Evaluation of the accuracy of several symptoms and domains in distinguishing patients diagnosed with fibromyalgia from healthy controls

C. Gómez-Perretta[1], Y. Triñanes[2], A.J. González-Villar[2], M.T. Carrillo-de-la-Peña[2]

[1]Research Foundation of La Fe Hospital, Valencia, Spain;
[2]Department of Clinical Psychology and Psychobiology, University of Santiago de Compostela, Spain.

Claudio Gómez-Perretta, MD, PhD
Yolanda Triñanes, MsC, PhDc
Alberto J. González-Villar, MsC, PhDc
María T. Carrillo-de-la-Peña, PhD

Please address correspondence to:
Dr María T. Carrillo-de-la-Peña,
Department of Clinical Psychology and Psychobiology,
University of Santiago de Compostela.
15702 Santiago de Compostela, Spain.
E-mail: mteresa.carrillo@usc.es

Received on February 12, 2015; accepted in revised form on May 4, 2015.

Clin Exp Rheumatol 2016; 34 (Suppl. 96): S14-S25.

Key words: fibromyalgia, core symptoms, clinical assessment, reference standard, algometry.

Funding: The research reported in this manuscript has been funded by the Spanish Ministry of Science and Innovation (ref. PSI2009-14555).

Competing interests: none declared.

ABSTRACT

Objective. *To assess the discriminative power of several symptoms and domains that may assist in the diagnosis of subjects with fibromyalgia (FM).*

Methods. *79 individuals with FM and 66 healthy controls participated in the study. The potential domains proposed by the American College of Rheumatology (ACR) criteria were considered (Wolfe et al., 1990). Binary logistic regression and area under a ROC curve (AUC) were used to rank the importance of the variables in distinguishing patients from pain-free controls. Z values were then calculated to compare the AUC values obtained for each variable with that which yielded the highest AUC (reference standard). For each measure, the cut-offs that maximise sensitivity and specificity were also calculated.*

Results. *The mean pressure pain threshold (PPT) yielded the highest discriminative power (AUC, 0.991) and was therefore chosen as the reference standard; considering an optimal cut-off ≤3.97, it correctly classified 95% of patients and 97% of controls. The discriminative powers of tender point count (cut-off ≥9), health-related quality of life (cut-off ≤63.27) and vitality (cut-off ≤46.97) were as good as that of the reference standard. Finally, items related to physical role and function, body pain, fatigue and memory loss showed adequate discriminative power, although slightly lower than that of the reference.*

Conclusion. *In addition to pain, health-related quality of life and fatigue/vitality were confirmed as the best predictors of individuals with FM. The study findings indicate that tender point count and especially pressure pain threshold (measured with an algometer) continue to be key issues in the clinical assessment of subjects with FM relative to pain-free controls.*

Introduction

As early as 1989, Yunus advocated that diagnosis of FM should not be made by exclusion but through positive assessment of a constellation of symptoms (1). In 1990, the American College of Rheumatology (ACR) published the first classification criteria for fibromyalgia (FM): tenderness in 11 out of 18 discrete regions (tender points) and widespread pain during 3 months (2).

Although the 1990 ACR criteria have been validated and shown to be reliable, it has been argued that patients with FM display increased sensitivity to pressure pain throughout the body, not only at tender sites, and that the tender point count is influenced by physical and/or mental states (3-5). Moreover, applying the 1990 ACR criteria in clinical practice may overestimate the importance of tenderness (*e.g.* by oversampling in women) (6, 7). It has also been suggested that the 1990 criteria are not reproducible in 36% of fibromyalgia patients over a 6-month period (8). Taking into account these criticisms and on the basis of previous findings, Wolfe and colleagues proposed provisional modified criteria in 2010 (9). Thus, un-refreshed sleep, fatigue, cognitive symptoms and widespread pain, along with a number of symptoms such as pain or cramp in the lower abdomen, depression and headache were proposed as key variables in the diagnosis of FM (9-12). Although disregarding tender points decreased the specificity of diagnosis, the new classification criteria correctly classified 88.1% of cases (9).

The ACR provided provisional endorsement of the 2010 FM diagnostic criteria proposed by Wolfe *et al.* (9), while waiting for external validation. Nevertheless, the modified criteria have not received full endorsement by the ACR, because this society has es-

tablished a policy that it will no longer endorse diagnostic criteria, in the belief that the final decision about any patient diagnosis and appropriate treatment should remain part of the physician-patient relationship.

The Outcome Measures in Rheumatology Clinical Trials (OMERACT) (11) conducted several Delphi exercises with both patients and clinicians to obtain a consensus of opinion about which domains should be assessed in clinical trials for FM (13-15). Along with the study of Boomershine (16), these provided a core group of relevant FM symptom domains. Despite criticisms regarding tender point evaluation, more than 70% of OMERACT participants agreed that tenderness and sleep disturbance should be measured in all FM clinical trials, while other variables such as dyscognition and depression should be measured in only some of them (11).

Thus, although many domains and tests are currently available for diagnosing FM, it is not known which domain best predicts the occurrence or absence of FM or what role tender point assessment plays in the diagnosis of FM. This is very important as study of the discriminative power of the various measurement instruments may assist in the diagnosis of FM and may be used to differentiate responders from non responders to treatment (17, 18).

Predictive models were used to assess the discriminative power of all variables. A predictive model has the advantage of describing the relationship between the dichotomous characteristic of interest (healthy subject or FM patient) and a set of independent (predictor or explanatory) variables. Measurement instruments or domains with the highest ability to predict inclusion in the FM group (the gold standard established by the clinicians) would be those of greatest diagnostic relevance and would demonstrate robust construct validity.

Finally, we identified the optimal cut-off point that misclassifies the smallest number of patients and controls for each variable, which may also have practical implications for the diagnosis of subjects with FM.

Material and methods

Design and participants

With the collaboration of a local patient's association, an invitation to take part in this study was sent out to 150 women, aged between 18 and 75 years old, with a confirmed clinical history of FM by experts in this area. Initially, inclusion criteria were diagnosis of FM and no other chronic pain disease or any disorder that could explain the principal symptoms of FM. However, this had to be modified as the source population comprised a majority of subjects with characteristic symptoms of other comorbid disorders. Reconfirmation of diagnosis and exploration of comorbidities were carried out by clinical interview.

To search for controls of the similar demographical characteristics but no major illness, an announcement was posted on bulletin boards of local health departments and women's associations. Seventy possible controls were recruited in this way.

The exclusion criteria for both groups included a chronic physical condition other than FM for the patients, moderate to severe psychiatric illnesses, substance abuse, pregnancy and lactation, active cancer and specifically any history of polymyalgia rheumatica, connective tissue diseases, endocrine myopathy or chronic fatigue syndrome (19). Because of the high comorbidity between FM and other stress-related disorders such as post-traumatic stress disorder (PTSD) (20) and mild depression (13-15), patients showing evidence of any of these syndromes were included in the study.

Finally, seventy-nine subjects with FM were accepted as study participants (mean age= 46.42 years, SD=9.8), after the diagnosis of FM was endorsed by specialised medical examination following the 1990 ACR criteria (2). Sixty-six controls of similar age were also recruited (mean age= 46.0 years, SD=11.4). Both groups were predominantly middle-aged Caucasians and most were from Galicia (NW Spain).

Both groups attended a clinical session that included a comprehensive interview, pressure algometry at the tender points and administration of a number of test instruments (see below).

Patients were not requested to stop using their medication, including analgesics, for the purpose of the study.

Written informed consent was obtained from all patients and healthy subjects before participation in the study. The study was approved by the Ethics Committee of the University of Santiago de Compostela, in accordance with the Declaration of Helsinki.

Clinical interview

All subjects (*i.e.* controls and subjects with FM) were evaluated by administration of a systematic questionnaire that included age, height, weight, family history of FM, existence of physical and psychological stressful events preceding the onset of the FM symptoms, marital status, level of education, occupation, work status and ethnic background, medical history, years since diagnosis of FM and specialty of the professional who made the diagnosis. Data on medication at the time of the study were also recorded.

A group of symptoms extracted from the American College of Rheumatology criteria was also assessed in the participants (2, 9). These symptoms also included variables that reflect limited functioning and emotional well-being (21, 22). According to the hierarchy proposed by Arnold *et al*. (23), we explored 1) the *core symptoms* of FM described as a triad that includes chronic widespread pain (in the four quadrants and axial skeleton) of long duration (≥3 months), fatigue and sleep disturbance; 2) other *key associated symptoms*: tenderness, stiffness, paraesthesia (tingling or sensation of numbness), loss of strength, headache, mood disturbances (*e.g.* depression and/or anxiety), memory deficit, impaired physical function and quality of life including social function; and finally 3) other *comorbid conditions* such as abdominal pain, constipation/diarrhea, swelling, frequent urination, menstrual cramp, skin changes and hypersensitivity (SSV) to psychological distress, noise, cold, weather changes, heat and humidity.

