Reprinted from
ARTHRITIS AND RHEUMATISM
Vol. 31, No. 3, March 1988

Printed in U.S.A.



# THE AMERICAN RHEUMATISM ASSOCIATION 1987 REVISED CRITERIA FOR THE CLASSIFICATION OF RHEUMATOID ARTHRITIS

FRANK C. ARNETT, STEVEN M. EDWORTHY, DANIEL A. BLOCH, DENNIS J. McSHANE, JAMES F. FRIES, NORMAN S. COOPER, LOUIS A. HEALEY, STEPHEN R. KAPLAN, MATTHEW H. LIANG, HARVINDER S. LUTHRA, THOMAS A. MEDSGER, JR., DONALD M. MITCHELL, DAVID H. NEUSTADT, ROBERT S. PINALS, JANE G. SCHALLER, JOHN T. SHARP, RONALD L. WILDER, and GENE G. HUNDER

**The revised criteria for the classification of rheumatoid arthritis (RA) were formulated from a computerized analysis of 262 contemporary, consecutively studied patients with RA and 262 control subjects with rheumatic diseases other than RA (non-RA). The new criteria are as follows: 1) morning stiffness in and around joints lasting at least 1 hour before maximal improvement; 2) soft tissue swelling (arthritis) of 3 or more joint areas observed by a physician; 3) swelling (arthritis) of the proximal interphalangeal, metacarpophalangeal, or wrist joints; 4) symmetric swelling (arthritis); 5) rheumatoid nodules; 6) the presence of rheumatoid factor; and 7) radiographic erosions and/or periarticular osteopenia in hand and/or wrist joints. Criteria 1 through 4 must have been present for at least 6 weeks. Rheumatoid arthritis is defined by the presence of 4 or more criteria, and no further qualifications (classic, definite, or probable) or list of exclusions are required. In addition, a "classification tree" schema is presented which performs equally as well as the traditional (4 of 7) format. The new criteria demonstrated 91–94% sensitivity and 89% specificity for RA when compared with non-RA rheumatic disease control subjects.**

In 1956, a committee of the American Rheumatism Association (ARA) proposed diagnostic criteria for rheumatoid arthritis (1). The criteria were formulated from the experiences of the 5 committee members, review of a recent epidemiologic survey (2,3), and analysis of 332 cases provided by interested physicians in 19 cities in the United States and Canada. Eleven criteria with 19 exclusions were proposed. "Definite" rheumatoid arthritis (RA) required at least 5 criteria and 6 weeks of joint symptoms. " Probable" RA required at least 3 criteria and at least 4 weeks of joint symptoms. A "possible" RA category with another set of formulae was also proposed (1).

In an attempt to improve specificity and simplicity, the same committee revised the ARA criteria

From the Rheumatoid Arthritis Criteria Subcommittee of the Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association.

Supported in part by NIH grant AM-21393 to the American Rheumatism Association Medical Information System.

Frank C. Arnett, MD: Professor of Internal Medicine, The University of Texas Health Science Center at Houston; Steven M. Edworthy, MD: Fellow of Immunology, Stanford University School of Medicine, Stanford, CA; Daniel A. Bloch, PhD: Senior Research Associate of Biostatistics in Immunology, Stanford University School of Medicine; Dennis J. McShane, MD: Assistant Professor of Immunology, Stanford University School of Medicine; James F. Fries, MD: Associate Professor of Medicine, Stanford University School of Medicine; Norman S. Cooper, MD: Professor of Pathology, New York University School of Medicine, New York, NY; Louis A. Healey, MD: Virginia Mason Clinic, Seattle, WA; Stephen R. Kaplan, MD: Professor of Medicine, Brown University Program in Medicine, Providence, RI; Matthew H. Liang, MD, MPH: Associate Professor of Medicine, Harvard Medical School, Boston, MA; Harvinder S. Luthra, MD: Associate Professor of Medicine, Mayo Medical School and Clinic, Rochester, MN; Thomas A. Medsger, Jr., MD: Professor of Medicine, University of Pittsburgh, Pittsburgh, PA; Donald M. Mitchell, MD: Professor of Medicine, University of Saskatchewan, Saskatoon, Saskatchewan, Canada; David H. Neustadt, MD: Clinical Professor of Medicine, University of Louisville School of Medicine, Louisville, KY; Robert S. Pinals, MD: Professor of Medicine, UMDNJ–Robert Wood Johnson Medical School, Newark, NJ, and The Medical Center at Princeton, Princeton, NJ; Jane G. Schaller, MD: Professor of Pediatrics, Tufts–New England Medical Center, Boston, MA; John T. Sharp, MD: Staff Physician, Tifton Medical Clinic, Tifton, GA; Ronald L. Wilder, MD, PhD: Senior Investigator, National Institute of Arthritis and Musculoskeletal and Skin Diseases; Gene G. Hunder, MD: Professor of Medicine, Mayo Medical School and Clinic.

Address reprint requests to the American Rheumatism Association, 17 Executive Park Drive, NE, Suite 480, Atlanta, GA 30329.

Submitted for publication September 16, 1987; accepted in revised form November 16, 1987.

in 1958 (4). The category "classic" rheumatoid arthritis was added to describe patients who fulfilled 7 of the original 11 criteria. In addition, the duration of joint symptoms in "probable" disease was increased from 4 to 6 weeks, and another exclusion was added. At the Third International Symposium on Population Studies of the Rheumatic Diseases in 1966, further recommendations were made (New York criteria) (5); however, these criteria, although more specific and detailed, were cumbersome and never gained wide application.

