

Development and Validation of a Case Ascertainment Tool for Ankylosing Spondylitis

MICHAEL H. WEISMAN,¹ LAN CHEN,² DANIEL O. CLEGG,³ JOHN C. DAVIS, JR.,⁴
ROBERT W. DUBOIS,⁵ PAMELA E. PRETE,⁶ LAURIE M. SAVAGE,⁷ LAURA SCHAFFER,⁸
MARIA E. SUAREZ-ALMAZOR,⁹ HSING-TING YU,⁵ AND JOHN D. REVEILLE¹⁰

Objective. Ankylosing spondylitis (AS) diagnosis is often delayed. The availability of effective biologic agents for treating AS has increased the importance of early diagnosis. We tested questions derived from a comprehensive literature review and an advisory board in a case-control study designed to identify patients with AS from among patients with chronic back pain (CBP).

Methods. Question items were cognitively tested among patients with AS, and then in case-control studies for validation and creation of a scoring algorithm and question item reduction. AS cases were recruited from a known database, and CBP subjects (controls) were recruited from clinics, employers, and from the SpineUniverse Web site. We used individual question items in a multivariate framework to discriminate between people with and without AS.

Results. Forty-three questions yielded 24 items for analyses; 12 of these were entered into a multivariate regression model. Individual items yielded odds ratios ranging from 0.07 to 30.31. Question items with a significant positive relationship to AS included male sex, neck or hip pain/stiffness, longer pain duration, decreased pain/stiffness with daily physical activity, pain relief within 48 hours of nonsteroidal antiinflammatory drugs, and diagnosis of iritis. The tool demonstrated a sensitivity of 67.4 and a specificity of 94.6. The tool was developed from clinically and radiologically diagnosed AS cases and therefore is designed to distinguish AS cases among CBP subjects. In addition, ~54% of the AS cases in the study were treated with biologic agents, which may impact questionnaire responses.

Conclusion. This tool can identify undiagnosed patients with AS and, potentially, those at an earlier stage in their disease course.

INTRODUCTION

Lower back and neck pain are common (1) and may produce disability, with great societal cost (2). Ankylosing spondylitis (AS), an inflammatory spondylarthritis (SpA), is among the multiple causes of lower back and neck pain and was recently demonstrated to be nearly as prevalent in the general population as rheumatoid arthritis (3,4). Unfortunately, the period between symptom onset and AS

diagnosis is typically lengthy; detection of sacroiliitis by standard radiography may lag 8–11 years after onset of inflammatory back pain (IBP) (5–7). Until recently, available therapies provided only a modest clinical impact. New, more effective biologic therapies can substantially reduce AS-associated pain, disability (8–11), and quality of life impairment (12,13). Such therapies emphasize the need for earlier diagnosis.

Although many people experience chronic back and neck pain, few have AS. Specialist evaluation of all pain

Supported by a research grant from the Spondylitis Association of America, with funding supplied by Centocor, Abbott, Amgen, Wyeth, and Pfizer.

¹Michael H. Weisman, MD: Cedars-Sinai Medical Center, Los Angeles, California; ²Lan Chen, MD, PhD: Penn Presbyterian Medical Center, Philadelphia, Pennsylvania; ³Daniel O. Clegg, MD: University of Utah, Salt Lake City; ⁴John C. Davis, Jr., MD, MPH: University of California, San Francisco; ⁵Robert W. Dubois, MD, PhD, Hsing-Ting Yu, MPH: Cerner LifeSciences, Beverly Hills, California; ⁶Pamela E. Prete, MD, FACP, FACR: Veterans Administration Healthcare System, Long Beach, California; ⁷Laurie M. Savage, BS: Spondylitis Association of America, Sherman Oaks, California; ⁸Laura Schaffer, RN, BSN: Seattle, Washington; ⁹Maria E. Suarez-Almazor, MD, PhD: University of Texas MD Anderson Cancer Center, Houston; ¹⁰John D. Reveille, MD: University of Texas Health Science Center at Houston.

Dr. Weisman has received research grants from Abbott, Centocor, and Wyeth, and has received consultant fees from serving on the advisory/data safety monitoring board for Amgen/Wyeth. Dr. Prete has received consultant fees, speaking fees, and/or honoraria (less than \$10,000) from UCB. Ms Schaffer owns stock in Abbott Laboratories. Dr. Suarez-Almazor has received consultant fees, speaking fees, and/or honoraria (less than \$10,000 each) from BMS, Genentech, and Roche.

Address correspondence to Michael H. Weisman, MD, 8700 Beverly Boulevard, Suite B-200, Los Angeles, CA 90048. E-mail: Weisman@cshs.org.

Submitted for publication March 23, 2009; accepted in revised form August 20, 2009.

patients is neither feasible nor cost effective. To improve case identification, we have developed a patient questionnaire to identify individuals with an increased likelihood of having AS as the etiology of their back pain. If successful, this instrument could identify patients with AS undiagnosed by current methods, providing a cost-effective group for further assessment (14).

MATERIALS AND METHODS

Design overview. The case ascertainment tool development had 3 phases (Figure 1). In phase I, potential question items were identified using a literature review, reviewed and revised by an expert panel, and cognitively tested among a small group of patients known to have AS. The panel included rheumatologists and nurses with clinical, academic, and community backgrounds from various US locations (Appendix A). In phase II, the modified question items were tested in a study and revised if necessary. Phase III was a subsequent case-control study for validation and creation of a scoring algorithm and question item reduction. The study received ethical approval from the Institutional Review Board (IRB) of Cedars-Sinai Medical Center (CSMC), the Western IRB, and The Committee for the Protection of Human Subjects, Office of Research Support Committees, at the University of Texas Health Science Center at Houston.

