The Fibromyalgia Impact Questionnaire: Development and Validation

CAROL S. BURCKHARDT, SHARON R. CLARK, and ROBERT M. BENNETT

Abstract. **An instrument has been developed to assess the current health status of women with the fibromyalgia syndrome. The Fibromyalgia Impact Questionnaire (FIQ) is a brief 10-item, self-administered instrument that measures physical functioning, work status, depression, anxiety, sleep, pain, stiffness, fatigue, and well being. We describe its development and validation. This initial assessment indicates that the FIQ has sufficient evidence of reliability and validity to warrant further testing in both research and clinical situations.** (*J Rheumatol 1991; 18: 728–33*)

Key Indexing Terms:

FIBROMYALGIA	FIBROSITIS	QUESTIONNAIRE	AIMS
HEALTH STATUS	FUNCTIONAL DISABILITY	DEPRESSION	PAIN

Several well established instruments are available to measure health status and functional disability in persons with rheumatic diseases[1–3]. Parts of 2 of these instruments, the Health Assessment Questionnaire (HAQ) and the Arthritis Impact Measurement Scales (AIMS), have been used in conjunction with other questions in descriptive studies of the severity and stability of fibromyalgia[4–6]. Although both instruments are reliable and valid for several rheumatic disease groups, no formal psychometric testing of either instrument has been reported for the population with fibromyalgia.

We report the development of a brief, self-administered instrument, the Fibromyalgia Impact Questionnaire (FIQ). It contains items comparable to the AIMS and HAQ as well as some unique items. We also compare the psychometric properties of the new instrument to the AIMS to determine whether a new short instrument would be a useful addition to outcome measurement in fibromyalgia. The AIMS was chosen for comparison purposes because it is a more comprehensive health status instrument than the HAQ disability index.

MATERIALS AND METHODS

Patients. Data for this study were obtained from 2 sources. The first sample, 64 outpatient women, between the ages of 24 and 66 (mean age 45 years), with a diagnosis of fibromyalgia and no other significant rheumatic diseases, was entered into the study during 1987–1988. The diagnosis of fibromyalgia was established by a rheumatologist (RMB) using criteria of 12 of 18 tender points and widespread, diffuse pain unexplained by other diagnoses. Although this study was initiated before the publication of the classification criteria for fibromyalgia, the criteria used were essentially the same[7]. Median time since diagnosis for the sample was 5 years. Sixty-eight percent were married and 62% were employed outside the home. Median education was 1 to 4 years of college and median income was between $15,000 and $20,000/year.

The second sample, 25 women with fibromyalgia diagnosed in the same manner, completed the FIQ and tender point testing in late 1989 as part of clinical evaluation in a fibromyalgia treatment program. The sample was derived from the same patient population as the first sample. There were no significant differences in age, education, income, employment or marital status; however, the patients had been diagnosed with fibromyalgia for a significantly shorter period of time (median 1 year).

Fibromyalgia impact questionnaire (FIQ). Initial construction of the FIQ was based conceptually on the premise that such an instrument should contain physical, psychological, social, and global well being components. Items for the instrument were derived from clinical interactions with patients, from preexisting work that has documented major characteristics of the syndrome, and from existing rheumatology health status instruments.

The first item focuses primarily on the patient's ability to do large muscle tasks and contains 10 subitems. The responses are scaled in a Likert format from 0 = *always able to do* to 3 = *never able to do*. The 10 subitems are added together and divided by the number of valid scores to yield one physical functioning score. The next 2 items ask patients to circle the number of days in the past week that they felt good and the number of days they missed work. The last 7 items — ability to do job, pain, fatigue, morning tiredness, stiffness, anxiety, and depression — are all measured by 100 mm anchored horizontal visual analog scales. The instructions for the first item and the 7 visual analogs ask patients to mark the category on the scale or the point on the line that best describes their abilities or feelings for the past week. Each item is standardized on a scale ranging from 0 to 10 with 10 indicating greater impairment. Table 1 shows each item in the scale and its scoring format.

The Arthritis Impact Measurement Scales (AIMS). The AIMS is a well known, psychometrically and clinically sound instrument for measuring the health status of patients with rheumatic disease[1,8]. It contains 9 scales measuring mobility, physical activity, activities of daily living, household activities, dexterity, social activity, pain, depression, and anxiety. Each scale score is standardized and ranges from 0 to 10 with 10 indicating increased disability. Using factor analysis, a 5-component model of health status has been identified with the first 4 scales listed above aggregated into a lower extremity physical functioning factor. The remaining scales form upper extremity, social interaction, symptom, and affect components[9]. The AIMS also con-

From the Division of Arthritis and Rheumatic Diseases Oregon Health Sciences University, Portland, OR, USA.