Participants answered questions about the occurrence of these variables on a 4-option scale (never, rarely, often or always).

Algometry at tender point sites

Pressure pain threshold (PPT) and tolerance (PPTol) were measured by algometry performed by Y.T. (24). This method provides a measure of tenderness that has been demonstrated to be accurate and reliable (25). For the examination, a pressure algometer (Wagner Force One®, Model FDI) was applied at the eighteen tender sites, according to the ACR criteria (2). The pressure was increased at an approximate rate of 1 kilogram per second, until the participant reported feeling pain. The pressure was then stopped, and the PPT was recorded in kilograms per square centimetre (kg/cm^2). The pressure continued to increase until the participant indicated inability to withstand pain (PPTol).

We calculated the mean PPT and PPTol of the algometry results for all 18 points. The tender point count (TPC) was determined as the number of tender points with a threshold measurement ≤4 kg/cm^2, which is roughly equal to the pressure required to blanch the examiner's thumbnail.

Instruments

• *Fibromyalgia Impact Questionnaire (FIQ)*

The Spanish version of the FIQ (26) validated and adapted to Spanish-speaking populations (S-FIQ) (27) was used in this study, only for the sample of patients. Although this version asks questions about aspects of work-related disability, the scoring was adapted for women who reported that they had never worked outside of the home.

• *SF-36*

The SF-36 is an instrument that measures health-related quality of life (HRQoL) across eight different dimensions: physical functioning (PF), role limitation because of physical health (RP), social functioning (SF), vitality (VT), body pain (BP), mental health (MH), role limitation because of emotional problems (RE) and general health (GH). The reliability and construct validity of this self-report questionnaire have been demonstrated (28-31). We used the Spanish version of the Short Form 36 Health Survey validated and adapted to Spanish-speaking populations in Spain (32).

• *Visual analogue scales (VAS)*

Visual analogue scales consisted of a straight horizontal line of fixed length 100 mm on which participants were asked to indicate their state concerning pain intensity, headache, fatigue, morning rigidity, and sadness in the last month. The ends are defined as the extreme limits of the parameter to be measured orientated from the left (best) to the right (worst). VAS are valid and reliable measures of FM symptomatology (33-37).

• *Beck Depression Inventory (BDI)*

The Beck depression inventory (BDI) consists of 21 multiple-choice items that are answered by choosing which of four statements best describes the way the patient feels. The BDI assesses the severity of current depressive symptoms (38).

• *Pittsburgh Sleep Quality Index (PSQI)*

Sleep quality was assessed with the Spanish version of the Pittsburgh Sleep Quality Index (PSQI), which has shown good reliability and validity when administered to patients with FM (39-41). The PSQI consists of 19 self-rating questions split into 7 components. Scores of 0–4 indicate good sleep quality, scores of 5–10, poor quality, and scores over 10, sleep disorder (39).

Statistical analysis

Characteristics of the patients and control groups were described by summary measures. Continuous variables were expressed as means ± standard deviation and intervals, and categorical variables as frequencies and percentages. To analyse between-group differences in the categorical variables, we performed Chi-square tests and measures of association for cross tabulations, depending on whether the variables were nominal (lambda) or ordinal (Somers'd (directional) or Kendall's Tau b (symmetrical)). As lambda may underestimate relationships when many values are clustered around one response, in this case we used Cramers V (42). We also used the Kolmogorov–Smirnov-Lilliefors (KS test) adjustment test for a normal distribution in the case of continuous variables. When the KS test indicated the absence of normal distribution, we applied the non-parametric Mann-Whitney U-test. Otherwise, we used the Student's test. The size of the influence was calculated following Cohen's d formula. We determined the probability of rejecting the null hypothesis when it is in fact false (power).

We also calculated Pearson's correlation coefficients for education and each of the other variables.

We used univariate binary logistic regression analysis to obtain odds ratios (ORs), in order to calculate the strength of association and the impact of confounding variables, 95%-confidence intervals (95%-CI) and probability values (*p*-value). To assess the possible confounding effect of education level, all regression models were adjusted for this variable.

The dichotomous dependent variable was coded as 0 (control subject) or 1 (FM patient). The following predictor variables were included for each univariate analysis: age, body mass index (BMI), education level, BDI, PSQI, PPT, PPTol, left trochanter PPT), we performed additional analyses to identify the tender point that best discriminated between the samples, which was the left trochanter (Fig. 1). TPC, SF-36 subscales, VAS and the other variables assessed in the clinical interview. Although the latter variables were scored on a 4-option scale, they were recoded in binary code, as 0 if the symptom was never or rarely present or 1 if often or always suffered. Only subjective memory loss was assessed on a 4-option scale as none of the controls obtained scores in the range indicating significant impairment. TPC was also binary recoded, as positive (≥11) or negative (<11).

The predicted probability scores derived from the regression analysis were used to construct receiver operating characteristic (ROC) curves for all symptoms or domains in order to determine sensitivity and specificity levels. For each of the curves, the best cut-off values that maximise sensitivity + specificity were reported.

The area under the ROC curve (AUC) represents the validity of a model. For each possible cut-off point or criterion value, we discriminated between the

two samples. The correctly classified FM subjects were defined as positive (TP = true positives), while cases classified as negative constituted the false negatives (FN). On the other hand, controls correctly classified as negative were referred to as the true negatives (TN), whereas controls classified as positive were considered as false positives (FP) (Fig. 2).

AUC values of between 0.7 and 0.8 indicate reasonable discrimination, and values exceeding 0.8 indicate good discrimination. The domain with the maximum AUC and minimal standard error was used as a reference standard, to calculate z values for comparing differences in AUC (43). A z test was constructed using standard errors and provided the data normalisation necessary for comparison of the different AUC values. Briefly, the numerator of the expression is the difference in area under the ROC curves (AUC) for the measure of interest, and the denominator is composed of the standard error of the difference in areas (SEi; see formula below). The z value is then compared with the normally distributed values included in a table, and values above some cut-off are taken as evidence that one AUC value is significantly higher than the other. The null hypothesis considered the equality of the areas, and otherwise, the alternative hypothesis was rejected or accepted. Finally, the probability of making a type 1 error was considered to be p=0.05 (one tailed).

$$Z = [A1 - A2] / \sqrt{(SE1^2 + SE2^2)}$$

Finally, we calculated the probability density function (pdf) of four continuous variables: the mean PPT and the three test instruments used with both patients and controls (*i.e.* BDI, PSQI, and SF-36).

We used SPSS, v. 21 software to carry out all statistical analyses.

Results

Demographic data

All FM patients and controls were Spaniards of Galician origin (except for one patient). Family history of FM revealed a familial aggregation of 31.4%. Thirty-one percent of patients reported a physical episode and 77.5% a psychological stressful event prior to the onset of FM symptoms, which were perceived as triggers of the syndrome. Eighty-five percent of patients had experienced FM symptoms for more than 2 years (see Table I).

Table I. Demographic data (percentages) for each sample and measures of association (directional) for cross tabulations for nominal (lambda, V Cramer) and ordinal (Somers'd) variables.

	Variable	Patients	Controls
Pathogenesis	Family history of FM	31.4 %	
	Physical triggering factors	31.0 %	
	Psychological triggering factors	77.5 %	
Years since FM diagnosis	> 2 years	85%	
	mean (SD)	5.23 (4.9)	
	interval	0.5-24	
Age distribution: mean±SD (interval)		46.42 ±9.8 [22–64]	46.0±11.4 [25–75]
BMI (ranks)	15-18.5	2 (2.56 %)	3 (4.62 %)
	18.6-25	38 (48.72 %)	36 (55.38 %)
	25.1-30	21 (26.92 %)	22 (33.85 %)
	30.1-35	14 (17.95 %)	3 (4.62 %)
	35.1-40	1 (1.28 %)	0 (.00 %)
	>40.1	2 (2.56 %)	1 (1.54 %)
BMI (kg/m^2) (*)		26.40 ± 5.0	24.35 ± 4.0
Marital status	Single	8 (10.26 %)	22 (33.85 %)
[Lambda: 0.46*]	Married	64 (82.05 %)	21 (32.31 %)
	Separated/divorced	4 (5.13 %)	15 (23.08 %)
	Widowed	2 (2.56 %)	7 (10.77 %)
Education	None	1 (1.28 %)	1 (1.54 %)
[Somers'd : - 0.23*]	Elementary	36 (46.15 %)	18 (27.69 %)
	Technical	6 (7.69 %)	13 (20.00 %)
	Graduated	18 (23.08 %)	12 (18.46 %)
	University	17 (21.79 %)	21 (32.31 %)
Occupation	Housewife	17 (21.79 %)	9 (15.00 %)
[Lambda: 0.05]	Part time	9 (11.54 %)	3 (5.00 %)
	Full time (not specialised)	23 (29.49 %)	13 (21.67 %)
	Full time (average)	16 (20.51 %)	17 (28.33 %)
	Full time (highly specialised)	11 (14.10 %)	14 (23.33 %)
	Student	2 (2.56 %)	4 (6.67 %)
Work Status	Active	38 (60.32 %)	45 (73.77 %)
[V Cramer: 0.46*]	Invalidity	19 (30.16 %)	0
	Unemployed	3 (4.76 %)	13 (21.31 %)
	Retired	3 (4.76 %)	3 (4.9 %)

BMI: Body Mass Index. *p<0.05.

