What Do Tender Points Measure? Influence of Distress on 4 Measures of Tenderness

FRANK PETZKE, RICHARD H. GRACELY, KAREN M. PARK, KIRSTEN AMBROSE, and DANIEL J. CLAUW

ABSTRACT. ***Objective.*** To examine the relationship between current pain, distress, and ascending and random measures of tenderness.

Methods. Manual tender point counts and dolorimeter measures of the pressure pain threshold were determined in a sample of 47 women representative of the general population with respect to tenderness. In addition, discrete pressure stimuli of varying intensities to the left thumb were applied in random fashion. Distress was measured with the Brief Symptom Inventory and the Beck Depression Inventory, and pain was evaluated with the Short Form McGill Pain Questionnaire.

Results. Only the random measure of tenderness was relatively independent of an individual's current psychological state. The respective correlation coefficients between measures of tenderness and psychological state were generally greatest for the manual tender point count and also significant for the dolorimeter measures. In contrast, all measures were highly correlated with ratings of spontaneous pain, again with the manual tender point count showing the strongest, and the random method the weakest, correlations. Linear regression analysis replicated the results of the correlational analysis.

Conclusion. As a measure of tenderness, the number of positive tender points is clearly influenced by an individual's distress. Other more sophisticated measures of tenderness that randomly present stimuli in an unpredictable fashion appear to be relatively immune to these biasing effects, although our results obtained in a research setting have yet to be replicated in clinical practice. (J Rheumatol 2003;30:567–74)

Key Indexing Terms:

TENDER POINTS	PAIN	TENDERNESS
PSYCHOPHYSICAL	DISTRESS	FIBROMYALGIA

To fulfill the criteria for fibromyalgia (FM) established by the American College of Rheumatology (ACR) in 1990, an individual must have both a history of chronic widespread pain involving all 4 quadrants of the body (and axial skeleton) and the presence of 11 of 18 "tender points" on physical examination[1]. These criteria were never intended to be strictly applied to individual patients as diagnostic criteria, and many persons with the clinical diagnosis of FM do not fulfill this definition.

The validity of the construct of tender points is one contentious aspect of the ACR definition[2]. A tender point is defined as an anatomic site where an individual complains of pain when approximately 4 kg of pressure is applied (about the amount of pressure required to blanch the examiner's nail). Although early studies suggested that patients with FM experienced tenderness only in these discrete regions, recent data show that individuals with FM display increased sensitivity to pain throughout the entire body[3]. The presence of tender points is not inherently abnormal since many people have some tender points, with the mean value in the general population ranging from one to 4, depending on the methodologies employed[4,5]. Tender points (e.g., mid-trapezius region, epicondyles, etc.) merely represent regions of the body where the general population is more tender. Thus, individuals who are more diffusely tender will generally have a greater number of tender points.

Using tender points as diagnostic criteria presents several potential problems. For example, it is now clear that the requirement for having 11 of 18 tender points to fulfill the ACR FM criteria is largely responsible for FM being a condition that is exceedingly more prevalent in women. The other component of the ACR definition, chronic pain in all 4 quadrants of the body plus the axial skeleton, only occurs in about 1.5 times as many women as men in the population. But women are roughly 10 times more likely than men to have 11/18 tender points[6]. Thus, any criterion that requires a

From the Chronic Pain and Fatigue Research Center, Georgetown University Medical Center, and the Department of Anesthesiology, George Washington University Medical Center, Washington, DC, and the Clinical Measurement and Mechanisms Unit, PNMB, NIDCR, National Institutes of Health, Bethesda, Maryland, USA.

Supported in part by Department of the Army grant DAMD 17-00-0018. Dr. Petzke was supported by a grant of the Deutsche Forschungsgemeinschaft (Pe 713/1-1).

F. Petzke, MD; K. Ambrose, MS; D.J. Clauw, MD, Chronic Pain and Fatigue Research Center, Georgetown University Medical Center; K.M. Park, MD, Department of Anesthesiology, George Washington University Medical Center; R.H. Gracely, PhD, Clinical Measurement and Mechanisms Unit, PNMB, NIDCR, National Institutes of Health.

Address reprint requests to Dr. D.J. Clauw, The University of Michigan Medical Center, 101 Simpson, Ann Arbor, MI 48109-0723, USA. E-mail: dclauw@med.umich.edu

Submitted January 29, 2002; revision accepted September 4, 2002.

certain number of tender points to establish the diagnosis of FM will be skewed towards identifying primarily women.

