

Comparison of the 1987 American College of Rheumatology and the 2010 American College of Rheumatology/European League against Rheumatism criteria for classification of rheumatoid arthritis in the Nurses' Health Study cohorts

Shanthini Kasturi · Barbara L. Goldstein ·
Susan Malspeis · Elizabeth W. Karlson ·
Karen H. Costenbader

Received: 27 March 2013 / Accepted: 6 September 2013 / Published online: 24 September 2013
© Springer-Verlag Berlin Heidelberg 2013

Abstract Performance of rheumatoid arthritis (RA) classification by the 2010 American College of Rheumatology (ACR)/European League against Rheumatism (EULAR) criteria, compared to the 1987 ACR criteria, has not been assessed in population-based cohorts in which disease identification is by mailed questionnaire. Women followed in the Nurses' Health Study and Nurses' Health Study II cohorts self-reported new doctor-diagnosed RA on biennial questionnaires. Two RA experts reviewed medical records of 128 new RA self-reports to obtain individual 1987 and 2010 criteria and arrived at a consensus opinion. We compared agreement in classification by the two criteria sets (κ), and calculated sensitivity and specificity, with reviewers' opinion as gold standard. Ninety-eight (77 %) participants were classified as RA by reviewers' consensus opinion; 98 (77 %) fulfilled 1987 criteria, while 79 (63 %) fulfilled 2010 criteria. Seventy-two (56 %) were classified as RA by both sets, 21 (16 %) by neither, 26 (20 %) by only 1987 criteria, and 9 (7 %) by only 2010 criteria. κ for concordance was 0.36 (95 % CI 0.20–0.53). Compared to reviewer's opinion, sensitivity and specificity were 0.93 and 0.77 for 1987 criteria, and 0.79 and 0.87 for 2010 criteria. Participants fulfilling 1987 criteria only were more likely to be seronegative. In these

prospective population-based cohorts, significant discordance between 1987 ACR and 2010 ACR/EULAR criteria for classifying RA exists. Using the 2010 ACR/EULAR criteria alone had decreased sensitivity, and seronegative RA cases would be excluded in particular. Combined use of both will be necessary to maximize inclusion and allow sensitivity analyses.

Keywords Rheumatoid arthritis · ACR RA criteria · ACR/EULAR RA criteria · Epidemiology

Introduction

In the past decade, diagnosis and treatment for rheumatoid arthritis (RA) have advanced greatly. Development and marketing of the anti-cyclic citrullinated peptide antibody (anti-CCP) assay, with high specificity for RA, has allowed earlier and more specific identification of new onset RA. Additionally, the institution of early therapy has resulted in improved outcomes. The 2010 American College of Rheumatology (ACR)/European League Against Rheumatism (EULAR) criteria for RA classification were developed to facilitate the identification and study of RA at earlier stages than the existing 1987 ACR criteria. The 2010 criteria incorporated anti-CCP assay results and measures of systemic inflammation, and did not include criteria such as nodules and radiographic erosions, which are seen primarily in individuals with long-standing RA. As the 2010 criteria become the new standard for classification of RA, ascertainment of the degree of overlap between the criteria is important in order to determine the extent to which previous research can be applied in newly classified patients.

The Nurses' Health Study (NHS) and Nurses' Health Study II (NHSII) are large prospective cohorts of

S. Kasturi
Department of Medicine, Beth Israel Deaconess Medical Center,
Harvard Medical School, Boston, MA, USA

B. L. Goldstein · S. Malspeis · E. W. Karlson ·
K. H. Costenbader (✉)
Division of Rheumatology, Immunology and Allergy,
Department of Medicine Brigham and Women's Hospital,
Harvard Medical School, 75 Francis St., Boston,
MA 02115, USA
e-mail: KCostenbader@partners.org

