

Heterogeneity Within the Fibromyalgia Population: Theoretical Implications of Variable Tender Point Severity Ratings

HILARY D. WILSON, TERENCE W. STARZ, JAMES P. ROBINSON, and DENNIS C. TURK

ABSTRACT. *Objective.* The American College of Rheumatology (ACR) tender point (TP) criterion is used in diagnosing fibromyalgia syndrome (FM). There has been little research investigating patterns of positive TP. We investigated response patterns of TP in a sample of patients with FM.

Methods. Manual TP survey data were available on 1433 patients with FM. Factor analysis was conducted on ACR TP and control (CON) points. Factor scores were cluster analyzed to identify subgroups based on TP scores. Subgroups were compared on demographic and psychosocial variables.

Results. Factor analysis resulted in 4 TP groupings: neck/shoulder girdle, gluteal/trochanteric, and upper extremity regions, and a set of CON TP. Cluster analysis revealed 3 clusters. Group 1 was high on all 3 TP regions and the CON set; Group 2 moderate on the 3 TP regions, low on the CON set; and Group 3 was relatively low on all 3 TP regions and the CON set. The group highest on the CON and TP regions reported the greatest pain (7.58 ± 1.23 ; $p < 0.001$), sleep disturbance (7.05 ± 1.61 ; $p < 0.001$), anxiety (10.14 ± 4.57 ; $p < 0.001$), and depression (8.42 ± 4.4 ; $p < 0.001$).

Conclusion. TP severity ratings varied among cluster groups, suggesting patients with FM are not homogeneous. Variations in TP severity provide information regarding the degree to which FM affects patients' quality of life. Patients with elevated scores on the CON TP demonstrated a general pattern reflecting lower thresholds for symptom reporting and, perhaps, disease severity. Research is needed to elucidate mechanisms underlying heterogeneity among the FM population. (First Release Nov 15 2009; J Rheumatol 2009;36:2795–801; doi:10.3899/jrheum.090432)

Key Indexing Terms:

FIBROMYALGIA

TENDER POINTS

SUBGROUPS

DIAGNOSIS

Although fibromyalgia syndrome (FM) has been formally recognized for more than 15 years, has a history that long precedes formal recognition, and is believed to be a prevalent disorder with an extensive literature, there is a great deal of concern about the diagnosis itself.

Prior to 1990 various investigators proposed a wide range of criteria for classifying FM. In 1990 a multicenter study was performed to create consistency in the classification criteria¹. The authors concluded that the classification of FM could be based on 2 criteria: (1) widespread pain (i.e., 3 of 4 quadrants of the body and axial) of at least 3 months' duration; and (2) pain reported following palpation of at least 11 of 18 (9 bilateral) specific locations throughout the body, the so-called tender points (TP). The American College of Rheumatology (ACR) criteria had a sensitivity of 88.4% and specificity of 81.1% when differentiating

patients with FM from those with low back pain and rheumatoid arthritis¹. Thus the criteria appeared to offer an acceptable level of accuracy in diagnosis, and more important, a consistent set of criteria to be used in both research and clinical practice.

Although the ACR criteria provided a basis for the standardization of FM classification, a number of issues have been raised challenging their validity. These can be grouped into 2 general sets, namely, methodological and conceptual. Several of the major methodological and conceptual issues are discussed below.

From a methodological standpoint, questions have been raised that the ACR criteria are dependent on binary responses by patients during palpation of each TP — a patient's score consists of the number of points (out of a total of 18) that he or she identifies as painful in response to pressure of 4 kg. The ACR suggest a minimum of 11 of 18 TP to meet the criteria for FM. However, this process creates a truncated distribution that may mask potentially important differences in ratings of TP severity.

An alternative strategy recommended to provide more sensitive assessments of TP is the Manual Tender Point Survey (MTPS) that consists of the mean of patients' ratings of the severity of pain on an 11-point scale (0 = no pain; 10 = worst pain possible) during palpation of each point². Data

From the Department of Anesthesiology and Pain Medicine, University of Washington, Seattle, Washington, USA.

Supported by NIH (NIAMS) grant AR 44724 awarded to D.C. Turk.

H.D. Wilson, PhD; T.W. Starz, MD; J.P. Robinson, MD, PhD; D.C. Turk, PhD.

Address correspondence to Dr. D.C. Turk, Department of Anesthesiology, Box 356540, University of Washington, Seattle, WA 98195.

E-mail: turkdc@u.washington.edu

Accepted for publication July 9, 2009.

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2009. All rights reserved.

from the MTPS indicate that when patients use an 11-point scale to rate their pain, it is possible to identify effects of treatments and make distinctions among different groups of FM patients that are masked when binary ratings are used³.

