

## SPECIAL ARTICLE

# DEVELOPMENT OF CRITERIA FOR THE CLASSIFICATION AND REPORTING OF OSTEOARTHRITIS

## Classification of Osteoarthritis of the Knee

R. ALTMAN, E. ASCH, D. BLOCH, G. BOLE, D. BORENSTEIN, K. BRANDT, W. CHRISTY, T. D. COOKE, R. GREENWALD, M. HOCHBERG, D. HOWELL, D. KAPLAN, W. KOOPMAN, S. LONGLEY, III, H. MANKIN, D. J. McSHANE, T. MEDSGER, JR., R. MEENAN, W. MIKKELSEN, R. MOSKOWITZ, W. MURPHY, B. ROTHSCILD, M. SEGAL, L. SOKOLOFF, and F. WOLFE

For the purposes of classification, it should be specified whether osteoarthritis (OA) of the knee is of unknown origin (idiopathic, primary) or is related to a known medical condition or event (secondary). Clinical criteria for the classification of idiopathic OA of the knee were developed through a multicenter study group. Comparison diagnoses included rheumatoid arthritis and other painful conditions of the knee, exclusive of referred or paraarticular pain. Variables from the medical history, physical examination, laboratory tests,

and radiographs were used to develop sets of criteria that serve different investigative purposes. In contrast to prior criteria, these proposed criteria utilize classification trees, or algorithms.

The diagnosis of osteoarthritis (OA) has most often been based on radiographic appearance, rather than clinical features. Radiographic criteria were proposed by Kellgren and Lawrence in 1957 (1), and those criteria were later accepted by the World Health Organization at a symposium held in Milan in 1961 (2). Lequesne has proposed sets of clinical criteria for OA in several specific joints (3,4).

In 1981, the American Rheumatism Association asked the Diagnostic and Therapeutic Criteria Committee to establish a subcommittee on OA. The subcommittee accepted as its charge the development of criteria for the classification of OA. The initial objectives were to standardize and clarify the clinical definition of OA, in order to promote consistency in the reporting of OA and in the interpretation of research concerning OA. This can be accomplished by 1) development of a classification of OA that includes recognized subsets; and 2) identification of OA subsets through the use of a combination of clinical and laboratory features.

For the purposes of this report, OA is defined as a heterogeneous group of conditions that lead to joint symptoms and signs which are associated with defective integrity of articular cartilage, in addition to related changes in the underlying bone and at the joint margins. Although articular cartilage is poorly innervated and defects in cartilage are not, in themselves, symptomatic, a clinical syndrome of symptoms, which often includes pain, may evolve from such defects.

From the Subcommittee on Classification Criteria of Osteoarthritis, a subcommittee of the Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association.

R. Altman, MD: University of Miami School of Medicine, Miami, FL; E. Asch, MD: New York University, New York, NY; D. Bloch, PhD: Stanford University, Palo Alto, CA; G. Bole, MD: University of Michigan, Ann Arbor; D. Borenstein, MD: George Washington University Medical Center, Washington, DC; K. Brandt, MD: Indiana University School of Medicine, Indianapolis; W. Christy, MD: University of Pittsburgh, Pittsburgh, PA; T. D. Cooke, MD: Queen's University, Kingston, Ontario, Canada; R. Greenwald, MD: Long Island Jewish Medical Center, New Hyde Park, NY; M. Hochberg, MD, MPH: Johns Hopkins University, Baltimore, MD; D. Howell, MD: University of Miami, Miami, FL; D. Kaplan, MD: Downstate Medical Center, Brooklyn, NY; W. Koopman, MD: University of Alabama at Birmingham; S. Longley, III, MD: University of Florida, Gainesville; H. Mankin, MD: Massachusetts General Hospital, Boston; D. J. McShane, MD: Stanford University, Palo Alto, CA; T. Medsger, Jr., MD: University of Pittsburgh, Pittsburgh, PA; R. Meenan, MD, MPH: Boston University School of Medicine, Boston, MA; W. Mikkelsen, MD: University of Michigan, Ann Arbor; R. Moskowitz, MD: Case Western Reserve University, Cleveland, OH; W. Murphy, MD: Mallinckrodt Institute of Radiology, Washington University, St. Louis, MO; B. Rothschild, MD: Menorah Medical Center, Kansas City, MO; M. Segal, PhD: Stanford University, Palo Alto, CA; L. Sokoloff, MD: State University of New York at Stony Brook; F. Wolfe, MD: University of Kansas, Wichita.

Address reprint requests to Roy D. Altman, MD, University of Miami, PO Box 016960 (VA 111), Miami, FL 33101.

Submitted for publication December 16, 1985; accepted in revised form April 21, 1986.



Until a diagnostic method is developed that will integrate the clinical findings with etiologic, biochemical, biomechanical, and histologic abnormalities of the syndrome we call OA, the subcommittee is compelled to employ commonly available diagnostic techniques in developing the classification criteria. This report details the subcommittee's approach to the classification of idiopathic (primary) osteoarthritis of 1 joint, the knee, utilizing findings from the medical history, physical examination, laboratory testing, and radiography.

## METHODS

**Classification of subsets of osteoarthritis.** The subcommittee's first objective was to develop a classification of OA subsets. There have been several published classifications of OA. The subcommittee proposes a classification system that separates patients with OA into 2 categories: 1) those with no presently known prior event or disease related to the OA (idiopathic); and 2) those with known events or disease associated with OA (secondary) (5) (Table 1). The classification is further divided according to anatomic (idiopathic) or etiologic (secondary) conditions.

Classifications presented in textbooks have generally followed this proposed pattern (6-8); each emphasizes certain aspects of OA. An ideal classification would include histologic findings, but tissue from the joint affected with OA is infrequently available. The classification proposed in 1977 by Mitchell and Cruess (9) is unique in that diagnostic groups are defined by the status of the cartilage matrix. Although their method focuses on etiologic factors, it omits clinically identifiable syndromes, especially those related to involvement of the spine and interphalangeal joints.

The proposed classification (Table 1) recognizes that OA may involve virtually any joint, and some of the less common sites of involvement are noted. It also takes into account that all OA may be secondary to phenomena not yet discovered; hence, the term "idiopathic OA" is used in lieu of the term "primary OA." For example, if it is discovered that interphalangeal OA has a distinct etiology, separate from other forms of OA, it may be readily repositioned in the secondary OA section.

