The Potential of ESSG Spondyloarthropathy Classification Criteria as a Diagnostic Aid in Rheumatological Practice

ELISA MUÑOZ GOMARIZ, AGUSTÍN CISNAL del MAZO, VERÓNICA PÉREZ GUIJO, ALEJANDRO ESCUDERO CONTRERAS, MARÍA del CARMEN MUÑOZ VILLANUEVA, and EDUARDO COLLANTES ESTÉVEZ

ABSTRACT. ***Objective.*** The European Spondylarthropathy Study Group (ESSG) criteria for the classification of spondyloarthropathies (SpA) were developed with the aim of unifying and facilitating international medical communication. We assessed the potential of ESSG criteria as a diagnostic aid for rheumatology practices in terms of sex and prevalence rate.

Methods. Data from 2 similarly designed and developed studies conducted in France and Spain were examined. Data were obtained from 3494 patients seen at rheumatology outpatient services (28 in each country). The sensitivity and specificity of each ESSG criterion (except the radiological one) were assessed in terms of sex and country. Patients were divided into 4 groups according to number of criteria present at the time of the study: Group 1 had neither inflammatory spinal pain (ISP) nor synovitis; Group 2 had ISP and/or synovitis; Group 3 ISP and/or synovitis plus one additional criterion; Group 4 ISP and/or synovitis plus more than one additional criterion. The predictive value was determined by using different prevalence rates.

Results. A prevalence of 27.6% for male and 8.0% for female patients was found at Spanish services; prevalence in French services was 9.1% males and 3.2% females. No significant differences in sensitivity and specificity for each sex between French and Spanish individuals were detected; the overall sensitivity and specificity were similar for men and women. By contrast, there were differences between patients from the 2 countries regarding individual ESSG criteria; thus, inflammatory spinal pain and synovitis were less specific in the female and male Spanish patients, respectively, relative to the French patients.

Conclusion. ESSG criteria can be used meaningfully to aid diagnosis when the prevalence of SpA exceeds 10% and the patient meets more than one of the additional criteria, or when prevalence exceeds 30% and the patient meets only one additional criterion. (J Rheumatol 2002;29:326–30)

Key Indexing Terms:
SPONDYLOARTHROPATHIES CLASSIFICATION CRITERIA DIAGNOSTIC CRITERIA
PREVALENCE OF DISEASE PERFORMANCE OF CLASSIFICATION CRITERIA

The criteria developed by the European Spondylarthropathy Study Group (ESSG) for classifying spondyloarthropathies (SpA)[1] have proved capable of identifying initial, atypical, and undifferentiated forms of this type of disease[2] (Table 1). Notwithstanding the good intrinsic performance (sensitivity and specificity) of these criteria, not all patients who meet them have SpA; whether or not they do depends on the prevalence of the disease in the population considered.

We used the predictive value (PV) as the likelihood of a patient meeting given criteria actually having (positive PV) or not having (negative PV) the disease. The PV is a statistical parameter that incorporates the prevalence of the disease. Although ESSG criteria have been found to perform quite well (with a high sensitivity and specificity) in classification in various populations[3-7], their usefulness as a SpA diagnostic aid[8,9] depends on the prevalence of SpA in the specific environment concerned.

The concepts of sensitivity/specificity and PV are rather different[10]; the most essential difference between them is probably that both sensitivity and specificity are inherent in the screening test. It makes no difference which population is used in testing them. Not so, however, for the PV, which changes with the prevalence of the disease in the population being studied; thus, when the estimated prevalence is close to 50%, if the test result is positive, the likelihood of a

From the Servicio de Reumatología, Hospital Universitario Reina Sofía; and Departamento de Bioestadística, Universidad de Córdoba, Córdoba, Spain.

E.M. Gomariz, BD, Professor of Biostatistics, University of Córdoba; A.C. del Mazo, MD, Clinical Methodologist; V. Pérez Guijo, MD, Rheumatologist; A.E. Contreras, MD, Rheumatologist; M.C.M. Villanueva, MD, Clinical Methodologist; E.C. Estévez, MD, Professor of Rheumatology and Chief of Rheumatology Service, Hospital Universitario Reina Sofía.