A more fundamental problem with tender points is that the number of such points is not a good measure of tenderness, if tenderness is defined as an individual's pressure pain threshold. For example, several population-based studies have shown that the number of positive tender points is highly correlated with a number of measures of distress[4,7]. However, distress is not necessarily consistently related to tenderness, since performing a tender point determination concurrently with evaluation of pressure pain thresholds using dolorimetry/algometry (a pressure gauge with a rubber probe attached), the tender point count correlates better with measures of distress than the pressure pain threshold[7].

Although dolorimetry provides an objective measure of the pressure pain threshold that is relatively independent of distress, it may still be vulnerable to factors that bias the outcome of the evaluation. Both tender point counts and dolorimeter determinations can be classified as an ascending evoked pain paradigm. In such paradigms, the stimulus intensity (applied either continuously or as discrete stimuli) is predictably increased, and the individual is asked to rate the intensity of the stimuli, either with a numerical or verbal descriptor, or as a dichotomous response (e.g., no pain or pain). These ascending paradigms are known to be vulnerable to response biases and to factors such as psychological status and reaction time[8,9]. Thus, these ascending paradigms may not generally give an objective measure of a subject's pain threshold. In particular, expectancy of a painful stimulus on the part of the patient as well as a generalized hypervigilance towards potentially noxious stimuli might play a prominent role in evoked pain measurements using predictable, ascending paradigms. Recent data even suggest that the examiner may be biased by a similar expectancy effect, especially in a scenario in which the stimulus is manually applied and controlled, as is the case for dolorimetry or for the manual palpation in the tender point count, where results are easily influenced by rate of stimulus application, examiner cues, etc.[10].

Many of the factors that may influence the outcome of ascending methods of pain testing can be minimized by experimental methods. The presentation of a randomized sequence of preselected stimuli of varying intensity, as used in fixed-stimuli or direct scaling paradigms, present stimuli in the range between pain threshold and tolerance, and require subjects to rate perceived pain intensity on specified scales. The results are expressed as stimulus response curves that express the magnitude of evoked pain sensation over a broad range of stimulus intensities. This methodology has been used by Bendtsen, *et al*, who applied predetermined manual pressure stimuli in random order to the trapezius and temporal muscle of FM patients and healthy controls. Their results showed a shift and change of the stimulus response curve toward higher pain ratings or increased tenderness in patients for the tender muscle (trapezius) only[11]. Kosek, *et al* used a similar method of individually chosen fixed stimuli for heat pain stimuli, and observed a shift of the stimulus response curve for all stimuli towards higher pain ratings or increased heat pain sensitivity in patients with FM compared to controls, irrespective of the site tested[12].

However, no study has evaluated the relative influence of factors such as distressed mood on ascending and random paradigms. To assess these effects, we developed a device to randomly administer discrete pressure stimuli, and compared the results obtained from this measure to the results obtained from tender point counts and dolorimetery. We compared these determinations to each other and to measures of mood and distress. In order not to bias these correlational analyses with extreme scores, we constructed a sample of patients with FM and of healthy volunteers that spanned the range of tenderness observed in the general population, with a normal distribution of tenderness as measured by dolorimetry. Because of the strong influence gender has on tenderness, we included only women in our sample.

MATERIALS AND METHODS

Study participants. Patient recruitment for the study included consecutive female clinic patients at the Georgetown University Medical Center between June and October 1998 presenting with both regional and widespread pain, and with an established diagnosis of FM. They were invited to participate in the study via mail a few weeks before a scheduled clinic visit. If they agreed to participate, a testing session was arranged after the scheduled clinic visit. Patients with concurrent inflammatory rheumatic conditions and serious medical conditions were excluded.

Controls were recruited through flyers and newspaper advertisements and compensated for their participation. Women with chronic medical conditions, longterm medications, or who were pregnant or breastfeeding were excluded from the control group.

Patients who agreed to participate in the study were allowed to continue their regular medication. However, all participants were advised not to take any opioid or non-opioid analgesics for 24 h prior to the testing session including over-the-counter analgesics. Prior to participation, patients and controls signed a consent form. The consent form and protocol were approved by the Georgetown University Medical Center Institutional Review Board.

Psychophysical testing. All subjects initially underwent a manual tender point count (MTPC) and were then familiarized with the testing environment. The pain testing equipment was shown and explained using a scripted text. Instructions for completing the different self-report measures also followed a standardized script. Additional information and explanations were provided if required.

