Candidate Early Predictors for Progression to Joint Damage in Systemic Juvenile Idiopathic Arthritis

CHRISTY SANDBORG, TYSON H. HOLMES, TZIELAN LEE, KATHLEEN BIEDERMAN, DANIEL A. BLOCH, HELEN EMERY, DEBORAH McCURDY, and ELIZABETH D. MELLINS

ABSTRACT. ***Objective.*** To assess if joint damage at 2 years after diagnosis in patients with systemic juvenile idiopathic arthritis (SJIA) can be predicted by clinical or laboratory features assessed up to 3 or 6 months after diagnosis.

Methods. Medical records from 70 children were retrospectively reviewed. The primary outcome measure was presence of joint damage at 2 years after diagnosis (JD2) as defined by presence of erosions or fusion in one or more joints. Potential predictor variables for JD2 in the first 3 and 6 months after diagnosis consisted of the highest observed white blood cell count, platelet count, erythrocyte sedimentation rate, active joint count, and presence of symptomatic pulmonary or cardiac disease or macrophage activation syndrome, and treatment data.

Results. The outcome of interest, JD2, was identified in 15/70 patients. Classification-tree analysis identified a pair of variables (highest observed platelet count and number of active joints) measured within the first 3 months after diagnosis that together predicted progression to JD2 with an estimated sensitivity of 87%, specificity of 82%, and positive predictive value of 57%. Multivariate logistic regression analyses at 3 months found that higher quantities of joints with active arthritis and early use of methotrexate (MTX) were factors significantly associated with increased odds of progression to JD2 (active joints odds ratio = 1.08, 95% CI 1.00–1.16, p = 0.04; MTX OR = 11.85, 95% CI 1.89–74.26, p = 0.01). Unsupervised cluster analysis identified 2 major phenotypes of patients at 3 months characterized by different ages at onset, acute phase markers, active joint counts, and presence of serositis. These phenotypes differed 3-fold in proportion of subjects progressing to JD2 ($p < 0.05$).

Conclusion. By 3 months after diagnosis, a clinical phenotype based on active joint count and platelet count may be prognostic of an increased risk of progression to JD2. Use of corticosteroids did not appear to change the risk of joint damage. In contrast, the presence of serositis appeared to be associated with decreased risk of joint damage. (First Release Sept 1 2006; J Rheumatol 2006;33:2322–9)

Key Indexing Terms:
SYSTEMIC JUVENILE ARTHRITIS | JOINT DAMAGE | OUTCOME
CLASSIFICATION TREES | PREDICTORS | CLINICAL PHENOTYPE

Systemic juvenile idiopathic arthritis (SJIA) is characterized by uniquely disparate outcomes: from remission with rare relapses and minimal if any residual problems, to a relapsing course with symptom-free intervals, to a progressive and often relentless destructive arthritis with chronic disability[1]. The global longterm outcome of children with SJIA depends upon the degree of growth impairment, psychosocial effects, and severity of disability due to joint damage and arthritis[2]. The key pathophysiologic abnormalities that are responsible for these adverse outcomes are persistent inflammation, joint damage, and secondary treatment side effects.

A major problem in care of children with SJIA is the inability to reliably predict with adequate specificity and sensitivity which patients require more aggressive therapy early in the disease course in order to prevent poor outcome. Three studies have evaluated the relationship of clinical measures at 6 months of disease to the outcomes of active erosive arthritis or functional disability at 2 or more years after onset[3-5]. These studies identified systemic features such as persistent fever, corticosteroid therapy, thrombocytosis, lymphadenopathy, and activity and distribution of arthritis, as possible predictors of poor outcomes. We sought to confirm and expand these findings, including whether the presence of more severe SJIA manifestations such as symptomatic cardiopulmonary serosi-

From the Department of Pediatrics and Department of Health Research and Policy (Biostatistics), Stanford University School of Medicine, Stanford, California; Department of Pediatrics, University of Washington Medical School, Seattle, Washington; and Department of Pediatrics, UCLA School of Medicine, Los Angeles, California, USA.

Supported by grants from the Arthritis Foundation, The Wasie Foundation, Lucile Packard Foundation for Children's Health, Oxnard Foundation, and Valley Foundation.

C.I. Sandborg, MD, Department of Pediatrics; T.H. Holmes, PhD, Department of Health Research and Policy (Biostatistics); T. Lee, MD, Department of Pediatrics; K. Biederman, RN, Department of Pediatrics; D.A. Bloch, PhD, Department of Health Research and Policy (Biostatistics), Stanford University School of Medicine; H. Emery, MD, Department of Pediatrics, University of Washington Medical School; D. McCurdy, MD, Department of Pediatrics, UCLA School of Medicine; E. Mellins, MD, Department of Pediatrics, Stanford University School of Medicine.

Address reprint requests to Dr. C. Sandborg, Department of Pediatrics, Room G310, Stanford University Medical Center, 300 Pasteur Drive, Stanford CA 94305-5208. E-mail: Sandborg@stanford.edu

Accepted for publication June 1, 2006.

tis or macrophage activation syndrome (MAS) identified a subset with different outcomes and whether accurate prediction can be obtained earlier (i.e., at 3 months of disease). In addition, we specifically evaluated whether early treatment with methotrexate (MTX), other disease modifying antirheumatic drugs (DMARD), or corticosteroids modified outcomes.