over > 230,000 women living across the US, who have been followed many years with data concerning lifestyle factors and development of disease collected on biennial questionnaires. Data from these cohorts have been used to examine relationships between multiple environmental, hormonal and lifestyle factors, as well as genetics, and the development of RA [3–5]. The correct identification and classification of incident RA cases in the NHS cohorts is thus essential to these ongoing studies, as it is to other large population-based cohorts being followed for incident RA [6–7]. To date, this has been done in a standardized two-step procedure, described below, based on the 1987 ACR criteria and reviewers' expert opinion. Approximately 40 % of past RA cases included in NHS cohort analyses have been seronegative. While the 1987 and the new 2010 criteria for the classification of RA have been compared in several different clinical early arthritis populations [8–13], it is not clear how newly reported RA cases in a population-based cohort study such as this should be classified, in particular as this is done on the basis of medical records received up to 2 years from the time of the initial report. We aimed to compare the 1987 ACR and the 2010 ACR/EULAR classification criteria in the NHS/NHSII cohorts, and to examine their performance characteristics and the characteristics of participants classified as having RA by the two different systems, in order to determine how to classify new cases in the future.

Materials and methods

The NHS is a prospective cohort study involving 121,700 female nurses in the USA, aged 30–55 years in 1976, followed since that time. The NHSII was established in 1989 and includes 116,608 female nurses aged 25–42 at cohort inception and followed since then. All participants completed a baseline questionnaire about their medical histories and lifestyles, and have been followed with biennial questionnaires to update exposures and new disease diagnoses [2]. The RA confirmation procedure is a two-stage process, in which all women newly reporting a doctor diagnosis of RA on a biennial questionnaire are asked to complete the Connective Tissue Screening Questionnaire (CSQ) [14] and to sign a medical records release form. For all women with signs and symptoms of RA on the CSQ and available medical records, these records are then independently reviewed by two board-certified rheumatologists (KC, EK) for both the 1987 ACR classification criteria for RA [15], the treating physician's ultimate diagnosis, the reviewer's ultimate diagnosis, and whether the treating physician was an ACR member rheumatologist or not.

In 2010, we began using the 2010 ACR/EULAR classification criteria for RA [1] as well. For the current study,

we reviewed the medical records of women in NHS and NHSII who newly self-reported doctor-diagnosed RA in 2009–2012. The criteria sets were only applied when no other diagnosis (e.g., gout, systemic lupus erythematosus, psoriatic arthritis) better explained the signs or symptoms. Results of testing for rheumatoid factor (RF), anti-CCP, C-reactive protein (CRP), erythrocyte sedimentation rate (ESR), and other features of RA were based on medical record review. We reviewed reports of radiographs in the medical record, both as documented by the physician and the reports themselves. Cases with ≥ 4 of 7 of the 1987 ACR criteria documented in the medical record were considered to have definite RA by 1987 criteria, and cases with a score of ≥ 6 of the 10 from the 2010 ACR/EULAR criteria were considered to have definite RA by 2010 criteria. Ultimate agreement between the two reviewers as to the diagnosis of RA has been by consensus. Demographic characteristics of the women at cohort baseline and at time of self-report of RA were obtained from NHS to NHSII cohort datasets. Using the NHS expert rheumatologists' opinion as the gold standard, we compared the sensitivity and the specificity, positive and negative predictive values of classification of incident RA by either or both classification systems.

Statistical analysis

We calculated the sensitivity and specificity, using the reviewers' opinion as the gold standard for the diagnosis of RA. Agreement between the two criteria sets was based on the κ coefficient. All statistical analyses involved use of SAS, version 9.2 (Cary, NC). The Partners' Institutional Review Board approved this study.

Results

Baseline characteristics of the women self-reporting new RA in the NHS and NHSII cohorts are described in Table 1. Participants in NHSII are younger, and eighty percentage of participants were seen by an ACR member rheumatologist. The characteristics of the women classified by one or the other criteria system differed from those meeting both sets of criteria (Table 2). Ninety-eight (77 %) of the 128 participants fulfilled the 1987 ACR criteria (69 % of NHS and 80 % of NHSII records reviewed), while 79 (63 %) fulfilled the 2010 ACR/EULAR criteria (59 % of NHS and 65 % of NHSII records reviewed). Seventy-two (56 %) met both sets of criteria, 21 (16 %) met neither, 26 (20 %) met only the 1987 ACR criteria, and 9 (7 %) met only the 2010 ACR/EULAR criteria. Discordance between the classification criteria thus occurred in 35 cases (27 %). Concordance between the two sets per the