Following familiarization, 3 further measures of tenderness were obtained: (1) mean pressure pain threshold at tender points determined by standard dolorimetry (Dolor-TP); (2) mean pressure pain threshold at the right and left thumb nailbeds using the same dolorimeter method (Dolor-THU); and (3) the area under the curve for a randomized fixed stimulus paradigm in which subjects rated the intensity of suprathreshold sensations evoked by discrete pressure stimuli applied to the left thumb nailbed (RAN). The time interval between the different testing procedures (MTPC, Dolor-TP, Dolor-THU, RAN) was 15 to 20 min.

MTPC. The MTPC is the most widely used measure of tenderness in clinical practice[1]. The method followed the ACR guidelines; the operator used

the dominant thumb to slowly increase pressure at a constant rate (1 kg/s) up to a maximum of 4 kg at 18 defined sites. Subjects were asked to state whether this pressure was painful and the number of painful sites was recorded.

Dolorimetry. Dolorimetry represents the most widely used method of pressure pain threshold assessment in FM research[6]. However, the protocols vary and there is no universal standard. In this study all participants received a dolorimeter examination at the 18 defined tender points, and at 4 control points (bilateral thumbs and anterior tibial muscles) using a standard dolorimeter (Chatillon). The same sequence and patient positioning instructions were used as for the MTPC. Pressure was increased at a rate of 1 kg/s using a 3.14 cm^2 hard rubber circular probe. Subjects were instructed to indicate when they first perceived pain. Pressure was increased up to 12 kg if necessary and 12 kg was recorded as pain threshold if pain was not reported. The mean tender point pain threshold for the 18 sites (Dolor-TP), and the mean pressure pain threshold for the 2 thumbs (Dolor-THU) were calculated for each subject.

Randomized fixed stimulus paradigm. Discrete pressure stimuli of 5 s duration were applied to the thumbnail by a 1 cm^2 hard rubber circular probe. The stimulator was positioned over the thumb by a plastic housing. A hydraulic system was activated by calibrated weights placed on a nonfixed platform. Valves controlled stimulus timing. The combination of valves and calibrated weights produced controlled, repeatable stimuli that approached a rectangular waveform.

The range of individual tolerance was determined by stimulation of the right thumb. Stimulation pressure for each subsequent stimulus was increased in 0.45 kg increments from a starting point of 0.45 kg to either pain tolerance or a maximum of 9.1 kg. The interstimulus interval was roughly 30 s. After each stimulus individual pain ratings were recorded on a 21-box combined numerical analog descriptor scale[9].

Using a fixed stimulus paradigm, a series of pressure stimuli within each subject's defined range were then applied in a random order to the left thumb. Subjects were told they would be receiving a different series of stimuli within the range of the previous series. Up to 7 stimuli (0.45, 0.91, 1.36, 1.82, 2.73, 3.64, and 4.54 kg) were twice presented in random order and ratings recorded. Preliminary psychophysical testing had indicated that this weight distribution led to at least 3 or 4 values that would fall between pain threshold and tolerance in most subjects. A measure of pain sensitivity was computed from the area under the curve defined by the mean of the 2 respective pain ratings for each stimulus (RAN).

Pain. General information on presence, onset, duration, and location of pain was collected on all subjects. Subjects completed the Short Form of the McGill pain questionnaire (SF-MG)[13], the 6-point verbal present pain intensity (PPI), and a visual analog scale for current pain intensity. In addition, a regional pain score was computed for each subject. The regional pain score was developed for patients with FM and constitutes the sum of pain intensities at 21 predefined sites on a body map, each rated using a 6-point (0–5) Likert scale. Because of high intercorrelations between all the different pain measures of clinical pain (Table 3), only the SF-MG pain questionnaire was included in the secondary analysis.

Distress. The Brief Symptom Inventory (BSI) was used as a measure of distress[14]. This 51-item instrument contains 9 subscales (somatization, obsessive-compulsive behavior, interpersonal sensitivity, depression, anxiety, hostility, phobic anxiety, paranoid ideation, and psychotism) and a Global Severity Index (GSI), which was used as a general measure for distress.

Depression. The Beck Depression Inventory (BDI) is a 21-item measure of the severity of current depressive symptoms, including both neurovegetative and cognitive symptoms of depression, and has been well validated in rheumatic diseases[15]. It was included as a secondary and more focused measure of negative affect.

Subsample selection. Subjects were grouped according to the duration and extent of their pain complaints and number of tender points: (1) healthy controls without any pain reports over the past month; (2) healthy controls with some pain report over the last month; (3) subjects with chronic regional pain syndromes of greater than 3 months' duration; (4) subjects with chronic widespread pain of greater than 3 months' duration with less than 11 tender points; and (5) patients with FM.