Literature review for identifying potential question items (phase I). We conducted a comprehensive PubMed literature search for English-language articles reporting symptoms, risk factors, and patient characteristics associated with AS, as well as prior case identification instruments, published from January 1995 to July 2005. Detailed AS search terms included ankylosing spondylitis and spondylarthropathy, as well as diagnosis terms such as diagnosis, diagnostic techniques, procedures, mass screening, and risk factors. Reference lists of relevant publications and review articles were reviewed to identify studies not captured by the PubMed search. References suggested by the expert panel were also reviewed. The abstracts were reviewed for relevance based upon the search strategy. If an abstract indicated that the article might contain relevant primary data or secondary information (e.g., a review article citing results from a population-based study), its full text was reviewed to extract pertinent information. The pool of candidate questions was constructed by domains (e.g., demographics, AS-associated symptoms, family history) and risk factors associated with AS.

Previously developed classifications and diagnostic criteria for AS were identified and evaluated for relevance to this study. The Rome criteria for AS from 1963 (15), the New York criteria for AS from 1968 (16), modified in 1984 (17), and the Mau et al criteria for determining early AS from 1990 (18) all require radiographic sacroiliitis in addition to patient-reported data to confirm a diagnosis of AS. Each instrument contains 3–6 clinical criteria, including duration of low back pain/stiffness, history of iritis, thoracic pain/stiffness, limited motion in the lumbar spine and/or chest expansion, and peripheral arthritis or heel

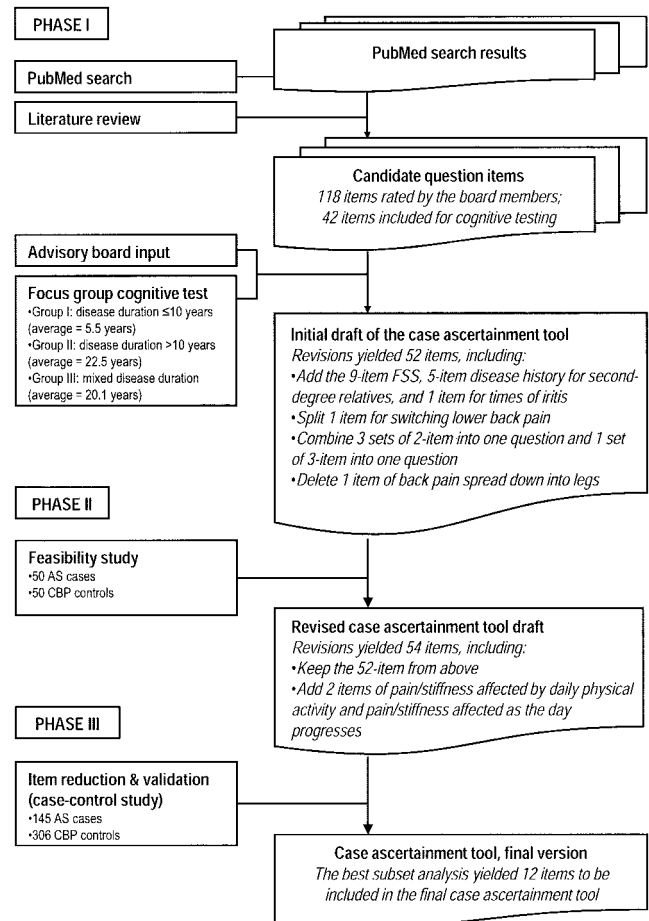


Figure 1. Case ascertainment tool development overview. FSS = Fatigue Severity Scale; AS = ankylosing spondylitis; CBP = chronic back pain.

pain, although the wording of each clinical criterion varied across these instruments. Questions were also identified from literature (e.g., morning stiffness, age at symptom onset, prescription of nonsteroidal antiinflammatory drugs [NSAIDs]) and the Calin et al criteria for AS (19), a 5-item, symptom-based questionnaire for differentiating IBP (i.e., AS) from mechanical or nonspecific types of pain.

Rudwaleit et al published 2 studies, one in 2004 and one in 2006, that proposed criteria for diagnosing axial SpA earlier in patients with IBP without radiographic sacroiliitis (14,20). The first study (14) suggested that the presence of IBP features (4 of the following: age at onset <40 years, duration of back pain >3 months, insidious onset, morning stiffness, and improvement with exercise) increased the probability of axial SpA from the 5% prevalence in patients with chronic back pain (CBP) to 14%, and therefore was chosen as the study entry criterion for identifying patients with axial SpA. Adding 2–3 clinical features (e.g., heel pain, uveitis, positive family history, or positive response to NSAIDs) increased the probability of axial SpA to 90%.

The second study evaluated IBP features and compared various combinations of features as classification and di-

agnostic criteria for AS (20). Several sets of combined clinical features, as opposed to individual features, conferred a better ability to differentiate AS from mechanical low back pain. A set of 4 parameters was identified: morning stiffness of >30 minutes' duration, improvement in back pain with exercise but not with rest, awakening because of back pain during the second half of the night only, and alternating buttock pain. Fulfillment of 2 of these would yield a sensitivity of 70.3% and a specificity of 81.2% with a positive likelihood ratio of 3.7; fulfillment of 3 or 4 would increase the positive likelihood ratio to 12.4.

Question items based on symptoms and risk factors with varying wording were identified from the literature search and reviewed by the expert panel. Each member rated the 118 potential question items for its ability to distinguish between a patient with AS and a patient with other CBP, using a 9-point scale (higher scores indicated greater ability). For decision making, the experts received information on the prevalence/sensitivity of the question item in reference studies, the data sources for each question item, and the number of times a specific question item was identified in the literature. Forty-two question items receiving an average score of ≥ 6 were selected for inclusion in the initial case ascertainment tool for patient cognitive testing.