The outcome chosen for this study was the presence of severe joint damage at 2 years after diagnosis (JD2), as measured by radiographic evidence of joint destruction (erosions or joint fusion). Radiographic joint damage is the endpoint of greatest biologic and clinical importance because early radiographic changes represent an aggressive disease course[6-8].

The 3 major goals of our study were to identify predictors of severe joint damage early in the disease course (by 3 or 6 months after diagnosis). Specifically, we sought (1) to identify a candidate prognostic model that, if validated, could be used by clinicians in making early treatment decisions and in designing clinical trials; (2) to identify general associations between early features of disease and joint damage not revealed in the prognostic model; and (3) to distinguish major phenotypes of patients early in their disease course that may predict JD2. We relied heavily upon multivariate statistical methods because diseases such as SJIA appear to have complex pathophysiology, with diverse and interdependent clinical and biologic variables.

MATERIALS AND METHODS

Patient population. Medical records of all patients with SJIA (n = 94) followed at the University of California at San Francisco Medical Center and Lucile Salter Packard Children's Hospital at Stanford University between January 1990 and March 2004 were reviewed, retrospectively. Criteria for inclusion in the study included definite diagnosis of SJIA, complete records available from the time of diagnosis, interval between symptom onset and diagnosis of SJIA < 5 months, and at least 2 years of followup after diagnosis. Definite diagnosis of SJIA is defined by the International League of Associations for Rheumatology (ILAR)[9,10] revised criteria for JIA and subtypes as onset before age 16 years of arthritis exceeding 6 weeks' duration, with or preceded by daily fever of at least 2 weeks' duration, and accompanied by one or more of the following: (1) evanescent, non-fixed erythematous rash; (2) generalized lymph node enlargement; (3) hepatomegaly or splenomegaly; or (4) serositis. All analyses were based on time since diagnosis of SJIA. The 24 patients who did not meet inclusion criteria were 4 who had a delay in diagnosis > 5 months, 14 who received their early care elsewhere, 5 without adequate data at 2 years, and 1 patient with comorbidity (serious joint infection) confounding radiological analysis. The excluded patients were similar in age at onset (7.2 yrs) and sex distribution (50% male) to the study cohort, but had an increased incidence of erosive joint disease (47%), found almost exclusively in those patients with late referral to pediatric rheumatology centers.

Data collection and management. Comprehensive clinical information was entered into the Stanford Juvenile Arthritis and Rheumatic Disease Database (JARDD), designed and maintained by the authors. Quality control of data entry included audit of all entered data from source documents. JARDD was queried for the following candidate predictors for analysis: age of onset, lag time between onset of symptoms and diagnosis, sex, presence of symptomatic pulmonary disease, presence of symptomatic cardiac disease, presence of MAS, highest observed white blood cell (WBC) count (cells/mm^3), highest observed platelet count (cells/mm^3), highest observed erythrocyte sedimentation rate (ESR; mm/min), highest observed active joint count, and any use of prednisone (including intravenous forms, PRED), MTX or tumor necrosis factor (TNF) inhibitor. A "highest observed" value was the largest value observed on a measure over the period of interest (i.e., 0–3 months or 0–6 months) and was chosen to reflect the most extreme representation of a given variable. Pulmonary involvement was defined as symptomatic pleuritis, pleural effusion, or pneumonitis confirmed by radiograph. Cardiac involvement was defined as presence of symptomatic pericarditis or myocarditis, confirmed by echocardiogram. MAS was diagnosed based on preliminary criteria for diagnosis[11]: (1) laboratory criteria including decreased platelet count ≤ 262,000, elevated aspartate aminotransferase > 59, decreased WBC count ≤ 4000, and fibrinogen ≤ 250 mg/dl; (2) clinical criteria including evidence of hemorrhage, central nervous system irritability, and hepatomegaly; and (3) histopathologic confirmation of hemophagocytosis. Diagnosis of MAS requires 2 laboratory criteria or 2 or more laboratory and/or clinical criteria, with histologic confirmation required only in doubtful cases. Active joint count was defined as the number of joints observed with swelling or limitation of motion with pain (range 0–67)[12]. Since all patients had fever, and > 90% had other features of SJIA such as rash, lymphadenopathy, and organomegaly, these variables were not included as they would not distinguish between clinical subsets. Persistence of these features was not evaluated as the effect of corticosteroids may mask the feature in certain patients, making the features less reliable as predictors. No data were missing for any of the candidate predictors in this data set.

The outcome variable, joint damage at 2 years, was also obtained from JARDD and was defined as the presence of joint erosions or fusion on plain radiograph in one or more joints. Patients with no clinical evidence of articular disease (no active, limited, or painful joints) at 2 years did not have radiographic examination to minimize patients' exposure to radiation. Patients with active or limited joints had affected joints evaluated radiographically. All radiographs were reviewed by the attending radiology faculty at the 2 institutions. Because of the retrospective design of this study, a portion of the actual radiographs for each patient were no longer available for review, so that subtler progressive changes such as joint space narrowing could not be evaluated. For this reason, presence of clearly defined and accepted bony changes (erosions or fusion) was used as the endpoint.