After completion of the testing phase a sample of the study population was randomly selected to approach a normal distribution of tenderness similar to the distribution of tenderness observed in the general population. Tenderness for selection was based on dolorimetry pain thresholds at tender points, since at present dolorimetry represents the gold standard of measurement of tenderness and is known to be normally distributed in the general population.

Statistical analysis. All statistical analyses were performed with SPSS 9.0 and MS Excel. Data are displayed as mean ± SEM unless indicated. Normality was evaluated by displaying the data as histograms with an overlying normal distribution curve and normal probability, with Q–Q plots, and tested statistically by Kolmogorov-Smirnov and Shapiro Wilk's tests. Pearson correlation coefficients were calculated between measures of tenderness and the various measures of distress, negative affect, and pain, as well as age. Partial correlation coefficients were calculated for the relation between the measures of tenderness controlling for distress. Age, pain intensity, and distress were entered stepwise as independent variables and the 4 measures of tenderness constituted the dependent variables in multiple linear regression analysis.

RESULTS

Study population. The sample of 47 subjects was selected out of 79 complete data sets of women, including 28 healthy controls with no pain, 3 healthy controls with pain, 3 patients with regional pain, 6 patients with chronic widespread pain, and 7 patients with FM (Table 1). The mean age of the population sample was 48.3 ± 1.3 years (range 22–68, median 49 yrs).

Normality. Both the distribution of tenderness scores determined by the dolorimetry pressure pain threshold at tender points (Dolor-TP) and the distribution of distress scores from the GSI of the BSI were evaluated for normality. The respective histograms and Q–Q plots are displayed in Figures 1 and 2. The distribution of Dolor-TP scores satisfied criteria for normality (Kolmogorov-Smirnov $p > 0.2$; Shapiro-Wilk $p > 0.67$). The distribution of GSI scores approached normality with less goodness of fit (Kolmogorov-Smirnov $p > 0.02$; Shapiro-Wilk $p > 0.09$) due to a high number of low scores.

Correlational analysis. The correlation coefficients for the measures of tenderness and the measures of distress and negative affect are summarized in Table 2. All 3 ascending paradigms, MTPC, Dolor-TP, and Dolor-THU, were associ-

Table 1. Selection of study population out of a total sample of 79 subjects.

Subject Groups	Total Sample n = 79 (%)	Population Sample n = 47 (%)
Healthy controls — no pain	28 (35)	28 (59.5)
Healthy controls — pain	3 (4)	3 (6.5)
Regional pain	3 (4)	3 (6.5)
Chronic widespread pain	6 (8)	6 (13)
Fibromyalgia	39 (49)	7 (14.5)