Setting and participants. *Cognitive testing (phase I).* A cognitive test was conducted to ensure that patients could easily and accurately understand the directions for completing the question items and the meaning of the questions within the tool. Comments regarding the flow of questions and font size and type were obtained. Typographical errors, wording/vocabulary, and meaning/comprehension were discussed for each item to ensure that they accurately captured/described patients' disease experience. Three focus groups of patients with AS ages ≥ 18 years were identified from the Prospective Study of Outcomes in AS database. This database consists of ~ 600 patients with AS with an average disease duration of 11.5 years, all of whom had to fulfill the modified New York criteria for AS (21). The focus groups (convened in Los Angeles) consisted of patients with disease durations of ≤ 10 years, of >10 years, and mixed. A representative proportion of women (25%) was ensured to match the proportion in the database. The expert panel qualitatively reviewed the cognitive testing results for tool revision: modifying wording, reordering question items, and in some cases splitting 1 question item into 2 or combining 2 into 1. Extreme fatigue was mentioned frequently among the focus groups but not included in the original question items cognitively evaluated; therefore, the 9-item Fatigue Severity Scale (22) was added to the instrument. The expert panelists held a teleconference discussion to attain consensus on question item modifications. Revision yielded 52 question items in the tool for testing in the phase II study.

Feasibility study (phase II). This phase pilot tested the question items developed in phase I among 100 patients (50 cases of AS and 50 CBP controls). Patients with AS

were randomly selected from the Prospective Study of Outcomes in AS database (cases), and CBP patients (controls) were identified through e-mail invitations via the CSMC intranet. CBP controls were required to have experienced back pain for ≥ 3 months. The inclusion criteria were as follows: English speakers, ≥ 18 years of age, and not participants in the focus groups in phase I. We attempted to screen patients who already had a physician diagnosis of AS among the CBP controls, but we encountered no cases. A study package including a cover letter, a consent form, a response card, and a paper copy of the case ascertainment tool were mailed to potential participants. Completed tools were mailed to the clinical study coordinator at CSMC, who organized a deidentified database for analysis. We screened and recruited 121 patients, but 18 were nonresponders and 3 were excluded due to poor data quality; 100 were included for phase II analysis. The mean \pm SD age of the 50 included AS cases was 33.4 ± 11 years, the average disease duration was 5.6 years, and $\sim 46\%$ were men. The CBP controls ($n = 50$) were older (mean \pm SD 43.5 ± 11.5 years) with longer disease duration (8.7 years), and fewer were men (42%).

Case-control study (phase III). Although the literature provides no well-established single method for estimating the sample size required for construct validation of a new disease case ascertainment tool, the literature is helpful in identifying reasonable ranges for sample size requirements. We ran simulations using data from 1 validation study on smoking among patients with chronic obstructive pulmonary disease, which estimated that the minimum number of cases that we needed to reproduce the analysis was 100 patients (23). Assuming a 15–25% AS prevalence (24) among our target population of IBP patients, a total of 420 AS and CBP patients was required to provide sufficient sample sizes among case and control groups.

AS cases were identified from the Prospective Study of Outcomes in AS database and from among patient members of the Spondylitis Association of America (via postal mail). The latter were screened using the same criteria (21) as were applied to the Prospective Study of Outcomes in AS cohort. CBP controls were identified from the following sources: intranet notice of the research study available to employees within Cerner Corporation and CSMC, signs and flyers posted at local primary care and chiropractor clinics, and postings on the SpineUniverse Web site (www.spineuniverse.com). SpineUniverse is an online source for patients who have back or neck problems, and our recruitment advertisement was posted on their e-newsletter. Recruitment methodology and patient inclusion/exclusion criteria for phase III were similar to those of phase II, with the addition that participants of this study phase could not have been participants in any previous study phase. As with phase I and II, a written consent was required for a patient to be included in the study. The data were then analyzed to develop and validate the case ascertainment tool's scoring algorithm.

Statistical analyses. Information collected from the focus groups (phase I) was qualitatively analyzed, and mod-

ifications were made, such as adjusting the flow and wording of question items for ease of understanding.

Feasibility study (phase II). A bivariate analysis was conducted to evaluate whether question item responses were significantly different between cases and controls. Categorical and continuous variables were analyzed using Fisher's exact test and the 2-sample *t*-test, respectively. To analyze how sets of question items differentiated between cases and controls, a multivariate logistic regression model was constructed. Best subset regression analysis was employed to reduce the size of the model without significant loss of discriminating power, obtaining a parsimonious set of question items. To ensure that question item responses were consistent and reliable, the tool's internal consistency reliability was calculated using Cronbach's alpha. The consistency of each question item was analyzed using 3 measures: the change in Cronbach's alpha when that question item was removed, the average of the question item's correlations with every other question item, and the correlation between the question item and the sum of all of the other question items.

The expert panel reviewed the bivariate, multivariate, and psychometric results, and their insights guided the ongoing revision of the question items. Although the sample size for phase II was small, there was enough variability in the data to perform a multiple logistic regression and obtain parameter estimates. The decrease in power suggests that small and moderate effects may be undetectable. However, the aim of phase II was not to formally test hypotheses, but to examine possible question items that exhibited larger effects and to qualitatively investigate other question items that may have exhibited a significant effect in the larger phase III study. Because most of the question items were performed in the direction as expected (data are available by request), the expert panel agreed to keep all of the question items to be tested in the phase III case-control study. Additionally, to clarify "pain/stiffness affected by physical activities," the expert panel suggested adding 2 question items: 1) how does daily physical activity affect the pain or stiffness in your lower back or buttocks? and 2) what happens to the amount of pain or stiffness in your lower back or buttocks as the day progresses?