Mean age and distribution did not differ between groups. The FM rate increased with age, reached two peaks at 45 and 50 years old, and then decreased. Regarding weight, 6.2% of the controls and 21.8% of patients were obese (BMI >30). The BMI differed significantly between groups.

Measures of association revealed statistically significant differences between groups in marital status, work status and level of education, but not occupation (Table I). Finally, level of education was significantly correlated with many of the symptoms and domains studied (data not shown). The regression models were therefore adjusted for education level.

Pharmacological therapy

Most of the patients were polymedicated. Mild analgesics (ethanamide, acetyl-salicylic acid and ibuprofen), opiates (fentanyl, tramadol) and antidepressants (duloxetine) plus sedatives (tetrazepam) were the most common currently prescribed drugs. The proportion of patients receiving each of the above treatments was 35.2%, 32.4% and 32.4%, respectively. Pregabalin (20%), gabapentin (1.5%) and other

Table II. Comparison of subjects according to myalgic scores.

Algometry (a) Tender Point Count [#)]	Mean (SD) 95% Interval Controls	FM	t (b)(*)	Cohen's d [95%Interval]	Power
PPT (a)	5.5±1.0 [5.2-5.7]	2.4±0.8 [2.3-2.6]	19.6	3.3 [2.8-3.8]	1.0
PPTol (a)	6.1±1.0 [5.8-6.3]	3.2±1.1 [2.9-3.4]	16.3	2.7 [2.3-3.2]	1.0
TPC [#]	4.2± 2.5 [3.5-4.8]	16±3.1 [15.3-16.7]	24.6	4.1 [3.5-4.7]	1.0

(a) [(kg/cm^2); (b) t value (Equal variances assumed) (*) p<0.000.

pharmacological or homeopathic drugs were also used. Regarding compliance with treatments, most patients reported good adherence and small changes in their medication.

Frequency of tender points and quantification

Table II shows the algometer data obtained at the tender points. FM patients reported lower mean pressure pain threshold and tolerance, and a greater number of tender points, than controls. The between-group differences were statistically significant for all these variables. Regarding the tender point count, 1 control had more than 11 tender points and 8 patients fewer than 11 tender points. As it has been suggested that a cut-off value of 11 tender points may be somewhat arbitrary (see review [44].), we decided to include patients with fewer than 11 tender points for assessment.

Figure 1 shows the mean PPT values at the eighteen tender sites, for each group. As the confidence intervals do not overlap, we concluded that the differences between both groups were highly statistically significant for the eighteen tender sites. Maximal between-group differences were observed in the trochanter area.

Clinical characterisation of the FM group

FM patients were usually polysymptomatic, with symptoms affecting several systems. All subjects diagnosed with FM reported multiple current complaints, the occurrence of which has often been associated with FM.

Table III shows the most frequent symptoms in FM and the results of the statistical tests of comparisons between groups. The prevalence of the symptoms considered was higher in the FM group than in the controls. The calculated z (Mann-Whitney U-test) and Kendall tau b-values showed highly statistically significant differences for the variables considered. The only exception was food intolerance. Most of the variables (continuous) exceeded the conventional level of power in sample size calculations at the 80% level, while the effects varied somewhat between variables, ranging between the medium, strong, and very strong categories.

Ranking of the main symptoms reported by FM subjects according to their effect size of Mann-Whitney U-test x Power revealed that the maximal significant difference corresponded to quality of life (SF-36 mean), vitality (SF-36), fatigue VAS, physical role and body pain (SF-36).

Sciatica and cervical pain radiating to the arm or head were only assessed in the FM group and were also highly prevalent.

Other symptoms that commonly appear in the FM syndrome, such as anxiety, insomnia, depression and loss of memory, were also highly prevalent in the FM sample and differed significantly between patients and controls (*e.g.* BDI scores: 19.38±9.53 in patients *vs.* 5.97±5.87 in healthy controls; PSQI: 12.97±4.47 *vs.* 5.56±3.55, p<0.001 for each test). Further symptoms described as accompanying FM, such as stress sensitivity, digestive problems, menstrual cramp, frequent urination, skin changes, and sensitivity (SSV) to environmental agents, were also reported as prevalent.

Table III also includes the cut-off values for each symptom and test derived from the regression analyses. Taking these values into account, the proportion of individuals affected in each group was provided for each variable.

FIQ and SF36 scores are shown in Table IV. In the FIQ, the highest scores obtained by the FM group corresponded to the following subscales: morning tiredness (8.15), fatigue (7.75), feel good (7.15), stiffness (7.10) and pain (6.91). The scores for physical disabil-

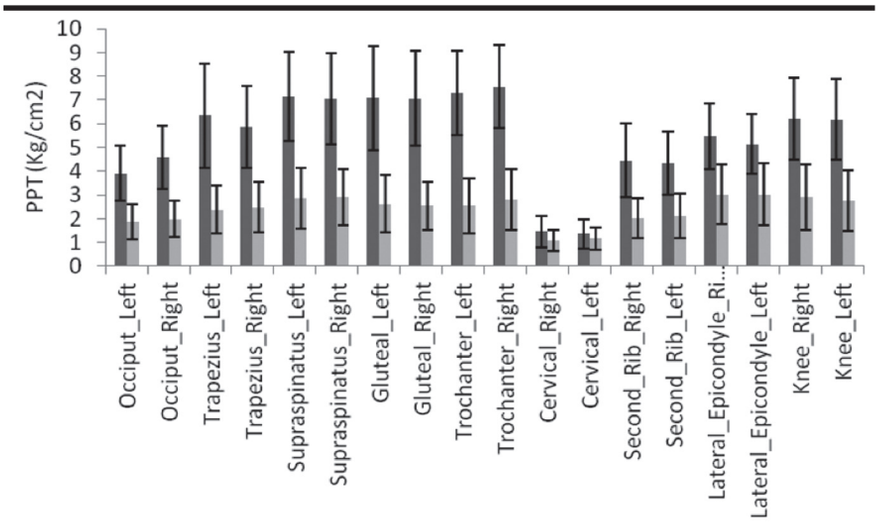

Fig. 1. Pressure Pain Threshold (PPT in kg/cm^2) at the eighteen tender sites (dark = controls; grey = FM subjects). Mean and standard deviation are represented.
Left trochanter point yielded the maximum statistical difference between groups [ANOVA (F = 358.38; p<0.0001).

Table III. Ranking of the main symptoms reported by FM subjects and statistical differences from controls according to Mann-Whitney U-test, effect size and power. For each group the proportion of subjects (%) clinically affected, according to the cut-off in the predictive logistic regression analysis, is shown.

Modality	Statistics (•)	Cohen's d [95%Interval]	Power (▲)	Groups	Cut-off	Affected (%)
Quality of Life ■	**8.89***	2.8 [2.3 - 3.3]	1.0 (n=124)	Controls	≤ 63.3	8.8
				FM		92.5
Vitality (b)	8.84*	2.7 [2.2- 3.2]	1.0 (n=124)	Controls	≤ 47.0	3.5
				FM		92.5
Fatigue†	8.73*	2.7 [2.2 - 3.1]	1.0 (n=126)	Controls	≥ 4.4	11.0
				FM		91.0
Physical Role (b)	8.70*	2.6 [2.1 - 3.1]	1.0 (n=124)	Controls	≤ 65.7	3.5
				FM		86.6
Body Pain (b)	8.63*	2.6 [2.1 - 3.1]	1.0 (n=124)	Controls	≤ 62.3	13.8
				FM		92.5
Pain†	8.33*	2.2 [1.8 - 2.7]	1.0 (n=126)	Controls	≥ 4.3	19.0
				FM		87.3
Physical Function (b)	8.14*	2.0 [1.6-2.5]	1.0 (n=124)	Controls	≤ 73.1	12.1
				FM		86.6
Social Function (b)	8.09*	2.1 [1.6 - 2.5]	1.0 (n=124)	Controls	≤ 80.9	21.1
				FM		89.6
Sleep disorder (c)	7.93*	1.8 [1.4 – 2.2]	1.0 (n=140)	Controls	≥ 8.5	19.0
				FM		85.9
Depression (a)	7.72*	1.7 [1.3 – 2.1]	1.0 (n=125)	Controls	≥ 11.5	15.9
				FM		79.4
Rigidity†	7.41*	1.7 [1.3 – 2.2]	1.0 (n=126)	Controls	≥ 5.4	25.0
				FM		79.0
Mental Health (b)	6.10*	1.3 [.89 -1.7]	.99 (n=124)	Controls	≤ 67.5	22.4
				FM		73.1
Headaches†	6.16*	1.2 [.8 – 1.5]	.99 (n=126)	Controls	≥ 3.1	23.3
				FM		66.7
Emotional Role (b)	6.00*	1,3 [.9 – 1.7]	.99 (n=124)	Controls	≤ 87.4	25.9
				FM		76.1
Mood (Sadness)†	5.63*	1.1 [.7 - 1.5]	.99 (n=126)	Controls	≥ 2.8	15.9
				FM		69.9

(a) Beck Depression Inventory; (b) SF-36 (subscales) ■SF-36 mean; (c) PSQI; †VAS ; •Mann-Whitney (z value); Effect size of Mann-Whitney U-test according to Cohen's d formula; ▲Power (α=.05); *p (two tailed) <0.001.

ity (3.77) and work missed (3.03) were within a lower range. The division in quartiles showed that 62% of patients scored higher than 61.75, 20.3% scored between 51.77 and 61.75, 10.1% scored within the range 41.84 - 51.76 and 7.6% scored less than 41.84. Between-group comparisons showed statistically significant differences in all SF36 subscales and total scale scores.