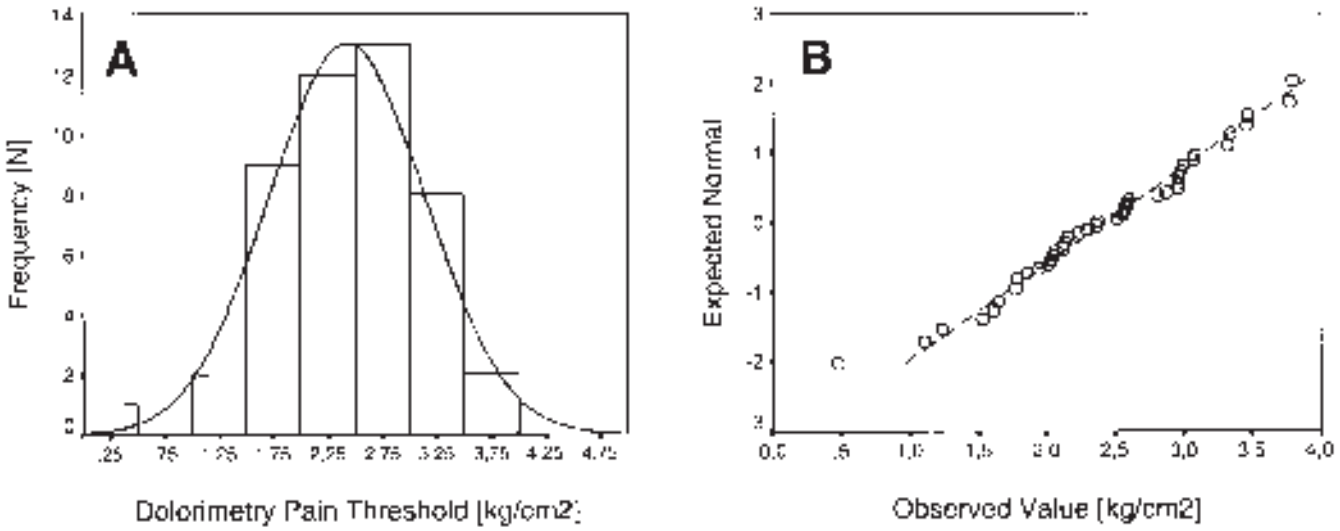

Figure 1. A. Histogram of the distribution of dolorimetry pain thresholds overlayed with a normal distribution. B. Q–Q or normal probability plot for dolorimetry pain threshold. For each data point the Q–Q plot shows the observed value and the value that is expected if the data were a sample from a normal distribution. The points cluster around a straight line if the data are from a normal distribution.

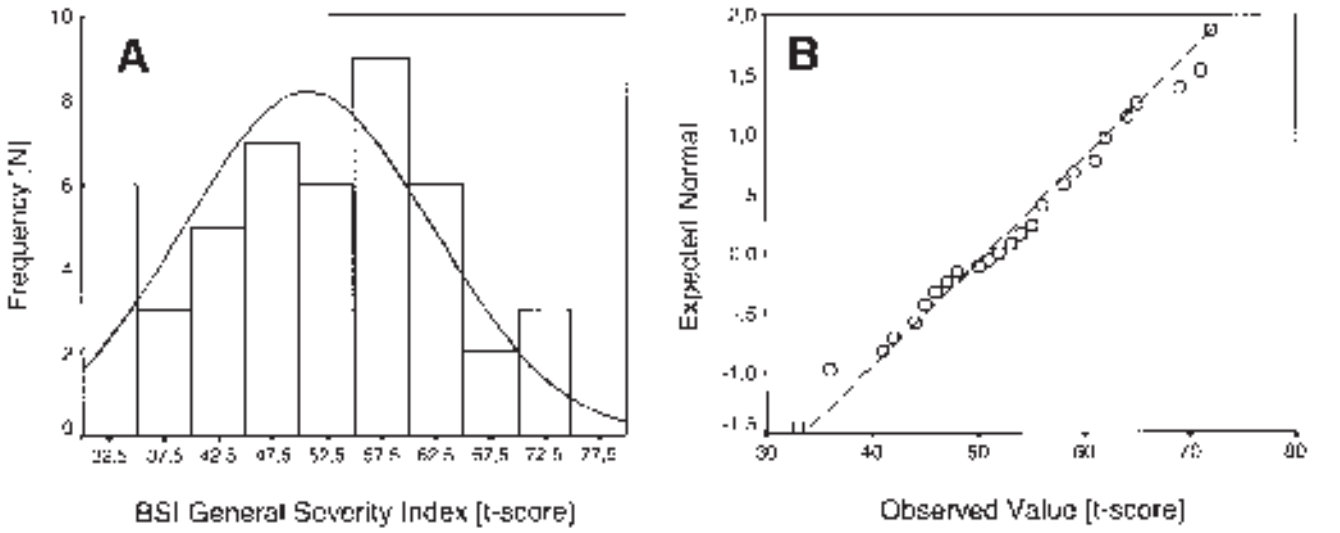

Figure 2. A. Histogram for distribution of BSI General Severity Index overlayed with a normal distribution. B. Q–Q or normal probability plot for BSI General Severity Index. For explanation see legend of Figure 1B.

ated with measures of distress and negative affect, while RAN showed essentially no association (with the exception of the BSI depression subscale). Overall, MTPC showed the most significant association with the distress measures, although all 3 ascending measures were generally associated with several measures of distress in the same order (MTPC > Dolor-TP > Dolor-THU).

The correlation coefficients for all measures of tenderness and the measures of clinical pain are displayed in Table 3. All correlations were significant and showed the same order of association as with measures of distress and negative affect. The partial correlation coefficients for the different measures of tenderness adjusting for distress are displayed in Table 4; the respective simple correlation coefficients are included in Table 5. All partial and simple correlations were significant, and only minimally different. As expected, there was a strong association between the MTPC and the dolorimetry measures, and a weak association between RAN and the other measures.