Final case-control study (phase III). To develop and validate the tool's final scoring algorithm, the data were divided into 2 samples. A random sample of 70% of the AS cases and 70% of the CBP controls formed the model development sample and was used to create the final tool and its scoring algorithm, while the remaining 30% formed the model validating sample, used to validate the scoring algorithm. To ensure that the random selection of our development sample and validation sample were balanced, we conducted bivariate analyses comparing the response to each question item between the 2 samples (data are available by request). None of the comparisons showed significant difference, which confirmed the comparability of the development and validation samples. Similar to the feasibility study (phase II), the model development sample was analyzed using bivariate, multivariate, and psychometric analyses. These analyses informed decisions regarding question items to include in the final tool

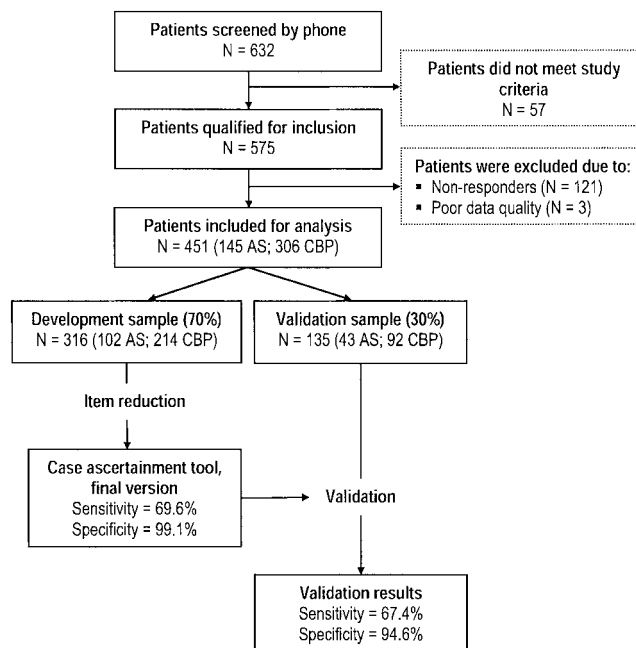


Figure 2. Phase III final case-control study overview. AS = ankylosing spondylitis; CBP = chronic back pain.

algorithm. Motivated by the best subset variable selection results and supplemented with clinical expertise, the expert panel selected a final model, and performance parameters including sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were examined. The receiver operating characteristic (ROC) curves were constructed using the development subsample. Using the regression coefficients of the selected logistic regression model, an algorithm to estimate a patient's predicted probability of having AS was created.

To validate the findings of the developed model, the case ascertainment tool score was calculated for the patients randomized to the model validation sample, and their true AS status was compared with the tool's predicted status. The sensitivity and specificity within this subsample were calculated and compared with the tool's target sensitivity and specificity of 70% and 99%, respectively. To explore the potential bias in our results of having a "contaminated" control group, that is, a CBP control group in which some of the patients may have actually had AS, we performed a sensitivity analysis during the validation phase under the assumption that 5% of the control sample had AS. Each CBP control was randomly reassigned as a patient with AS with a 5% probability, and ROC statistics were calculated on this reassigned sample. This process was repeated 1,000 times, and average ROC statistics were calculated across these 1,000 samples.

All analyses were performed using SAS, version 9.1 (SAS Institute, Cary, NC).

RESULTS

The phase III analysis included 451 people: 145 AS cases and 306 CBP controls (Figure 2). The mean \pm SD age of the

Table 1. Bivariate relationships between candidate variables and study diagnosis (development sample, n = 316)*

Question items and response categories	AS	CBP	P
Sex, male	61.8	37.4	< 0.001
Pain/stiffness location			
Back	86.3	97.2	< 0.001
Neck	72.5	48.1	< 0.001
Hip	62.7	39.3	< 0.001
Other regions	47.1	31.3	0.009
No pain/stiffness	2.0	0	0.104
Age at onset, mean \pm SD years	24.5 \pm 9.1	33.2 \pm 12.2	< 0.001
Pain onset description			0.030
Suddenly	25.5	15.9	
Come and go	43.1	57.0	
Gradual	30.4	27.1	
No pain	1.0	0	
Pain duration, mean \pm SD months	272.6 \pm 161.8	137.4 \pm 123.1	< 0.001
Upper back/rib cage pain	84.3	61.7	< 0.001
Numbness/tingling spread into legs	51.0	70.6	0.001
Pain around the heel of foot	52.9	40.2	0.040
Pain/stiffness improve with physical activity	86.3	57.0	< 0.001
Morning pain >30 minutes	70.6	78.5	0.159
Pain worse at night/early morning	79.4	70.1	0.103
Pain/stiffness due to fall/sprain	22.5	44.4	< 0.001
Pain switches from side to side	80.4	82.2	0.756
Pain switches from buttock to buttock	55.9	45.3	0.092
Impact of exercise on pain/stiffness			< 0.001
Decreases	64.7	43.9	
No change	22.5	21	
Increases	12.7	35.0	
Impact of daily physical activity on pain/stiffness			< 0.001
Decreases	58.8	26.2	
No change	21.6	27.1	
Increases	19.6	46.7	
Pain/stiffness as day progresses			0.003
Decreases	37.3	22.4	
No change	28.4	23.8	
Increases	34.3	53.7	
Pain/stiffness when inactive			0.029
Decreases	8.8	21.0	
No change	31.4	28.5	
Increases	58.8	50.0	
Missing/unknown	1.0	0.5	
Use of NSAIDs	75.5	65.6	0.290
Pain reduction within 48 hours after taking NSAIDs	93.5	77.9	0.003
Stomach pain prior to back pain	18.6	8.9	0.011
Diagnosed with iritis	42.2	2.3	< 0.001
Disease history			
Patient self			
Urethritis	8.8	3.3	0.052
Colitis/enteritis/IBD/Crohn's disease	20.6	8.9	0.006
Dactylitis	1.0	0.5	0.542
Enthesitis/synovitis/tenosynovitis	12.7	1.9	< 0.001
Psoriasis	10.8	5.6	0.076
Balanitis/cervicitis	2.0	1.9	1.000
First-degree relatives			
AS	24.5	3.3	< 0.001
Reactive arthritis	16.7	19.6	0.285
Colitis/enteritis/IBD/Crohn's disease	11.8	11.2	0.186
Psoriasis	5.9	9.3	0.107
Iritis	13.7	0.5	< 0.001
Second-degree relatives			
AS	11.8	4.7	0.077
Reactive arthritis	13.7	22.9	0.163
Colitis/enteritis/IBD/Crohn's disease	4.9	10.3	0.279
Psoriasis	8.8	6.5	0.732
Iritis	2.0	0	0.175
Fatigue summary score, mean \pm SD	39.8 \pm 13.5	40 \pm 14.5	0.888