Thus, the ARA revised criteria for rheumatoid arthritis (4) have been extensively used for nearly 30 years. They have assisted in the development of a more uniform vocabulary, improved communication, allowed more effective teaching, and enabled research results from different locations to be more directly comparable. Analysis of their function, while pointing out particular problems, has generally been supportive (6). On the other hand, over the same period, clinical knowledge has expanded remarkably. Many other forms of arthritis previously misdiagnosed or included within the spectrum of RA have now been separately classified. Numerous examples exist and include the HLA–B27–associated spondylarthropathies, the pseudorheumatoid form of calcium pyrophosphate dihydrate deposition disease, polymyalgia rheumatica, Lyme disease, and others (7,8).

Several general clinical observations relevant to the criteria have also been made. In practice, the distinction between "definite" and "classic" RA has not proven useful, and the 2 terms are usually merged into a single cumbersome phrase "definite or classic" RA. There is general agreement that many patients previously classified as having "probable" RA have a different disease. Rheumatoid factor (RF) has assumed a pivotal role in subcategorizing patients into "seropositive" and "seronegative" groups, despite incomplete knowledge of the role of RF in disease pathogenesis, its frequent absence in early stages of disease, its suppression by disease-modifying drugs, and its occurrence in hidden form. In addition, other criticism emphasizes the observation that 3 of the criteria require invasive procedures (mucin clot, nodule biopsy, and synovial biopsy) that are very rarely performed; the criteria are too sensitive, yet not specific enough, for population studies; and the "exclusions" are unwieldy, impractical, and somewhat circular (6,8).

Accordingly, in 1983, the ARA appointed a subcommittee of the Diagnostic and Therapeutic Criteria Committee to review and, if necessary, revise the criteria for RA. After a critical assessment of current knowledge, the subcommittee formulated 2 hypotheses:

1. That new sets of criteria could be constructed, possibly with fewer criteria, which would be more sensitive and, in particular, more specific than the old ones, and
2. That there was an unacceptably broad spectrum of disease identified by the old criteria, and thus, a stricter definition of RA was required.

The subcommittee addressed these issues by study and computer analysis of contemporary adult patients with RA and control subjects.

## PATIENTS AND METHODS

**RA patients.** Both patients with RA and control subjects with rheumatic diseases other than RA (non-RA) were prospectively and consecutively studied. Only adults with disease onset after the age of 16 were included. Study patients were submitted by 9 subcommittee members (158 patients) and by 32 other rheumatologists recruited from both university and private practice settings (366 patients). Each physician enrolled 20 consecutive patients in the study, 10 RA patients and 10 non-RA control subjects. A record of all patients seen over the 6-month recruitment period was maintained to ensure a complete sample and to confirm that patients were, indeed, entered consecutively. To broaden the spectrum of diseases studied, collaborators were chosen to represent a variety of practice settings.

Patients had a clinical diagnosis of RA without regard to the presence or absence of specific criteria. Five RA subjects and 5 control subjects from each investigator were "established" patients, and 5 patients and 5 control subjects were "new" to that practice. It should be noted that new cases did not necessarily indicate early-onset disease.

**Control subjects.** The next consecutive patient (after an RA patient in a practice) with a rheumatic disease other than RA, as defined in the *Primer on the Rheumatic Diseases* (7), was designated a control subject. In this manner, a control group was assembled. Excluded from study were patients who had localized periarticular diseases such as tendinitis, low back pain, or painful shoulder. Patients with generalized conditions, such as polymyalgia rheumatica and fibromyalgia, were to be included.

**Certainty estimation.** All subjects, both patients and controls, had the degree of "certainty" of their diagnosis estimated by the investigator. A 10-cm analog scale was presented, with one end of the scale representing absolute certainty that the patient did *not* have rheumatoid arthritis and the other end absolute certainty that he or she *did*. It was anticipated that RA patients would cluster at one end of the scale and non-RA control subjects at the other. RA and non-RA patients with similar clinical manifestations might be identified somewhere in the middle of the scale. In analyzing potential criteria sets, it was determined that misclassification of a patient with clinically ambiguous disease would be

Table 1. Demographic characteristics of patients with rheumatoid arthritis (RA) and control subjects with rheumatic diseases other than RA

| | n | Age (mean ± SD) | Disease duration (mean ± SD) | % female | Race (%) | | |
|---|---|---|---|---|---|---|---|
| | | | | | White | Black | Other |
| Rheumatoid arthritis | 262 | 51.2 ± 17.2 | 7.7 ± 8.6 | 73 | 85 | 8 | 7 |
| Control subjects | | | | | | | |
| Osteoarthritis | 83 | 63.9 ± 10.7 | 9.2 ± 7.9 | 82 | 88 | 10 | 2 |
| Systemic lupus erythematosus | 53 | 39.0 ± 13.2 | 7.5 ± 7.7 | 94 | 62 | 30 | 8 |
| Psoriatic arthritis | 10 | 46.4 ± 18.7 | 18.4 ± 19.5 | 60 | 100 | 0 | 0 |
| Other | 116 | 47.4 ± 17.9 | 5.9 ± 7.7 | 60 | 90 | 4 | 6 |
| Total | 262 | 53.2 ± 19.0 | 7.7 ± 8.8 | 74 | 84 | 11 | 5 |

less damaging than misclassification of a patient with more classic disease.

**Data collection.** The data collection form included the individual old ARA (4) and New York (5) criteria for RA, as well as items considered by subcommittee members to be potential contributions to the new criteria (Delphi method).