Statistical analysis. Analyses addressed the study's primary, secondary, and tertiary objectives in 3 steps. Step 1: We used classification trees to estimate that particular subset of our candidate predictors that together best predict progression to JD2. Step 2: Using logistic regression, we estimated if and how each of our candidate predictors was associated with progression to JD2. This second step was included because a candidate variable could be associated with JD2 and yet not be among that best subset of predictors of JD2 identified in step 1. Step 3: Unsupervised cluster analysis allowed us to use our candidate predictors to categorize patients into distinct post-diagnosis phenotypes. We then compared rates of JD2 among these phenotypes to assess if these phenotypes, although derived without reference to the outcome, nevertheless differed in risk of progression to JD2. For all 3 steps, separate analyses were run using candidate predictors measured during 0–3 months or 0–6 months post-diagnosis to identify the earliest yet most significant candidate predictors. Statistical significance is defined by attained significance levels of $p < 0.05$.

Classification trees[13] were employed to identify that combination of candidate predictors that are most predictive of either having or not having JD2. Classification trees are constructed by repeated splits of groups of patients into pairs of descendant subgroups. Beginning with the entire sample, each candidate predictor is tentatively used to split all patients into 2 groups: the particular split that best succeeds in producing a high proportion of patients that progress to JD2 in one group and a low proportion in the other is chosen to serve as the first split. Each of these 2 resultant groups is then separately split by this same procedure. Repeated splitting in this way produces a full tree. The full tree is then "pruned back" to give a final reduced tree of minimum cross-validated misclassification risk. The most descendant subgroups of the final trees (see Figure 1) provide predictions of outcome. Sensitivity (proportion of patients progressing to JD2 that are correctly classified) and

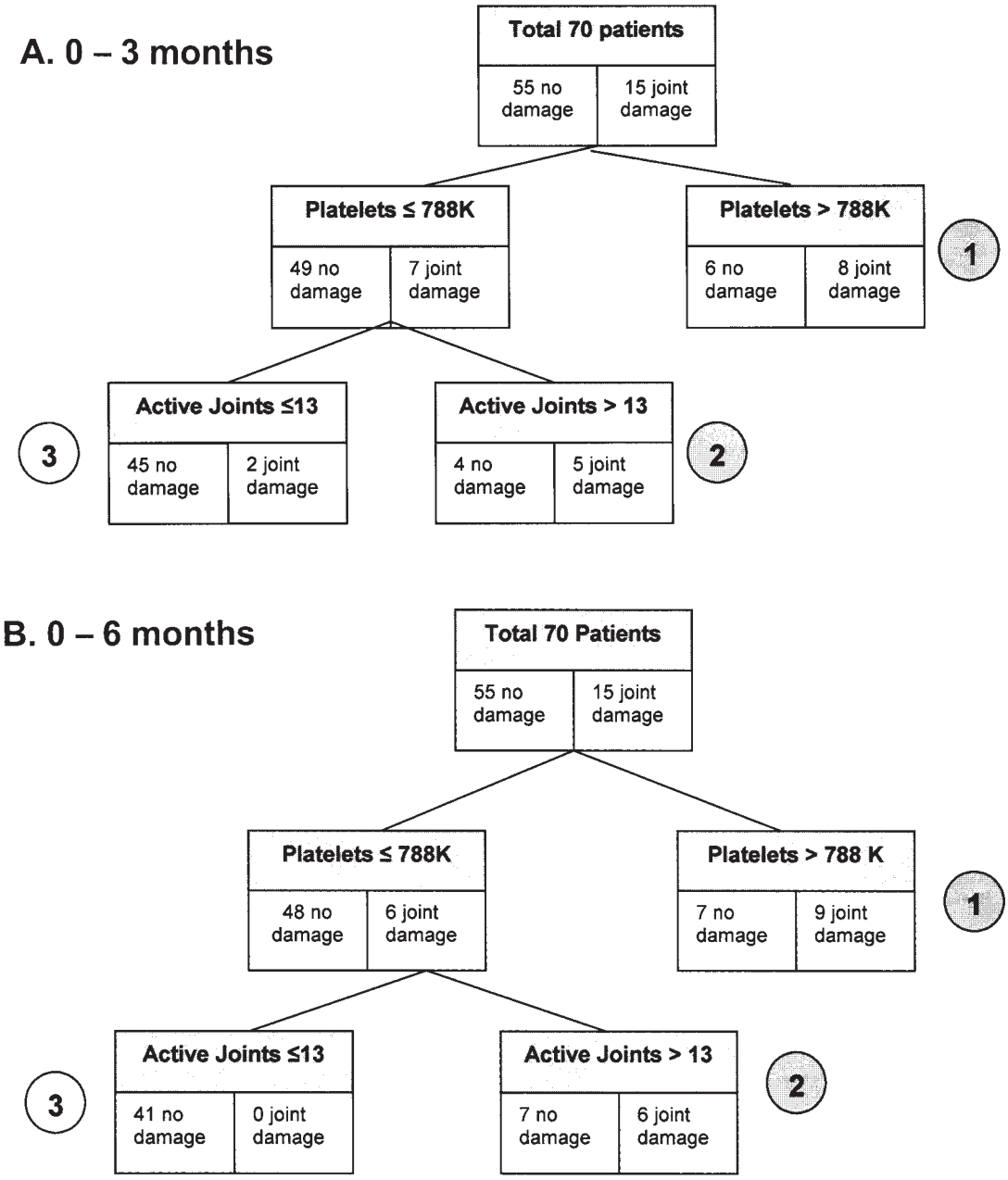

Figure 1. Classification trees for predicting progression to joint damage at 2 years (JD2) from the 0–3 month data (A) and the 0–6 month data (B). White circles denote groups that are predicted not to progress to JD2. Shaded circles denote groups that are predicted to progress to JD2. Numbers of subjects with and without joint damage are shown for each split.