Diagnostic test assessment by ROC curve regression analysis

The mean PPT was the domain with the maximum ROC area (0.991) and minimal standard error (0.005) (Table V). The ROC curve showed a cut-off index at a score of 3.97 kg/cm^2, when sensitivity was 0.95 and specificity 0.97. A tender point count score of ≥9 was 0.92 sensitive (SV) and 0.97 specific (SP), and yielded an area under the ROC curve (AUC) value of 0.989. The AUC value for the trochanter at the left side representative of the most affected region was slightly lower (0.988).

As the mean PPT was the domain with the maximum AUC and minimal standard error, this index was chosen as the reference standard, and the other domains and symptoms described in Table V were classified according to their ¨z¨ differences for this parameter. This yielded four groups of variables. The variables in each group are shaded in different tones of grey in Table V.

The first group mainly comprised variables related to physical components, except for memory complaint and social function. In addition to the algometer data, the domains with the maximum AUC corresponded to the SF-36 total score (AUC=0.965; SV=0.93; SP=0.91) and some of its subscales: Vitality (AUC=0.961; SV=0.93; SP=0.97); Physical role (AUC=0.951; SV=0.87; SP=0.91) and Body pain (AUC=0.947; SV=0.93; SP=0.86). For the SF-36 total scale, the cut-off score was 63.27. The VAS for fatigue and pain also yielded high AUC values. For VAS fatigue, a score >4.4 was 0.91 sensitive and 0.89 specific (AUC=0.950) and for VAS pain, the cut-off was around 4.3 (SV=0.87; SP=0.81; AUC=0.930). Finally, classification of subjects according to memory function was excellent (AUC=0.924; SV=0.89; SP=0.81).

The AUC values were between 0.8 and 0.9 for the following symptoms, which constituted a second group of relevant clinical symptoms: depression (BDI), sleep disturbance (PSQI), SF-36 emotional role, VAS morning rigidity, stress SSV, cervical pain, humidity SSV, mental health, headaches VAS, sensitivity to weather change and emotional role. From this group, the BDI yielded the highest AUC value of 0.902 (SV=0.81; SP= 0.84) and a cut-off index at score of 11.5. The AUC for the PSQI was 0.891(cut-off score = 8.5; SV = 0.86; SP

Table III *(continued)*. Ranking of less commonly reported symptoms [Kendall's Tau b measures of association between groups], and proportion of clinically affected subjects per group (%).

Modality	Kendall's Tau b	Groups	% Affected (1)
Chest pain	**0.80***	Controls	0
		FM	80.8
Loss strength	**0.80***	Controls	6.8
		FM	87.2
Memory loss	**0.80***	Controls	0.0
		FM	80.7
Back pain	0.79*	Controls	15.25
		FM	93.6
Anxiety	0.79*	Controls	13.6
		FM	92.2
Stress SSV	0.73*	Controls	20.6
		FM	92.2
Cervical pain	0.72*	Controls	23.7
		FM	93.6
Humidity SSV	0.69*	Controls	20.3
		FM	88.5
Weather changes SSV	0.63*	Controls	19.0
		FM	81.8
Paraesthesia	0.58*	Controls	11.9
		FM	70.5
Noise SSV	0.58*	Controls	22.0
		FM	80.5
Constipation / Diarrhoea	0.57*	Controls	3.4
		FM	5.5
Menstrual cramps	0.54*	Controls	17.6
		FM	71.9
Cold SSV	0.52*	Controls	30.5
		FM	82.1
Abdominal pain	0.45*	Controls	15.3
		FM	60.3
Swelling	0.44*	Controls	17.0
		FM	60.3
Frequent urination	0.30*	Controls	10.2
		FM	35.9
Hot sensitivity	0.35*	Controls	25.4
		FM	61.5
Skin changes	0.29*	Controls	11.9
		FM	37.2
Food intolerance	NS	Controls	5.3
		FM	9.0
Cervical pain to (▲)	Head	FM	85.9
	Arm		90.1
Sciatica (▲)		FM	67.6

(1) Measured according to Likert scale of frequency: never present (0), rarely (1), often (2), always (3), but finally coded into binary variables where 0 (non-affected) = symptom never or rarely present, and 1 (affected) = often or always present. ▲Only FM group was analysed *p<0.001

= 0.81). Thus, the second group consisted of a mixture of variables related to the emotional domain, insomnia, localised pain, stiffness and environmental sensitivity to external agents.

Sensitivity to noise and feeling of sadness (VAS) yielded AUC values at the top of the third group (*i.e.* values of 0.7–0.8). Paraesthesia, menstrual cramp, constipation/ diarrhoea, cold SSV, abdominal pain and swelling were also classified in this group.

The remaining variables, situated at the bottom of the classification showed low discriminative power (AUC <0.7). The relationship between the outcome variable (for the FM or Control group) and each of the predictors are also described in Table V (see *Odds ratios*). Within the tenderness measures, a decrease of one unit in the mean PPT (reference standard) increased the number of patients diagnosed with FM -by 33.33 (OR: 0.03; 95% Cl: 0.01–0.1), while one unit increase in the number of tender point counts decreased the number of patients diagnosed with FM by 50% (OR: 2.0; 95% Cl: 1.5–2.6). Higher SF-36 scores substantially decreased the likelihood of FM. Variables such as depression (BDI), insomnia (PSQI) and memory deficit were associated with the outcome variable: each increase of one unit in BDI increased the odds by 1.28 times (OR: 1.3; 95% Cl: 1.2–1.4), while an increase of one unit of PSQI increased the odds by 1.46 (OR: 1.5; 95% Cl: 1.3–1.6).

Probability density function analysis

Figure 2 shows the probability density function (pdf) for the 4 variables studied: mean PPT, BDI, PSQI and SF-36. In general, there was a slight overlap between the two groups, which indicated the existence of false positive and negative diagnoses. However, because of the small number of false diagnoses, the threshold modality was almost completely successful in differentiating the two groups.

Discussion

The main aim of this study was to assess the discriminative power of some recognised domains and instruments used to aid diagnosis of FM, to establish the optimal cut-off points to prevent misclassification of patients and healthy subjects, and to compare the efficacy of these instruments with that of a reference standard.

To accomplish these objectives, it was first necessary to have a gold standard (diagnosis by the clinician) against which to compare the accuracy of the various measures. Thus, the role of qualified specialists in the proper diagnosis of the sample of patients was critical. It was also essential to carry out a thorough examination of participants and appropriate selection of the indices to be studied. In this respect, the degree to which the FM patients were affected was examined by the FIQ, revealing that most patients were included within the highest quartile (maximum degree of severity of the illness).

The ROC analyses revealed that mean pressure pain threshold (PPT) was the most accurate measure for discriminating between individuals with FM and healthy subjects. We therefore chose this as the reference standard. Although examination of tender points has re-

Table IV. Data from the Fibromyalgia Impact Questionnaire (FIQ) administered to FM patients and from the Short Form Health Survey (SF-36) administered to FM patients and controls.

FM Group FIQ Scores (n= 63)	
Subscales	Mean
Physical impairment	3.77 ± 2.13
Feel good	7.15 ± 2.89
Work missed	3.03 ± 3.83
Do work	5.85 ± 3.19
Pain	6.91 ± 2.36
Fatigue	7.75 ± 2.39
Morning tiredness	8.15 ± 2.13
Stiffness	7.10 ± 2.45
Anxiety	6.85 ± 2.82
Depression	5.19 ± 3.04
Total (mean)	6.18 ± 1.67
Total (sum)	63.77 ± 16.30
SF-36 Scores FM Group (n= 67)	
Subscales	Mean
General Health	28.28 ± 18.10
Physical Function	48.81 ± 21.15
Physical Role	34.89 ± 24.86
Emotional Role	65.30 ± 26.14
Social Function	48.69 ± 26.12
Body Pain	29.33 ± 20.48
Vitality	23.51 ± 17.46
Mental Health	52.71 ± 21.57
SF-36 mean	40.98 ± 15.44
SF-36 Scores Control Group (n= 58)	
Subscales	Mean
General Health	72.33 ± 18.19
Physical Function	88.45 ± 17.65
Physical Role	89.44 ± 16.65
Emotional Role	92.06 ± 12.85
Social Function	91.81 ± 14.16
Body Pain	81.05 ± 19.03
Vitality	69.40 ± 16.04
Mental Health	78.32 ± 17.23
SF-36 mean	79.54 ± 11.76

As proposed in the original paper, the FIQ item (feel good) score was reversed, in such a way that when the patient answered 0 days for "felt good during the past week" the score considered was 7; and then, it was multiplied by 1.43 (to obtain a maximum score of 10).
There was a significant difference between groups for all subscales and the mean score of the SF-36.