At a conceptual level, fundamental questions have been raised about the biological significance of TP. Early theories suggested that FM was a soft-tissue disorder with multiple areas of muscle tenderness and chronic widespread pain^{4,5}. In contrast, myofascial trigger points, which were considered distinct from FM tender points, resulted in referred pain to a distant but specific site. Accordingly, the 18 survey sites (standard TP) designated in the ACR criteria ranged across muscle bellies or tendon insertions. Support for this theory was provided by a widely cited study that demonstrated control sites not located over muscle bellies or tendons did not differ between FM and control groups⁴.

Since 1990, many FM investigators have postulated that central nervous system (CNS) sensitization plays a fundamental role in the pathogenesis of FM, so that patients with FM are more sensitive than others to a wide range of sensory stimulation^{6,7}. A corollary of this hypothesis is that FM patients should show generalized tenderness, rather than selective tenderness over specific muscles. Thus, the CNS sensitization model implies that both control and test sites should be more sensitive, whereas the earlier model stressed the importance of test sites.

Both the CNS sensitization and the soft-tissue disorder models postulate high covariation among responses to the 18 survey sites. However, some researchers have found evidence for patterns in sensitivity among these sites. The original multicenter study that examined many locations found that 18 specific sites reliably differentiated patients with FM from patients with low back pain and rheumatoid arthritis¹. Moreover, Turk, *et al* studied tenderness among individuals with recent whiplash injuries⁸. They found that 46% of the cohort met the ACR tender point criterion for a diagnosis of FM. However, the muscular sensitivity of these individuals was not randomly distributed, but rather was clustered in neck and shoulder girdle muscles. These results support the view that FM patients might show clustering in their sensitivity to palpation of different survey sites, and that these clusters are meaningfully related disorders that affect specific regions of their bodies.

Given the conceptual and methodological issues surrounding FM TP, it is noteworthy that very little research has been done to explore the relationship among TP within FM patients.

The purpose of our study was to utilize data from a large multicenter trial to explore patterns of variability within the MTPS. Specifically, factor analysis was used to explore the dimensional structure within the 18 ACR TP and 3 control TP. Then cluster analysis was used to determine whether there are differing patterns of responses based on these dimensions identified within the MTPS. We hypothesized

that FM patients will show clustering in their sensitivity to palpation of different survey TP sites.

MATERIALS AND METHODS

Sample. Demographic and baseline data from 2 randomized, double-blind, multicenter, placebo-controlled pharmaceutical trials were used in our current study. The first study was conducted at 79 research sites throughout the United States and consisted of 748 FM patients⁹, and the second was conducted at 84 research sites throughout the United States and consisted of 731 FM patients¹⁰. All participants met ACR criteria for FM. MTPS data were available on 1433 (96.8%) of the combined 1479 patients, and factor analysis was conducted on this set of patients. Following factor analysis, the 1433 participants were randomly divided into 2 groups utilizing the SPSS random sample generator. The first sample (n = 690) was utilized in the initial cluster analysis, and the second sample (n = 743) was utilized to cross-validate the cluster solution.

Measures. *The Manual TP Survey (MTPS).* The MTPS was developed as a standardized TP procedure to identify positive TP and determine individual TP pain severity². Patients report TP pain severity on a scale of 0 (no pain) to 10 (most pain possible). In addition to the standard 18 TP included in the ACR criteria, 3 control TP (i.e., mid-forehead, left thumbnail, right dorsum forearm) are assessed.

The Hospital Anxiety and Depression Scale (HADS). The HADS¹¹ was developed to measure symptoms of anxiety and depression among nonpsychiatric medical inpatients. This self-report questionnaire contains 14 items that generate scores for depressive (0–7) and anxious (0–7) symptoms.

The Fibromyalgia Impact Questionnaire (FIQ). The FIQ¹² is a self-report instrument designed to measure the overall influence of FM over multiple dimensions, such as function, pain, and psychological distress. The FIQ consists of 20 items and is scored from 0 to 100, with higher scores associated with greater effect of FM. The first section consists of 11 questions related to physical functioning, and the average of these 11 items is known as the physical functioning scale of the FIQ, and is reported in the current study.

Mean pain score. Patients were asked to select the number that best described their FM pain during the past 24 h, on an 11-point scale with 0 being “no pain” and 10 being “worst possible pain.”

Mean sleep score. Patients were asked to select the number that best described the quality of their sleep during the past 24 hours, on a scale from 0 to 10, 0 being the “best possible sleep” and 10 the “worst possible sleep.”