Certain previous criteria were "dissected" into components, with the hope of establishing more discriminating definitions. For example, questions were raised concerning how many weeks of morning stiffness should be required as a minimum, whether the stiffness was in muscles, joints, or elsewhere, how long it should last each day, and whether the daily duration was better stated as "until improvement began" or "until maximum improvement." Each of these issues was addressed on the data form. Collaborators were required to complete all questions for each patient, with a notation of "not known" if the information was not available. All forms were carefully reviewed by 3 subcommittee members, and questions or ambiguities were referred to the original physicians for clarification. Data were double entered into the computer, and all errors were corrected.

**Data analysis.** Univariate comparisons of the accuracy of individual potential criteria were made using chi-square tests for 2 × 2 tables. Subanalyses compared new patients with new controls, established patients with established controls, and the combined groups with each other. Data were also examined by investigator and by type of practice (private, institutional). Since no significant differences were found in any of these subcomparisons, all patients and all controls were combined for further analyses.

Two different statistical methods were used to develop classification criteria. The first method included procedures that have been previously used in criteria studies (9). In this method, combinations of the variables which were most sensitive and specific to the classification of RA were selected by means of Boolean algebra, using union and intersection operations. The resulting rule of classification was in the form "if, for a given subject, at least x out of a list of y characteristics are present, then classify the subject as having RA." A second method was also used to classify subjects: the technique of "classification trees" or "recursive partitioning" as defined by Breiman et al (10). A simplified overview of this technique, as well as a discussion of the relative merits of the 2 classification methods, is given by Altman et al (9). Briefly, in this method of classification, the first variable selected was the one that most effectively divided the subject population into those with RA and those without RA, after testing all candidate variables at all variate values represented in the sample of 524 RA and non-RA patients. The procedure was repeated for the 2 resulting subgroups and then again on the subgroups resulting from this second split, and so on. The most appropriate tree was determined by cross-validation.

Finally, the specificity of each of the 2 classification methods was tested against 137 consecutive subjects, age range 52–74 years, who were enrolled in a prospective study of musculoskeletal aging at the Stanford University Multipurpose Arthritis Center (11).

## RESULTS

Table 1 presents the demographic characteristics of the 262 RA patients and 262 controls. As expected, the lupus patients were younger and the osteoarthritis patients were older than the RA patients. Nevertheless, considering all controls, the mean ages of RA and non-RA patients were similar, as was the disease duration. Sex and racial percentages were virtually identical.

The diagnoses of the control patients (Table 2) represented a cross-section of rheumatic diseases.

Table 2. Diagnoses of control subjects with other rheumatic diseases

| Disease category | No. of patients | % of patients |
|---|---|---|
| Osteoarthritis | 83 | 32.1 |
| Systemic lupus erythematosus | 53 | 20.2 |
| Fibromyalgia | 24 | 9.2 |
| Ankylosing spondylitis | 14 | 5.4 |
| Psoriasis | 10 | 3.8 |
| Reiter's syndrome | 7 | 2.7 |
| Gout | 8 | 2.8 |
| Arteritis (general) | 10 | 3.8 |
| Polymyalgia rheumatica | 6 | 2.3 |
| Temporal arteritis | 3 | 1.1 |
| Scleroderma | 6 | 2.3 |
| Mixed connective tissue disease | 4 | 1.4 |
| Other | 34 | 13.0 |
| Total | 262 | 100 |

Osteoarthritis was the most frequent diagnosis in controls and systemic lupus erythematosus was the second most frequent. The seronegative spondylarthropathies and other connective tissue diseases were less well represented.

Table 3 lists some of the items which the subcommittee selected as being potentially important disease discriminators. Accuracy is the mean of sensitivity and specificity values; this number serves as a rough gauge of the relative effectiveness of the potential criteria. Pain on motion in the distal interphalangeal (DIP) and hip joints were highly inaccurate, with scores of 50.3 and 51.7, respectively. All other variables included in Table 3 had an accuracy significantly higher than 50% at $P = 0.0001$, except for synovial biopsy ($P = 0.02$). RF had not been determined in 12 RA patients and 55 controls. Nodules were biopsied in 44 RA patients; however, no control sub-