* Values are the percentage unless otherwise indicated. AS = ankylosing spondylitis; CBP = chronic back pain; NSAIDs = nonsteroidal antiinflammatory drugs; IBD = inflammatory bowel disease.

Table 2. Predictive ability of reduced question item set (development sample, n = 316)*

Question item and response categories	OR (95% CI)
Sex, male	3.45 (1.51–7.91)
Pain/stiffness location	
Neck	3.49 (1.54–7.94)
Hip	3.54 (1.52–8.23)
Other regions	2.57 (1.07–6.15)
Age at onset, years	0.93 (0.89–0.97)
Pain duration, months	1.00 (1.00–1.01)
Numbness/tingling spread into legs	0.36 (0.16–0.81)
Pain/stiffness due to fall/sprain	0.25 (0.10–0.62)
Impact of exercise on pain/stiffness	
No change	Reference
Decreases	0.21 (0.06–0.82)
Increases	0.07 (0.01–0.33)
Impact of daily physical activity on pain/stiffness	
No change	Reference
Decreases	8.31 (1.92–36.00)
Increases	2.76 (0.69–11.10)
NSAID use and response to the treatment	
Did not use NSAIDs	Reference
Relief within 48 hours	1.39 (0.58–3.33)
Not relief within 48 hours	0.12 (0.02–0.81)
Diagnosed with iritis	30.31 (7.26–126.49)

* Hosmer-Lemeshow goodness-of-fit test: $P = 0.633$ ($\chi^2 = 6.123$, 8 df); Cox and Snell $R^2 = 0.499$; Nagelkerke $R^2 = 0.6970$; Omnibus test: $P < 0.001$ (likelihood ratio test: $\chi^2 = 218.33$, 15 df). OR = odds ratio; 95% CI = 95% confidence interval; NSAID = nonsteroidal antiinflammatory drug.

AS cases was 48.6 ± 13.5 years, their average disease duration was 23.1 years, and 62% were men. The CBP controls were younger (mean \pm SD 43.9 ± 12.1 years), had a shorter average disease duration (8.6 years), and fewer were men (37%). Bivariate results from the development sample are shown in Table 1, which presents the proportion of people with AS and the proportion of people with CBP who gave a positive response to each question item. Of the 43 question items tested, 24 showed statistical differences between these 2 groups ($P < 0.05$).

Based on the output from the best subset regression analysis (data are available by request), the final regression model (Table 2) included 12 question items for the case ascertainment tool (Figure 3). Using the validation sample, the tool demonstrated a sensitivity of 67.4% and a specificity of 94.6%, comparable with the threshold determined from the development sample (sensitivity = 69.6%, specificity = 99.1%, ROC = 0.937). Results from the sensitivity analysis showed an average sensitivity of 61.5% and an average specificity of 94.6% generated across 1,000 samples of randomly reassigned CBP controls (see Statistical analyses for details).

DISCUSSION

The need to identify patients with AS among those with CBP is of clinical relevance because effective disease-

specific therapies have become available that can improve pain and stiffness and reduce disability (25–27). We developed a case ascertainment tool that performed well in differentiating the less commonly encountered patients with possible AS from a large pool of individuals with low back pain from a mechanical cause. Our tool, containing 12 patient-reported question items, achieved an ROC of 0.937, a sensitivity of 70%, and a specificity of 99%. The strongest discriminators were male sex, pain/stiffness in the neck and/or hip, pain/stiffness decreases with daily physical activity, and a history of iritis.

Selecting the optimal case ascertainment tool threshold requires tradeoffs: increasing instrument sensitivity increases the probability of identifying patients with disease (and the ability to institute disease-altering therapies). However, this would increase the patient care burden on the rheumatologist due to additional unnecessary referrals, reduce quality of life and cost for those with false-positive results, and raise the economic burden (additional medical evaluations) from the payer perspective. Results from our model revealed that sensitivities ranging from 69.6% to 90.2% were associated with specificities of 99.1% to 79.9%, respectively (Table 3). Lowering sensitivities reduced the portion of false-positive results seen by the provider. For instance, within this sensitivity/specificity range for the CBP population, the proportion of false-positive results seen by providers ranged from 20.3% to 80.9%. Using a clinically relevant, minimally sufficient sensitivity of 70% and specificity of 99% (a set of parameters agreed upon by the expert panel), the tool would have a predicted probability threshold of 66.86%.

This case ascertainment tool could be used in a variety of populations with differing AS prevalences (pretest probabilities), and therefore display alternate performance characteristics. Among patients with back pain, it is estimated that AS has a 5% prevalence (14,28). In this population, our tool would be expected to achieve a 79.7% PPV (98.4% NPV). Assuming that the tool performs similarly in the general population (AS prevalence at a conservative estimate of 0.5%) (3), it would likely achieve 27.2% PPV (99.8% NPV).