Analyses. All data analyses were conducted utilizing SPSS for Windows, version 14 (SPSS, Chicago, IL, USA). Principal components analysis, with oblique rotation, was conducted on the 21 TP [18 ACR and 3 controls (i.e., mid-forehead, left thumbnail, right dorsum forearm)] at baseline. A TP was considered to load significantly on a factor if it met 2 criteria: (1) the loading on the factor was > 0.39; (2) the point did not cross-load on a second factor, where cross-loading was defined as a difference of less than 0.15 between the loading of a TP on 2 different factors.

Cluster analysis was performed to identify TP profiles among the set of patients. The k-means clustering procedure, which allocates data points into a specified number of clusters based on the centroids of each data point, was used to classify patients into unique clusters¹³. The number of clusters retained was based on 2 criteria: stability (i.e., reproducibility) and interpretability. A solution was considered stable if the centroids produced in the second sample were within one-half SD of the centroids produced in the first sample. The cluster groups must also be interpretable, which refers to the alignment of the clusters with clinical reports and experience of working with FM patients.

Cluster analysis was first performed on Sample 1 (n = 690), and then cross-validated on Sample 2 (n = 743). Following determination of cluster groups, data from Sample 1 were used to determine the external validity of the cluster analysis through significance tests that compared groups defined by the cluster solution on a set of relevant clinical variables¹⁴. Chi-square

tests of significance were utilized on categorical variables and analysis of variance (ANOVA) on continuous variables.

RESULTS

Demographic data for the total sample ($n = 1433$) and the 2 subsamples are presented in Table 1. The total cohort was predominantly female (94.6%), overweight [mean body mass index (BMI) = 30.75 ± 7.31], and middle-aged (age 49.39 ± 11.24 yrs). On average, patients had been experiencing FM symptoms for $95.64 (\pm 2.51)$ months. Statistical analyses comparing the 2 subsamples revealed no significant differences on any of the variables, suggesting samples were each representative of the total cohort.

The frequency and average TP scores for the total cohort are presented in Table 2. Paired *t*-tests revealed no significant differences between left and right TP of each respective pair of TP, thus only right TP are presented. Each of the 18 TP was positive more than 90% of the time when utilizing a minimum score of 1.

Factor analysis. Principal components analysis was conducted on 21 TP, first forcing a one-factor solution. The total amount of variance accounted for by one factor was 40.17%. Since the one-factor solution did not account for a majority of the variance, a principal components analysis, with oblique rotation, was conducted on the 21 TP (18 ACR and 3 controls). The number of factors was selected based on evaluation of the scree plot, size of the eigenvalues (> 1 considered acceptable for factor retention), and interpretability. Although a 5-factor solution was supported by the eigenvalue > 1 criteria, evaluation of the scree plot, in addition to clinical interpretation, led us to retain the 4-factor solution. The first 3 groups are associated with body regions — neck + shoulder TP formed Factor 1 (NS), the gluteal/trochanteric TP Factor 2 (GT), extremity TP (lateral epicondyle) Factor 3 (UE), and Factor 4 consisted of all 3 control points (CON). The 2 knee TP cross-loaded on both the GT and the UE factor, and thus were deleted. All respective right and left TP loaded on the same factor (Table 3).

Evaluation of the component correlation matrix revealed that although the factors were well delineated using oblique rotation, some were moderately correlated with one another. The highest correlations were between the NS and GT ($r =$

Table 2. Frequency and average tender point severity for individual tender points and factor scores for total sample ($n = 1433$).

| Tender Point (right) | Frequency | Mean (SD) |
|------------------------|-------------|-------------|
| Occiput | 1323 (90.3) | 4.86 (2.78) |
| Trapezius | 1410 (96.2) | 6.13 (2.54) |
| Supraspinatus | 1374 (93.8) | 5.71 (2.75) |
| Low cervical | 1389 (94.8) | 5.85 (2.67) |
| Second rib | 1400 (95.6) | 6.01 (2.60) |
| N/S factor mean | | 5.68 (1.98) |
| Greater trochanter | 1385 (94.5) | 6.48 (2.83) |
| Gluteal | 1353 (92.4) | 6.31 (2.76) |
| G/T factor mean | | 6.58 (2.31) |
| Lateral epicondyle | 1363 (93.0) | 5.67 (2.78) |
| UE factor mean | | 5.61 (2.63) |
| Thumbnail | 541 (36.9) | 1.41 (2.34) |
| Mid-forehead | 613 (41.8) | 1.51 (2.27) |
| Dorsum forearm | 950 (64.8) | 2.95 (2.97) |
| CON factor mean | | 1.96 (1.98) |
| Knee (not in solution) | 1384 (94.5) | 6.18 (2.82) |

N/S: neck/shoulder; G/T: gluteal/trochanteric; UE: upper extremity; CON: control.