Address reprints requests to Prof. E.C. Estévez, Servicio de Reumatología, Hospital Provincial, Hospital Universitario "Reina Sofía," Avda. Menéndez Pidal, s/n, E-14004 Córdoba, Spain.
E mail: ecollantes@sofia.hrs.sas.cica.es

Submitted October 20, 2000; revision accepted August 8, 2001.

Table 1. European Spondylarthropathy Study Group (ESSG) classification criteria for spondyloarthropathy[1].

- Inflammatory spinal pain
  or
- Synovitis: Asymmetric or predominantly in the lower limbs and one or more of the following:
- Alternating buttock pain
- Sacroiliitis
- Enthesopathy
- Positive family history
- Psoriasis
- Inflammatory bowel disease
- Urethritis, cervicitis, or acute diarrhea occurring within one month before arthritis.

patient having the disease (positive predictive value) roughly approximates the specificity of the test. On the other hand, if the test result is negative the likelihood of a patient not having the disease (negative predictive value) roughly approximates the sensitivity of the test.

This is usually not the case with SpA, owing to their low prevalence in the general population. Indeed, the prevalence of SpA is directly correlated with that of the HLA-B27 antigen in the population[11]. The highest prevalence of ankylosing spondylitis (AS), 4.5%, has been found in Canadian Haida First Nations people, where 50% of the population is B27 positive. In the European and US populations, the prevalence of AS is estimated to be 0.25 to 1.4%[12]. The results of a recent study in Brittany, France, suggest that SpA is as common in women as it is in men, with a prevalence of 0.16–0.90% in the former and 0.05–0.77% in the latter[13].

In clinical practice, ESSG criteria are applied to patients seen at rheumatology practices, where SpA prevalence is much higher than in the general population.

We assessed the potential of the ESSG criteria as a diagnostic aid in terms of sex and prevalence rates of patients seen at Spanish and French rheumatology outpatient departments.

MATERIALS AND METHODS

Data were obtained from 2 multicenter studies conducted on 3494 patients seen at 28 French[3] and 28 Spanish rheumatology services[4]. Diagnosis of SpA and designation of control patients were decided by the head researcher based on his/her personal assessment, with no reference to ESSG criteria. Consensus was discussed before starting the study, but actual interobserver variation in labelling subjects as SpA or non-SpA has not been assessed. For one week each of the 56 experts categorized all consecutive new and followup patients as SpA or non-SpA.

Of the 1460 Spanish patients, 218 (133 men, 85 women) were diagnosed as definite SpA (each patient met the diagnostic criteria for the basic nosological entity) and 1242 (306 men, 936 women) were used as controls (individuals having other rheumatic diseases). Of the 2034 French patients, 122 (82 men, 40 women) were diagnosed as SpA and 1912 (757 men, 1155 women) were used as controls.

Independently, each patient was categorized as having: (1) no inflammatory spinal pain (ISP) and no synovitis; or (2) either ISP or synovitis; or (3) ISP and/or synovitis plus one additional criterion; or (4) ISP and/or synovitis plus 2 or more additional criteria. The patients in Groups 3 and 4 met ESSG criteria (Table 1). All data were collected by a researcher unaware of the diagnosis.

The study assessed the sensitivity and specificity of each criterion, in terms of sex and country; comparisons were based on the chi-squared test with the Yates correction or on Fisher's exact test when any expected value was ≤ 5.

Comparisons will be made of the sensitivity and specificity values of the criterion set by sex and country. If there are no significant differences the global performance of these criteria will be calculated using all the data; otherwise the calculations will be for each subcategory.

Because radiological changes develop slowly (e.g., radiological evidence for sacroiliitis usually appears 3–7 years after the onset of the disease)[2], the performance of the criteria was determined with exclusion of the radiological criterion.

If for a newly referred (not yet diagnosed) patient his or her pretest probability of having SpA is known (or can be estimated) then the posttest probability or predictive value can be calculated easily. Pretest probability is defined as the probability of having the target disorder before a diagnostic test result is known; it can be calculated as the proportion (or prevalence) of a mix of new and followup patients who have SpA, out of all the patients with the symptoms(s), both those with and without the disorder. Also included in calculating the denominator of prevalence were 138 French patients (63 men, 75 women) and 89 Spanish patients (42 men, 47 women) whose symptoms were only suggestive of SpA (possible spondyloarthropathy, i.e., patients with signs or symptoms suggesting SpA), but who failed to meet the diagnostic criteria for any specific SpA type[1]. This prevalence, as well as by sex and country, was calculated for patients under and over 35 years of age at the time of diagnosis.