Table V. Comparison of the areas under two ROC curves (Mean threshold (PPT) *vs*. others) according to z values. ORs -95% CIs and cut-off values are also shown.

	Area	SE	Z value (*)	OR(95% CI) (▲)	Cut-off	SV	SP
Threshold (PPT)	**.991**	**.005**	**0**	**.03***	**(.01 -- .1)**	**3.97**	**.95**	**.97**
TP count	.989	.01	.22	2*	(1.5–2.6)	9	.92	.97
Left trochanter	.988	.01	.35	.12*	(.05–.28)	4.5	.96	.92
Tolerance (PPTol)	.980	.01	1.04	.05*	(.02–.15)	4.72	.92	.95
Quality of Life (1)	.965	.02	1.44	.85*	(.8–.9)	63.3	.93	.91
Vitality (1´)	.961	.02	1.64	.88*	(.8–.9)	47.0	.93	.97
Physical role (1´)	.951	.02	1.91*	.91*	(.8–.94)	65.7	.87	.91
Fatigue VAS	.950	.02	1.94*	2.3*	(1.78–3.0)	4.4	.91	.90
Body pain (1´)	.947	.02	1.98*	.91*	(.89–.94)	62.3	.93	.86
General health (1´)	.948	.02	2.11*	.90*	(.87–.93)	52.5	.91	.85
TP count ▫	.941	.02	2.21*	569*	(69-4673)		.90	.99
Physical function (1´)	.925	.03	2.22*	.91*	(.9–.94)	73.6	.87	.88
Memory loss (▪)	.924	.03	2.30*	11.9*	(5.5–25.5)		.89	.81
Pain VAS	.930	.02	2.75*	2.2*	(1.7–2.9)	4.3	.87	.81
Social function (1´)	.913	.03	2.98*	.92*	(.9–.94)	80.9	.90	.79
Loss of strength (b)	.902	.03	3.02*	90*	(26.7–304)		.87	.93
Lumbar pain (b)	.892	.03	3.05*	78*	(24.6–248)		.94	.85
Anxiety (b)	.893	.03	3.11*	75*	(24.7–230.7)		.92	.87
Depression (2)	.902	.03	3.33*	1.3*	(1.2–1.4)	11.5	.81	.84
Sleep disturbance (3)	.891	.03	3.57*	1.5*	(1.3–1.6)	8.5	.86	.81
Rigidity-Morning stiffness VAS	.882	.03	3.58*	1.6*	(1.4–1.8)	5.4	.79	.75
Stress SSV(b)	.858	.04	3.65*	45*	(15–129)		.92	.79
Cervical pain (b)	.849	.04	3.80*	47*	(16–139)		.94	.76
Humidity SSV (b)	.841	.04	4.01*	30*	(11.7–77)		.89	.80
Mental health (1´)	.821	.04	4.34*	.94*	(.9–.96)	67.5	.73	.78
Headaches VAS	.817	.04	4.42*	1.4*	(1.3–1.7)	3.2	.67	.76
Weather change (b)	.814	.04	4.49*	19*	(8–46)		.82	.81
Emotional role (1´)	.804	.04	4.63*	.93*	(.9–1.0)	87.4	.76	.75
Noise SSV (b)	.792	.04	4.81*	14.6*	(6–34)		.81	.78
Mood-Sadness VAS	.790	.04	4.82*	1.6*	(1.3–2)	2.79	.70	.84
Paraesthesia (b)	.793	.04	4.90*	17.8*	(7–45)		.70	.88
Menstrual cramp (b)	.772	.04	4.94*	12*	(5–28.8)		.72	.83
Constipation (b)	.775	.04	5.35*	40*	(9.1–176)		.58	.97
Cold SSV (b)	.758	.04	5.38*	10*	(4.7–23)		.82	.70
Abdominal pain (b)	.725	.04	6.0*	8*	(3.6–19.6)		.60	.85
Swelling (b)	.717	.04	6.18*	7*	(3.3–16.8)		.60	.83
Hot SSV (b)	.678	.05	6.61*	4.6*	(2–9.7)		.62	.74
Frequent urination (b)	.629	.05	7.65*	5*	(1.9–13)		.36	.90
Skin changes (b)	.627	.05	7.69*	4.4*	(1.8–11)		.37	.88
BMI	.617	.05	7.93*	1.1*	(1–1.2)	23.7	.65	.46
Education level (b)	.591	.05	8.28*	.46*	(.23–.92)		.47	.71

SF-36 Health Survey (1´) SF-36 subscale (2) BDI (3) PSQI; SV: Sensitivity SP: Specificity, ▫Binary: < 11 and ≥ 11 tender points; ▪Variable included on a 4-option Likert scale of frequency; (b) Variables converted to binary ("often" or "permanently" affected); ▲List of variables with adjusted coefficients (by Level of education): Memory Loss, BMI, BDI, PSQI, Pain, Morning stiffness, Fatigue, Sadness, Headaches, General Health, Physical Function, Physical Role, Vitality, Quality of Life, Lumbar pain, Paraesthesia, Swelling, Loss of Strength, Weather Change SSV. *p (one tailed) ≤0.05.

ceived several criticisms, our findings support the usefulness of pressure algometry in the assessment of FM and the conceptualisation of it as a disorder of pain regulation (central sensitisation), as indicated by increased sensitivity to painful stimuli and lowered pain thresholds (3, 11, 16, 45-49). This is consistent with recent suggestions that algometer scores could be used as a marker of peripheral neuro-sensitivity (3, 45), included in the clinical diagnosis and research [15] and used as a potential biomarker in FM (11, 47).
The mean PPT scores reported here were similar to the values obtained in a survey of fibromyalgia patients (50). Nevertheless, we found that some patients had fewer than 11 tender points. Indeed, our results indicate that the 11-point cut-off is arbitrary, as we found that 9 sensitive points are sufficient for correct classification of FM and corroborate the idea that there is no fixed range for diagnosis confirmation (3, 4, 51). Similarly, a recent study reported that a cut-off at 8 may assist in the identification of patients with chronic widespread pain and fibromyalgia and supported the use of tender point examination as a valid measure of pain hypersensitivity in the clinical setting (52). Some authors have also stated that the tender point count discriminated maximally at a count of ≥6, whereas the ACR criterion (≥11) displayed reduced sensitivity (4). Tender