0.42), UE ($r = 0.39$), and CON factors ($r = 0.37$). Remaining correlations were all less than 0.3 (GT/UE, $r = 0.29$; GT/CON, $r = 0.25$; UE/CON factor, $r = 0.25$).

Based on factor analysis, the MTPS (including the 18 ACR TP and 3 control TP) revealed 4 modestly correlated factors. This result supports the possibility that there is some underlying structure within the MTPS, such that NS, GT, and UE TP tend to group somewhat independently. The correlations among the NS and the remaining 3 factors ranged from 0.37 to 0.42, however, suggesting that although the TP formed unique dimensions, they shared some common variance. To further explore the relationship among TP, cluster analysis, which aims to maximize both within-homogeneity of FM patients and between-homogeneity of manual TP, was conducted utilizing the factor scores from the factor analysis.

Development of TP profiles. The *k*-means clustering procedure was conducted with 4 factor mean scores (NS, GT, UE, and CON) as the clustering variables, with iterations of 2, 3, 4, 5, and 6-cluster solutions. The cluster analyses were repeated on the second sample for cross-validation, with

Table 1. Demographic characteristics of patients with fibromyalgia (FM) for Sample 1 ($n = 690$) and Sample 2 ($n = 743$). Values are expressed as mean (SD) unless otherwise indicated.

| Variable | Sample 1 ($n = 690$) | Sample 2 ($n = 743$) | Total ($n = 1433$) | p |
|---------------------------|---------------------------|---------------------------|-------------------------|------|
| Average age, yrs | 48.95 (11.55) | 49.80 (10.95) | 49.39 (11.24) | 0.15 |
| Female, n (%) | 664 (94.7) | 718 (94.5) | 1382 (94.6) | 0.83 |
| Duration of FM, mo | 115.77 (97.01) | 115.85 (94.42) | 95.64 (2.51) | 0.99 |
| Height, m | 1.64 (0.08) | 1.64 (0.08) | 1.64 (0.08) | 0.71 |
| BMI, weight/height, m^2 | 30.57 (7.20) | 30.92 (7.42) | 30.75 (7.31) | 0.35 |

BMI: body mass index.

Table 3. Factor loadings and communalities (h^2 ; loadings below 0.45 are replaced by 0).

| Item | NS | GT | UE | CON | h^2 |
|----------------------------|-----------|-----------|-----------|------|-----------|
| Occiput (R)/(L) | 0.69/0.69 | 0/0 | 0/0 | 0/0 | 0.59/0.58 |
| Trapezius (R)/(L) | 0.77/0.79 | 0/0 | 0/0 | 0/0 | 0.63/0.64 |
| Supraspinatus (R)/(L) | 0.69/0.69 | 0/0 | 0/0 | 0/0 | 0.60/0.59 |
| Low cervical (R)/(L) | 0.68/0.65 | 0/0 | 0/0 | 0/0 | 0.59/0.59 |
| Second rib (R)/(L) | 0.61/0.59 | 0/0 | 0/0 | 0/0 | 0.59/0.59 |
| Greater trochanter (R)/(L) | 0/0 | 0.81/0.80 | 0/0 | 0/0 | 0.69/0.68 |
| Gluteal (R)/(L) | 0/0 | 0.72/0.72 | 0/0 | 0/0 | 0.62/0.65 |
| Lateral epicondyle (R)/(L) | 0/0 | 0/0 | 0.69/0.69 | 0/0 | 0.65/0.66 |
| Thumb nail (L) | 0/0 | 0/0 | 0/0 | 0.84 | 0.67 |
| Mid-forehead | 0/0 | 0/0 | 0/0 | 0.72 | 0.59 |
| Dorsum forearm (R) | 0/0 | 0/0 | 0/0 | 0.72 | 0.60 |

NS: neck/shoulder girdle factor; GT: gluteal/trochanteric factor; UE: upper extremity factor; CON: control factor; h^2 : portion of variance in each variable accounted for by the solution.

successful replication with up to 3 groups. Comparisons between centroids indicated the mean factor scores of Sample 2 were within one-half SD of the mean factor scores for 11 of 12 comparisons for the 3-cluster solution in Sample 1. In addition, meaningful differences were identified on external variables of interest (Table 4), and thus the 3-cluster solution was retained (Figure 1).

Cluster 1: High on NS, GT, UE and high on CON (HH). Factor means for each of the clusters are presented in Figure 1. Examination of Figure 1 suggests that Cluster 1, which included 223 respondents (32.3%), is highest on all 3 TP regions (NS, GT, UE) and the CON set. All average factor scores were greater than one-half SD from their respective means (NS = 7.53 ± 1.25 , GT = 8.12 ± 1.44 , UE = 7.59 ± 1.70 , CON = 3.76 ± 1.91), suggesting they were different from the average TP scores. With respect to all other clusters, the first cluster scored at least 25% higher than the other clusters on all standard TP regions (NS, GT, and UE), and over 2.5 times higher on the control set, and was thus labeled the high standard TP/high control factor (HH) group.