**Table 3.** Comparison of sensitivity, specificity, and accuracy of criteria for rheumatoid arthritis

| Criterion* | No. of patients | Sensitivity (%) | Specificity (%) | Accuracy |
|---|---|---|---|---|
| Historical information | | | | |
| Morning stiffness | 522 | 91.2 | 40.4 | 65.8 |
| Morning stiffness (>1 hour)† | 509 | 81.1 | 57.3 | 69.2 |
| ARA stiffness criterion | 521 | 89.7 | 46.3 | 68.0 |
| Pain on motion in | | | | |
| DIPs | 523 | 26.1 | 74.4 | 50.3 |
| PIPs | 524 | 86.6 | 60.3 | 73.5 |
| MCPs | 524 | 88.2 | 67.5 | 77.9 |
| Wrists | 524 | 90.8 | 64.9 | 77.9 |
| Elbows | 524 | 64.9 | 73.7 | 69.3 |
| Shoulders | 523 | 77.1 | 55.9 | 66.5 |
| Hips | 523 | 36.3 | 67.0 | 51.7 |
| Knees | 523 | 80.9 | 36.8 | 58.9 |
| Ankles | 524 | 68.7 | 69.1 | 68.9 |
| MTPs | 524 | 74.4 | 71.4 | 72.9 |
| Physical examination | | | | |
| Swelling (fluid/synovium) | | | | |
| DIPs | 523 | 79.4 | 83.5 | 81.5 |
| PIPs | 523 | 79.3 | 84.0 | 81.7 |
| MCPs | 523 | 86.6 | 84.0 | 85.3 |
| Wrists | 523 | 80.8 | 86.6 | 83.7 |
| Elbows | 524 | 42.7 | 90.8 | 66.8 |
| Knees | 523 | 64.9 | 71.6 | 68.3 |
| Ankles | 522 | 53.6 | 88.1 | 70.9 |
| ≥3 joint areas†‡ | 507 | 90.7 | 84.0 | 87.4 |
| Swelling (fluid/synovium) | | | | |
| Wrist, MCP, or PIP† | 523 | 96.6 | 74.8 | 85.7 |
| Wrist or MCP‡ | 524 | 95.0 | 77.1 | 86.1 |
| Wrist and MCP‡ | 524 | 72.5 | 93.5 | 83.0 |
| Symmetric swelling (arthritis) | 524 | 92.7 | 79.3 | 86.0 |
| PIPs | 524 | 71.4 | 86.6 | 79.0 |
| MCPs | 524 | 76.3 | 87.4 | 81.9 |
| Wrists | 523 | 69.7 | 90.8 | 80.3 |
| Any region(s)†‡ | 523 | 94.3 | 74.3 | 84.3 |
| Subcutaneous nodules† | 519 | 43.4 | 97.7 | 70.6 |
| Laboratory and radiologic findings | | | | |
| Abnormal serum RF†‡ | 457 | 80.4 | 87.0 | 83.7 |
| Synovial fluid | 123 | 74.4 | 73.3 | 74.0 |
| Synovial biopsy | 53 | 42.4 | 95.0 | 68.7 |
| Nodule biopsy | 44 | 27.3 | – | – |
| Radiographic changes (ARA)†‡ | 410 | 77.2 | 93.7 | 85.5 |
| Radiographic changes (NY) | 400 | 69.4 | 89.7 | 79.5 |
| Erosions of hand region | 353 | 63.5 | 93.8 | 78.7 |

* ARA = American Rheumatism Association; DIPs = distal interphalangeal joints; PIPs = proximal interphalangeal joints; MCPs = metacarpophalangeal joints; MTPs = metatarsophalangeal joints; RF = serum rheumatoid factor; NY = New York criterion.
† Criteria selected for "at least 4 out of 7" criteria set.
‡ Criteria selected by classification tree.

**Table 4.** Comparison of sensitivity, specificity, and accuracy of the 1958 American Rheumatism Association (ARA) and the 1966 New York criteria for rheumatoid arthritis

| Criterion | Defined | Sensitivity (%) | Specificity (%) | Accuracy |
|---|---|---|---|---|
| ARA criteria | | | | |
| Morning stiffness* | 521 | 89.7 | 46.3 | 68.1 |
| Joint pain | 524 | 98.1 | 30.9 | 64.5 |
| One joints swollen | 509 | 97.3 | 63.9 | 80.7 |
| Two joints swollen | 509 | 93.8 | 77.8 | 85.9 |
| Symmetric swelling* | 523 | 92.7 | 79.3 | 86.0 |
| Rheumatoid nodules* | 519 | 43.4 | 97.7 | 70.7 |
| Serum rheumatoid factor* | 457 | 80.4 | 87.0 | 83.4 |
| Mucin clot | 123 | 74.4 | 73.3 | 74.0 |
| Synovial biopsy | 53 | 42.4 | 95.0 | 62.3 |
| Nodule biopsy | 44 | 27.3 | – | – |
| Radiographic findings* | 410 | 77.2 | 93.7 | 84.9 |
| New York criteria | | | | |
| Joint pain | 524 | 99.6 | 23.3 | 68.8 |
| Joint swelling | 524 | 91.2 | 83.6 | 87.4 |
| Serum rheumatoid factor | 457 | 80.4 | 87.0 | 83.4 |
| Radiographic findings | 400 | 69.4 | 89.7 | 78.8 |

* Retained in new criteria essentially in similar form.

jects had biopsies performed. When items had comparable accuracy scores, the subcommittee generally elected the more specific formulation rather than the more sensitive one. Accuracy varied from 50% for pain on motion in DIP joints to 87% for the criterion "3 or more joints swollen for 6 or more weeks." Some criteria techniques, synovial fluid examination, synovial biopsy, and nodule biopsy, were seldom performed in patients and almost never performed in control subjects; therefore, accuracy could not be reliably determined. Because of the method of patient selection, the infrequency of these observations must be considered reasonably typical of contemporary practice.

Table 4 presents the comparison of the earlier ARA (4) and New York (5) criteria. Five of the existing ARA criteria were retained in similar form in the revised criteria presented here. The New York criterion of joint swelling showed great accuracy. This definition is in many ways parallel to the new criterion "3 or more swollen joints for 6 or more weeks"; however, the complex formulation of the New York criterion made its clinical application difficult.

Table 5 presents the 1987 revised criteria (with definitions) for classification of RA. As noted, 5 were retained from the 1958 ARA criteria (4). More precise definitions have been provided for several. The criteria "arthritis of 3 or more joint areas" and "arthritis of hand joints" conceptually replace the second, third, and fourth criteria (pain on motion or tenderness in a joint, swelling in at least 1 joint, and swelling of at least one other joint) of the 1958 version.