Our tool differs from previous ones. Compared with the New York criteria for AS (a 5-item, symptom-based questionnaire) (16,17), our tool covers additional aspects (e.g., patient's sex, location of pain/stiffness, responsiveness to NSAIDs). Instead of using binary question items (yes/no), several of our question items allow gradations of patients' input. For example, for "how does exercise affect the pain or stiffness," patients can select an answer that best describes their disease experience from several options (i.e., it decreases, increases, or does not change the pain/stiffness). Two question items (i.e., duration of low back pain/stiffness and history of iritis) from this 5-item, symptom-based questionnaire remained in our final tool. Although all questions in our screener were patient reported, most of the previous AS criteria required the presence of radiographic sacroiliitis to confirm a diagnosis of AS, except for the Calin et al criteria for AS (19).

Our case ascertainment tool for AS performed well (area under the curve [AUC] = 0.937) compared with other commonly used screening tests: mammography for breast

Question item	Response categories	Item score
What is your gender?	male	1.2397
	female	0
Have you experienced pain or stiffness that lasted for at least 3 months? If so, please indicate the location(s).	Neck yes/no	Yes = 1.2502
	Hip yes/no	Yes = 1.2644
	Other regions yes/no	Yes = 0.9421
Approximately how old were you when you first had pain or stiffness in your back that lasted for at least 3 months?	in years	-0.0747 × (number of years)
Approximately how long have you had back pain or stiffness?	in months	0.00374 × (number of months)
Have you felt numbness or tingling that spread into or down your leg(s) that you think or have been told might have been caused by your back pain or stiffness?	yes/no	Yes = -1.0214
Is the pain or stiffness due to fall, sprain, or other incidents, such as twisting or lifting?	yes/no	Yes = -1.3775
How does exercise affect the pain or stiffness in your lower back or buttocks? Select the one that best describes your experience.	▪ It decreases the pain or stiffness	-1.5437
	▪ It does not change the pain or stiffness	0
	▪ It increases the pain and stiffness	-2.6988
	▪ I don't have pain or stiffness in the lower back or buttocks	0
How does daily physical activity affect the pain or stiffness in your lower back or buttocks? Select the one that best describes your experience.	▪ It decreases the pain or stiffness	2.1178
	▪ It does not change the pain or stiffness	0
	▪ It increases the pain and stiffness	1.0141
	▪ I don't have pain or stiffness in the lower back or buttocks	0
Do you take any NSAID medication(s)? If so, do they help reduce your back pain or stiffness within 48 hours?	▪ Yes, they help reduce my back pain or stiffness within 48 hours	0.3293
	▪ No, they do not help reduce my back pain or stiffness within 48 hours	-2.1489
	▪ I don't take NSAID medication	0
Have you been diagnosed with iritis?	yes/no	Yes = 3.4113
Scoring algorithm		
1. Assign an item score for each of the patient's responses.		
2. Take the sum of the patient's item scores. Let x be the sum of the patient's item scores.		
3. Let y be the patient's transformed score. We calculate y as follows:		
$y = \frac{e^{x-1.0242}}{1 + e^{x-1.0242}} \times 100 \quad (\text{Note: } -1.0242 \text{ is the intercept of the logistic regression mode.})$		
4. If y is greater than or equal to 66.86, then the case ascertainment tool result is positive for AS.		

Figure 3. Final set of question items for the ankylosing spondylitis (AS) case ascertainment tool and scoring algorithm. NSAID = nonsteroidal antiinflammatory drug.

cancer (AUC = 0.74) (29), Papanicolaou smear for cervical cancer (AUC = 0.93) (30), and the combination of exercise echocardiography and ²⁰¹Tl single-photon-emission computed tomography for coronary artery disease (AUC = 0.78) (31).

Sensitivity	Specificity	CBP population†		General population‡	
		PPV	NPV	PPV	NPV
69.6	99.1	79.7	98.4	27.2	99.8
70.6	97.2	57.0	98.4	11.2	99.8
75.5	94.9	43.6	98.7	6.9	99.9
80.4	89.7	29.2	98.9	3.8	99.9
85.3	84.6	22.5	99.1	2.7	99.9
90.2	79.9	19.1	99.4	2.2	99.9

* Values are the percentage. CBP = chronic back pain; PPV = positive predictive value; NPV = negative predictive value.
 † Ankylosing spondylitis prevalence 5%.
 ‡ Ankylosing spondylitis prevalence 0.5%.

We used a self-identified population of CBP patients as the control group. This could produce a bias; although it is unlikely given the low population prevalence of AS, it is possible that some of the control patients had underlying AS. To evaluate the robustness of our results to this issue, we performed a sensitivity analysis that assumed that 5% of the control patients had AS. Even with this hypothetical contamination in the control sample, our results remained very similar.

The patients with AS in our study had an average disease duration of 21.8 years, and their disease experience may not accurately reflect earlier-stage AS-related symptoms. We therefore evaluated the impact of disease duration on patients' experience with AS by comparing responses between a group of patients with ≤10 years' disease duration and a group with >10 years' experience. After adjusting for multiple comparisons, there were no statistically significant differences between the groups.

In our AS sample, ~54% were receiving biologic treatments (tumor necrosis factor α inhibitors) for active AS; these agents are expected to alter patients' symptoms, possibly affecting the performance of the tool. To investigate

this potential challenge to validity, we compared responses from patients who were receiving biologic agents with those who were not. None of the responses to the 12 question items differed significantly between the 2 groups.