Cluster 2: Moderate on NS, GT, UE, and low on CON (ML). The second cluster comprised 38.8% ($n = 268$) of the sample, and scores on the standard TP factors were within one-half SD of the means on all standard TP factors (NS = 5.3 ± 1.39 , GT = 6.29 ± 1.83 , UE = 6.02 ± 1.50), and lower than one-half SD on the control factor (CON = 1.12 ± 1.16).

Thus, the second cluster was considered moderate on the TP standard regions, and low on the control set (ML).

Cluster 3: Low on NS, GT, UE, and CON (LL). About 28.8% of respondents ($n = 199$) were included in the third cluster group. The scores on both the standard TP regions and CON factor for this cluster were more than one-half SD lower than the mean (NS = 4.27 ± 1.74 , GT = 4.19 ± 2.19 , UE = 2.34 ± 1.55 , CON = 1.15 ± 1.28), and was thus labeled the low standard TP/low control TP (LL) group.

Validation of cluster solution. In addition to cross-validating the cluster solution on a separate sample, significance tests were conducted that compared the clusters obtained in Sample 1 on variables of clinical importance¹⁴. Results are described below.

Demographic variables. Table 4 presents demographic characteristics for the 4 cluster profiles in Sample 1. Statistical comparisons indicated there were no significant differences among the 3 clusters in duration of FM ($F_{2,684} = 1.29$, $p = 0.28$) or BMI ($F_{2,687} = 0.50$, $p = 0.61$). Age was statistically different among groups ($F_{2,684} = 5.68$, $p = 0.004$, $h^2 = 0.02$), with post-hoc tests indicating that the LL cluster (46.7 ± 13.16 yrs) was slightly younger than the ML (49.59 ± 10.81 yrs) and the HH clusters (50.27 ± 10.69 yrs).

Psychosocial variables. Psychosocial characteristics for the cluster groups are presented in Table 5. Overall, the HH group was characterized by the greatest degree of pain and

Table 4. Demographic characteristics of patients with fibromyalgia (FM) by cluster group for Sample 1 ($n = 690$). Values are expressed as mean (SD) unless otherwise indicated.

| Variable | HH, $n = 223$ | ML, $n = 268$ | LL, $n = 199$ | Total, $n = 690$ | p | h^2 |
|---------------------------|---------------|-----------------|----------------|------------------|-------|-------|
| Average age, yrs | 50.27 (10.69) | 49.59 (10.81) | 46.70 (13.16) | 48.98 (11.58) | 0.004 | 0.01 |
| Female, % | 215 (96.4) | 251 (94.4) | 185 (93.4) | 651 (94.8) | 0.37 | |
| Duration of FM, mo | 121.1 (92.25) | 115.71 (101.76) | 106.15 (91.94) | 114.71 (96.01) | 0.28 | |
| BMI, weight/height, m^2 | 30.45 (7.26) | 30.91 (7.6) | 30.27 (6.55) | 30.58 (7.2) | 0.61 | |

h^2 : portion of variance in each variable accounted for by the solution. BMI: body mass index.