It is difficult to place calcium deposition diseases within this classification system because these diseases may be idiopathic or secondary. We have elected to classify the calcium deposition diseases as a form of secondary OA, although this decision is admittedly arbitrary. Forestier's ankylosing hyperostosis (diffuse idiopathic skeletal hyperostosis, DISH) is empirically classified as an idiopathic OA variant. Further research may establish this as a separate disease.

**Criteria selection.** Because of differences in the clinical presentation of OA in different joints, the subcommittee initially focused on idiopathic OA of the knee. Twenty-three historical, physical, and laboratory features were considered worthy of further evaluation (5). The sensitivity and specificity of these variables were initially determined by the "Delphi" technique of opinion sampling (10,11). The Delphi procedure was selected because it is designed to generate a

**Table 1.** Classification for subsets of osteoarthritis\*

I. Idiopathic
A. Localized
1. Hands: e.g., Heberden's and Bouchard's nodes (nodal), erosive interphalangeal arthritis (non-nodal), scapho-metacarpal, scaphotrapezoid
2. Feet: e.g., hallux valgus, hallux rigidus, contracted toes (hammer/cockup toes), talonavicular
3. Knee
a. Medial compartment
b. Lateral compartment
c. Patellofemoral compartment (e.g., chondromalacia)
4. Hip
a. Eccentric (superior)
b. Concentric (axial, medial)
c. Diffuse (coxae senilis)
5. Spine (particularly cervical and lumbar)
a. Apophyseal
b. Intervertebral (disc)
c. Spondylosis (osteophytes)
d. Ligamentous (hyperostosis [Forestier's disease, or DISH])
6. Other single sites: e.g., shoulder, temporomandibular, sacroiliac, ankle, wrist, acromioclavicular
B. Generalized: includes 3 or more areas listed above (Kellgren-Moore, see ref. 18)
1. Small (peripheral) and spine
2. Large (central) and spine
3. Mixed (peripheral and central) and spine
II. Secondary
A. Post-traumatic
B. Congenital or developmental diseases
1. Localized
a. Hip diseases: e.g., Legg-Calvé-Perthes, congenital hip dislocation, slipped capital femoral epiphysis, shallow acetabulum
b. Mechanical and local factors: e.g., obesity (?), unequal lower extremity length, extreme valgus/varus deformity, hypermobility syndromes, scoliosis
2. Generalized
a. Bone dysplasias: e.g., epiphyseal dysplasia, spondylo-apophyseal dysplasia
b. Metabolic diseases: e.g., hemochromatosis, ochronosis, Gaucher's disease, hemoglobinopathy, Ehlers-Danlos disease
C. Calcium deposition disease
1. Calcium pyrophosphate deposition disease
2. Apatite arthropathy
3. Destructive arthropathy (shoulder, knee)
D. Other bone and joint disorders: e.g., avascular necrosis, rheumatoid arthritis, gouty arthritis, septic arthritis, Paget's disease, osteopetrosis, osteochondritis
E. Other diseases
1. Endocrine diseases: e.g., diabetes mellitus, acromegaly, hypothyroidism, hyperparathyroidism
2. Neuropathic arthropathy (Charcot joints)
3. Miscellaneous: e.g., frostbite, Kashin-Beck disease, Caisson disease

\* DISH = diffuse idiopathic skeletal hyperostosis.



**Table 2.** Composition of osteoarthritis comparison group\*

Condition	No. of patients
Rheumatoid arthritis	55
Other diagnoses (n = 52)	
Meniscal, ligamentous, or cruciate abnormalities	12
Arthralgia or fibromyalgia (fibrositis)	11
Patella laxa	5
Osteonecrosis	4
Connective tissue disease	4
Synovitis, type undetermined	3
Gouty arthritis	3
Rheumatoid variant	3
Septic arthritis	2
Osteochondritis dissecans	1
Synovitis, other	4

\* There were 107 patients with knee pain who did not have osteoarthritis of the knee. These composed the comparison group.

consensus of expert opinion in situations of uncertainty, by the use of anonymity, feedback, and iteration. The list of 23 variables was mailed to the subcommittee members. They independently rated each variable as to its specificity and sensitivity. The results were collated, and a table was established that included means, standard deviations from the mean, and medians. This table and the list of variables were recirculated to the members for 3 rounds. As expected, the mean ratings for the variables changed slightly, and the standard deviations narrowed. The Delphi technique clarified the thinking of the subcommittee members and produced a reasonable consensus about the potential classification criteria.

The subcommittee expanded a core set of Delphi findings by directly defining 85 historical, physical, laboratory, and radiographic features to incorporate in a prospective data collection effort. A protocol and worksheet were designed. Consecutive patients with OA and a comparison population, all of whom met the following criteria, were enrolled from each of the contributing centers: 1) knee pain, irrespective of quality, duration, or periodicity; 2) pain of articular origin, not referred from the hip, spine, or para-articular regions such as the anserine, prepatellar, or infrapatellar bursae; 3) available current radiographs of the knee; 4) absence of features of secondary OA of the knee, regardless of cause, as defined in Table 1.

Centers submitted protocols for 20–25 patients with knee pain. About half of the patients had symptomatic idiopathic OA of the knee. The remainder composed the comparison group, which included patients with knee pain of other origins, including rheumatoid arthritis (RA), Reiter's syndrome, psoriatic-associated arthritis, fibromyalgia (fibrositis), or any knee pain syndrome (Table 2). All patient data and radiographs were coded.

The standard against which the classification criteria were judged was the clinical diagnosis of OA. All data forms were reviewed independently by 3 members of the subcommittee, for verification of the clinical diagnosis. If the reviewers disagreed with the submitted clinical diagnosis, the center coordinator was contacted, and the diagnosis was discussed. If the consensus was that the case represented

secondary OA, the data were excluded from the analysis. Coded radiographs were read blindly and independently by one of the subcommittee members (WM), a musculoskeletal radiologist. The readings and diagnoses made by the radiologist were compared with those submitted by the center coordinator.

Upon verification of the diagnosis and interpretation of the radiographs, data were entered into the American Rheumatism Association Medical Information System (ARAMIS) computer at Stanford University, Palo Alto, CA, via remote terminal. ARAMIS was employed for data management and analysis because of experience with the development of criteria for rheumatic disease (12–14). Accuracy of data entry was verified by review of 20% of the entries, which were randomly selected. Also, specific criteria in all records were examined for accuracy of entry (e.g., diagnosis, age, and rheumatoid factor [RF] titer).

