann Rheum Dis* 2012;71:386-9.
11. Jung SJ, Kang Y, Ha YJ, Lee KH, Lee SW, Lee SK, et al. Application of the 2010 ACR/EULAR classification criteria for rheumatoid arthritis in Korean patients with undifferentiated arthritis. *Scand J Rheumatol* 2012;41:192-5.
12. Kaneko Y, Kuwana M, Kameda H, Takeuchi T. Sensitivity and specificity of 2010 rheumatoid arthritis classification criteria. *Rheumatology (Oxford)* 2011;50:1268-74.
13. Kennish L, Labitigan M, Budoff S, Filopoulos MT, McCracken WA, Swearingen CJ, et al. Utility of the new rheumatoid arthritis 2010 ACR/EULAR classification criteria in routine clinical care. *BMJ Open* 2012;2. pii: e001117.
14. Raja R, Chapman PT, O'Donnell JL, Ipenburg J, Frampton C, Hurst M, 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.
15. Reneses S, Pestana L, Garcia A. Comparison of the 1987 ACR criteria and the 2010 ACR/EULAR criteria in an inception cohort of patients with recent-onset inflammatory polyarthritis. *Clin Exp Rheumatol* 2012;30:417-20.
16. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009;6:e1000097.
17. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *Fam Pract* 2004;21:4-10.
18. Sakellariou G, Scirè CA, Zambon A, Caporali R, Montecucco C. Performance of the 2010 classification criteria for rheumatoid arthritis: a systematic literature review and a meta-analysis. *PLoS One* 2013;8:e56528.