Figure 1 shows the best of several classification trees derived using the computer program CART (12). Many different combinations of variables were evaluated using tree structures; however, none of the variables performed as well as those shown in Figure 1. In this system, observations concerning arthritis of three or more joint areas, arthritis of hand joints, and symmetric arthritis must be made for the classification procedure to be applied. When data such as hand radiographs and results of serum rheumatoid factor tests are not available, the methodology allows selection of a surrogate variable which most closely divides the study population in a manner that best approxi-

**Table 5.** The 1987 revised criteria for the classification of rheumatoid arthritis (traditional format)*

| Criterion | Definition |
|---|---|
| 1. Morning stiffness | Morning stiffness in and around the joints, lasting at least 1 hour before maximal improvement |
| 2. Arthritis of 3 or more joint areas | At least 3 joint areas simultaneously have had soft tissue swelling or fluid (not bony overgrowth alone) observed by a physician. The 14 possible areas are right or left PIP, MCP, wrist, elbow, knee, ankle, and MTP joints |
| 3. Arthritis of hand joints | At least 1 area swollen (as defined above) in a wrist, MCP, or PIP joint |
| 4. Symmetric arthritis | Simultaneous involvement of the same joint areas (as defined in 2) on both sides of the body (bilateral involvement of PIPs, MCPs, or MTPs is acceptable without absolute symmetry) |
| 5. Rheumatoid nodules | Subcutaneous nodules, over bony prominences, or extensor surfaces, or in juxtaarticular regions, observed by a physician |
| 6. Serum rheumatoid factor | Demonstration of abnormal amounts of serum rheumatoid factor by any method for which the result has been positive in <5% of normal control subjects |
| 7. Radiographic changes | Radiographic changes typical of rheumatoid arthritis on posteroanterior hand and wrist radiographs, which must include erosions or unequivocal bony decalcification localized in or most marked adjacent to the involved joints (osteoarthritis changes alone do not qualify) |

* For classification purposes, a patient shall be said to have rheumatoid arthritis if he/she has satisfied at least 4 of these 7 criteria. Criteria 1 through 4 must have been present for at least 6 weeks. Patients with 2 clinical diagnoses are not excluded. Designation as classic, definite, or probable rheumatoid arthritis is *not* to be made. See Table 3 for definitions of abbreviations.

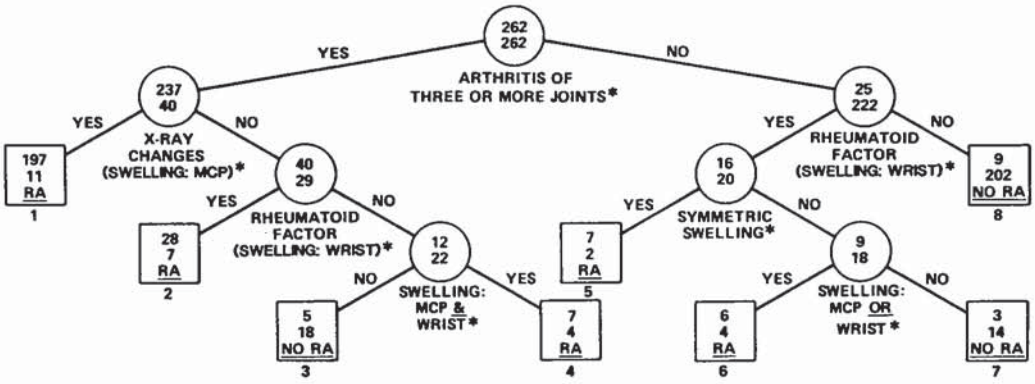


**Figure 1.** Schematic representation of classification tree for rheumatoid arthritis (RA). Circles contain numbers of subjects with RA (upper value) and numbers of controls without RA (lower value). Boxes specify whether subjects can be classified as having RA or as not having RA (NO RA). Parentheses indicate surrogate variables that can be used when another variable (radiograph or rheumatoid factor test result) is unavailable. Numbers under the boxes are the subset numbers (see Table 7 for explanation of subsets). * = a clinical criterion that must have been observed by a physician and present for 6 or more weeks; MCP = metacarpophalangeal joint(s).

mates the original variable. Thus, swelling of the metacarpophalangeal (MCP) joints is the variable that divides the patient populations in a manner similar to the variable of radiographic changes, and thus, may be used as a surrogate when radiographs are not available. Similarly, swelling of a wrist is a surrogate variable for serum RF. Another classification tree analysis forced rheumatoid factor as the first partitioning variable and yielded a sensitivity of 95% and a specificity of 87%.

Table 6 defines the 1987 criteria for classification of rheumatoid arthritis using the tree method. The definitions are similar to those used in the traditional formulation. Note that morning stiffness (57% specific) and rheumatoid nodules (43% sensitive) are not present in the tree structure. Classification tree analysis divides the entire population of patients and controls into 8 subsets: 5 represent patients classified as having RA (Table 7 and Figure 1). The great majority of RA patients are correctly identified, with high accuracy, by the simple requirement of arthritis of 3 or more joint areas together with typical radiographic findings on hand radiographs. When combined with the second subset requirement, an even greater proportion of patients are correctly classified. Thus, a simple specification of RA as arthritis of 3 or more joint areas for 6 or more weeks with either radiographic changes or serum RF provides a powerful approximation of the performance of an entire criteria set. Conversely, arthritis of less than 3 joint areas together with an absence of RF excludes the great majority of RA patients.

Table 8 summarizes the relative performance of the old and new criteria sets in the population of RA patients and controls studied by the subcommittee. The classification tree system, which had only 45 misclassifications, was the best, followed by the new criteria in traditional (4 of 7) format, which had 51 misclassifications. The old ARA and New York criteria proved to be substantially less accurate. The performances of each of the classification methods were also tested in the 47 RA and 51 non-RA patients whose disease duration was less than 1 year (Table 8). The classification tree was more accurate than the traditional format, although both showed reasonable specificity, 90% and 88%, respectively.