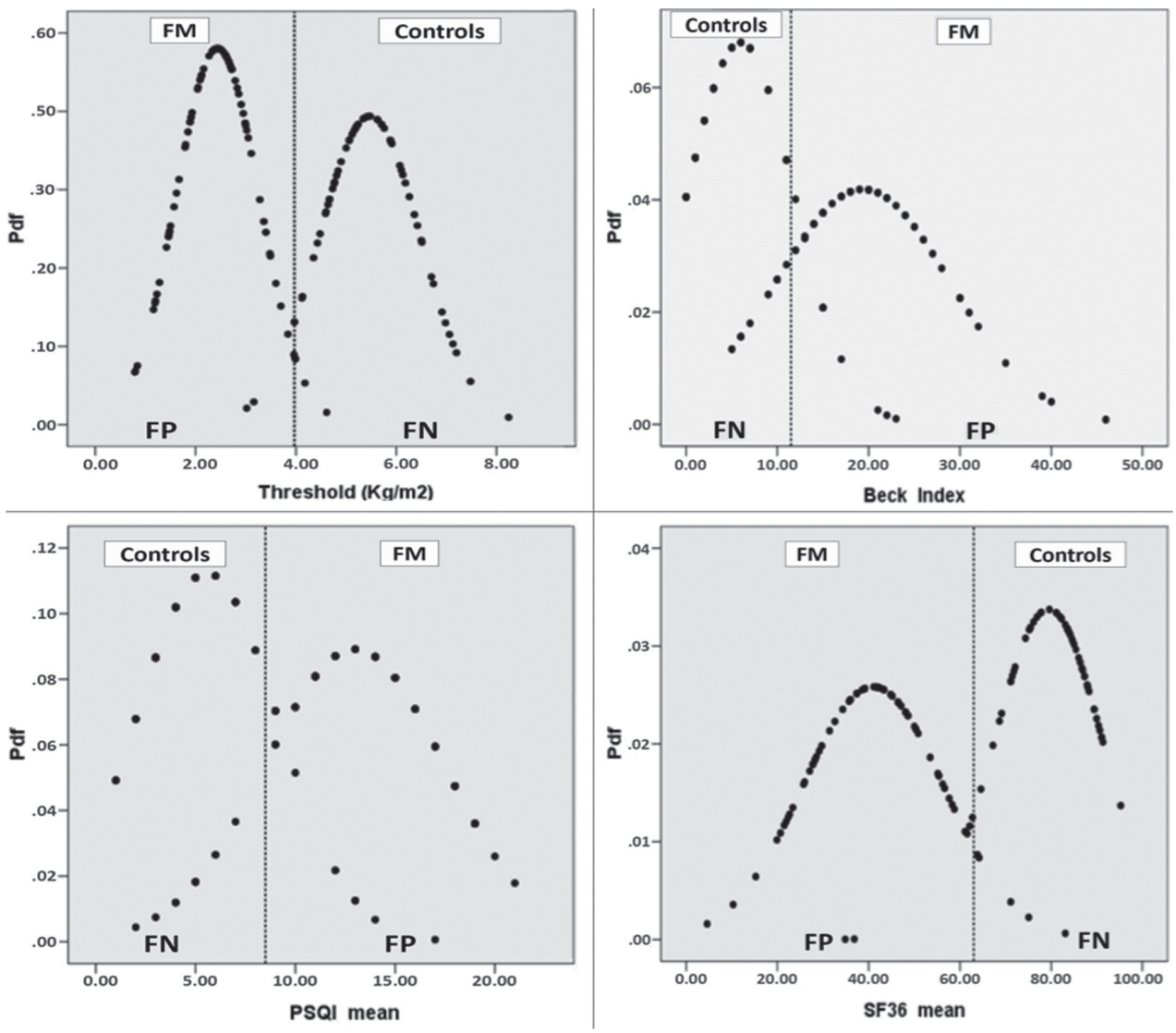

Fig. 2. Normal probability density function (Pdf) by group, of the continuous random variables (Threshold, Beck, PSQI and SF-36). Cut-off points were calculated for Threshold (PPT): 3.97 Kg/cm2; PSQI (8.5); SF36 (63.3) and BDI (11.5).
FN: False Negative fraction (FM subjects classified as negative); FP: False Positive fraction (controls classified as positive).
The other areas represent the diseases (FM sample) or controls correctly classified as negative.

points have been reported as being considerably influenced by distress, while dolorimetry results are only mildly influenced by this variable, thus being a more valid measure than the tender point count (53). Thus, pressure threshold measured with an algometer appears to be better than the tender point count for diagnosis of FM.

The individual analysis of each tender point revealed that the left trochanter is the anatomical localisation that best discriminates patients from controls. Nevertheless, we used the mean PPT across the 18 tender sites considered, because FM is clinically characterised by widespread reductions in pain thresholds, as opposed to lowered PPT localised at any specific body region. Therefore, the mean value should be a more representative measure of the widespread reductions in thresholds of experimental pain and of generalised clinical pain that occur in FM patients (54).

Having identified the variable with the maximum discriminative ability, *i.e.* mean PPT, the analyses enabled classification of symptoms in 4 groups, according to their AUC values and their relative differences from the reference variable.

Apart from the algometer scores, the first group (AUC>.90) included clinical pain measured by both the visual analogue scale and the SF-36 subscale body pain, confirming that pain assessment is central to FM diagnosis. Fatigue, loss of strength, memory loss and poorer quality of life accurately discriminated patients from healthy controls. This pattern of results is consistent with previous findings and suggests that FM may not be only defined by widespread pain (6, 9, 55, 56).

As suggested in previous findings, fatigue appears to be one of the core symptoms of FM and is probably the most limiting condition because of how it interferes with daily activities and affects mood (34, 44, 57, 58). In fact, FM subjects presented important functional limitations due to the disorder, both in

the physical and psychological domains, showing reduced general well-being assessed by the SF-36 questionnaire.

We also found that memory complaints differentiate well between groups and seem to be an important variable in FM diagnosis (59). As the healthy controls had higher levels of education than FM patients and as this variable may mask mild cognitive deficits (60), we performed additional analyses to clarify possible confounding effects and observed that the positive predictive value of memory loss was maintained.

The second group of variables were similar to those proposed by Wolfe and colleagues and other studies as symptoms accompanying the disorder (9, 11). In particular, for depression, anxiety, sleep disturbance, morning stiffness, hypersensitivity to stressful events and environmental sensitivity, and other symptoms such as headache or pain or cramp in the lower abdomen. The AUC was between 0.80 and 0.90. Of all these variables, depression showed the highest discriminative power (AUC=0.902). In this study, 79.4% of FM patients and 15.9% of controls scored above the 11.5 cut-off in the BDI. Previous studies have reported prevalence rates of depression in FM patients of between 20% and 80% (61-64) and even correlation between BDI scores of FM women and their spouses (65). Sleep quality also accurately discriminated between patients and controls. A cut-off >8.5 in the PSQI correctly classified most of the FM patients, the majority of whom were defined as poor sleepers. Non-restorative sleep has also been frequently reported in FM patients (66, 67).

A third group of variables included sensitivity to noise and cold, sadness and other symptoms such as paraesthesia, menstrual cramp, excessive swelling and abdominal dysregulation. Although these variables are not included as top domains, our study findings echoed some earlier descriptions of a subgroup of FM patients with elevated psychological distress (68), exaggerated adrenergic reactivity to chilling (69) and a perceptual style of amplification and noise intolerance (70, 71).

Finally, we found one group of variables with very low discriminative power: sensitivity to cold, frequent urination and body mass index. In relation to the latter parameter, this may have practical implications for health and in particular for FM. Previous research suggests that excessive weight is a common comorbidity factor and may be related to the severity of FM (15, 72-74). In the present study, we observed statistically significant differences in mean weight between groups but found that BMI differentiated poorly between FM patients and controls.

In summary, in line with previous studies (11, 28), tenderness (tender point count and PPT measured by dolorimeter), pain (SF-36 Body Pain and VAS), fatigue (SF-36 Vitality, VAS), quality of life and multidimensional function/ health related quality of life (SF-36 physical Component Scores) were the most accurate variables for diagnosis of FM. Interestingly, our findings lead us to conclude that tender point examination, and specially dolorimetry continue to be key factors in the clinical assessment of FM. Although pressure algometry provides quantitative measures, the scores are subjective and caution is especially advised when interpreting the results. Moreover, tenderness is not a uniform measure of the FM status and thus distinct subgroups of patients with fibromyalgia can be classified depending on the degree of tenderness (51).

As in other studies that have used SF-36, our findings highlighted that impairments in body pain and vitality are central features of FM (75), and they also stressed the importance of waking unrefreshed, cognitive symptoms, anxiety and depression, morning rigidity and susceptibility to stress, some of which were recently proposed as key domains in FM diagnosis (9).

Finally, the cut-offs obtained should be useful for planning future research to assess the test measures that maximise the probability of adequate diagnosis.

One strength of this study was the common and particular origin of the samples, which eliminates possible confounding aspects such as possible ethnic differences in the prevalence and severity of pain (76, 77). We also measured pressure pain threshold, in addition to the tender point count, in both subjects with FM and healthy participants, and assessed the possible confounding role of education level. Therefore, comparison with pain-free controls made possible to indirectly assess the state of nociceptive processing in fibromyalgia patients (46). The potential bias in the selection of controls was minimised, since this sample was just used to fulfill the role of a common reference, given that the main objective of this paper was to compare the discriminative power of the variables under study.

Because PPT was used as the reference standard, the procedure used to obtain this should not be compromised. In order to overcome possible biased estimations by treatment suppression, described in previous literature, we allowed patients to continue taking their medication (46). Given the huge difference between patients and controls, the analgesic effect might be considered of little relevance in the estimation of PPT.

However, there are some limitations to the study. Regarding the sample, although FM is more prevalent in women (60), male participants should be included in this type of study. Despite encouraging the selection of patients without overt *comorbid* disorders such as osteoarthritis, migraine headaches, irritable bowel syndrome, restless leg syndrome or temporomandibular pain disorder, it was almost impossible to rule out these conditions in our sample. In fact, some of the patients had symptoms that could meet the criteria for the above comorbidities. Some authors have recently claimed that the 18 predetermined sites of examination for tender points in fibromyalgia syndrome are also myofascial trigger points, and thus it should be recognised that a considerable overlap and/or interaction between myofascial pain syndrome and FM may exist (78-80). In this respect, the role of physical and psychiatric comorbidity continues to be a common challenge in the diagnosis of FM (81).