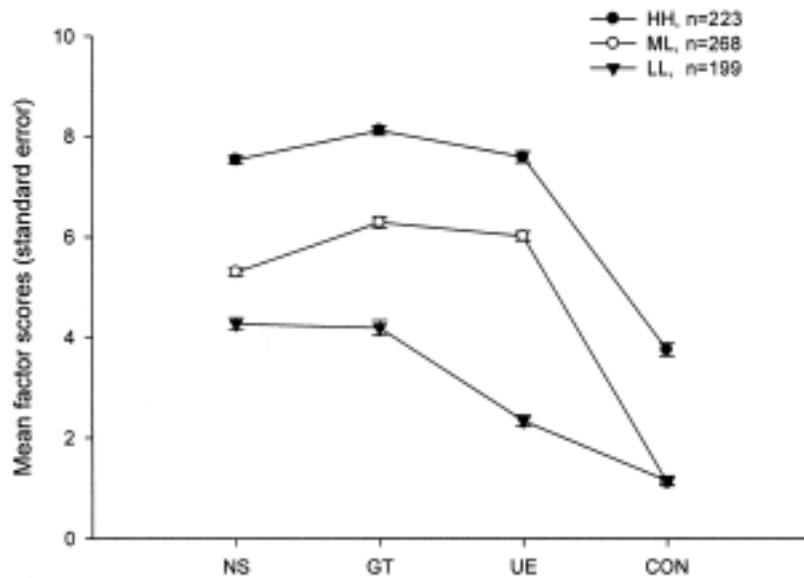


Figure 1. Mean scores on the NS (neck/shoulder), GT (gluteal/trochanteric), UE (upper extremity), CON (control) factors for the 3 groups in the cluster solution. The HH (high/high) group was at least 25% higher than all other clusters on all standard TP sets (NS = 7.53 ± 1.25, GT = 8.12 ± 1.44, and UE = 7.59 ± 1.70), and over 2.5 times higher on the control domain (3.76 ± 1.91). Scores on the standard TP sets for the ML (moderate/low) cluster (NS = 5.3 ± 1.39, GT = 6.29 ± 1.83, UE = 6.02 ± 1.50) were all within one-half SD of the mean for all groups, and lower than one-half SD on the control factor (CON = 1.12 ± 1.16). Scores on the LL (low/low) group were more than one-half SD lower than the mean for both the standard and control factors (NS = 4.27 ± 1.74, GT = 4.19 ± 2.19, UE = 2.34 ± 1.55, CON = 1.15 ± 1.28).

Table 5. Psychosocial characteristics by cluster group for Sample 1 (n = 690). Values are expressed as mean (SD).

| Variable | HH (n = 223) | ML (n = 268) | LL (n = 199) | Total (n = 690) | p | h ² |
|-------------------------------|--------------|--------------|--------------|-----------------|---------|----------------|
| Pain score (0–10) | 7.58 (1.23) | 6.69 (1.19) | 6.42 (1.35) | 6.9 (1.34) | < 0.001 | 0.13 |
| Sleep score (0–10) | 7.05 (1.61) | 6.36 (1.53) | 6.0 (1.54) | 6.48 (1.61) | < 0.001 | 0.07 |
| FIQ physical impairment (0–3) | 1.35 (0.69) | 1.20 (0.7) | 1.09 (0.64) | 1.22 (0.69) | 0.001 | 0.02 |
| HADS Anxiety | 10.14 (4.57) | 8.72 (4.36) | 8.96 (4.23) | 9.25 (4.43) | < 0.001 | 0.02 |
| HADS Depression | 8.42 (4.4) | 7.53 (4.37) | 6.83 (3.84) | 7.62 (4.28) | < 0.001 | 0.02 |
| Manual tender point count | 17.79 (0.64) | 17.06 (1.47) | 15.72 (2.31) | 16.91 (1.78) | < 0.001 | 0.21 |

FIQ: Fibromyalgia Impact Questionnaire; HADS: Hospital Anxiety and Depression Scale; h²: portion of variance in each variable accounted for by the solution.

psychosocial impairment, whereas the LL group was characterized by the least amount of all 3 groups.

FIQ physical impact scores differed significantly among groups (F 2,685 = 7.57, $p = 0.001$, $h^2 = 0.02$), with post-hoc results indicating that the HH (1.35 ± 0.69) group reported being significantly more functionally disabled compared to the LL group (1.09 ± 0.64; $p < 0.001$). Mean pain (F 2,684 = 79.01, $p < 0.001$, $h^2 = 0.13$) and sleep scores (F 2,684 = 60.48, $p < 0.001$, $h^2 = 0.07$) also varied among clusters, with post-hoc tests indicating the HH group differed from both the ML and the LL groups (HH > ML ≥ LL, all p values < 0.001). A main effect for the HADS anxiety (F 2,685 = 6.96, $p < 0.001$, $h^2 = 0.02$) was also detected, with post-hoc tests indicating that the HH group (10.14 ± 4.57) was more anxious, compared to both the ML (8.72 ± 4.36) and LL (8.96 ±

4.23) groups (all p values < 0.001). The HADS depression subscale also varied significantly (F 2,685 = 7.42, $p < 0.001$, $h^2 = 0.02$), with the HH group (8.42 ± 4.4) reporting more mood compared to both the ML (7.53 ± 4.37) and LL group (6.83 ± 3.84; HH > ML ≥ LL, all significant p values < 0.001).

DISCUSSION

The factor analysis of the MTPS revealed 4 groups of TP: neck/shoulder region (NS), gluteal/trochanteric region (GT), upper extremity region (UE), and a set of control (CON) TP. In order to consider individual response patterns to the MTPS cluster analysis, a classification analysis that groups individual cases (in this study patients with FM) rather than individual variables (in this study TP severity ratings) was

conducted. Three unique response patterns were identified, such that compared to other FM patients in the sample: (1) one group was high on all 3 standard TP regions and the control set of TP (HH); (2) one group was moderate on 3 standard TP regions and low on the control TP set (ML); and (3) one group was relatively low on all 3 sets of standard TP regions and the control set (LL) (Figure 1).