To test specificity against a normal population, the new ARA criteria in both traditional and tree formats were tested on 137 consecutive normal individuals aged 52–74 who participated in a longitudinal study of human aging and who had given a rheumatologic history and undergone a physical examination and radiographic studies of the hands (11). No subjects from this group would have been classified as having RA by either the traditional or the classification tree formulations. Only 18 positive findings were detected (3 patients had swelling of the proximal interphalangeal [PIP] joints, 3 had symmetric arthritis, and 12 patients had morning stiffness). Thirteen subjects had 1 criterion, 1 patient had 2 (symmetric arthritis and PIP joint swelling), and 1 patient had 3 (symmetric arthritis, PIP joint swelling, and morning stiffness). No

Table 6. The 1987 classification tree criteria and definitions for rheumatoid arthritis (RA)*

| Criterion | Definition |
|---|---|
| 1. Arthritis of 3 or more joint areas | At least 3 joint areas simultaneously have had soft tissue swelling or fluid (not bony overgrowth alone) observed by a physician. The 14 possible joint areas are right or left PIP, MCP, wrist, elbow, knee, ankle, and MTP joints |
| 2. Arthritis of hand joints<br>Wrist<br>MCP<br>MCP or wrist<br>MCP and wrist | Soft tissue swelling or fluid (not bony overgrowth alone) of the specified area observed by a physician. Where 2 areas are specified, involvement must have been simultaneous |
| 3. Symmetric swelling (arthritis) | Simultaneous involvement of the same joint areas (as defined in 1) on both sides of the body (bilateral involvement of PIPs, MCPs, or MTPs is acceptable without absolute symmetry) |
| 4. Serum rheumatoid factor | Demonstration of abnormal amounts of serum rheumatoid factor by any method for which the result has been positive in <5% of normal control subjects |
| 5. Radiographic changes of rheumatoid arthritis | Radiographic changes typical of rheumatoid arthritis on posteroanterior hand and wrist radiographs which must include erosions or unequivocal bony decalcification localized in or most marked adjacent to the involved joints (osteoarthritis changes alone do not qualify) |

* A patient is said to have RA if he/she is included in 1 of the 5 RA subsets listed in Table 7 and has a clinical diagnosis of RA by his/her physician. Criteria 1, 2, and 3 must have been present for at least 6 weeks. See Table 3 for definitions of abbreviations.

radiographic findings of RA were identified in these subjects.

The study of consecutive patients from multiple practices inevitably includes some in whom the clinical diagnosis of RA or another rheumatic disease was uncertain. Table 9 shows that contributing physicians assigned an average overall certainty score of 89 for RA patients and, conversely, for non-RA patients a score of 6 on a scale of 0–100, where 100 represented certain RA and 0 represented certain non-RA. When one looks at those cases and controls who were correctly classified, the corresponding certainty scale scores were 91 and 5 by both classification methods. However, when examining those patients "incorrectly" classified, the certainty scores shifted toward the mid-region of the scale for both cases and controls, suggesting uncertainty about the original clinical diagnosis. Thus, the analog scale seemed to accurately identify patients who were difficult to classify by any means. Of the 17 RA patients misclassified by the tree technique, 12 were considered to have "possible" or "probable" rheumatoid arthritis by their physicians. Thus, classification accuracy as presented here importantly understates the true accuracy.

## DISCUSSION

The 1987 ARA criteria for adult RA using either a traditional formulation or a classification tree performed better than the older ARA criteria (4) or the New York criteria (5). The revised classification criteria for RA, by either method, requires fewer criteria than previously, and less extensive and costly investigations are necessary for their rigorous application since invasive requirements have been deleted and radiographic requirements are restricted to interpretation of a single posteroanterior radiograph of both hands. Eliminating exclusions renders the criteria more suitable for epidemiologic studies, removes some logical inconsistencies associated with exclusions (i.e., patients could not have 2 diagnoses), and increases simplicity and teaching value. More explicit definitions of criteria such as morning stiffness (formerly listed without a duration or location requirement) have improved performance and, thus, confidence that patients are correctly classified. At the same time, the new criteria preserve continuity by retaining many features of the old criteria set and, in fact, closely parallel the old definition of definite or classic RA. Finally, the new criteria, unlike the old (4), are purposefully formulated to facilitate the classification of RA rather than clinical diagnosis.

Five major changes have been made in the new criteria. First, the designation of "probable" RA has been dropped. The subcommittee's initial distrust of this designation was reinforced by examination of the data, which revealed that the great majority of misclassified patients had been so designated clinically. The subcommittee suggests that terms such as undifferentiated polyarthritis, undifferentiated oligoarthritis, or undifferentiated monarthritis are preferable to describe patients previously designated as having probable RA. Thus, such patients are not given a diagnosis that suggests a particular disease process when in fact their disease status is highly uncertain.