As suggestions for future research, the inclusion of a different clinical control group (*e.g.* patients with rheumatoid arthritis or unipolar depression) would help to identify symptoms and domains

that are specific to the disorder. Moreover, as pressure pain threshold may be biased by idiosyncratic factors, it would be useful to include objective measures of experimental pain such as assessment of nociceptive reflexes. Finally, a follow-up study should be carried out to check whether tender point examination is stable over time.

Although at the present moment the diagnosis of FM continues to be a challenge due to the lack of objective markers, studies of this type may be of great value in the clinical management of the disease.

Acknowledgements

This work was supported by the Spanish Ministry of Science and Innovation (ref.PSI2009-14555) and by the Spanish Ministry of Economy and Competitiveness (ref. PSI2013-45818-R).

References

1. YUNUS M: Fibromyalgia syndrome: new research on an old malady. *Brit Med* 1989; 298: 474-5.
2. WOLFE F, SMYTHE HA, YUNUS MB *et al*.: The American College of Rheumatology 1990 criteria for the classification of fibromyalgia: Report of the Multicenter Criteria Committee. *Arthritis Rheum* 1990; 33: 160-72.
3. WOLFE F: The relation between tender points and fibromyalgia symptom variables: evidence that fibromyalgia is not a discrete disorder in the clinic. *Ann Rheum Dis* 1997; 56: 268-71.
4. KATZ RS, WOLFE F, MICHAUD K: Fibromyalgia diagnosis: A comparison of clinical, survey, and American College of Rheumatology criteria. *Arthritis Rheum* 2006; 54: 169-76.
5. PETZKE F, GRACELY RH, PARK KM *et al*.: What do tender points measure? Influence of distress on 4 measures of tenderness. *J Rheumatol* 2003; 30: 567-74.
6. CROFFORD LJ, CLAUW DJ: Fibromyalgia: where are we a decade after the American College of Rheumatology classification criteria were developed? *Arthritis Rheum* 2002; 46: 1136-38.
7. CLAUW DJ, CROFFORD LJ: Chronic widespread pain and fibromyalgia: what we know, and what we need to know. *Best Pract Res Clin Rheumatol* 2003; 17: 685-701.
8. BIDARI A, GHAVIDEL-PARSA B, GHALEHBAGHI B: Reliability of ACR criteria over time to differentiate classic fibromyalgia from nonspecific widespread pain syndrome: a 6-month prospective cohort study. *Mod Rheumatol* 2009; 19: 663-9.
9. WOLFE F, CLAUW D, FITZCHARLES MA *et al*.: The American College of Rheumatology preliminary diagnostic criteria for fibromyalgia and measurement of symptom severity. *Arthritis Care Res* 2010; 62: 600-10.
10. BENNETT RM, JONES J, TURK DC *et al*.: An internet survey of 2596 people with fibromyalgia. *BMC Musculoskelet Disord* 2007; 8: 27.
11. MEASE PJ, ARNOLD LM, CHOY EH *et al*.: OMERACT Fibromyalgia Working Group. Fibromyalgia syndrome module at OMERACT 9: domain construct. *J Rheumatol* 2009; 36: 2318-29.
12. WILSON HD, ROBINSON JP, TURK DC: Toward the identification of symptom patterns in people with fibromyalgia. *Arthritis Rheum* 2009; 61: 527-34.
13. MEASE PJ, CLAUW DJ, ARNOLD LM *et al*.: OMERACT 7 Workshop Fibromyalgia Syndrome. *J Rheumatol* 2005; 32: 2270-7.
14. MEASE PJ, ARNOLD LM, BENNETT R *et al*.: Fibromyalgia syndrome (OMERACT 8). *J Rheumatol* 2007; 34: 1415-25.
15. MEASE PJ, ARNOLD LM, CROFFORD LJ *et al*.: Identifying the clinical domains of fibromyalgia: Contributions from clinician and patient Delphi exercises. *Arthritis Care Res* 2008; 59: 952-60.
16. BOOMERSHINE CS: A comprehensive evaluation of standardized assessment tools in the diagnosis of fibromyalgia and in the assessment of fibromyalgia severity. *Pain Res Treat*. Available at: http://www.hindawi.com/journals/prt/2012/653714/ Accessed 2012
17. ARNOLD LM, WILLIAMS DA, HUDSON JI *et al*.: Development of responder definitions for fibromyalgia clinical trials. *Arthritis Rheum* 2012; 64: 885-94.
18. WILLIAMS DA, ARNOLD LM: Measures of fibromyalgia: Fibromyalgia Impact Questionnaire (FIQ), Brief Pain Inventory (BPI), Multidimensional Fatigue Inventory (MFI-20), Medical Outcomes Study (MOS) Sleep Scale, and Multiple Ability Self-Report Questionnaire (MASQ). *Arthritis Care Res* (Hoboken) 2011; 63 (Suppl. 11): S86-97.
19. AARON LA, BUCHWALD D: Chronic diffuse musculoskeletal pain, fibromyalgia and comorbid unexplained clinical conditions. *Best Pract Res Clin Rheumatol* 2003; 17: 563-74.
20. GREENFIELD S, FITZCHARLES MA, ESDAILE JM: Reactive fibromyalgia syndrome. *Arthritis Rheum* 1992; 35: 678-81.
21. WHITE KP, SPEECHLEY M, HARTH M *et al*.: Comparing self-reported function and work disability in100 community cases of fibromyalgia syndrome versus controls in London, Ontario. *Arthritis Rheum* 1999; 42: 76-83.
22. BERNARD A, PRINCE A, EDSALL P: Quality of life issues for fibromyalgia patients. *Arthritis Care Res* 2000; 13: 42-50.
23. ARNOLD LM, CLAUW DJ, McCARBERG BH: Improving the recognition and diagnosis of fibromyalgia. Fibro Collaborative. *Mayo Clin Proc* 2011; 86: 457-64.
24. Wagner Force One (Model FDI) dolorimeter. Available at: http://www.wagnerinstruments.com/force_gauges/tension_meters_tension_gauges.php
25. COTT A, PARKINSON W, BELL MJ *et al*.: Interrater reliability of the tender point criterion for fibromyalgia. *J Rheumatol* 1992; 19: 1955-9.
26. BURCKHARDT CS, CLARK SR, BENNET RM: The Fibromyalgia Impact Questionnaire: development and Validation. *J Rheumatol* 1991; 18: 728-33.
27. MONTERDE S, SALVATA I, MONTULLA S *et al*.: Validación de la versión española del Fibromyalgia Impact Questionnaire. *Rev Esp Reumatol* 2004; 31: 507-13.
28. CHOY EH, ARNOLD LM, CLAUW DJ *et al*.: Content and criterion validity of the preliminary core dataset for clinical trials in fibromyalgia syndrome. *J Rheumatol* 2009; 36: 2330-4.
29. TALAMO J, FRATER A, GALLIVAN S *et al*.: Use of the short form 36 (SF36) for health status measurement in rheumatoid arthritis. *Br J Rheumatol* 1997; 36: 463-9.
30. WARE JE, KOSINSKI M, TURNER-BOWKER DM *et al*.: How to Score Version 2 of the SF-12® Health Survey (With a Supplement Documenting Version 1). Lincoln, RI: Quality Metric Incorporated 2002.
31. BIRTANE M, UZUNCA K, TAŞTEKIN N *et al*.: The evaluation of quality of life in fibromyalgia syndrome: a comparison with rheumatoid arthritis by using SF-36 Health Survey. *Clin Rheumatol* 2007; 26: 679-84.
32. VILAGUTA G, FERRERA M, RAJMIL L *et al*.: The Spanish version of the Short Form 36 Health Survey: a decade of experience and new developments. *Gac Sanit* 2005; 19: 135-50.
33. LEFAUCHEUR JP, DROUOT X, KERAVEL Y *et al*.: Pain relief induced by repetitive transcranial magnetic stimulation of pre-central cortex. *Neuroreport* 2001; 12: 2963-5.
34. CRAWFORD BK, PIAULT EC, LAI C *et al*.: Assessing fibromyalgia-related fatigue: content validity and psychometric performance of the Fatigue Visual Analog Scale in adult patients with fibromyalgia. *Clin Exp Rheumatol* 2011; 29 (Suppl 69): S34-43.
35. GASTON-JOHANSSON F: Measurement of pain: the psychometric properties of the Pain-O-Meter, a simple, inexpensive pain assessment tool that could change health care practices. *J Pain Symptom Manag* 1996; 12: 172-81.
36. FINSET A, WIGERS SH, GÖTESTAM KG: Depressed mood impedes pain treatment response in patients with fibromyalgia *J Rheumatol* 2004; 31: 976-80.
37. BOOMERSHINE CS, CROFFORD LJ: A symptom based approach to pharmacological management of fibromyalgia.http://www.ncbi.nlm.nih.gov/pubmed/19337283 *Nat Rev Rheumatol* 2009; 5: 191-9.
38. BECK AT, WARD CH, MENDELSON M *et al*.: An inventory for measuring depression. *Arch Gen Psychiat* 1961; 4: 561-71.
39. BUYSSE DJ, REYNOLDS CF, MONK TH *et al*.: The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res* 1989; 28: 193-213.
40. ROYUELA A, MACÍAS JA: Propiedades clinimétricas de la versión castellana del cuestionario de Pittsburgh. *Vigilia-Sueño* 1997; 9: 81-94.
41. SIERRA JC, DELGADO-DOMINGUEZ C, CARRETERO-DIOS H: [Internal structure of Dysfunctional Beliefs and Attitudes about Sleep Scale in a Spanish shift workers sample]. *Rev Neurol* 2006; 43: 454-60.
42. CRAMER H: Mathematical Methods of Statistics. Princeton: Princeton University 1946; Press, p282. ISBN 0-691-08004-6.