The predictive value or posttest probability (i.e., the likelihood of a patient belonging to one of the 4 groups or categories actually having SpA) was calculated at different prevalence rates, using the following expression:

$$\text{Predictive value} = \text{posttest OR}/(1 + \text{posttest OR})$$

where

$$\text{Posttest OR} = \text{likelihood ratio} \times [\text{prevalence}/(1 - \text{prevalence})]$$

The likelihood ratio (LR) is the likelihood that a given test result would be expected in a patient with the target disorder compared to the likelihood that the same result would be expected in a patient without the target disorder. The LR, defined as sensitivity/(1 – specificity), was calculated for each of these 4 categories or groups, by excluding the sacroiliitis criterion.

All data were centralized and computed using Epi-Info v. 6.04 and Epidat software. Data recorded as uncertain were discarded.

RESULTS

Table 2 summarizes the characteristics (diagnosis and mean age) of the 340 patients with SpA and the 3154 control individuals (patients with other rheumatic diseases). No significant differences among the data for each pathology and sex were detected between French and Spanish patients with SpA; there were, however, significant differences ($p < 0.001$) among pathologies in the control group within each sex, except for male patients with arthrosis.

Table 3 compares the sensitivity and specificity of each individual ESSG criterion in the Spanish and French patients. Inflammatory spinal pain and synovitis were less specific ($p < 0.001$) in Spanish women and men, respectively, than in their French counterparts. The overall sensitivity and specificity was similar for sex and country, so the likelihood ratio (Table 4) was calculated as a whole.

Table 2. Characteristics of patients with SpA and controls.

	Men, n (%)		Women, n (%)	
	France	Spain	France	Spain
SpA	82(100)	133(100)	40 (100)	85 (100)
Ankylosing spondylitis	57(69.5)	78 (58.6)	18 (45.0)	27 (31.8)
Psoriatic arthritis	17 (20.7)	35 (26.3)	13 (32.5)	45 (52.9)
Reactive arthritis	10 (12.2)	13 (9.8)	4 (10.0)	4 (4.7)
Inflammatory bowel disease	3 (3.7)	2 (1.5)	3 (7.5)	1 (1.2)
Undifferentiated SpA	3 (3.7)*	5 (3.8)	2 (5.0)	9 (10.6)*
Age, yrs, mean ± SD	42.6 ± 14.5	42.7 ± 14.2	40.9 ± 13.9	45.5 ± 12.9
Min-max	12–78	15–74	19–71	21–74
Control group	757 (100)	306 (100)	1155 (100)	936 (100)
Inflammatory rheumatism	112 (14.8)	91 (29.7)	277 (24.0)	289 (30.9)
Mechanical back pain	343 (45.3)	63 (20.6)	406 (35.2)	196 (20.9)
Osteopathy	113 (14.9)	20 (6.5)	154 (13.3)	88 (9.4)
Arthrosis	67 (8.9)	24 (7.8)	141 (12.2)	187 (20.0)
Miscellaneous	169 (22.3)*	125 (40.8)*	250 (21.6)*	329 (35.1)*
Age, yrs, mean ± SD	53.3 ± 17.2	54.1 ± 15.4	55.7 ± 17.0	54.2 ± 14.8
Min-max	10–90	12–89	2–90	6–90

* Overall figure exceeded 100% because some patients exhibited more than one pathology.

When calculating prevalence, the number of patients with SpA was divided by the total number of patients including those with possible SpA. A prevalence of 27.6% [95% confidence interval (CI) 23.7–31.9%] and 8.0% (95% CI 6.4–9.8%) for Spanish male and female patients, respectively, was found. The prevalence among French patients was 9.1% (95 CI% 7.3–11.2%) for men and 3.2 (95% CI 2.3–4.3%) for women. If only those patients who were under 35 years of age at the time of diagnosis are considered, the prevalence in Spain was 48.9% (95% CI 40.7–57.6%) for men and 15.4% (95% CI 10.8–21.4%) for women; prevalence in France was lower, with 22.5% (95% CI 16.8–29.0%) for men and 11.0% (95% CI 7.2–15.7%) for women. On the other hand, prevalence among male and female Spanish patients who were over age 35 years when diagnosed was 19.5% (95% CI 15.6–24.2%) and 6.3% (95% CI 4.8–8.2%), respectively; similarly, the figure for France was 7.0% (95% CI 5.1–9.5%) for men and 1.9% (95% CI 1.1–3.1%) for women.