Table 1 Characteristics of participants in the NHS/NHSII cohorts newly self-reporting doctor-diagnosed RA in 2009–2012

	NHS (<i>n</i> = 39)	NHSII (<i>n</i> = 89)
Mean age (SD)	71.3 (6.9)	54.2 (4.7)
White race (%)	39 (100)	84 (94)
Parous (%)	39 (100)	76 (85)
Postmenopausal (%)	24 (62)	23 (26)
Mean body mass index (SD)	28.1 (5.9)	28.8 (6.2)
Ever smoker (%)	26 (67)	36 (40)
Mean alcohol Intake, grams/day (SD)	4.4 (5.0)	3.8 (5.2)
Seen by an ACR member rheumatologist (%)	27 (69)	75 (84)

ACR American College of Rheumatology

kappa statistic was 0.36 (95 % CI 0.20–0.53). Notably, participants who met the 1987 ACR criteria only were more likely to have involvement of 1–10 small joints (62 vs. 11 %, $p < 0.018$), and to have a negative RF and negative anti-CCP (69 vs. 11 %, $p < 0.005$), than those meeting only the 2010 ACR/EULAR criteria. Women who only met the 2010 criteria were less likely to have symmetric arthritis (44 vs. 96 %, $p < 0.002$).

For the 128 women self-reporting new RA, there was report of a radiograph in the medical record for 95 (74 %). Of those who had a radiograph, 31 had erosions. Of the 56 women who were seronegative for RF or anti-CCP, 11 had radiographic erosions. All 11 fulfilled the 1987 ACR criteria for RA, only three fulfilled the 2010 ACR criteria for RA, and ultimately nine were thought to have RA. On the other hand, of the 72 women who were seropositive, 20 had radiographic erosions. All 20 fulfilled the 1987 ACR criteria, 19 of 20 fulfilled the 2010 criteria, and all 20 were ultimately thought to have RA (Chi squared test for reviewer's impression of RA among seropositive vs. seronegative women with erosions, $p = 0.049$). Of the 31 patients with erosions, 100 % were thought to have RA by their treating rheumatologists.

Using the NHS rheumatologist medical record reviewers' opinion as the gold standard for RA classification, sensitivity was 0.93 for the 1987 ACR criteria compared with 0.79 for the 2010 ACR/EULAR criteria (Table 3). Specificity of the 2010 ACR/EULAR criteria was 0.87, greater than the 0.77 specificity of the 1987 ACR criteria. The positive and negative predictive values when compared to the reviewers' opinion were highest when either criteria set was fulfilled. NHS rheumatologist reviewers ultimately agreed with the non-ACR rheumatologists'

Table 2 RA classification by individual criteria among participants in the NHS/NHSII cohorts newly self-reporting doctor-diagnosed RA in 2009–2012