Linear regression. The correlation matrix for the variables included in the regression analysis is shown in Table 5. Age was excluded as it did not significantly correlate with any other measure. A stepwise linear regression model with clinical pain (SF-MG) and distress (GSI) as independent vari-

Table 2. Correlation coefficients for measures of tenderness and measures of distress and negative affect (BSI subscales, Global Severity Index, and BDI).

		MTPC	Dolor-TP	Dolor-THU	RAN
Somatization	r	0.51	–0.47	–0.33	0.15
	p	0.0003	0.0008	0.023	0.33
Obsessive-compulsive	r	0.42	–0.35	–0.30	0.00
	p	0.003	0.015	0.04	0.98
Interpersonal sensitivity	r	0.44	–0.46	–0.24	0.21
	p	0.002	0.001	0.10	0.16
Anxiety	r	0.30	–0.29	–0.18	0.16
	p	0.04	0.05	0.23	0.27
Hostility	r	0.42	–0.31	–0.24	0.14
	p	0.004	0.04	0.09	0.36
Depression	r	0.35	–0.43	–0.39	0.44
	p	0.02	0.003	0.007	0.002
Phobic anxiety	r	0.40	–0.27	–0.02	–0.01
	p	0.005	0.07	0.87	0.97
Paranoid ideation	r	0.23	–0.20	–0.08	0.11
	p	0.11	0.17	0.60	0.46
Psychotism	r	0.13	–0.11	0.02	0.02
	p	0.40	0.47	0.92	0.91
Global Severity Index	r	0.48	–0.38	–0.34	0.20
	p	0.0006	0.008	0.02	0.17
Beck Depression Inventory	r	0.39	–0.29	–0.33	0.23
	p	0.006	0.05	0.02	0.11

Table 3. Correlation coefficients for measures of tenderness and measures of pain intensity.

		MTPC	Dolor-TP	Dolor-THU	RAN
Regional Pain Score	r	0.68	–0.55	–0.55	0.38
	p	0.0001	0.0001	0.0001	0.008
VAS	r	0.80	–0.65	–0.55	0.47
	p	0.0001	0.0001	0.0001	0.0008
SF-MG PPI	r	0.68	–0.55	–0.45	0.39
	p	0.0001	0.0001	0.0013	0.008
SF-MG Total	r	0.60	–0.50	–0.36	0.38
	p	0.0001	0.0004	0.013	0.009

VAS: visual analog scale.

Table 4. Partial correlation matrix for the different measures of tenderness controlling for the BSI-GSI score.

		MTPC	Dolor-TP	Dolor-THU	RAN
MTPC	r	—			
	p				
Dolor-TP	r	–0.67	—		
	p	0.0001			
Dolor-THU	r	–0.52	0.70	—	
	p	0.0001	0.0001		
RAN	r	0.32	–0.36	–0.46	—
	p	0.031	0.014	0.001	

ables showed that only clinical pain contributed significantly to the variance of all 4 measures of tenderness (MTPC: $R^2 = 0.36$ > Dolor-TP: $R^2 = 0.25$, > Dolor-THU: $R^2 = 0.13$ = RAN: $R^2 = 0.14$; see Table 6). Including only distress in the model explained significant portions of the variance of the 3 ascending measures (with MTPC: $R^2 = 0.23$ > Dolor-TP: $R^2 = 0.15$ > Dolor-THU: $R^2 = 0.12$). In contrast, distress did not significantly contribute to the variance of the randomized fixed stimulus paradigm measure (RAN: $R^2 = 0.04$; see Table 7).

Table 5. Correlation matrix for variables included in the linear regression analysis.

		MTPC	Dolor-TP	Dolor-THU	RAN	BSI-GSI	SF-MG Total	Age
MTPC	r							
	p	—						
Dolor-TP	r	–0.73						
	p	0.0001	—					
Dolor-THU	r	–0.59	0.74					
	p	0.0001	0.0001	—				
RAN	r	0.35	–0.33	–0.47				
	p	0.015	0.022	0.0008	—			
BSI-GSI	r	0.48	–0.38	–0.34	0.20			
	p	0.0006	0.008	0.02	0.17	—		
SF-MG Total	r	0.60	–0.50	–0.36	0.37	0.57		
	p	0.0001	0.0004	0.013	0.009	0.0002	—	
Age	r	–0.09	–0.06	–0.10	0.02	0.13	0.10	
	p	0.54	0.68	0.53	0.88	0.37	0.51	

Table 6. Results of regression analysis (R-square, regression coefficient B, standard error of B, significance, 95% confidence interval for B) for the 4 dependent tenderness measures with pain (SF-MG) and distress (GSI) as independent variables, with only pain being a significant contributor.