One interesting finding from the feasibility study (phase II) was the lack of significant differences between patients with AS and CBP controls on pain/stiffness affected by physical activities. The results for patients with AS were as expected: most reported that pain/stiffness were decreased by physical activities. However, CBP group results were unexpected (56% reported that pain/stiffness were decreased by physical activities). It was suggested that CBP respondents may have attributed physical activities to physical therapy. Therefore, 2 additional question items were added, impact of exercise on pain/stiffness and impact of pain/stiffness as the day progresses, to capture "daily activities" rather than "physical activities" and also to assess pain/stiffness as the day progressed from waking in the morning. The original item (impact of daily physical activities) and 1 added item (impact of exercise) were included in the final tool, but they showed opposite predictability of AS: impact of daily physical activities scored positive, whereas impact of exercise scored negative. It is generally believed that exercise typically alleviates pain/stiffness in patients with AS (i.e., "exercise decreases pain/stiffness" should be scored positive). We suspect that respondents may have interpreted "exercise" as intensive physical exercise, as opposed to "daily physical activities" as light physical activities (e.g., walking). Also, the odds ratio (OR) of 0.214 for the impact of exercise on decreasing pain/stiffness, compared with the OR of 8.313 for the impact of daily physical activities on decreasing pain/stiffness, infers that the latter question item better describes the AS disease experience.

Today, most patients experience AS symptoms for 5–10 years before a correct diagnosis is made and many patients are never diagnosed, in part because these patients resemble the large pool of individuals with CBP. Our results suggest that a set of patient-reported question items can identify a group of patients with a high likelihood of having AS. We believe that this tool could substantially improve the possibility of making a diagnosis of this disease. Patients with a high likelihood of AS could be identified to undergo further evaluation (e.g., with HLA-B27 genetic testing or a magnetic resonance imaging scan). Published studies have shown that patients with AS have a 50–90% likelihood of having an HLA-B27 antigen, and the association between AS and HLA-B27 varies in different ethnic and racial groups (32). Combining the patient questionnaires with HLA-B27 testing could provide a sensitive and cost-effective approach to initial diagnostic assessment.

In conclusion, it is our aim that this case ascertainment tool might identify patients earlier in their disease, such that they could begin therapy. However, we realize the limitation of the study design, in that we only studied patients with AS with an established diagnosis who were likely to be at a later disease stage. Therefore, this tool needs further validation in an early disease stage population. This AS case ascertainment tool is intended for a patient to use as a Web-based tool with a built-in scoring

algorithm. If the tool's characteristics (e.g., sensitivity and specificity) are replicated in future studies, it could be made available to patients through the Internet or in the offices of primary care providers.

ACKNOWLEDGMENTS

The authors would like to thank Bonnie B. Dean, MPH, PhD, of Cerner LifeSciences, Beverly Hills, California, for assistance in designing the study and writing the manuscript; Brian Calimlim, MS, of Cerner LifeSciences, Beverly Hills, California, for programming, statistical analysis, and writing the manuscript; and Monica Choi, BA, for expert study coordination efforts.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be submitted for publication. Dr. Weisman had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Weisman, Chen, Clegg, Davis, Dubois, Prete, Savage, Schafer, Suarez-Almazor, Yu, Reveille.

Acquisition of data. Weisman, Chen, Davis, Dubois, Prete, Savage, Yu, Reveille.

Analysis and interpretation of data. Weisman, Chen, Clegg, Davis, Dubois, Prete, Schafer, Suarez-Almazor, Yu, Reveille.

ROLE OF THE STUDY SPONSOR

Neither the primary sponsor (the Spondylitis Association of America) nor the secondary sponsors (Centocor, Abbott, Amgen, Wyeth, and Pfizer) played a role in the study design, data collection, data analysis and management, or writing of the manuscript. The publication of this article was not contingent upon the approval of either the primary or the secondary sponsors.

REFERENCES

1. Walker BF. The prevalence of low back pain: a systematic review of the literature from 1966 to 1998. *J Spinal Disord* 2000;13:205–17.
2. Dagenais S, Caro J, Haldeman S. A systematic review of low back pain cost of illness studies in the United States and internationally. *Spine J* 2008;8:8–20.
3. Helmick CG, Felson DT, Lawrence RC, Gabriel S, Hirsch R, Kwoh CK, et al, and the National Arthritis Data Workgroup. Estimates of the prevalence of arthritis and other rheumatic conditions in the United States: part I. *Arthritis Rheum* 2008; 58:15–25.
4. Sieper J, Braun J. New treatment strategies in ankylosing spondylitis: proceedings of the Ankylosing Spondylitis Workshop; 2002 Jan 18–19; Berlin, Germany. *Ann Rheum Dis* 2002;61 Suppl 3:iii1–2.
5. Feldtkeller E, Khan MA, van der Heijde D, van der Linden S, Braun J. Age at disease onset and diagnosis delay in HLA-B27 negative vs. positive patients with ankylosing spondylitis. *Rheumatol Int* 2003;23:61–6.
6. Khan MA. Ankylosing spondylitis: introductory comments on its diagnosis and treatment. *Ann Rheum Dis* 2002;61 Suppl 3:iii3–7.
7. Mau W, Zeidler H, Mau R, Majewski A, Freyschmidt J, Stangel W, et al. Clinical features and prognosis of patients with possible ankylosing spondylitis: results of a 10-year followup. *J Rheumatol* 1988;15:1109–14.