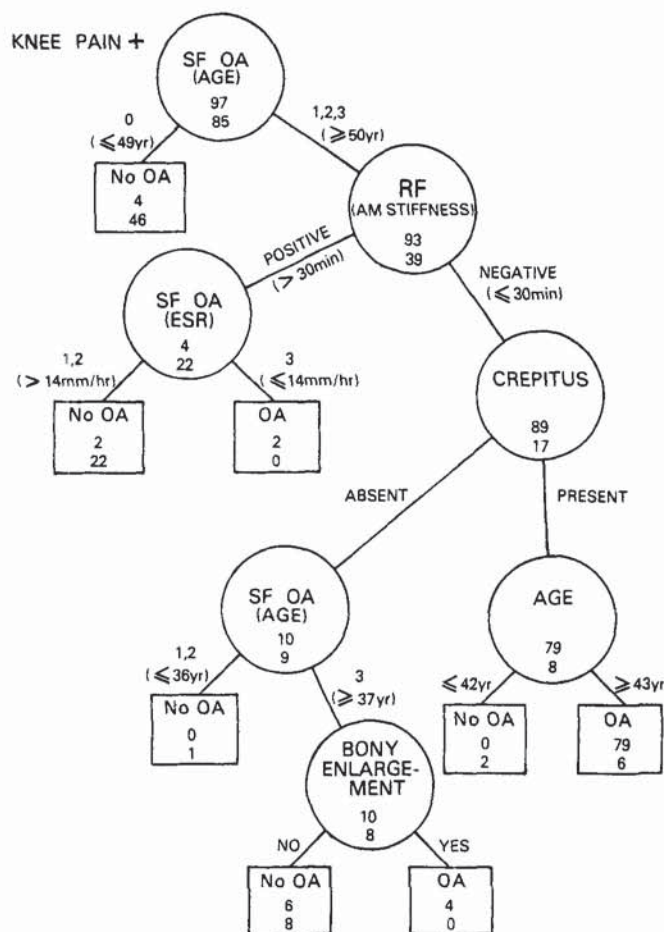
The items assessed included 28 historical features; 16 physical examination findings; 3 laboratory test results, including 10 synovial fluid measurements; and 28 features derived from radiographs of the knee that included weight-bearing anteroposterior and lateral views, with skyline (sunset) patellar views. Each of these items was analyzed (by univariate techniques) for sensitivity by calculating the mean value, for continuous variables, and the proportion present, for dichotomous variables. A variable was included in subsequent analyses if it discriminated between OA and the comparison group at a level of  $P < 0.05$ . Other variables were included if other published studies had found them to discriminate and/or if the Delphi exercise had found them to discriminate. Using this approach, 42 variables from the original list of 85 were considered to be potentially important for the classification of OA of the knee.

**Data analysis.** Two different statistical methods were used to develop classification criteria. The first method included procedures that have been used in prior criteria studies (12–14). Variables were separately analyzed by stepwise logistic regression techniques (15). Nine clinical variables demonstrated significance (data not presented). These 9 clinical and 28 radiographic variables were subjected to correlation analysis and to stepwise logistic regression as a group (15). Correlations indicate the degree of association between variables. The variables utilized in the final traditional analysis were variables "significantly" associated with predicting OA of the knee.

Correlations ranged from  $-0.25$  to  $+0.59$ , with none close to  $\pm 1$  (data not presented). Results of stepwise logistic regression indicated that 7 of the 9 clinical characteristics were significant at  $P \leq 0.003$ , while the other 2 (no palpable effusions, no palpable fever) were close to being significant at the 0.05 level. Therefore, all 9 items remained candidates for the classification of subjects. Despite the finding that osteophytes were the sole predictor of OA by these methods, as determined by logistic regression analysis, 6 radiographic characteristics were considered candidates for classification of subjects.

The last step of this analytic method was to select combinations of the variables which were most sensitive and specific to classification of OA of the knee. Composite criteria (combinations) were derived by the combination of individual candidate variables by means of Boolean algebra, using union and intersection operations. The resulting clas-





**Figure 1.** Osteoarthritis of the knee classification tree, for clinical and laboratory criteria. The classification tree is derived from recursive partitioning analysis (see Methods). The circles depict variables to be used in the classification algorithms. Each circle includes the variable, the alternative, or surrogate, variable (in parentheses), and 2 numbers. When a particular variable is unavailable, then its surrogate variable may be used. The upper numbers represent the subjects with osteoarthritis (OA), and the lower numbers represent the subjects without OA (from the comparison group). The cutpoints at which the variables are split are shown on the arms of the circles, with the splits of the surrogate variables noted in parentheses, i.e., the age split is  $\leq 49$  years and  $\geq 50$  years; the rheumatoid factor (RF) split is positive and negative; the morning stiffness (AM stiffness) split is  $>30$  minutes duration and  $\leq 30$  minutes duration; the erythrocyte sedimentation rate (ESR) split is  $>14$  mm/hour and  $\leq 14$  mm/hour. For example, to be included in the clinical and laboratory algorithm, a synovial fluid analysis result indicative of OA (SF OA) may contain up to 3 findings (clear, viscous, and/or white blood cell count  $<2,000/\text{mm}^3$ ) for the purposes of the algorithm; the split of the variable specifies the number of findings required at each level. The boxes specify whether the subject can be classified as having OA (OA) or as not having OA (No OA). The overall sensitivity of this method is 88%; the overall specificity is 93%.

sification rule is of the form "if a subject has knee pain, and at least X out of a list of Y characteristics are present, then classify the subject as having OA of the knee."

There are serious deficiencies with the aforementioned method of classification (see Discussion). Therefore, a second method was used to classify subjects: the technique of "classification trees," or "recursive partitioning." The definitive reference is *Classification and Regression Trees*, by Breiman et al (16). We present below a brief and simplified overview of this technique.

A classification tree is grown as follows. For each node (the initial node contains the entire sample), examine every allowable split of each variable. Select the most discriminating, or "best," of these splits and create 2 new "daughter" nodes. "Best" is assessed in terms of a goodness-of-split criterion (see chapter 4, ref. 16). The splitting criterion is such that one recursively creates smaller and smaller nodes of progressively increased purity, i.e., nodes progressively contain a larger and larger proportion of either OA subjects only or non-OA subjects only. For example, the tree diagram for clinical plus laboratory characteristics (Figure 1) reveals that the "best" first split was for synovial fluid findings of OA (SF OA). If the subject has none of the 3 SF OA findings, then he or she is classified as not having OA. If 1 or more SF OA findings are present, then the subject is further classified according to whether the RF is positive or negative.