specificity (proportion of patients not progressing to JD2 that are correctly classified) were calculated[14]. Classification trees were constructed using the RPART package[15], which was run in S-Plus® 6.1 for Windows, Professional Edition, Release 1.

For our second objective, logistic regression analyses were used to identify which of the 0–3 month or 0–6 month demographic, clinical, laboratory, and treatment variables had statistically significant associations with JD2. Separate univariate and multivariate logistic regression analyses were used for this purpose, with the demographic, clinical, laboratory, and treatment variables serving as candidate predictors, and presence or absence of JD2 serving as the dependent variable. Univariate regressions consisted of regressing JD2 on each candidate predictor separately, while multivariate regression regressed JD2 on all candidate predictors together in one model.

For our third objective, we used unsupervised cluster analysis to identify subsets of patients of markedly different phenotypes at 0–3 months or 0–6 months, based on all candidate predictors (demographic, clinical, laboratory, and treatment). Separately for each time period (0–3 and 0–6 months), an initial set of several phenotypes was identified using hierarchical cluster analysis[16], which were then reduced to 2 major phenotypes using k-means clustering (k = 2)[17]. To ensure that results were not affected by measurement units,

each variable was standardized in preparation for hierarchical cluster analysis by subtracting the variable's sample mean from each observation and dividing this difference by the variable's sample standard deviation. Rates of JD2 were compared between the 2 phenotypes via continuity-corrected chi-square analyses.

RESULTS

Characteristics of the study sample. A total of 70 patients with SJIA who met study criteria were identified (Table 1). The mean lag time between onset of symptoms and diagnosis of SJIA was 6.8 weeks (SD 6.1). Patients were diagnosed between 1990 and 2004 when MTX was in wide use, but fewer patients had access to TNF inhibitors or other biologics. Because only 1 patient in this cohort started taking a different DMARD other than MTX (sulfasalazine) in the first 6 months after diagnosis, MTX was the only DMARD evaluated in these analyses. Highest observed values for WBC count, ESR, platelet count, and active joint counts were obtained over the 0–3 or 0–6 month periods from an average of 0.8 measurements per patient per month. The variation among subjects of the highest observed WBC count was due to factors other than steroid use as the correlation between WBC count and prednisone or methylprednisolone pulses was weak (r = 0.35). Race/ethnicity is included in Table 1 and is similar to proportions of different racial and ethnic groups in the 0–18-year age group in the Northern California Bay Area at last census, reflecting the diversity of this population. Due to small numbers of subjects in each group, race/ethnicity is not included in any analysis other than the descriptive analyses of Table 1.

The outcome variable (JD2, erosions and/or fusions) was observed in 15 of the 70 patients studied. Radiographic erosions were noted in 14 patients; 12 had multiple joints with erosions (7 bilateral hips, 8 bilateral wrists, then variably, fingers, knees, shoulders) and 2 patients had single joints involved (hip and wrist, respectively). One patient had fusion of the posterior processes of the cervical spine only. One patient had fusion of posterior processes of the cervical spine as well as erosions in other joints. Forty-seven patients (67%) had no evidence of active arthritis or joint damage at 2 years. Forty-five patients who had quiescent disease (no evidence of active arthritis or limitation of motion) and a normal joint examination did not have radiographs. All 23 patients with active arthritis or limitation of motion had radiographs of affected joints at 2 years, except for 2 patients. One patient's family refused radiographs because the patient had been doing well and was undergoing a brief flare of disease that subsequently resolved. The other did not have radiographs until 3 years after diagnosis, and these were normal. Two patients with quiescent disease and no limited joints had normal radiographs. In the 19 patients with active arthritis at 2 years, the mean number of active joints was 8.0 (SD 6.8). To evaluate if the 2-year outcome was an accurate measure of eventual joint damage, we examined a subset of patients in this cohort (N = 34) that were followed for at least 4 years. Baseline characteristics of this subset were not statistically different from the larger cohort (mean age 6.4 yrs, male:female 1:1, highest

Table 1. Demographic, laboratory, clinical, and therapeutic characteristics.