8. Braun J, Sieper J, Breban M, Collantes-Estevez E, Davis J, Inman R, et al. Anti-tumour necrosis factor α therapy for ankylosing spondylitis: international experience. *Ann Rheum Dis* 2002;61 Suppl 3:iii51–60.
9. Gorman JD, Sack KE, Davis JC Jr. Treatment of ankylosing spondylitis by inhibition of tumor necrosis factor α . *N Engl J Med* 2002;346:1349–56.
10. Van den Bosch F, Kruithof E, Baeten D, Heressens A, de Keyser F, Mielants H, et al. Randomized double-blind comparison of chimeric monoclonal antibody to tumor necrosis factor α (infliximab) versus placebo in active spondylarthropathy. *Arthritis Rheum* 2002;46:755–65.
11. Braun J, Brandt J, Listing J, Zink A, Alten R, Golder W, et al. Treatment of active ankylosing spondylitis with infliximab: a randomised controlled multicentre trial. *Lancet* 2002;359:1187–93.
12. Rudwaleit M, Sieper J. Infliximab for the treatment of ankylosing spondylitis. *Expert Opin Biol Ther* 2005;5:1095–109.
13. Toussirot E, Wendling D. Current guidelines for the drug treatment of ankylosing spondylitis. *Drugs* 1998;56:225–40.
14. Rudwaleit M, van der Heijde D, Khan MA, Braun J, Sieper J. How to diagnose axial spondyloarthritis early. *Ann Rheum Dis* 2004;63:535–43.
15. Kellgren JH, Jeffrey MR, Ball L. The epidemiology of chronic rheumatism. Oxford (UK): Blackwell Scientific; 1963.
16. Bennett PH, Burch TA. Population studies of the rheumatic diseases. Amsterdam (The Netherlands): Excerpta Medica Foundation; 1968.
17. Van der Linden S, Valkenburg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis: a proposal for modification of the New York criteria. *Arthritis Rheum* 1984;27:361–8.
18. Mau W, Zeidler H, Mau R, Majewski A, Freyschmidt J, Stangel W, et al. Evaluation of early diagnostic criteria for ankylosing spondylitis in a 10 year follow-up. *Z Rheumatol* 1990;49:82–7.
19. Calin A, Porta J, Fries JF, Schurman DJ. Clinical history as a screening test for ankylosing spondylitis [abstract]. *JAMA* 1977;237:2613–4.
20. Rudwaleit M, Metter A, Listing J, Sieper J, Braun J. Inflammatory back pain in ankylosing spondylitis: a reassessment of the clinical history for application as classification and diagnostic criteria. *Arthritis Rheum* 2006;54:569–78.
21. Goie The H, Steven MM, van der Linden SM, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis: a comparison of the Rome, New York and modified New York criteria in patients with a positive clinical history screening test for ankylosing spondylitis. *Br J Rheumatol* 1985;24:242–9.
22. Krupp LB, LaRocca NG, Muir-Nash J, Steinberg AD. The fatigue severity scale: application to patients with multiple sclerosis and systemic lupus erythematosus. *Arch Neurol* 1989;46:1121–3.
23. Price DB, Tinkelman DG, Halbert RJ, Nordyke RJ, Isonaka S, Nonikov D, et al. Symptom-based questionnaire for identifying COPD in smokers. *Respiration* 2006;73:285–95.
24. Sieper J, Braun J, Rudwaleit M, Boonen A, Zink A. Ankylosing spondylitis: an overview. *Ann Rheum Dis* 2002;61 Suppl 3:iii8–18.
25. Sieper J, Rudwaleit M. Early referral recommendations for ankylosing spondylitis (including pre-radiographic and radiographic forms) in primary care. *Ann Rheum Dis* 2005;64:659–63.
26. Borenstein DG. Chronic low back pain. *Rheum Dis Clin North Am* 1996;22:439–56.
27. Deyo RA, Weinstein JN. Low back pain. *N Engl J Med* 2001;344:363–70.
28. Underwood MR, Dawes P. Inflammatory back pain in primary care. *Br J Rheumatol* 1995;34:1074–7.
29. Stoutjesdijk MJ, Boetes C, Jager GJ, Beex L, Bult P, Hendriks JH, et al. Magnetic resonance imaging and mammography in women with a hereditary risk of breast cancer. *J Natl Cancer Inst* 2001;93:1095–102.
30. Ogilvie GS, Patrick DM, Schulzer M, Sellors JW, Petric M, Chambers K, et al. Diagnostic accuracy of self collected vaginal specimens for human papillomavirus compared to clinician collected human papillomavirus specimens: a meta-analysis. *Sex Transm Infect* 2005;81:207–12.
31. Olmos LI, Dakik H, Gordon R, Dunn JK, Verani MS, Quinones MA, et al. Long-term prognostic value of exercise echocardiography compared with exercise ^{201}Tl , ECG, and clinical variables in patients evaluated for coronary artery disease. *Circulation* 1998;98:2679–86.
32. Akkoc N, Khan MA. Epidemiology of ankylosing spondylitis and related spondyloarthropathies. In: Weisman MH, Reveille JD, van der Heijde D, editors. *Ankylosing spondylitis and the spondyloarthropathies*. Philadelphia: Mosby; 2006. p. 117–31.

APPENDIX A: EXPERT ADVISORY BOARD MEMBERS

Michael H. Weisman, Director of the Division of Rheumatology, Cedars-Sinai Medical Center, Professor of Medicine at the David Geffen School of Medicine at the University of California, Los Angeles, Professor Emeritus of Medicine at the University of California, San Diego School of Medicine; John C. Davis, Jr., Associate Professor of Medicine, Associate Director Clinical Trials Center, Director Lupus Center, University of California, San Francisco; Maria E. Suarez-Almazor, University of Texas MD Anderson Cancer Center; Daniel O. Clegg, Professor of Medicine, Chief, Division of Rheumatology, Harold J., Ardella T., and Helen T. Stevenson Presidential Chair in Rheumatology, University of Utah School of Medicine; John D. Reveille, Professor and Director, Division of Rheumatology and Clinical Immunogenetics, The University of Texas Houston Health Science Center at Houston; Lan Chen, Clinical Assistant Professor, Rheumatology Division, University of Pennsylvania; Pamela E. Prete, Professor of Medicine in Residence, University of California Irvine, Chief, Rheumatology, Veterans Administration, Long Beach, CA; Laura Schafer, registered nurse.