An appropriately sized tree is derived by initially growing a very large tree. The tree is then iteratively pruned, thereby creating a sequence of trees. The best tree is selected from this sequence, by the use of cross-validated estimates of misclassification rates (for details, see chapter 3, ref. 16).

Considerable information can be extracted in cases with missing values, by the use of "surrogate" variables rather than using conventional methods (see Discussion). The first surrogate for a given non-terminal node is the variable that best reproduces the actual split of that node. The second surrogate does second best, and so on. For example, in the clinical, laboratory, and radiograph, or x-ray, tree (Figure 2), patients without osteophytes or spurs noted on radiographs could be further classified using the variable of SF OA. The first surrogate for SF OA is the variable of age. Those cases missing SF OA, i.e., not having a synovial fluid analysis available, are sent to the left or right daughter node according to the patient's age,  $\leq 39$  years or  $\geq 40$  years, respectively.

After selection of a tree, the majority rule classifies cases in every terminal node as either having OA or not having OA. For example, in the clinical tree (Figure 3), 108 subjects have crepitus,  $\leq 30$  minutes of morning stiffness, and an age of at least 38 years. One hundred of these subjects have OA, and 8 do not have OA. Subjects with these 3 characteristics are classified as having OA.

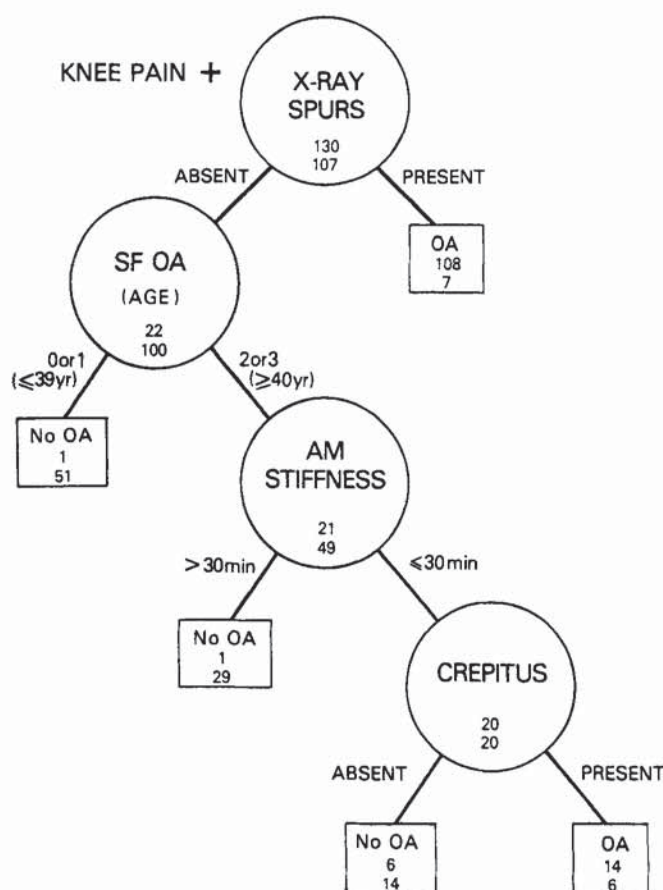
Subsequent independent data analysis was performed by a subcommittee member (DJM) who had not been involved with the initial analysis. Validity testing of the developed classification criteria was performed in a single clinic (by subcommittee member FW) on an additional 100 consecutive patients with knee pain from OA or from other rheumatic causes.



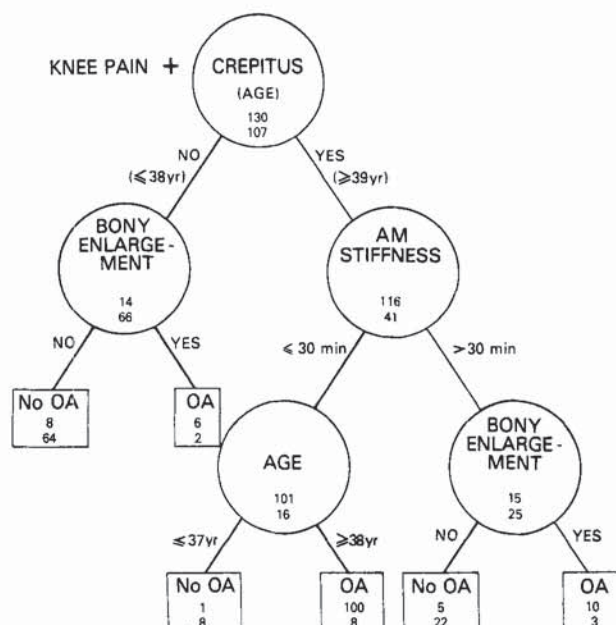
## RESULTS

The study included 264 patients from 14 contributing centers. The reviewers changed the diagnostic category of 27 of the patients (10.2%) to secondary OA, which was related to calcium pyrophosphate deposition disease in 13, RA in 8, and other diseases such as gout or fracture in 6. After these exclusions, the remaining 237 patients were analyzed. One hundred thirty were diagnosed as having idiopathic OA of the knee. The remaining 107 patients, 55 of whom had RA and 52 of whom had other diseases, served as the comparison group (Table 2).

The historical features, physical examination findings, laboratory test results, and radiographic findings are presented in Tables 3-6. Individuals with idiopathic OA were older ( $P < 0.001$ ). Although morning stiffness was common in the group with OA, it was



**Figure 2.** Osteoarthritis of the knee classification tree, for clinical, laboratory, and radiographic (x-ray) criteria. The overall sensitivity of this method is 94%; the overall specificity is 88%. See Figure 1 for explanation and definitions.



**Figure 3.** Osteoarthritis of the knee classification tree, for clinical criteria. The overall sensitivity of this method is 89%; the overall specificity is 88%. See Figure 1 for explanation and definitions.

most often present for  $\leq 30$  minutes (Table 3), while those patients with RA had stiffness for  $> 30$  minutes. Seventy-nine percent of the non-RA comparison group had  $\leq 30$  minutes of morning stiffness. Presence of pain for more than 2 weeks of the previous month implied some chronicity for all groups in this survey.

Even without prior criteria for classification, the clinical diagnosis of OA was previously recorded for 77% of patients with OA and for only 8% of the patients in the comparison group. At some point in the course of their disease, 30% of the patients with OA had used a cane. Similarly, a cane was used by 24% of patients with RA, but by only 4% of the non-RA comparison group. A walker or wheelchair was used by 3% of patients with OA, 10% of the patients with RA, and 2% of the non-RA comparison group.