43. HANLEY JA, MCNEIL BJ: A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology* 1983; 148: 839-43.
44. HARTH M, NIELSON WR: The fibromyalgia tender points: use them or lose them? A brief review of the controversy. *J Rheumatol* 2007; 34: 914-22.
45. SCHWEINHARDT P, FITZCHARLES MA, BOOMERSHINE C *et al*.: Fibromyalgia as a Disorder Related to Distress and its Therapeutic Implications. *Pain Res Treat* 2012; 2 pages (Editorial).
46. BROWN CA, EL-DEREDY W, JONES AK: When the brain expects pain: common neural responses to pain anticipation are related to clinical pain and distress in fibromyalgia and osteoarthritis. *Eur J Neurosci* 2013; 12: 1-10.
47. DADABHOY D, CROFFORD LJ, SPAETH M *et al*.: Biology and therapy of fibromyalgia. Evidence-based biomarkers for fibromyalgia syndrome. *Arthritis Res Ther* 2008; 10: 211.
48. STAUD R: Biology and therapy of fibromyalgia: pain in fibromyalgia syndrome. *Arthritis Res Ther* 2006 ; 8: 208.
49. LEE Y, NASSIKAS N, CLAUW D: The role of the central nervous system in the generation and maintenance of chronic pain in rheumatoid arthritis, osteoarthritis and fibromyalgia. *Arthritis Res Ther* 2011; 13: 2-11.
50. WOLFE F, ROSS K, ANDERSON J *et al*.: The prevalence and characteristics of fibromyalgia in the general population. *Arthritis Rheum* 1995; 38: 19-28.
51. BUSKILA D: Developments in the scientific and clinical understanding of fibromyalgia. *Arthritis Res Ther*. 2009; 11: 242.
52. AMRIS K, WAEHRENS EE, JESPERSEN A *et al*.: The relationship between mechanical hyperalgesia assessed by manual tender point examination and disease severity in patients with chronic widespread pain: A cross-sectional study. *Int J Rheumatol* 2014; 2014: 417596.
53. GIESECKE T, WILLIAMS DA, HARRIS RE *et al*.: Sub-grouping of fibromyalgia patients on the basis of pressure-pain thresholds and psychological factors. *Arthritis Rheum* 2003; 48: 2916-22.
54. HOOTEN WM, ROSENBERG CJ, ELDRIGE JS *et al*.: Knee extensor strength is associated with pressure pain thresholds in adults with fibromyalgia. *PLoS One* 2013; 8: e59930.
55. GOLDENBERG DL: Diagnosis and differential diagnosis of fibromyalgia. *Am J Med* 2009; 122: S14-21.
56. KIM SM, LEE SH, KIM HR: Applying the ACR preliminary diagnostic criteria in the diagnosis and assessment of fibromyalgia. *Korean J Pain* 2012; 25: 173-82.
57. WOLFE F, HAWLEY DJ, WILSON K: The prevalence and meaning of fatigue in rheumatic disease. *J Rheumatol* 1996; 23: 1407-17.
58. YAMAMOTO K, SHIOTA S, OHNO S *et al*.: A diagnosis of depression should be considered in patients with multiple physical symptoms in primary care clinics. *Tohoku J Exp Med* 2013; 229: 279-85.
59. SCHMIDT-WILCKE T, WOOD P, LÜRDING R: Cognitive impairment in patients suffering from fibromyalgia. an underestimated problem. *Schmerz* 2010; 24: 46-53.
60. STERN Y: Cognitive reserve. *Neuropsychologia* 2009; 47: 2015-28.
61. FIETTA P, MANGANELLI P: Fibromyalgia and psychiatric disorders. *Acta Biomed* 2007; 78: 88-95.
62. ARNOLD LM, CROFFORD LJ, MARTIN SA *et al*.: The effect of anxiety and depression on improvements in pain in a randomized, controlled trial of pregabalin for treatment of fibromyalgia. *Pain Med* 2007; 8: 633-8.
63. BOYER AL, MIRA MAP, CALATAYUD NP *et al*.: Comparing fibromyalgia patients from primary care and rheumatology settings: clinical and psychosocial features. *Rheum Int* 2009; 29: 1151-60.
64. DOS SANTOS EB, QUINTANS JUNIOR LJ, FRAGA BP *et al*.: An evaluation of anxiety and depression symptoms in fibromyalgia. *Rev Esc Enferm USP* 2012; 46: 590-6.
65. YENER M, ASKIN A, SOYUPEK F *et al*.: The evaluation of anxiety and depression status in spouses of sexually active reproductive women with fibromyalgia. *Clin Exp Rheumatol* 2015; 33 (Suppl. 88): 20-4.
66. AMERICAN ACADEMY OF SLEEP MEDICINE: International classification of sleep disorders: Diagnostic and coding manual. 2.ed. Westchester, IL: American Academy of Sleep Medicine; 2005.
67. THEADOM A, CROPLEY M, HUMPHREY KL: Exploring the role of sleep and coping in quality of life in fibromyalgia. *J Psychosom Res* 2007; 62: 145-51.
68. VERBUNT JA, PERNOT DH, SMEETS RJ: Disability and quality of life in patients with fibromyalgia. Health Qual Life Outcomes. Available at: http://www.hqlo.com/content/6/1/8. Accessed January 22 2008
69. BENNETT RM, CLARK SR, CAMPBELL SM *et al*.: Symptoms of Raynaud's syndrome in patients with fibromyalgia. A study utilizing the Nielsen test, digital photoplethysmography, and measurements of platelet α2-adrenergic receptors. *Arthritis Rheum* 1991; 34: 264-9.
70. MCDERMID AJ, ROLLMAN GB, MCCAIN GA: Generalized hypervigilance in fibromyalgia: evidence of perceptual amplification. *Pain* 1996; 66: 133-44.
71. CARRILLO-DE-LA-PEÑA MT, VALLET M, PÉREZ MI, GÓMEZ-PERRETTA C: Intensity dependence of auditory-evoked cortical potentials in fibromyalgia patients: a test of the generalized hypervigilance hypothesis. *J Pain* 2006; 7: 480-7.
72. YUNUS MB, ARSLAN S, ALDAG JC: Relationship between body mass index and fibromyalgia features. *Scand J Rheumatol* 2002; 31: 27-31.
73. OKIFUJI A, BRADSHAW DH, OLSON C: Evaluating obesity in fibromyalgia: neuroendocrine biomarkers, symptoms, and functions. *Clin Rheumatol* 2009; 28: 475-78.
74. KIM CH , LUEDTKE CA, VINCENT A *et al*.: Association of body mass index with symptom severity and quality of life in patients with fibromyalgia. *Arthrit Care Res* 2012; 64: 222-8.
75. HOFFMAN DL, DUKES EM: The health status burden of people with fibromyalgia: a review of studies that assessed health status with the SF-36 or the SF-12. *Int J Clin Pract* 2008; 62: 115-26.
76. EDWARDS RR, DOLEYS DM, FILLINGIM RB *et al*.: Ethnic Differences in pain tolerance: Clinical implications in a chronic pain. *Psychosom Med* 2001; 63: 316-23.
77. CAMPBELLA CM, EDWARDS RR, ROGER B *et al*.: Ethnic differences in responses to multiple experimental pain stimuli. *Pain* 2005; 113: 20-6.
78. CAKIT BD, TASKIN S, NACIR B, UNLU I, GENC H, ERDEM HR: Comorbidity of fibromyalgia and cervical myofascial pain syndrome. *Clin Rheumatol* 2010; 29: 405-11.
79. GE HY: Prevalence of myofascial trigger points in fibromyalgia: the overlap of two common problems. *Curr Pain Headache Rep* 2010; 14: 339-45.
80. GE HY, WANG Y, DANNESKIOLD-SAMSOE B, GRAVEN-NIELSEN T, ARENDT-NIELSEN L: The predetermined sites of examination for tender points in fibromyalgia syndrome are frequently associated with myofascial trigger points. *J Pain* 2010; 11: 644-51.
81. GIACOMELLI C, SERNISSI F, SARZI-PUTTINI P, DI FRANCO M, ATZENI F, BAZZICHI L: Fibromyalgia: a critical digest of the recent literature. *Clin Exp Rheumatol* 2013; 31 (Suppl. 79): S153-7.