Table 5 and Figure 1 show the likelihood that a patient

Table 3. Performance of each ESSG criterion in men and women. A patient met the criteria if at least one major criterion and one additional criterion were present.

	Men				Women			
	Sensitivity, %		Specificity, %		Sensitivity, %		Specificity, %	
Criterion	France	Spain	France	Spain	France	Spain	France	Spain
Major criteria								
Inflammatory spinal pain	69.5	78.2	92.7	93.8	62.5	57.6	93.5	89.5*
Synovitis	46.3	35.3	93.3	85.6*	55.0	51.8	95.8	94.2
Additional criteria								
Positive family history	28.0	33.8	97.8	97.7	27.5	38.8	96.1	96.5
Psoriasis	20.7	22.6	97.6	98.7	35.0	50.6	98.1	98.9
Inflammatory bowel disease	3.7	2.3	99.6	100	10.0	3.5	99.5	99.6
Urethritis, cervicitis, or diarrhea	20.7	10.5	98.7	99.3	12.5	11.8	99.3	98.8
Alternating buttock pain	39.0	47.4	97.6	94.1	37.5	37.6	97.6	95.2
Enthesopathy	50.0	49.6	93.4	95.1	52.5	42.4	91.3	93.6
Criterion set	69.5	75.2	96.0	95.8	75.0	77.6	96.7	95.6
France + Spain	73.0		96.0		76.8		96.2	

*Significant difference ($p < 0.001$) versus France.

Table 4. Likelihood ratio [sensitivity/(1 – specificity)] of ESSG criteria. Group 1: Neither inflammatory spinal pain (ISP) nor synovitis; Group 2: ISP and/or synovitis; Group 3: ISP and/or synovitis plus one additional criterion; Group 4: ISP and/or synovitis plus more than one additional criterion.

Group	SpA, n	Controls, n	Likelihood Ratio 95% CI
1	51	2730	0.17 (0.13–0.22)
2	36	302	1.10 (0.79–1.53)
3	79	92	7.96 (6.02–10.52)
4	174	30	53.80 (37.12–77.96)

not meeting (Groups 1 and 2) or meeting (Groups 3 and 4) ESSG criteria will develop SpA. The prevalence (%) should be read and interpreted as pretest likelihood of having SpA for newly referred, not yet diagnosed patients.

DISCUSSION

Diagnostic and classification criteria play central roles in clinical rheumatology practice; the criteria are useful insofar as they allow like to be compared with like, such that results of clinical and other studies may be compared on a direct basis. While very few rheumatic diseases lack a criterion system, the criteria are not always properly understood[9] or

Table 5. Likelihood of having SpA (predictive value or posttest probability) determined by using different prevalence rates or pretest probability. Group 1: Neither inflammatory spinal pain (ISP) nor synovitis; Group 2: ISP and/or synovitis; Group 3: ISP and/or synovitis plus one additional criterion; Group 4: ISP and/or synovitis plus more than one additional criterion.

	SpA Likelihood, % (95% CI)				
Prevalence	1%	2%	3%	4%	5%
Group 1	0.2 (0.1–0.2)	0.3 (0.3–0.4)	0.5 (0.4–0.7)	0.7 (0.5–0.9)	0.9 (0.7–1.1)
Group 2	1 (0.8–1.5)	2 (2–3)	3 (2–5)	4 (3–6)	5 (4–7)
Group 3	7 (6–10)	14 (11–18)	20 (16–25)	25 (20–30)	30 (24–36)
Group 4	35 (27–44)	52 (43 –61)	62 (53–71)	69 (61–76)	74 (66–80)
Prevalence	10%	20%	30%	40%	50%
Group 1	2 (1.5–2.5)	4 (3–5)	7 (5–9)	10 (8–13)	15 (12–18)
Group 2	11 (8–15)	22 (16 –28)	32 (25–40)	42 (34–50)	52 (44–60)
Group 3	47 (40–54)	67 (60 –72)	77 (72–82)	84 (80–88)	89 (86–91)
Group 4	86 (80–90)	93 (90–95)	96 (94–97)	97 (96–98)	98 (97–99)