On physical examination (Table 4), malalignment, particularly varus deformity, was more often present in OA patients than in patients of the comparison groups. When present in OA patients, the signs of inflammation were considerably less frequent and less severe than were those in patients of the comparison groups. In OA, periosteal joint margin bony tenderness on palpation was more frequent than was parapatellar synovial tenderness. Palpable bony enlargement at the joint margin was more common in

Table 3. Medical history findings (frequency) in 237 patients with knee pain\*

	OA group (n = 130)	Comparison group			P
		Total (n = 107)	RA (n = 55)	Other (n = 52)	
<b>Demographic data</b>					
Age (years), mean $\pm$ SE†‡	62 $\pm$ 1	47 $\pm$ 2	51 $\pm$ 2	44 $\pm$ 2	<0.001
Race (white:black)	2.1:1	4.3:1	4.4:1	4.1:1	<0.05
Sex (% male)	24	31	25	37	—
<b>Remote history</b>					
Previous diagnosis of OA	77	8	6	10	<0.001
History of significant trauma	18	11	2	20	—
<b>Joint distribution</b>					
Affected knee (right:left)	1.5:1	1.5:1	1.5:1	1.5:1	—
Bilateral symptoms	80	67	84	50	<0.05
Symptoms in other joints	50	56	93	17	—
Hands	30	51	87	12	<0.001
Hips	25	24	39	8	—
Back	26	10	11	8	<0.01
Feet	18	50	87	10	<0.001
<b>Pain characteristics</b>					
Sudden onset	29	33	21	45	—
Frequency during previous 30 days (no. days)	24	21	24	18	—
On weight-bearing	95	81	85	77	<0.01
Reduced with rest	65	62	67	58	—
Severity (0–3+) (mean $\pm$ SE)	1.8 $\pm$ 1	1.7 $\pm$ 1	1.9 $\pm$ 1	1.5 $\pm$ 1	—
<b>Other findings</b>					
History of swelling	70	75	87	62	—
Morning stiffness†‡	74	73	96	48	—
None	26	27	4	52	—
1–15 minutes	41	16	9	23	—
16–30 minutes	22	5	5	4	—
>30 minutes	12	52	82	21	—
Click on motion‡	62	35	32	37	<0.001
Buckling with activity	54	36	35	38	<0.01
Impairment of function (global)	71	49	62	35	<0.001
Walking	86	71	85	56	<0.01
Climbing stairs	95	78	91	65	<0.001
Rising from seated position	91	70	83	56	<0.001
Locking with activity‡	30	11	9	13	<0.001
Need for ambulatory aids	34	24	41	6	—
Benefit from NSAIDs	84	72	87	54	<0.05

\* Values are percentages except where noted otherwise. P values are for OA group versus comparison group, assessed by *t*-test or chi-square as appropriate. OA = osteoarthritis; RA = rheumatoid arthritis; NSAIDs = nonsteroidal antiinflammatory drugs.

† Included in final traditional analysis.

‡ Included in recursive partitioning analysis.

patients with OA than in the comparison groups. Care was taken to differentiate bony enlargement and paraarticular muscle atrophy.

Range of motion of the knee was reduced in 34% of patients with OA, 37% of patients with RA, and 25% of patients in the non-RA comparison group. Reduced function was defined as a loss of at least 4° of extension or no greater than 100° of flexion. Crepitus was demonstrated slightly more often by active motion, such as squatting, than by passive motion, performed by the examiner. The examiners attempted to distinguish synovial crepitus from bony crepitus, by clinical judgment. However, the examiners thought that tibiofemoral crepitus could not be readily sepa-

rated when patellofemoral crepitus was present; therefore, only bony crepitus was recorded. A surprising number of patients with RA (71%) had interphalangeal changes of osteoarthritis.

As might be expected, the Westergren erythrocyte sedimentation rate (ESR) was higher in patients with RA (Table 5). Serum RF was present (titer >1:40 by latex agglutination test) in 5% of patients with OA. One patient had a titer of 1:20; 4 patients had titers of >1:80 (data not shown). The latter 4 OA patients were from a single contributing center, and there was no apparent explanation for such a result (e.g., other connective tissue disease, chronic infection, age, etc.).

The results of only 62 SF analyses were re-



**Table 4.** Physical examination findings (frequency) in 176 patients with knee pain\*

	OA group (n = 116)	Comparison group			P
		Total (n = 50)	RA (n = 55)	Other (n = 52)	
Signs of joint inflammation					
Erythema	5	20	26	13	—
Palpable increase in temperature†‡	14	48	67	27	<0.001
Palpable effusion†	43	59	80	37	<0.05
Just appreciated	28	12	17	8	
Over joint contours	15	46	63	30	<0.001
Synovial tenderness					
Bony tenderness†‡	24	58	78	38	<0.001
Change in structure or function	63	18	17	20	<0.001
Alignment†					
Normal	53	76	65	87	—
Valgus	24	17	25	10	<0.05
Varus	22	7	10	4	<0.001
Limp	54	43	49	37	—
Bony enlargement†‡	55	5	8	2	<0.001
Range of motion (degrees)					
Extension (mean ± SE)	5 ± 2	9 ± 3	9 ± 4	9 ± 6	<0.05
Flexion (mean ± SE)	122 ± 3	130 ± 5	128 ± 3	132 ± 5	<0.05
Instability					
Mediolateral	36	26	33	19	—
Anteroposterior	16	16	19	14	—
Other					
Crepitus					
Passive motion	89	42	50	33	<0.001
Active motion	89	37	42	32	<0.001
Interphalangeal OA	49	39	71	8	—

\* Values are percentages except where noted otherwise. P values are for OA group versus comparison group, assessed by *t*-test or chi-square as appropriate. OA = osteoarthritis; RA = rheumatoid arthritis.

† Included in final traditional analysis.

‡ Included in recursive partitioning analysis.

ported. When SF was obtained, the volume was similar in all groups. However, the tendency of OA fluid to be less inflammatory was demonstrated by its frequent clear appearance, high viscosity, low white blood cell count, and predominance of mononuclear cells.