The theory that patients with FM might show clustering in their sensitivity to palpation of different TP sites was not supported by the cluster analysis, as indicated by the relative parallel lines between the NS, GT, and UE factors. Notably, however, 3 groups were identified based on severity differences among the standard TP factors (NS, GT, UE). The HH group rated the standard TP regions an average of 7.52 ± 1.46 , the ML group averaged 5.87 ± 1.57 , while the LL group averaged 3.65 ± 1.83 . This supports the notion that there are varying degrees of severity in FM¹⁵⁻¹⁷.

The CON factor was generally less tender than other TP factors, confirming the results of the 1990 ACR criteria¹. It is notable that the scores for the ML (1.12 ± 1.16) and LL groups (1.15 ± 1.28) on the control factor are within the range of scores reported for patients with chronic headache (control point scores averaged between 0.17 and 1.39 on each of the 3 control points)², suggesting that for 2 of the 3 cluster groups, sensitivity at control sites is comparable to a chronic pain population, without chronic widespread pain as a defining characteristic.

The HH group, however, averaged scores 3 times as high on the control factor (3.76 ± 1.91). This group also scored the highest scores on all self-report measures administered, including the anxiety and depression measure, pain and sleep analog scales, and the FIQ physical functioning scale. This is consistent with previous reports of increased muscle sensitivity and increased pain and distress^{18,19}. One possibility to account for this finding is that they have the generalized sensitivity postulated by those who view FM as driven by CNS sensitization. However, another possibility is that responses of participants were influenced, at least to some extent, by a "response set" such that what is really being measured is the degree to which a person will agree or disagree with an item, regardless of content²⁰. It has been suggested that the acquiescence is associated with item ambiguity²¹, so that when respondents are unsure of an item they will tend to answer in the affirmative. A study that alters the scale presentation and balances the number of positive and negative symptom items, may provide insight into the role cognitive bias may play in response to self-report questionnaires in general.

The ML group scored moderately on the 3 standard TP regions and low on the control set. This differential response between the ACR TP and control TP is more compatible with the early theory that FM patients have some special sensitivity in muscles. The LL group, however, reported similar severity scores among the standard and control TP.

This pattern of responses is congruent with the generalized theory of sensitivity but with a lower degree of severity compared to the HH group.

Overall, our results provide partial support for the general sensitivity and muscle tenderness theories. Both theories may be valid, as FM may not consist of a homogeneous group of patients but rather there may be subgroups based on important characteristics²². Given the differing patterns of standard and control TP response between the HH and the ML/LL groups, it may be postulated that different mechanisms are involved in the etiology and maintenance of FM symptoms within the FM population. This notion is consistent with the view of researchers that FM may be a broader condition with specific subsets²³⁻²⁶. Potential subsets may include people with regional or generalized symptoms, FM secondary to reactive disease, or FM coexisting with other diseases. Yunus describes evidence that supports the interaction of central sensitivity and psychosocial characteristics to account for the wide and variable symptom presentation among FM patients²⁵. Additional research comparing these subgroups on various psychophysical and biological markers may provide insight into the unique mechanisms of FM, and further explore the hypothesis that different mechanisms are involved in subsets of patients diagnosed with FM.

Our study has several limitations. Most notably, no data on healthy controls or non-FM rheumatology patients were available, thus no comparison may be made regarding the TP scores compared to other groups. Important information could be gleaned from a study design that included both healthy controls and patients with arthritis. The datasets analyzed were obtained from 2 large-scale pharmaceutical trials, which exclude patients with any unstable psychological or medical disorders, other painful disorders, and any evidence of inflammatory rheumatic disease. Thus, the sample may not be representative of the complete FM clinical population. Additionally, although the dataset analyzed in our current study was large and the cluster analysis solution was replicated in the reserved sample, the clusters should be replicated in a second independent sample to verify the findings. Multiple analyses were performed to identify associations between the 3 clusters described in our study and variables such as sleep and pain severity. As a result, it is possible that some of the associations described above may be spurious. Since this was an exploratory study, we decided not to make any adjustment for the performance of multiple analyses. However, observation of the results does indicate that levels of statistical significance generally exceeded $p < 0.001$ and those may be reasonably valid.

The variation in TP severity will provide information on how severely a patient's condition is affecting their quality of life. Additional research exploring differences among these groups on other variables of interest, including variability in diffuse noxious inhibitory control, genotype, and psychological variables (e.g., self-efficacy and coping

styles) may provide insight into treatment options for different subgroups of patients with FM.

ACKNOWLEDGMENT

The authors express appreciation to Pfizer Inc. for making the data available.