Demographic	Mean (SD) or Percentage	
Age at onset, yrs	6.74 (3.83 SD)	
Interval between symptom onset to diagnosis, weeks	6.8 (6.2 SD)	
Sex	35 male/35 female	
Ethnicity, %		
Caucasian, non-Hispanic	53 (37/70)	
Caucasian, Hipanic	33 (23/70)	
Asian	11 (8/70)	
African American	3 (2/70)	
Clinical	0–3 Months (n = 70)	0–6 Months (n = 70)
Highest observed WBC/mm^3	21.3 (10.3 SD)	22.8 (11.5 SD)
Highest observed platelet count, cells/mm^3	611,000 (193,000 SD)	648,000 (209,000 SD)
Highest observed ESR, mm/min	93 (31 SD)	93 (31 SD)
Highest observed active joint count	7.7 (10.3 SD)	10.5 (12.6 SD)
Pulmonary disease, %	17 (12/70)	17 (12/70)
Cardiac disease, %	11 (8/70)	11 (8/70)
MAS, %	10 (7/70)	10 (7/70)
Taking prednisone, any dose, %	54 (38/70)	61 (43/70)
Taking MTX, %	23 (16/70)	37 (26/70)
Taking TNF inhibitors, %	1 (1/70)	6 (4/70)
Joint outcomes at 2 years after diagnosis		
No active joint counts or joint limitation, %		67 (47/70)
Active joint count > 0, %		27 (19/70)
Active joint count and joint damage, %		16 (11/70)
Joint damage, %		21 (15/70)

WBC: white blood cell count, ESR: erythrocyte sedimentation rate, MAS: macrophage activation syndrome, TNF: tumor necrosis factor.

platelet count 633,000, highest active joint count 11.7). Among these, all 14 patients with joint damage at 4 years already had radiographic evidence of joint damage by 2 years. Of the remaining 22 patients, 11 were in remission, and the remainder had no bony changes on radiographs. This suggests that in this cohort, JD2 was not overlooked in patients with normal joint examinations at 2 years, even though radiographs were not obtained. No patient with one or more active joints and normal radiographs at 2 years subsequently developed joint damage at 4 years.

Exploratory analysis of prognostic variables to identify subsets of patients that differ in rates of progression to JD2. The classification-tree algorithm seeks to identify the combination of candidate predictors that best predict outcome. Classification trees were generated using all 12 original candidate predictors as potential predictors. The 2 trees, one for 0–3 and one for 0–6 months, that gave the lowest cross-validated misclassification risk are shown in Figure 1. The sequence of splits that produced each tree reads from the top to the bottom in each panel of the figure.

Splitting begins with 70 patients, 55 with no JD2 and 15 with JD2 for the 0–3 month tree (Figure 1A) and for the 0–6 month tree (Figure 1B). For the 0–3 and the 0–6 month trees, 3 prediction groups (denoted by numbers 1 through 3) are identified based on just 2 candidate predictors, highest observed platelet count and highest observed active joint count. Using this algorithm, 13/15 patients who developed JD2 were correctly predicted at 3 months after diagnosis. Groups labeled by shaded circles are predicted to progress to JD2, while groups labeled by white circles are not. The first split is at a highest observed platelet count of 788,000 cells/mm^3. Patients with platelet counts > 788,000/mm^3 over 0–3 months (Group 1) are predicted to progress to JD2, which was true for 8 out of 14 patients in this data set (6 patients were incorrectly classified as progressing to JD2, but did not). Patients with highest observed platelet counts ≤ 788,000/mm^3 are split again at a highest observed active joint count of 13. Those with joint counts > 13 and a highest observed platelet count not exceeding 788,000/mm^3 are predicted to progress to JD2 (Group 2), as was seen in 5 out of 9 patients (4 patients were classified incorrectly). Patients assigned to Group 3 were predicted to not progress to JD2, which was true for 45 of 47 patients (2 patients who did progress to joint damage were included in this group). The tree predicts that patients with highest observed platelet counts ≤ 788,000/mm^3 will not proceed to JD2 if they have no more than 13 affected joints. In terms of overall predictive performance of the 0–3 month tree, 87% of those patients with JD2 were correctly classified to progress to JD2 (sensitivity) and 82% of those without JD2 were correctly classified to not progress to JD2 (specificity). While positive predictive value (PPV) was only 57%, it was over 2-fold greater than the percentage of patients with JD2 (15/70, 21%), indicating the gain in predictive value due to the tree.

The optimal predictive tree obtained for 0–6 months (Figure 1B) consisted of the same 2 candidate predictors with the same thresholds obtained for 0–3 months. In the 0–6 month tree, all 15 patients who ultimately developed joint damage were correctly classified. Therefore, for both trees, sensitivity was high in that all or almost all patients with an outcome of permanent joint damage were identified. Fourteen patients who did not develop joint damage were incorrectly predicted to develop joint damage, for a specificity of 75% (similar to the 0–3 month tree). PPV was lower, at 52%.