	R^2	B	SE	p	95% CI Lower	95% CI Upper
MTPC	0.36	0.57	0.11	< 0.0001	0.34	0.79
Dolor-TP	0.25	–0.06	0.015	< 0.0001	–0.09	–0.03
Dolor-THU	0.13	–0.16	0.063	0.013	–0.29	–0.04
RAN	0.14	2.06	0.75	0.009	0.54	3.58

Table 7. Results of regression analysis (R-square, regression coefficient B, standard error of B, significance, 95% confidence interval for B) for the 4 dependent tenderness measures with distress (GSI) as independent variable.

	R^2	B	SE	p	95% CI Lower	95% CI Upper
MTPC	0.23	0.24	0.06	0.001	0.11	0.37
Dolor-TP	0.15	–0.02	0.009	0.008	–0.04	–0.007
Dolor-THU	0.12	–0.08	0.033	0.019	–0.15	–0.014
RAN	0.04	0.59	0.42	0.169	–0.26	1.44

DISCUSSION

These preliminary data suggest that measures derived from the randomized stimulus paradigm evaluation of pressure pain sensitivity were relatively independent of an individual's current psychological state. These data suggest that tenderness *per se* is not necessarily associated with distress, even though the number of tender points certainly is. Our study also replicates previous findings in population-based samples showing that dolorimeter determinations are less influenced by psychological factors than tender point counts[3,4,6].

We found a relatively consistent relationship between these 4 measures of tenderness, and again this corroborates what has been noted in population-based studies. The MTPC correlated most strongly with measures of psychological distress, with dolorimetery less strongly related to these measures. The dolorimetry value at the thumbnail was generally less influenced by psychological factors than the mean value from all 18 tender points, perhaps because this is a neutral site, where subjects rarely have clinical pain and have no expectation of tenderness. In all calculations, the random pain measures were the least influenced by the current psychological state of the subject.

These data obviously do not indicate that random measures of pressure pain threshold are superior to tender point counts. Indeed, clinical pain ratings were related to the

various tenderness measures in the same order as distress measures, with the MTPC being most highly related, followed by dolorimetry and the random measures. What these data do suggest is that the random pressure method is measuring something quite different from the tender point count measures, and somewhat different from a dolorimeter examination performed at all 18 tender point measures.

This study was not designed to evaluate the relationship between clinical pain magnitude and sensitivity to painful pressure. Arguably the best measure of the validity of these new random tenderness measures will be in clinical trials. An ideal measure of tenderness would be one that improves in conjunction with improvements in clinical pain (something not necessarily observed with either tender points or dolorimetry), and does not change when there is an independent change in psychological state. Such an ideal measure of tenderness should also be stable and reliable, characteristics that also need to be addressed in further studies. This measure then could be used in the multidimensional assessment of various, although mostly musculoskeletal, pain states as a valid indicator of physiologic pain sensitivity.

Certain choices in study design may have resulted in potential limitations of this study. We chose to use only women in our sample because men are much less tender than women and combining sexes would have required a much larger sample size to obtain meaningful data. Comparison among the different measures of tenderness is potentially complicated by the differing scales (an ordinal scale from 0 to 18, a continuous scale from 0 to 12, and an area under the curve) as well as by the different mode of stimulus application (thumb pressure, pressure gauge with a 3.14 cm^2 footplate, stimulator with 1 cm^2 footplate). However, in a study by Smythe, *et al*[16], dolorimeter footplate size had a significant effect on mean values but resulted in almost identical slopes and intercepts in regression analysis. Similarly, scale length only had relevant distorting effects when the scale maximum was less then 11. Thus the use of differing scales and stimulus application likely does not challenge the validity of the relationships between each of these measures of tenderness and the measures of psychological status. Finally, there is no consensus on a scale of distress, although the BSI has been used in many previous studies of this construct, and is a well validated instrument.

One final potential criticism is that our selection of the subsample could have been biased. We included all the healthy controls, with or without pain, that we recruited, all of the patients with chronic widespread pain, and 7 patients with FM. We based this subset on a normalized curve for dolorimetry. The 7 patients with FM roughly reflect the prevalence of FM in the female population in this age range.