REFERENCES

1. Wolfe F, Smythe HA, Yunus MB, Bennett RM, Bombardier C, Goldenberg DL, 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. Okifuji A, Turk DC, Sinclair JD, Starz TW, Marcus DA. A standardized manual tender point survey. I. Development and determination of a threshold point for the identification of positive tender points in fibromyalgia syndrome. *J Rheumatol* 1997;24:377-83.
3. Turk DC, Okifuji A, Sinclair JD, Starz TW. Differential responses by psychosocial subgroups of fibromyalgia syndrome patients to an interdisciplinary treatment. *Arthritis Care Res* 1998;11:397-404.
4. Campbell SM, Clark S, Tindall EA, Forehand ME, Bennett RM. Clinical characteristics of fibrositis. I. A "blinded," controlled study of symptoms and tender points. *Arthritis Rheum* 1983;26:817-24.
5. Smythe HA, Moldofsky H. Two contributions to understanding of the "fibrositis" syndrome. *Bull Rheum Dis* 1977;28:928-31.
6. Geisser ME, Glass JM, Rajcevska LD, Clauw DJ, Williams DA, Kileny PR, et al. A psychophysical study of auditory and pressure sensitivity in patients with fibromyalgia and healthy controls. *J Pain* 2008;9:417-22.
7. Petzke F, Khine A, Williams D, Groner K, Clauw DJ, Gracely RH. Dolorimetry performed at 3 paired tender points highly predicts overall tenderness. *J Rheumatol* 2001;28:2568-9.
8. Turk D, Robinson J, Burwinkle T. Prevalence of fibromyalgia tender points following whiplash injury. *J Pain* 2006;7 Suppl 2:S27:693.
9. Mease PJ, Russell IJ, Arnold LM, Florian H, Young JP Jr, Martin SA, et al. A randomized, double-blind, placebo-controlled, phase III trial of pregabalin in the treatment of patients with fibromyalgia. *J Rheumatol* 2008;35:502-14.
10. Arnold LM, Russell IJ, Diri EW, Duan WR, Young JP Jr, Sharma U, et al. A 14-week, randomized, double-blinded, placebo-controlled monotherapy trial of pregabalin in patients with fibromyalgia. *J Pain* 2008;9:792-805.
11. Zigmond A, Snaith RP. The Hospital Anxiety and Depression Scale. *Acta Psychiatr Scand* 1983;67:361-70.
12. Burckhardt CS, Clark SR, Bennett RM. The Fibromyalgia Impact Questionnaire: development and validation. *J Rheumatol* 1991;18:728-33.
13. Milligan GW, Cooper MC. An examination of procedures for determining the number of clusters in a data set. *Psychometrika* 1985;50:159-79.
14. Aldenderfer MS, Blashfield RK. Cluster analysis. Lewis-Beck MS, editor. Newbury Park, NJ: Sage Publications; 1984.
15. Tastekin N, Birtane M, Uzunca K. Which of the three different tender points assessment methods is more useful for predicting the severity of fibromyalgia syndrome? *Rheumatol Int* 2007;27:447-51.
16. Bennett R. The Fibromyalgia Impact Questionnaire (FIQ): a review of its development, current version, operating characteristics and uses. *Clin Exp Rheumatol* 2005;23 Suppl 39:S154-62.
17. 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.
18. McBeth J, Macfarlane GJ, Benjamin S, Morris S, Silman AJ. The association between tender points, psychological distress, and adverse childhood experiences: a community-based study. *Arthritis Rheum* 1999;42:1397-404.
19. Petzke F, Gracely RH, Park KM, Ambrose K, Clauw DJ. What do tender points measure? Influence of distress on 4 measures of tenderness. *J Rheumatol* 2003;30:567-74.
20. Jackman MR. Education and prejudice or education and response-set? *Am Sociol Rev* 1973;38:327-39.
21. Bass BM. Authoritarianism or acquiescence? *J Abnorm Psychol* 1955;51:616-23.
22. Thieme K, Turk DC. Heterogeneity of psychophysiological stress responses in fibromyalgia syndrome patients. *Arthritis Res Ther* 2006;8:R9.
23. Schochat T, Raspe H. Elements of fibromyalgia in an open population. *Rheumatology* 2003;42:829-35.
24. Muller W, Schneider EM, Stratz T. The classification of fibromyalgia syndrome. *Rheumatol Int* 2007;27:1005-10.
25. Yunus MB. Fibromyalgia and overlapping disorders: the unifying concept of central sensitivity syndromes. *Semin Arthritis Rheum* 2007;36:339-56.
26. Wessely S, White PD. There is only one functional somatic syndrome. *Br J Psychiatr* 2004;185:95-6.