Associations of candidate predictors with outcome JD2. While classification trees can identify a subset of candidate predictors that best predict outcome, they may not reveal all associations between candidate predictors and the outcome. For this reason, logistic regression analyses were used to test for association between each candidate predictor and JD2 in the first 3 months or first 6 months after diagnosis. Although multivariate logistic regression provides a more comprehensive analysis, especially in the presence of interdependent candidate predictors, we include findings of the univariate logistic regression analyses because the finding that a candidate predictor has a statistically significant association with outcome in univariate regression as well as in multivariate regression indicates an association that is more robust to choice of statistical model. Because a large quantity of candidate predictors may lead to a spuriously good fit of a regression model[18], 7 of the 12 candidate predictors were collapsed to 3 new composite candidate predictors for a total of 8, as follows. Subanalysis on candidate predictors alone revealed that 2 of the major manifestations of systemic extraarticular disease — pulmonary and cardiac disease — were correlated with each other; therefore, a single new indicator variable for presence of any severe serositis (cardiac or pulmonary) was used. Similarly, subanalysis revealed that WBC count, ESR, and platelet count were correlated with each other; therefore, a single-count variable was created that assumed the values from 0 to 3 depending on how many of the following criteria were met: ESR > 70 mm/min, platelet count > 500,000 cells/mm^3, or WBC > 20,000 cells/mm^3. Thresholds for WBC count, ESR, and platelet count were set using clinical consensus of the authors prior to examination of the study data. Multivariate regression analysis used all candidate predictors simultaneously in one model.

For both time periods, the number of active joints was significantly associated with an increase in the odds of JD2 ($p \leq 0.04$) in the univariate and the multivariate analyses (Table 2). The estimated odds ratio of 1.06 (95% CI 1.00–1.12) for highest observed quantity of active joints over 0–3 months indicates that the addition of a single active joint increases the odds of progressing to JD2 by 6%, so that increased risk is considerable if multiple joints are active. For example, relative to a patient with no active joints, the odds of progressing to JD2 in a patient with 15 active joints increases by 2.3-fold and with 30 active joints increases by 5.3-fold. In the multi-

Table 2. Univariate and multivariate logistic regression analyses of outcome of joint damage at 2 years (JD2) on early variables present in the first 3 to 6 months of disease. Bold type indicates $p < 0.05$.

	Joint Damage at 2 Years After Onset			
	0–3 Months		0–6 Months	
Candidate Variable	OR (95% CI)	p	OR (95% CI)	p
Univariate				
Cardiac/pulmonary serositis	0.28 (0.04–1.79)	0.17	0.27 (0.04–1.76)	0.17
Highest observed active joint count	**1.06 (1.00–1.12)**	**0.04**	**1.06 (1.02–1.11)**	**0.01**
Highest observed WBC/ESR/platelets	1.68 (0.87–3.27)	0.12	1.84 (0.93–3.64)	0.08
MAS	0.58 (0.06–5.47)	0.63	0.70 (0.07–6.77)	0.75
Prednisone use	1.00 (0.28–3.54)	0.99	1.42 (0.35–5.85)	0.62
MTX or TNF inhibitor use	3.19 (0.96–10.56)	0.06	2.06 (0.71–5.91)	0.18
Age of onset	0.91 (0.77–1.07)	0.23	0.90 (0.76–1.06)	0.22
Male sex	1.56 (0.48–5.07)	0.46	1.62 (0.49–5.28)	0.42
Multivariate				
Cardiac/pulmonary serositis	0.16 (0.02–1.21)	0.08	**0.10 (0.01–0.92)**	**0.04**
Highest observed active joint count	**1.08 (1.00–1.16)**	**0.04**	**1.07 (1.00–1.14)**	**0.04**
Highest observed WBC/ESR/platelets	2.53 (0.90–7.13)	0.08	2.07 (0.76–5.62	0.15
MAS	0.09 (0.003–2.61)	0.16	0.15 (0.01–3.05)	0.21
Prednisone use	0.85 (0.17–4.27)	0.84	0.81 (0.14–4.86)	0.82
MTX or TNF inhibitor use	**11.85 (1.89–74.26)**	**0.01**	3.67 (0.79–16.92)	0.82
Age of onset	0.92 (0.73–1.17)	0.51	0.99 (0.79–1.25)	0.95
Male sex	5.06 (0.80–31.96)	0.08	4.54 (0.82–25.06)	0.08

WBC: white blood cell count, ESR: erythrocyte sedimentation rate, MAS: macrophage activation syndrome, TNF: tumor necrosis factor.

variate regression analyses, the odds of JD2 decreased significantly with the presence of pulmonary and cardiac serositis measured over the 0–6 month period ($p = 0.04$), and the odds of JD2 increased significantly with use of MTX ($p = 0.01$) in the first 3 months, but this association was not detected using the data over 0–6 months. This association is important, because it demonstrates that MTX did not appear to decrease the chance of progressing to joint damage; similarly, no association (positive or negative) with prednisone was seen.

Patient phenotypes based on candidate predictors. A third method used to analyze this cohort was cluster analysis, a technique for identifying groups (phenotypes) of patients based on degree of similarity on multiple attributes (here, demographic, clinical, laboratory and treatment candidate predictors). Results of unsupervised clustering (hierarchical cluster analysis followed by k-means) are given in Table 3. For 0–3 months and for 0–6 months, the most pronounced differences between the major phenotypes were pulmonary involvement, platelet count, use of MTX, age at onset, and quantity of active joints. Although cluster analysis used only early variables and did not utilize the outcome variable, the proportion of patients who progressed to JD2 differed significantly between the 2 major phenotypes for 0–3 months ($p = 0.0454$) and for 0–6 months ($p = 0.0165$), as assessed by chi-square analysis.