Table 6 shows the radiographic findings in the patients. A comparison of the results of the radiographic evaluation from each center coordinator and of the evaluation by the musculoskeletal radiologist (WM) is depicted in Table 7. The interpretation most concordant between the center and the radiologist was that of osteophytes, particularly medial compartment osteophytes. The presence of osteophytes seemed to best differentiate OA from non-OA (Table 6). Although common in OA patients, radiographic evidence of narrowing of the joint space of 1 or more compartments of the knee was present in 62% of RA patients, and thus did not distinguish these groups. Combining medial and lateral osteophytes improved sensitivity without altering specificity. Combinations of the different radiographic findings provided no higher speci-

ficity for OA than did the finding of osteophytes as a single indicator.

**Interpretation of survey.** It was believed that no single set of classification criteria could satisfy all circumstances to which the criteria for OA of the knee would be applied. For that reason, the subcommittee elected to design separate sets of classification criteria that might be utilized under different circumstances:

1. Clinical examination and laboratory tests. May be useful for clinical reports not involving radiographs (e.g., screening in an office practice). The results of at least 1 laboratory test (RF, SF OA, or ESR) must be available.

2. Clinical examination, laboratory tests, and radiographs. May be useful for clinical trials in which radiographs are routinely obtained. This set of criteria is expected to be used most commonly.

3. Clinical examination alone. May be useful for population survey (e.g., a nursing home) or other epidemiologic studies.



**Table 5.** Laboratory findings (frequency) in patients with knee pain\*

	OA group	Comparison group			P
		Total	RA	Other	
ESR (mm/hour)†	25 ± 2 (86)	37 ± 3 (75)	47 ± 4 (42)	24 ± 5 (33)	<0.05
Serum rheumatoid factor	74 (74)	64 (68)	73 (40)	54 (28)	
Negative or 1:20	95 (70)	51 (35)	20 (8)	96 (27)	<0.001
Positive at ≥1:40‡	5 (4)	49 (33)	80 (32)	4 (1)	<0.001
Synovial fluid†					
Volume (ml)	20 ± 5 (30)	27 ± 5 (32)	27 ± 7 (21)	28 ± 6 (11)	<0.05
Clear	83	27	20	40	<0.001
Viscous	77	32	28	43	<0.001
Total WBC/mm <sup>3</sup>	942 ± 298	17,396 ± 3,645	18,293 ± 4,281	16,051 ± 6,695	<0.001
(mean ± SE)					
Polymorphonuclear cells	37	68	73	61	<0.01
Mononuclear cells	47	24	23	25	

\* Values are percentages except where noted otherwise. Values in parentheses are numbers of subjects tested. *P* values are for OA group versus comparison group, assessed by *t*-test. OA = osteoarthritis; RA = rheumatoid arthritis; ESR = erythrocyte sedimentation rate (Westergren); WBC = white blood cell count.

† Included in final traditional and recursive partitioning analyses.

‡ Those with titers <1:40 are not presented.

Criteria for classification derived by the traditional method are provided in Table 8. For classification criteria involving the clinical examination and laboratory results, knee pain plus at least 5 of the following 9 clinical or laboratory findings should be present: age >50 years, morning stiffness <30 minutes duration, crepitus on active motion, tenderness of the bony margins of the joint, bony enlargement noted on examination, a lack of palpable synovial warmth, ESR <40 mm/hour, a negative or low RF titer (<1:40), and

synovial fluid suggestive of OA (viscous, clear, and/or white blood cell count <2,000 cells/mm<sup>3</sup>). To be included in this classification, subjects must have had at least 1 laboratory test performed.

Using the results of clinical examination and radiography (Table 8), the most sensitive and specific combination identified included that of knee pain, osteophytes, and 1 of the following: age >50 years, morning stiffness <30 minutes duration, or crepitus on active motion of the knee (i.e., with weight-bearing, such as squatting).

For classification criteria involving results of the clinical examination only (Table 8), the patient should have knee pain plus at least 3 of the following 6 clinical findings: age >50 years, morning stiffness <30 minutes duration, crepitus on active motion, tenderness of the bony margins of the joint, bony enlargement noted on examination, and a lack of palpable warmth of the synovium.

The difficulty in classification of OA of the knee by the above schema is emphasized by the fact that in each instance, increased sensitivity is achieved at the sacrifice of specificity. For this reason, an algorithm was developed, by recursive partitioning (see Methods), for each of the classification groups (Figures 1–3). These algorithms include variables that appear in the criteria developed by traditional means, but they are quite different. For example, the criteria for SF OA include 3 findings: clear appearance, viscous quality, and white blood cell count <2,000 cells/mm<sup>3</sup>. The traditional testing required only 1 of those findings. Recursive partitioning requires 1, 2, or 3 of these findings at levels specified below.

**Table 6.** Radiographic findings (frequency) in patients with knee pain\*

	OA group (n = 130)	Comparison group			P
		Total (n = 107)	RA (n = 55)	Other (n = 52)	
Osteophytes (spurs)	91	17	20	13	<0.001
Subchondral sclerosis	80	32	42	21	<0.001
Subchondral cysts	38	13	20	6	<0.05
Joint space narrowing	84	41	62	19	<0.001
Loss of bone stock (attrition)	55	17	29	4	<0.001
Malalignment	63	20	25	13	<0.001
Combined criteria†	89	34	56	10	<0.001

\* Radiographic findings of the medial and/or lateral tibiofemoral joint. Values are percentages. *P* values are for OA group versus comparison group, assessed by *t*-test. All 6 categories were included in traditional and recursive partitioning analysis. OA = osteoarthritis; RA = rheumatoid arthritis.

† See reference 1.



In Figures 1–3, surrogate variables are available in several areas of the tree. Though not equal in diagnostic value to the primary variable, they can be used in the absence of a primary variable. If the surrogate of age is used high in the tree, then lower in the tree, the variable of age has already been decided.

Recursive partitioning, as in Figure 1, allowed the distinction of patients with OA from those in comparison groups, when the results of the clinical examination and at least 1 laboratory test were available. That is, OA could be diagnosed by any of the following 3 combinations of variables or surrogate variables, according to their position on the tree: 1) knee pain with 1 or more SF findings of OA (surrogate: age  $\geq 50$  years), a positive RF titer, and all 3 SF findings of OA (surrogate: ESR  $\leq 14$  mm/hour); 2) knee pain with at least 1 SF finding of OA, negative RF, crepitus, and age  $\geq 43$  years; 3) knee pain with at least 1 SF finding of OA, negative RF, no crepitus, all 3 SF findings of OA, and bony enlargement. This set of criteria yielded 88% sensitivity and 93% specificity.