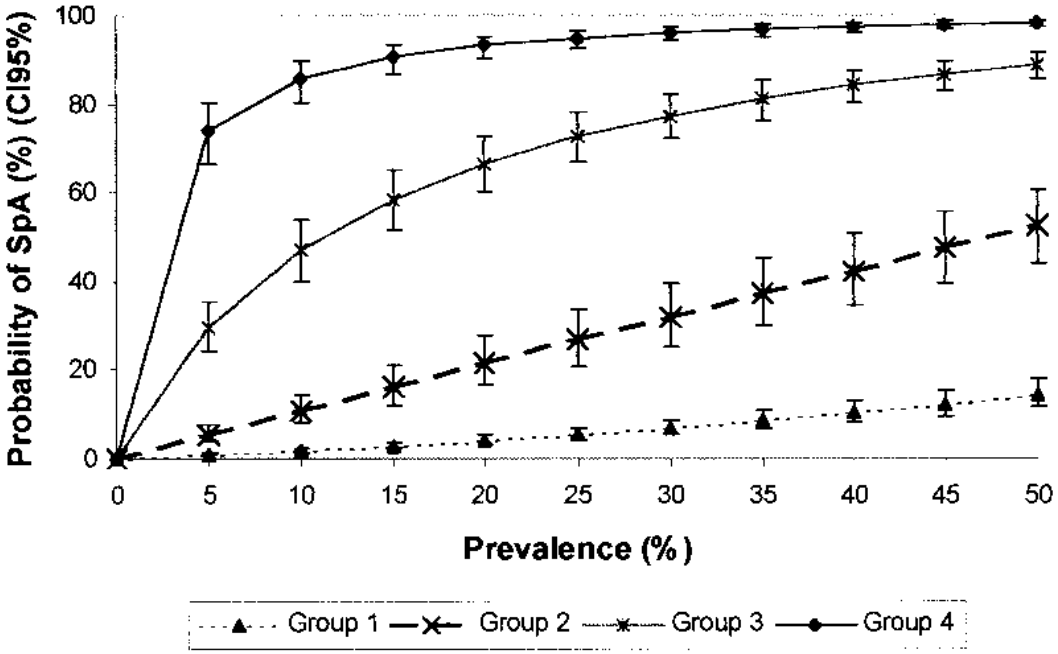

Figure 1. Likelihood of SpA (predictive value or posttest probability) at different prevalence rates or pretest probability. Group 1: Have neither inflammatory spinal pain nor synovitis. Group 2: Have inflammatory spinal pain and/or synovitis. Group 3: Have inflammatory spinal pain and/or synovitis plus one additional criterion. Group 4: Have inflammatory spinal pain and/or synovitis plus more than one additional criterion.

applied, most often through confusion or an inaccurate knowledge of the 2 concepts, diagnosis and classification.

Classification criteria use the "gold standard" to include patients in scientific studies[14] in order to ensure comparability in the identification and standardization of patients with related pathologies. However, they were not developed, nor are they suitable, for establishing a reliable clinical diagnosis, which can only be reached after a highly complex cognitive process very often dependent on subtle findings, but never on just a few criteria. Diagnostic and classification criteria are essentially different.

Despite their excellent sensitivity and specificity, the ESSG criteria were developed primarily as a "set of classification criteria," an instrument for clinical use intended to help one decide whether a patient with a given clinical picture can be classified as the carrier of a SpA even if the specific underlying nosological entity cannot be identified at the time. The usefulness of these criteria for diagnostic purposes has been discussed[15]. The proven success of these essentially classifying criteria and their widespread acceptance[4-7] have led to their misuse as diagnostic criteria. We examined the characteristics of the ESSG classification criteria obtained in France and Spain; for this purpose, we applied the criteria to the patients in both studies a posteriori and considered the prevalence of SpA in the 2 environments. The differences between the 2 populations suggest the presence of differences in the pathological spectrum dealt with at hospital rheumatology outpatient departments in the 2 countries — although the interobserver differences in the interpretation of what constitute signs and symptoms of SpA may also contribute to the differences observed, since the interobserver error in determining SpA/non–SpA has not been assessed.