DISCUSSION

The results demonstrate that there are characteristic profiles of demographic and early clinical, laboratory, and treatment variables that distinguish groups of patients with SJIA early in their disease course, and that these clinical phenotypes can potentially be used to identify patients at risk for progression to severe joint damage. Introducing treatment variables also allowed us to explore whether certain treatments were associated with protection or progression to severe joint damage. Using 3 different analytic methods, classification trees, regression analyses, and unsupervised cluster analysis, strong themes emerged that were consistent across these diverse methods, indicating the robustness of these conclusions.

The most important finding in this study, using classification-tree analysis, is the potential that a very high proportion of patients who will develop JD2 can be identified in the first 3 months after diagnosis. Among 0–3 month candidate predictors, the combination of highest observed platelet counts and highest observed quantity of active joints correctly classified 13/15 patients who progressed to JD2. This result was robust in that the same combination of candidate predictors over the 6-month period correctly identified all 15 patients. This prediction was sensitive (87%) and specific (82%), but 10/55 patients were incorrectly classified as at risk for developing JD2. The optimal balance of risks between overidentifying at-risk patients and underidentifying patients with JD2 depends on how use of a prognostic tool might change treatment decisions. We and others have found classification-tree analysis is very helpful in treatment decisions and diagnostic algorithms for diseases as diverse as brain cancer, addiction, and rheumatic diseases, as well as areas of study including public health and epidemiology[19-22].

Unsupervised cluster analysis identified 2 major clinical phenotypes based on subjects' early features. Phenotypes dif-

Table 3. Two major phenotypes identified by unsupervised cluster analysis within each time period (0–3 months and 0–6 months).

Phenotype	Mean Age at Onset, yrs	Male, %	Highest Observed WBC	Highest Observed Platelets	Highest Observed ESR	Treated with MTX, %	Treated with Prednisone, %	Treated with TNF, Inhibitor, %	Highest Observed no. of Active Joints	MAS, %	Cardiac, %	Pulm, %	JD2, %
0–3 Months (n = 70)													
1, N = 33	5.7	52	23,700	782,000	99.0	27	61	NA	11.2	12	12	9	33
2, N = 37	7.7	51	19,200	460,000	86.9	19	49	NA	4.6	8	11	24	11
0–6 Months (n = 69)													
1, N = 34	5.1	50	24,600	813,000	101.3	47	71	6	14.1	6	15	9	36
2, N = 35	8.4	51	20,600	482,000	84.4	29	51	6	6.8	11	9	26	9

Pulm: pulmonary, JD2: patients with joint damage 2 years after diagnosis. WBC: white blood cell count, ESR: erythrocyte sedimentation rate, MAS: macrophage activation syndrome, TNF: tumor necrosis factor.

fer on several features, including highest observed quantity of active joints and highest observed platelet counts (Table 3), which were also identified as prognostic by classification trees, as well as age and pulmonary involvement. These phenotypes appear to be prognostic, as they differ in rates of JD2 within each time period (0–3 and 0–6 months), even though the phenotypes were derived without use of JD2 as a variable.

The question of whether early and aggressive therapy might influence progression to JD2 was important in this study. The positive association of MTX usage with increased risk of JD2 (Tables 2 and 3) suggests that the pediatric rheumatologists who were caring for these children correctly identified patients at risk for progression to JD2 early and then initiated early DMARD therapy. These decisions were most likely based on "clinical impression" derived from the same clinical and laboratory features that formed the candidate predictors for this study. If so, then at least a portion of patients at risk for JD2 were correctly identified early and treated with standard aggressive therapy; despite early detection, progression to JD2 still occurred in these patients. This conclusion is consistent with observations by others that SJIA is more resistant to MTX than other types of rheumatoid arthritis[3]. Interestingly, no treatment variable (MTX, prednisone, or TNF inhibitors) possessed a strong enough association with JD2 to be selected as a predictor in the classification-tree analysis.

One major issue in this and other studies of outcome in SJIA is how to define a measurable outcome that is meaningful to patients. We chose to use evidence of the stringent criteria of severe joint damage occurring in the first few years of disease as the primary outcome, because the literature suggests that destructive synovitis is associated with longterm disability in adults and children[6-8]. Distinguishing permanent joint damage from active arthritis and pain is also important from the viewpoint of therapies. Functional status tools such as the instrument most commonly used in children, the Child Health Assessment Questionnaire[23], are more responsive to pain and disease activity in disease and therefore may not be as sensitive in detecting actual joint damage until very late[24].

Three other studies have examined early predictors of outcome in SJIA, and although the measured outcomes were different in each study, clear parallels exist between these reports and our findings[3-5]. These studies suggested that persistent active polyarticular arthritis and thrombocytosis at 6 months were associated with poorer outcomes as defined by erosive arthritis or functional status. Our data may permit earlier evaluation of prognostic risk (at 3 months), and further, suggest that patients with severe systemic disease (as defined by presence of symptomatic pulmonary or cardiac involvement) may be a unique subgroup that is less likely to develop joint damage.