Classification by recursive partitioning, utilizing results of clinical examination, laboratory tests, and radiography, is shown in Figure 2. The classifica-

**Table 7.** Frequency of radiographic findings, as interpreted by the contributing center coordinator and the study radiologist\*

	OA group		Comparison group		P	
	CC	SR	CC	SR	CC	SR
Osteophytes						
Medial	75	75	8	9	<0.001	<0.001
Lateral	73	62	7	6	<0.001	<0.001
Patellar	89	81	14	13	<0.001	<0.001
Subchondral sclerosis						
Medial	65	47	24	14	<0.001	<0.001
Lateral	51	19	15	14	<0.001	<0.05
Patellar	49	11	11	3	<0.001	—
Subchondral cysts						
Medial	29	4	7	6	<0.001	—
Lateral	26	2	8	4	<0.001	—
Patellar	20	4	13	2	<0.001	—
Joint space narrowing						
Medial	74	65	31	32	<0.001	<0.001
Lateral	41	25	21	25	<0.01	—
Patellar	67	24	33	10	<0.001	<0.001
Loss of bone stock (attrition)						
Medial	42	27	16	6	<0.001	0.001
Lateral	21	9	12	5	—	—
Effusion	27	33	39	39	—	—
Sharpened tibial spine	57	56	16	25	<0.001	<0.001
Malalignment	55	43	18	4	<0.001	<0.001

\* Values are percentages. P values are for OA group versus comparison group, assessed by *t*-test. OA = osteoarthritis; CC = center coordinator; SR = study radiologist.

**Table 8.** Criteria for classification of idiopathic osteoarthritis (OA) of the knee\*

Clinical and laboratory	Clinical and radiographic	Clinical†
Knee pain + at least 5 of 9: Age >50 years Stiffness <30 minutes Crepitus Bony tenderness Bony enlargement No palpable warmth ESR <40 mm/hour RF <1:40 SF OA	Knee pain + at least 1 of 3: Age >50 years Stiffness <30 minutes Crepitus + Osteophytes	Knee pain + at least 3 of 6: Age >50 years Stiffness <30 minutes Crepitus Bony tenderness Bony enlargement No palpable warmth
92% sensitive 75% specific	91% sensitive 86% specific	95% sensitive 69% specific

\* ESR = erythrocyte sedimentation rate (Westergren); RF = rheumatoid factor; SF OA = synovial fluid signs of OA (clear, viscous, or white blood cell count <2,000/mm<sup>3</sup>).

† Alternative for the clinical category would be 4 of 6, which is 84% sensitive and 89% specific.

tion of OA can be made with 83% sensitivity and 93% specificity, if the patient has knee pain and the knee radiograph demonstrates osteophytes. If knee pain is present but osteophytes are absent, the classification of OA can still be made if the patient has 2 of 3 SF findings of OA (surrogate: age  $\geq 40$  years), plus morning stiffness  $\leq 30$  minutes duration, plus crepitus on active motion of the knee. This combination demonstrated 94% sensitivity and 88% specificity.

With recursive partitioning, clinical examination separated OA patients from the comparison groups in 3 ways (Figure 3). The classification of OA of the knee could be made in the presence of knee pain without crepitus on motion (surrogate: age  $\leq 38$  years) if the patient had bony enlargement on examination. Alternately, crepitus on active motion, morning stiffness >30 minutes duration, and bony enlargement could be used. A third method included crepitus on active motion, morning stiffness  $\leq 30$  minutes duration, and age  $\geq 38$  years. Overall, in this model, the algorithm was 89% sensitive and 88% specific.

Validity testing of the developed criteria in a clinical setting, using 50 patients with other rheumatic causes of knee pain, yielded a sensitivity of over 90% for the traditional and the tree methods. Specificity ranged from 60–85% for the traditional method, versus 71–84% for the tree method; the results support the greater value of the latter technique.

The subcommittee recommends use of the clas-



sification tree algorithms because: 1) recursive partitioning is the more appropriate method of classification (see Discussion), 2) the algorithms give high sensitivity and specificity, and 3) cross-validated sensitivity and specificity rates for the classification trees were considerably higher than were those obtained for the traditional method of analysis via a test sample. In addition, they allow the inclusion of patients with early OA in trials from which such patients would previously have been excluded because of the lack of radiographic changes of OA.

## DISCUSSION

There are many more articles about the deficiencies of clinical classification criteria than there are articles about ways to improve them. These classification criteria are not designed—and should not be used—to replace the clinical diagnosis of OA. Instead, these criteria are valuable in reporting series of cases, because they assure consistency and improve communication. These proposed criteria are not perfect and will exclude some patients with OA, as well as include some who do not have OA. Such criteria are under constant criticism for their “imperfection” and will probably require refinement in the future.

The development of criteria for the classification of OA presents a particular problem because of the nonspecific nature of the disease, the high proportion of asymptomatic patients, and the lack of a “diagnostic test.” Interpretation of pathologic specimens would be an excellent supplement to the definition of OA; however, biopsy is obviously impractical in most patients. Arthroscopy of the knee is a procedure that allows direct visualization of knee cartilage, and the subcommittee will address the classification of knee OA by this newer technique in the future. Radiography remains the most available method of detecting changes in articular cartilage (joint space narrowing) and tissue reaction about the joint (synovial effusions, bony changes at joint margin, alterations of subchondral bone). However, radiographs do not define the clinical syndrome(s) since 40% of patients with radiographic changes of OA are asymptomatic (17).

The traditional methods of deriving criteria have 3 major deficiencies: 1) subsets of criteria of more (or less) importance are given equal weight; 2) some subsets that classify patients are not represented in the study sample; and 3) it is often difficult to know how to classify subjects when data are missing.

The problem with giving equal weight to all subsets of criteria can be illustrated. Thirty-three of

107 controls were incorrectly classified as having OA of the knee, using the clinical “at least 3 out of 6” rule; this results in 69% specificity. Consider the 10 patients who satisfied only the 3 criteria of age >50 years, presence of crepitus on motion, and no palpable increase in temperature. In this subgroup, 7 were incorrectly classified as having OA (30% specificity). In contrast, among the 42 patients who met only the 5 criteria of age >50 years, morning stiffness <30 minutes duration, presence of crepitus on motion, presence of bony enlargement, and no palpable increase in temperature, none were incorrectly classified as having OA (100% specificity). Various combinations of potential criteria differed considerably in their specificity.