Second, the concepts of definite and classic rheumatoid arthritis have been replaced by "rheu-

**Table 7.** The 1987 classification tree criteria for rheumatoid arthritis (RA)*

| RA subsets† | No. of patients RA/non-RA | % correctly classified | % RA patients in subset | Non-RA subsets† | No. of patients RA/non-RA | % correctly classified | % of non-RA patients in subset |
|---|---|---|---|---|---|---|---|
| 1. Arthritis of 3 or more joint areas and positive findings on hand radiographs | 197/11 | 95 | 75 | 3. Arthritis of 3 or more joint areas, specifically not involving MCP *and* wrist joint areas; serum rheumatoid factor–negative and negative findings on hand radiographs | 5/18 | 78 | 7 |
| 2. Arthritis of 3 or more joints and serum rheumatoid factor–positive; negative findings on hand radiographs | 28/7 | 80 | 11 | | | | |
| 4. Arthritis of 3 or more joint areas including MCP and wrist joints; serum rheumatoid factor–negative and negative findings on hand radiographs | 7/4 | 64 | 3 | 7. Seropositive, but asymmetric oligoarthritis not involving wrist *or* MCP joint areas | 3/14 | 82 | 5 |
| 5. Arthritis involving a single joint area symmetrically and serum rheumatoid factor–positive | 7/2 | 78 | 3 | 8. Seronegative oligoarthritis | 9/202 | 96 | 77 |
| 6. Oligoarthritis of MCP or wrist joints and serum rheumatoid factor–positive | 6/4 | 60 | 2 | | | | |

* See Table 3 for definitions of abbreviations. Missing data rules: If radiographs are not available, substitute MCP swelling; if results of tests for rheumatoid factor are not available, substitute wrist swelling; if the clinical data from routine tests are not available, the patient's disease cannot be classified.
† These subset numbers also appear under the subset boxes in Figure 1.

matoid arthritis." The committee was unable to find significant differences in the physician-rated certainty of diagnoses between patients designated as having definite disease and those rated as having classic disease, or differences in certainty scores with the new criteria based on the number of criteria (more than 4) present. Thus, these terms were deemed unnecessary.

Third, the criteria that involved the invasive techniques of joint aspiration, synovial biopsy, or biopsy of a rheumatoid nodule were deleted. The absence of such findings in control subjects makes it impossible to rigorously apply them as criteria. In addition, they are not helpful in the diagnosis of the vast majority of RA patients. Since these procedures were not used in practice, it is unlikely that they will be missed.

Fourth, the new ARA criteria are deliberately more specific; they have many features of the old ARA criteria (4) and the New York criteria (5). The requirement of 3 or more joints involved for 6 or more weeks

**Table 8.** Comparison of the relative performances of criteria for rheumatoid arthritis

| Criteria | Sensitivity (%) | Specificity (%) | No. of patients misclassified |
|---|---|---|---|
| 1958 ARA* | 92 | 85 | 61 |
| 1966 New York† | 98 | 76 | 69 |
| 1966 New York‡ | 81 | 94 | 64 |
| 1987 new ARA§ | 91.2 | 89.3 | 51 |
| 1987 new classification tree¶ | 93.5 | 89.3 | 45 |

* Criteria of mucin clot, synovial biopsy, and nodule biopsy excluded. At least 5 of 8 criteria must be present. ARA = American Rheumatism Association.
† At least 2 of 4 criteria must be present.
‡ At least 3 of 4 criteria must be present.
§ At least 4 of 7 criteria must be present; early onset of disease (<1 year): sensitivity 80.9%; specificity 88.2%.
¶ Early onset of disease (<1 year): sensitivity 85%; specificity 90%.

**Table 9.** Analysis of misclassification of rheumatoid arthritis (RA) patients and control subjects by certainty scale average values*

| | Overall | Correctly classified | Incorrectly classified |
|---|---|---|---|
| Classification tree method | | | |
| Patients (RA) | (262) | (245) | (17) |
| | 89 | 91 | 59 |
| Control subjects (non-RA) | (262) | (234) | (28) |
| | 6 | 5 | 14 |
| Classification by 4 of 7 criteria | | | |
| Patients (RA) | (262) | (239) | (23) |
| | 89 | 91 | 67 |
| Control subjects (non-RA) | (262) | (234) | (28) |
| | 6 | 5 | 14 |

* Scoring of analog certainty scale: 0 = absolutely not RA; 100 = absolutely RA. Numbers in parentheses are the numbers of patients.

is essentially a consolidation of the previous second, third, and fourth criteria of the old ARA set (4). The inclusion of wrist, MCP, or PIP joints increases specificity since these joints are typically affected in RA. These 2 combined criteria increase specificity and decrease the need for additional criteria.

Fifth, the requirement for exclusions has been deleted. Four conditions (systemic lupus erythematosus, psoriatic arthritis, mixed connective tissue disease, and Reiter's syndrome) appear likely to have substantial numbers of patients who might fulfill the requirements of the new criteria, and caution should be observed in these circumstances. In addition, several other rheumatic disorders that might cause classification difficulties were not represented with significant frequency in this study. One example is polymyalgia rheumatica, in which a symmetric polyarthritis of hands and wrists may occasionally be seen (13,14). The absence of criteria 5, 6, and 7 in these elderly patients, as well as their excellent clinical response to low-dose corticosteroid therapy, may discriminate them from patients with seropositive RA. Another disorder likely to require further study is primary Sjögren's syndrome with polyarthritis (15). Normal subjects (0 of 150 in this study) will very seldom fulfill the new criteria, and very low occurrence rates are expected with the other control conditions studied. The advantage of elimination of a long exclusion list is believed to greatly outweigh any occasional misclassification occasioned by this change.

The validity of the concept of seronegative RA remains unproven (16,17). The subcommittee carefully and extensively reviewed the available clinical, epidemiologic, radiographic, and immunogenetic literature regarding this distinction (16–18). While certain studies show an HLA–DR4 association only with seropositive RA, an equal number have found an excessive occurrence rate of DR4 in seronegative patients as well (18). Thus, the subcommittee was unable to discriminate meaningfully between seronegative and seropositive rheumatoid arthritis. Nevertheless, several classification schemes that require the presence of serum RF were tested in this study population. None performed as well as the current recommendations in either the traditional or tree formats. Thus, the new criteria retain serum RF as a criterion that is equal in importance to the other 6.