Our data set was employed as an exploratory sample (in contrast to a validation sample) for purposes of the analyses; as such, the estimated sensitivity, specificity, and positive predictive values from the classification-tree analysis are possibly overestimated. Validation of these 3 measures of predictive performance will require application of the classification trees identified here to additional patients in an independent data set. Our results suggest that it may be possible to predict progression to JD2 in SJIA with good specificity and high sensitivity and do so with economy, as only 2 candidate predictors need to be measured. The 3-month tree potentially has greater predictive accuracy than any other model offered in the literature to date. Such findings should facilitate early identification of those patients at risk for more severe disease and allow timely intervention to prevent joint damage. Refinement of our results and the predictive trees using other data sets could assist prognostication and treatment decisions regarding progression to joint damage in SJIA.

ACKNOWLEDGMENT

We thank David Kieffer and Vaishali Tenkale for assistance with database development, and Sara Alexander, Uma Srinivasan, Uma Maikappan, and Sanie Zehra for technical support and data entry.

REFERENCES

1. Lomater C, Gerloni V, Gattinara M, Mazzotti J, Cimaz R, Fantini F. Systemic onset juvenile idiopathic arthritis: a retrospective study of 80 consecutive patients followed for 10 years. J Rheumatol 2000;27:491-6.

2. Sandborg C, Nepom B, Mellins E. Juvenile arthritis. In: Rich R, Fleisher T, Shearer W, Kotzin B, Schroeder B, editors. Clinical immunology. 2nd ed. London: Mosby; 2001.
3. Schneider R, Lang BA, Reilly BJ, et al. Prognostic indicators of joint destruction in systemic-onset juvenile rheumatoid arthritis. J Pediatr 1992;120:200-5.
4. Modesto C, Woo P, Garcia-Consuegra J, et al. Systemic onset juvenile chronic arthritis, polyarticular pattern and hip involvement as markers for a bad prognosis. Clin Exp Rheumatol 2001;19:211-7.
5. Spiegel LR, Schneider R, Lang BA, et al. Early predictors of poor functional outcome in systemic-onset juvenile rheumatoid arthritis: a multicenter cohort study. Arthritis Rheum 2000;43:2402-9.
6. Magni-Manzoni S, Rossi F, Pistorio A, et al. Prognostic factors for radiographic progression, radiographic damage, and disability in juvenile idiopathic arthritis. Arthritis Rheum 2003;48:3509-17.
7. Scott DL, Smith C, Kingsley G. Joint damage and disability in rheumatoid arthritis: an updated systematic review. Clin Exp Rheumatol 2003;21:S20-7.
8. Welsing PM, van Gestel AM, Swinkels HL, Kiemeney LA, van Riel PL. The relationship between disease activity, joint destruction, and functional capacity over the course of rheumatoid arthritis. Arthritis Rheum 2001;44:2009-17.
9. Petty RE, Southwood TR, Baum J, et al. Revision of the proposed classification criteria for juvenile idiopathic arthritis: Durban 1997. J Rheumatol 1998;25:1991-4.
10. Petty RE, Southwood TR, Manners P, et al. International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. J Rheumatol 2004;31:390-2.
11. Ravelli A, Magni-Manzoni S, Pistorio A, et al. Preliminary diagnostic guidelines for macrophage activation syndrome complicating systemic juvenile idiopathic arthritis. J Pediatr 2005;146:598-604.
12. Cassidy JT, Levinson JE, Bass JC, et al. A study of classification criteria for a diagnosis of juvenile rheumatoid arthritis. Arthritis Rheum 1986;29:274-81.
13. Brieman L, Friedman JH, Olshen RA, Stone CJ. Classification and regression trees. Boca Raton, FL: Chapman and Hall/CRC; 1984.
14. Le CT. Applied categorical data analysis. New York: John Wiley & Sons; 1998.
15. Therneau TM. An introduction to recursive partitioning using the RPART routines. Rochester, MN: Mayo Foundation; 1997.
16. Jobson JD. Applied multivariate data analysis. Springer Texts in Statistics. New York: Springer-Verlag; 1991.
17. Hartigan JA, Wong MA. A K-means clustering algorithm. Appl Statist 1979;28:100-8.
18. Harrell FE Jr. Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis. New York: Springer-Verlag; 2001.
19. Elkashef A, Holmes TH, Bloch DA, et al. Retrospective analyses of pooled data from CREST I and CREST II trials for treatment of cocaine dependence. Addiction 2005;100 Suppl 1:91-101.
20. Gaspar LE, Scott C, Murray K, Curran W. Validation of the RTOG recursive partitioning analysis (RPA) classification for brain metastases. Int J Radiat Oncol Biol Phys 2000;47:1001-6.
21. Lemon SC, Roy J, Clark MA, Friedmann PD, Rakowski W. Classification and regression tree analysis in public health: methodological review and comparison with logistic regression. Ann Behav Med 2003;26:172-81.
22. Hunder GG. Classification/diagnostic criteria for GCA/PMR. Clin Exp Rheumatol 2000;18:S4-5.
23. Singh G, Athreya BH, Fries JF, Goldsmith DP. Measurement of health status in children with juvenile rheumatoid arthritis. Arthritis Rheum 1994;37:1761-9.
24. Oen K, Reed M, Malleson PN, et al. Radiologic outcome and its relationship to functional disability in juvenile rheumatoid arthritis. J Rheumatol 2003;30:832-40.