There are 42 distinctly different ways to choose “at least 3 out of 6” criteria. An example of a subset with no observed cases within this study sample is the subset having only the 3 criteria of crepitus on motion, morning stiffness <30 minutes duration, and bony enlargement. Nevertheless, the method leads one to classify such subjects as having OA of the knee, on clinical grounds, even though there are no data to support this decision.

Among the 182 subjects with at least 1 laboratory value defined, 7 were missing at least 3 of the 9 clinical plus laboratory values (6 subjects had 6 defined values, and 1 subject had 4 defined values). These subjects had a defined value for either SF OA or ESR, and none had an RF test result recorded. Three of the subjects met 5 or more criteria; the other 4 subjects met 4 criteria. Using traditional methods, how should one classify these subjects? The “5 out of 9” rule would classify 3 subjects as having OA and the other 4 as not having OA. Since these subjects have only 1 laboratory value recorded, and other variate values are missing (out of the list of 9), one could argue that they should be classified by the clinical “3 out of 6” rule. In this case, all 7 subjects would be correctly classified as having OA. The classification tree algorithms circumvent this difficulty, by use of surrogate splits. In fact, all 7 subjects were correctly classified as having OA of the knee, using either the clinical only tree or the clinical plus laboratory tree. It should be emphasized that in this survey, knee pain was first differentiated from paraarticular pain, such as that due to anserine bursitis or spinal radiculopathy. Also, in contrast to acute knee pain, pain had been present for most days of the preceding month, i.e., 79% of those with OA had pain for  $\geq 14$  days.

Although these classification criteria were not developed in a large population, their application in a



population survey appears to be a logical extension of their development and an appropriate means of validation.

It is suspected that those parameters which measure response to a therapeutic program differ from those used for classification. A search for therapy-directed parameters may become an objective of this subcommittee in the future.

These criteria are directed toward the classification of idiopathic OA of the knee. They appear to detect secondary OA, but they do not appear to separate idiopathic OA from secondary OA.

The classification criteria presented here are a product of clinical consensus and analysis of a prospectively gathered data set. Further refinement must await results of other clinical and experimental research and independent testing of these proposed criteria.

Criteria for the classification of OA of the hip and for OA of the hand are being developed. At a later date, it may become evident that there are features common to OA of these 3 sites, and that would allow the development and use of 1 set of classification criteria. However, we believe the unique features of these 3 sets of joints will require different sets of criteria.

## ACKNOWLEDGMENTS

The subcommittee would like to express sincere gratitude to James F. Fries, MD and Byron W. Brown, PhD of ARAMIS and Stanford University for their guidance in study design, and to Rose Bauman for coordination of the program and preparation of the manuscript.

## REFERENCES

1. Kellgren JH, Lawrence JS: Radiological assessment of osteoarthritis. *Ann Rheum Dis* 16:494-501, 1957
2. The Epidemiology of Chronic Rheumatism, Atlas of Standard Radiographs. Vol. 2. Oxford, Blackwell Scientific, 1963
3. Lequesne M: La coxarthrose: criteres de diagnostic, etiologie sur 200 cas, role de la dysplasie congenitale, *Epidemiology of Osteoarthritis*. Edited by J. Peyron. Paris, Ciba-Geigy, 1981, pp 198-210
4. Lequesne M: Clinical features, diagnostic criteria, functional assessments and radiological classifications of osteoarthritis (excluding the spine). *Rheumatology* 7:1-10, 1982
5. Altman RD, Meenan RF, Hochberg MC, Bole GG Jr, Brandt K, Cooke TDV, Greenwald RA, Howell DS, Kaplan D, Koopman WJ, Mankin H, Mikkelsen WM, Moskowitz R, Sokoloff L: An approach to developing criteria for the clinical diagnosis and classification of osteoarthritis: a status report of the American Rheumatism Association Diagnostic Subcommittee on Osteoarthritis. *J Rheumatol* 10:180-183, 1983
6. Howell DS: Etiopathogenesis of osteoarthritis, *Osteoarthritis: Diagnosis and Management*. Edited by RW Moskowitz, DS Howell, VM Goldberg, HJ Mankin. Philadelphia, WB Saunders, 1984, pp 139-142
7. Brandt KD: Osteoarthritis: clinical patterns and pathology, *Textbook of Rheumatology*. Second edition. Edited by WN Kelley, ED Harris Jr, S Ruddy, CB Sledge. Philadelphia, WB Saunders, 1985, pp 1432-1448
8. Moskowitz RW: Clinical and laboratory findings in osteoarthritis, *Arthritis and Allied Conditions*. Tenth edition. Edited by DJ McCarty. Philadelphia, Lea & Febiger, 1985, pp 1408-1432
9. Mitchell NS, Cruess RL: Classification of degenerative arthritis. *Can Med Assoc* 117:763-765, 1977
10. Milholland AV, Wheeler SG, Heieck JJ: Medical assessment by a Delphi group opinion technique. *N Engl J Med* 288:1272-1275, 1973
11. Dalkey NC: A Delphi study of factors affecting the quality of life, *The Delphi Method: Techniques and Application*. Edited by HA Linstone, M Turoff. Reading, MA, Addison Wesley, 1975, pp 387-398
12. Subcommittee for Scleroderma Criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee: Preliminary criteria for the classification of systemic sclerosis (scleroderma). *Arthritis Rheum* 23:581-590, 1980
13. Willkens RF, Arnett FC, Bitter T, Calin A, Fisher L, Ford DK, Good AE, Masi AT: Reiter's syndrome: evaluation of preliminary criteria for definite disease. *Arthritis Rheum* 24:844-849, 1981
14. Tan EM, Cohen AS, Fries JF, Masi AT, McShane DJ, Rothfield NF, Schaller JG, Talal N, Winchester RJ: The 1982 revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum* 25:1271-1277, 1982
15. Dixon WJ, Brown MB, Engelman L, Frane JW, Hill MA, Jennrich RI, Toporek JD, editors: *BMDP Statistical Software*. Berkeley, University of California Press, 1983, pp 152-156, 290-330
16. Breiman L, Friedman JH, Olshen RA, Stone CJ: *Classification and Regression Trees*. Belmont, CA, Wadsworth, 1984
17. Vital Health Statistics: Prevalence of osteoarthritis in adults by age, sex, race, and geographic area, United States 1960-1962, US Department of Health, Education, and Welfare publication No. 15, Series 11. Government Printing Office, June 1966
18. Kellgren JH, Moore R: Generalized osteoarthritis and Heberden's nodes. *Br Med J* 1:181-187, 1952