Our results add strength to the notion that requiring that an individual have a certain number of tender points to meet criteria for FM may have unwittingly created a distress syndrome[1]. Because of the historical linkage between tender points and FM, FM and distress have become inextricable. One could take an extreme view of this conundrum and state that because tender points are so correlated with distress, and rarely improve in conjunction with clinical pain measures in therapeutic trials, perhaps we should define FM strictly on the basis of chronic widespread pain. Population based studies suggest that such criteria might identify primarily individuals with idiopathic chronic widespread pain in younger age groups, but that with advancing age this pain is more likely to be multifocal pain localized to the joints, and represent conditions such as osteoarthritis.

Population-based studies of chronic widespread pain also indicate that even if the requirement for tender points were abandoned, individuals with chronic widespread pain alone would still have a higher than expected rate of psychological comorbidities, especially if the pain is truly widespread[4,17]. However, the association between chronic widespread pain alone and psychological comorbidities is considerably less than when chronic widespread pain and the presence of tender points are combined.

We chose to interpret our findings as a further clarification of what is measured by a tender point count. In clinical practice, the tender point count may be a very useful measure of tenderness that captures an overall evaluation of physiological and psychological dysfunction. But these data should serve as a further reminder to clinicians that rigidly requiring 11 tender points before diagnosing FM may be inadvisable.

For research purposes, there is nearly unanimity that the ACR criteria have served a useful purpose in standardizing subject selection. The tender point requirement in the definition has been largely responsible for defining a group of patients who display hyperalgesia and allodynia[18-20]. These constructs have been extremely useful in attempting to elucidate the mechanisms operative in FM. In research settings, pain pressure sensitivity should be evaluated with the least confounded measures possible: at a minimum a dolorimeter examination, and perhaps more sophisticated measures similar to those employed in this study.

REFERENCES

1. 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.
2. Cohen ML, Quintner JL. Fibromyalgia syndrome, a problem of tautology. Lancet 1993;342:906-9.
3. Granges G, Littlejohn G. Pressure pain threshold in pain-free subjects, in patients with chronic regional pain syndromes, and in patients with fibromyalgia syndrome. Arthritis Rheum 1993;36:642-6.
4. Croft P, Schollum J, Silman A. Population study of tender point counts and pain as evidence of fibromyalgia. BMJ 1994;309:696-9.
5. Wolfe F, Ross K, Anderson J, Russell IJ, Hebert L. The prevalence

and characteristics of fibromyalgia in the general population. Arthritis Rheum 1995;38:19-28.
6. Wolfe F, Ross K, Anderson J, Russell IJ. Aspects of fibromyalgia in the general population: sex, pain threshold, and fibromyalgia symptoms. J Rheumatol 1995;22:151-6.
7. 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.
8. Gracely RH, Dubner R, McGrath PJ. Narcotic analgesia: fentanyl reduces the intensity but not the unpleasantness of painful tooth pulp sensations. Science 1979;203:1261-3.
9. Gracely RH, Kwilosz DM. The Descriptor Differential Scale: applying psychophysical principles to clinical pain assessment. Pain 1988;35:279-88.
10. Ohrbach R, Crow H, Kamer A. Examiner expectancy effects in the measurement of pressure pain thresholds. Pain 1998;74:163-70.
11. Bendtsen L, Norregaard J, Jensen R, Olesen J. Evidence of qualitatively altered nociception in patients with fibromyalgia. Arthritis Rheum 1997;40:98-102.
12. Kosek E, Ekholm J, Hansson P. Sensory dysfunction in fibromyalgia patients with implications for pathogenic mechanisms. Pain 1996;68:375-83.
13. Melzack R. The short-form McGill pain questionnaire. Pain 1987;30:191-7.
14. Derogatis LR, Melisaratos N. The Brief Symptom Inventory: an introductory report. Psychol Med 1983;13:595-605.
15. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. Arch Gen Psychiat 1961;4:561-71.
16. Smythe HA, Gladman A, Dagenais P, Kraishi M, Blake R. Relation between fibrositic and control site tenderness; effects of dolorimeter scale length and footplate size. J Rheumatol 1992;19:84-9.
17. Macfarlane GJ, Croft PR, Schollum J, Silman AJ. Widespread pain: is an improved classification possible? J Rheumatol 1996; 23:1628-32.
18. Russell IJ. Neurochemical pathogenesis of fibromyalgia. Z Rheumatol 1998;57 Suppl 2:63-6.
19. Yunus MB. Towards a model of pathophysiology of fibromyalgia: aberrant central pain mechanisms with peripheral modulation. J Rheumatol 1992;19:846-50.
20. Bennett RM. Emerging concepts in the neurobiology of chronic pain: evidence of abnormal sensory processing in fibromyalgia. Mayo Clin Proc 1999;74:385-98.