Characteristic, <i>n</i> (%)	Reviewers' opinion (<i>n</i> = 98)	Both criteria sets (<i>n</i> = 72)	Only ACR 1987 criteria (<i>n</i> = 26)	Only ACR/EULAR 2010 criteria (<i>n</i> = 9)	Neither criteria (<i>n</i> = 21)
1987 ACR criteria					
Morning stiffness >1 h	72 (73.5)	60 (83.3)	17 (65.4)	3 (33.3)	3 (14.3)
Arthritis ≥ 3 joint areas	95 (96.9)	71 (98.6)	26 (100)	7 (77.8)	5 (23.8)
Hand arthritis	94 (95.9)	71 (98.6)	26 (100)	6 (66.7)	9 (42.9)
Symmetric arthritis	93 (94.9)	70 (97.2)	25 (96.2)	4 (44.4)	2 (9.5)
Nodules	6 (6.1)	4 (5.6)	2 (7.7)	0 (0)	0 (0)
RF positive	64 (65.3)	56 (77.8)	4 (15.4)	6 (66.7)	1 (4.8)
Radiographic erosions	29 (29.6)	22 (30.6)	9 (34.6)	0 (0)	0 (0)
2010 ACR/EULAR criteria					
1 Large joint	2 (2)	0 (0)	2 (7.7)	0 (0)	3 (14.3)
2–10 Large joints	2 (2)	0 (0)	0 (0)	2 (22.2)	3 (14.3)
1 Small joint	14 (14.3)	8 (11.1)	6 (23.1)	4 (44.4)	7 (33.3)
2–10 Small joints	43 (43.9)	29 (40.3)	16 (61.5)	1 (11.1)	5 (23.8)
>10 Large joints	35 (35.7)	35 (48.6)	1 (3.8)	2 (22.2)	0 (0)
Negative RF and anti-CCP	26 (26.5)	13 (18.1)	18 (69.2)	1 (11.1)	15 (71.4)
Low-positive RF or anti-CCP	28 (28.6)	25 (34.7)	4 (15.4)	2 (22.2)	3 (14.3)
High-positive RF or anti-CCP	38 (38.8)	32 (44.4)	0 (0)	6 (66.7)	1 (4.8)
Abnormal CRP or ESR	63 (64.3)	45 (62.5)	14 (53.8)	7 (77.8)	9 (42.9)
Normal CRP and ESR	23 (23.5)	19 (26.4)	6 (23.1)	1 (11.1)	11 (52.4)
> 6 weeks of symptoms	96 (98)	72 (100)	24 (92.3)	9 (100)	18 (85.7)

RF rheumatoid factor, Anti-CCP anti-cyclic citrullinated peptide antibody, CRP C-reactive protein, ESR erythrocyte sedimentation rate

Table 3 Sensitivity and specificity of two criteria systems for RA classification compared to the gold standard of the reviewers' opinion among participants in the NHS/NHSII cohorts newly self-reporting doctor-diagnosed RA in 2009–2012

	1987 ACR criteria	2010 ACR/EULAR criteria	Either 1987 or 2010	Both 1987 and 2010
Sensitivity	0.93	0.79	0.99	0.72
Specificity	0.77	0.87	0.67	0.97
Positive predictive value	0.93	0.95	0.91	0.99
Negative predictive value	0.77	0.55	0.95	0.52

diagnoses slightly more than the ACR rheumatologists' diagnoses (kappa coefficient 0.70 vs. 0.54).

Discussion

The 2010 ACR/EULAR criteria were developed with the intent of identifying early RA in particular. While past studies have examined the performance of these new criteria in the classification of early arthritis patients in several clinical settings [8–13], no studies have examined their performance in the setting of a population-based cohort study in which the identification and classification of RA is done on the basis of mailed questionnaires and medical record review, with an unavoidably longer time course than that of a clinical interaction. As we were not certain how the new criteria would compare to the older criteria and whether in transitioning entirely to the new criteria, we would misclassify or exclude new cases, or include a different type of RA, we performed a side-by-side review of the medical records of women newly self-reporting RA, for the 1987 and the 2010 criteria, as well as an expert reviewer opinion, which has always been our gold standard. We have found that agreement between the two criteria sets was not high (kappa 0.36), and a substantial proportion of seronegative cases, in particular, would not be captured by use of the 2010 criteria alone. The best positive and negative predictive values are found when both criteria sets are used.

Several past clinical studies have also documented high levels of disagreement in RA classification by the two sets of criteria, from 21 to 34 % [8–11], and we found 27 % disagreement. As has been reported in clinical cohorts [10–12], the cases that would not be captured in the NHS cohorts with use of the new classification criteria alone were more often seronegative. Use of the 2010 criteria exclusively would thus primarily miss seronegative RA, which may have otherwise been identified with the 1987

criteria. All prior studies in the NHS and NHSII cohorts have included both seropositive and seronegative populations and allowed for stratified analyses of predictors of different phenotypes of RA. In adding cases identified by the 2010 criteria to those classified by the 1987 criteria, we will essentially continue to identify the same RA cases comparable to those identified since cohort inception, while increasing sensitivity for early RA and anti-CCP positive cases. Our gold standard of the expert rheumatologist reviewers' opinion is slightly problematic in that our opinions are influenced by our knowledge of the criteria. The reviewers were also limited by the need to rely on medical records, which at times were incomplete and did not offer insight into the treating physicians' thought processes. However, in the absence of a true gold standard, rheumatologist diagnosis has been and continues to be the accepted outcome against which both old and new criteria have been compared [8, 10, 13]. Thus, these data support use of both the 1987 and 2010 criteria for identification of RA study populations in the NHS and NHSII cohorts, and other studies following population-based cohorts over many years may choose to do the same.