The new classification criteria have been presented in both traditional and tree formats, thus offering additional insights into clinical subsets of RA. Either may be used to classify patients; however, it is recommended that the format used be designated in studies and reports. The classification tree yields slightly better accuracy than the traditional format, especially in early disease, but both techniques represent major improvements over previous formulations.

Disease criteria which are descriptive reflect our current understanding of these disorders. Elucidation of specific pathogenetic mechanisms may at some point permit classification to be based directly on disease biology. However, these new criteria for RA will necessarily serve to improve understanding, classification, and comparability of patients with rheumatoid arthritis until other methods of achieving this purpose are available.

## ACKNOWLEDGMENTS

The subcommittee wishes to thank Dee Simpson of the American Rheumatism Association Medical Information System for data management and technical assistance and Margaret Kelley for preparation of the manuscript. We also thank the following investigators for participating in the study: Drs. Larry G. Anderson, Herbert S. B. Baraf, Kenneth L. N. Blocka, James R. Carpenter, Lewis Cogen, Marc D. Cohen, David H. Collier, David A. Cooley, Richard A. Gilman, Peter Hasselbacher, Bonnie Hepburn, Leroy H. Hunninghake, Joseph M. Isaacson, Phoebe R. Krey, Edward V. Lally, Aldo V. Londino, Jr., Tariq Mahmood, Janet J. Markland, Marc L. Miller, George L. Morton, Lawrence M. Mulhern, Thaddeus A. Osial, Jr., Paulding Phelps, Frank Roberts, Alexander P. Ruggieri, Sharon Stotsky, Paul D. Schneider, Randall M. Stevens, Clarence R. Thompson, Noranna B. Warner, Arthur L. Weaver, and Bernard L. Zimmerman.

## REFERENCES

1. Ropes MW, Bennett GA, Cobb S, Jacox R, Jessar RA: Proposed diagnostic criteria for rheumatoid arthritis. Bull Rheum Dis 7:121–124, 1956
2. Cobb S, Merchant WR, Warren JE: An epidemiologic look at the problem of classification in the field of arthritis. J Chronic Dis 2:50–54, 1955
3. Cobb S, Thompson DJ, Rosenbaum J, Warren JE, Merchant WR: On the measurement of prevalence of arthritis and rheumatism from interview data. J Chronic Dis 3:134–139, 1956
4. Ropes MW, Bennett GA, Cobb S, Jacox R, Jessar RA: 1958 revision of diagnostic criteria for rheumatoid arthritis. Bull Rheum Dis 9:175–176, 1958
5. Bennett PH, Wood PHN (editors): Population Studies of the Rheumatic Diseases. Amsterdam, Excerpta Medica, 1968, pp 477–478
6. Mitchell DM, Fries JF: An analysis of the American Rheumatism Association criteria for rheumatoid arthritis. Arthritis Rheum 25:481–487, 1982
7. Classification of the rheumatic diseases, Primer on the

Rheumatic Diseases. Eighth edition. Atlanta, Arthritis Foundation, 1983, pp 36–37
8. Harris ED Jr: Rheumatoid arthritis: the clinical spectrum, Textbook of Rheumatology. Second edition. Edited by WN Kelley, ED Harris Jr, S Ruddy, CB Sledge. Philadelphia, WB Saunders, 1985, pp 915–950
9. Altman R, Asch E, Bloch D, Bole G, Borenstein D, Brandt K, Christy W, Cooke TD, Greenwald R, Hochberg M, Howell D, Kaplan D, Koopman W, Longley S III, Mankin H, McShane DJ, Medsger T Jr, Meenan R, Mikkelsen W, Moskowitz R, Murphy W, Rothschild B, Segal M, Sokoloff L, Wolfe F: Development of criteria for the classification and reporting of osteoarthritis: classification of osteoarthritis of the knee. Arthritis Rheum 29:1039–1049, 1986
10. Breiman L, Friedman JH, Olshen RA, Stone CJ: Classification and Regression Trees. Belmont, CA, Wadsworth, 1984
11. Lane NE, Bloch DA, Wood PD, Fries JF: Aging, long-distance running, and the development of musculoskeletal disability. Am J Med 82:772–780, 1987
12. Cart TM: California Statistical Software. Lafayette, CA, 1984
13. Chou C-T, Schumacher HR Jr: Clinical and pathologic studies of synovitis in polymyalgia rheumatica. Arthritis Rheum 27:1107–1117, 1984
14. Sheets PK, Healey LA: Are polymyalgia rheumatica and seronegative rheumatoid arthritis the same? (abstract) Arthritis Rheum (suppl) 30(1):S27, 1987
15. Castro-Poltroniere A, Alarcon-Segovia D: Articular manifestations of primary Sjögren's syndrome. J Rheumatol 10:485–488, 1983
16. Calin A, Marks SH: The case against seronegative rheumatoid arthritis. Am J Med 70:992–994, 1981
17. Masi AT, Feigenbaum SL: Seronegative rheumatoid arthritis: fact or fiction? Arch Intern Med 143:2167–2172, 1983
18. Olsen NJ, Callahan LF, Brooks RH, Nance EP, Kaye JJ, Stastny P, Pincus T: Associations of HLA-DR4 with rheumatoid factor and radiographic severity in rheumatoid arthritis. Am J Med (in press)