From the study, a projection was developed on the basis of the actual prevalence of SpA in each specific environment (private practice, outpatient department, patients admitted to hospital) that should be of use for future diagnoses.

For a prevalence of SpA above 10% in the environments considered, a patient in Group 4 can be diagnosed with SpA with an error less than 20%; on the other hand, reaching the same conclusion for a patient in Group 3 requires that the prevalence be higher than 30%. For Spanish rheumatology services, the likelihood of a patient under 35 years of age (SpA prevalence = 48.9%) in Group 3 actually having SpA will be 87% (95% CI 84–90%); that of a patient in Group 4 will be 98%. On the other hand, for French patients of the same age classified in Groups 3 and 4, the likelihood will be 70% (95% CI 64–95%) and 94% (95% CI 91–96%), respectively (the prevalence among the French population studied was 22.5%).

These diagnostic tools should be applied to newly referred patients, as it is necessary to know the pretest probability (or prevalence) for age and sex in order to calculate the posttest probability. From the foregoing it follows that for ESSG criteria to be useful as a diagnostic aid for SpA, appropriate epidemiological data for each major rheumatic disease in the environment concerned must be available.

ACKNOWLEDGMENT

The authors are grateful to Prof. Bernard Amor, Prof. Maxime Dougados, and Dr. Veronique Listrat for their collaboration with data from the French study, and Justin Spoliar for technical assistance.

REFERENCES

1. Dougados M, van der Linden SM, Juhlin R, et al. The European Spondylarthropathy Study Group preliminary criteria for the classification of spondylarthropathy. Arthritis Rheum 1991;34:1218-27.
2. Dougados M. How to diagnose spondylarthropathy? [editorial]. Clin Exp Rheumatol 1993;11:1-3.
3. Amor B, Dougados M, Listrat V, et al. Evaluation des critères des spondylarthropathies d´Amor et de l´European Spondylarthropathy Study Group (ESSG). Ann Med Interne 1991;142:85-9.
4. Collantes E, Cisnal A, Muñoz E. Assessment of two systems of spondyloarthropathy diagnostic and classification criteria (Amor and ESSG) by a Spanish multicenter study. J Rheumatol 1995;22:246-51.
5. Boyer GS, Templin DW, Goring WP. Evaluation of the European Spondylarthropathy Study Group preliminary classification criteria in Alaskan Eskimo populations. Arthritis Rheum 1993;36:534-8.
6. Ertuk M, Alaca R, Tosun E, Duruoz MT. Evaluation of the Amor and ESSG classification criteria for spondylarthropathies in a Turkish population. Rev Rhum [Engl Ed] 1997;64:293-300.
7. Baddoura R, Awada H, Okais J, Habis T, Attoui S, Abi Saab M. Validation of the European Spondylarthropathy Study Group and B. Amor criteria for spondylarthropathies in Lebanon. Rev Rhum [Engl Ed] 1997;64:459-64.
8. Dougados M. Diagnostic features of ankylosing spondylitis [editorial]. Br J Rheumatol 1995;34:301-5.
9. Amor B, Dougados M, Listrat V, et al. Are classification criteria for spondyloarthropathy useful as diagnostic criteria? Rev Rhum [Engl Ed] 1995;62:10-5.
10. Gordis L. The role of epidemiology in the study of rheumatic diseases. In: Lawrence RC, Schulmann LE, editors. Epidemiology of the rheumatic diseases. New York: Gower Medical Publishing Ltd.; 1984.
11. Khan MA. A worldwide overview: the epidemiology of HLA-B27 and associated spondylarthritides. In: Calin A, Taurog JD, editors. The spondylarthritides. Oxford: Oxford University Press; 1998:17-26.
12. Arnett FC. Ankylosing spondylitis. In: Koopman WJ, editor. Arthritis and allied conditions. 13th ed. Baltimore: Williams & Wilkins; 1997:1197-209.
13. Saraux A, Guedes C, Allain J, et al. Prevalence of rheumatoid arthritis and spondyloarthropathy in Brittany, France. J Rheumatol 1999;26:2622-7.
14. Arnett FC. Revised criteria for the classification of rheumatoid arthritis. Bull Rheum Dis 1989;38:1-6.
15. Dougados M. Classification and diagnosis of seronegative spondylarthropathies. Scand J Rheumatol 1999;28:336-9.