Acknowledgments This work was supported by the National Institutes of Health [Grants AR59073 (Costenbader), CA87969, CA49449, CA50385, CA67262, AR49880 (Karlson), AR52403 (Karlson)].

Conflict of interest None of the authors have any competing interests to declare.

References

1. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO et al (2010) 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Arthritis Rheum* 62: 2569–2581
2. Colditz GA, Manson JE, Hankinson SE (1997) The Nurses' Health Study: 20-year contribution to the understanding of health among women. *J Womens Health* 6:49–62
3. Karlson EW, Chibnik LB, Tworoger SS, Lee IM, Buring JE, Shadick NA et al (2009) Biomarkers of inflammation and development of rheumatoid arthritis in women from two prospective cohort studies. *Arthritis Rheum* 60:641–652
4. Karlson EW, Chang SC, Cui J, Chibnik LB, Fraser PA, De Vivo I et al (2010) Gene-environment interaction between HLA-DRB1 shared epitope and heavy cigarette smoking in predicting incident rheumatoid arthritis. *Ann Rheum Dis* 69:54–60
5. Costenbader KH, Chang SC, Laden F, Puett R, Karlson EW (2008) Geographic variation in rheumatoid arthritis incidence among women in the United States. *Arch Intern Med* 168: 1664–1670
6. Mikuls TR, Saag KG, Criswell LA, Merlino LA, Kaslow RA, Shelton BJ et al (2002) Mortality risk associated with rheumatoid arthritis in a prospective cohort of older women: results from the Iowa Women's Health Study. *Ann Rheum Dis* 11:994–999

7. Peschken CA, Robinson DB, Hitchon CA, Smolik I, Hart D, Bernstein CN, El-Gabalawy HS (2012) Pregnancy and the risk of rheumatoid arthritis in a highly predisposed North American Native population. *J Rheumatol* 12:2253–2260
8. Britsemmer K, Ursum J, Gerritsen M, van Tuyl L, van Schaar-denbarg D (2011) Validation of the 2010 ACR/EULAR classification criteria for rheumatoid arthritis: slight improvement over the 1987 ACR criteria. *Ann Rheum Dis* 70:1468–1470
9. Bykerk VP, Jamal S, Boire G, Hitchon CA, Haraoui B, Pope JE et al (2012) The Canadian Early Arthritis Cohort (CATCH): patients with new-onset synovitis meeting the 2010 ACR/EULAR classification criteria but not the 1987 ACR classification criteria present with less severe disease activity. *J Rheumatol* 39:3071–3080
10. Cader MZ, Filer A, Hazlehurst J, de Pablo P, Buckley CD, Raza K (2011) Performance of the 2010 ACR/EULAR criteria for rheumatoid arthritis: comparison with 1987 ACR criteria in a very early synovitis cohort. *Ann Rheum Dis* 70:949–955
11. Fautrel B, Combe B, Rincheval N, Dougados M (2012) Level of agreement of the 1987 and 2010 ACR/EULAR rheumatoid arthritis classification criteria: an analysis based on ESPOIR cohort data. *Ann Rheum Dis* 71:386–389
12. Kaneko Y, Kuwana M, Kameda H, Takeuchi T (2011) Sensitivity and specificity of 2010 rheumatoid arthritis classification criteria. *Rheumatology* 50:1268–1274
13. van der Linden MP, Knevel R, Huizinga TW, van der Helm-van Mil AH (2011) 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. *Arthritis Rheum* 63:37–42
14. Karlson EW, Sanchez-Guerrero J, Wright EA, Lew RA, Daltroy HL, Katz JN et al (1995) A connective tissue disease screening questionnaire for population studies. *Ann Epidemiol* 5:297–302
15. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS et al (1988) The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 31:315–324