C.S. Burckhardt, PhD, Assistant Professor of Medicine (Research); S.R. Clark, PhD, Assistant Professor of Medicine (Research); R.M. Bennett, MD, FRCP, Professor of Medicine.

Address reprint requests to Dr. R.M. Bennett, Division of Arthritis and Rheumatic Diseases, Department of Medicine-L329A, Oregon Health Sciences University, 3181 SW Sam Jackson Park Rd., Portland, OR 97201-3098.

Submitted January 2, 1990 revision accepted January 16, 1991.

Table 1. *The Fibromyalgia Impact Questionnaire*

	Always	Most times	Occasionally	Never
1. Were you able to:				
a. Do shopping	0	1	2	3
b. Do laundry with a washer and dryer	0	1	2	3
c. Prepare meals	0	1	2	3
d. Wash dishes/cooking utensils by hand	0	1	2	3
e. Vacuum a rug	0	1	2	3
f. Make beds	0	1	2	3
g. Walk several blocks	0	1	2	3
h. Visit friends/relatives	0	1	2	3
i. Do yard work	0	1	2	3
j. Drive a car	0	1	2	3

2. Of the 7 days in the past week, how many days did you feel good?
1 2 3 4 5 6 7

3. How many days in the past week did you miss work because of your fibromyalgia? (If you don't have a job outside the home leave this item blank.) 1 2 3 4 5

4. When you did go to work, how much did pain or other symptoms of your fibromyalgia interfere with your ability to do your job?

No problem — Great difficulty

5. How bad has your pain been?

No pain — Very severe pain

6. How tired have you been?

No tiredness — Very tired

7. How have you felt when you got up in the morning?

Awoke well rested — Awoke very tired

8. How bad has your stiffness been?

No stiffness — Very stiff

9. How tense, nervous or anxious have you felt?

Not tense — Very tense

10. How depressed or blue have you felt?

Not depressed — Very depressed

Note: Complete scoring information may be obtained from Dr. Bennett on request.

tains other questions and scales that can be used as patient reported global measures of impact and severity as well as demographics.

Procedures. The FIQ, AIMS, and demographic information were collected by questionnaires mailed at 1-week intervals over 6 weeks to the first sample. Patients filled out the FIQ each week, for a total of 7 data collections points. The AIMS was collected at the initial data point and at 3 and 6 weeks. Demographics were collected at the initial data point only. The time interval for retesting the FIQ was short because our purpose was to determine the test-retest reliability of the instruments, not the stability or instability of the syndrome. Although there is some evidence that fibromyalgia symptoms may be stable from month to month[6], whether they fluctuate at shorter intervals (i.e., week to week) has not been established. We made the assumption that the syndrome would not change dramatically from week to week and reasoned that any discrepancies in the test-retest correlations would be more attributable to measurement error than actual change in syndrome symptoms. Thirteen of the 64 patients in the first sample had tender point testing done during clinic visits that occurred within 1 week of the initial, 3-week, and 6-week data collection points. Tender points were assessed by a trained rater using thumb pressure to 18 designated sites. Each point was scored as 0 = no pain and 1 = any indication of pain, either verbal or nonverbal.

Data from the second sample were collected at only one point through a mailed questionnaire and a clinic visit in which tender points were assessed in the same manner as described above.

Data analysis. Calculation of means and standard deviations for each item in the FIQ and each scale of the AIMS was completed before the other analyses. The test-retest reliability of the FIQ items and AIMS scales was estimated by Pearson correlations between each of the data points — 1 week apart for the FIQ and 3 weeks apart for the AIMS. To assess which items of the AIMS yielded useful information in patients with fibromyalgia (content validity), we looked at the percent of patients who answered each of the physical function items contained within the mobility, physical activity, dexterity, household activity and activities of daily living scales with a response denoting impairment. A cutoff criterion of $\geq 25\%$ impairment responses was set to indicate a valid item, one that addressed a common problem for patients with fibromyalgia. Content validity of the FIQ was assessed through calculation of the percent of missing data. We did not expect to have much missing data since in the preliminary stages of the instrument's development, patients had been asked to generate relevant items.

Before comparing the FIQ to the AIMS, we gathered some evidence for construct validity of the AIMS by correlating the standardized scale scores

with 2 measures of syndrome severity: patient global assessment of current syndrome impact (the impact analog scale on the AIMS) and patient global assessment of syndrome activity (one item on the AIMS). Then construct validity of the FIQ was estimated in 3 ways. First, the physical functioning item correlated with the lower extremity physical functioning component of AIMS and the pain, depression, and anxiety visual analogs were correlated with their comparable AIMS scales. Second, the FIQ item scores were correlated with 3 measures of symptom severity, the 2 described above and number of active tender points. Third, principal components factor analysis with varimax rotation was performed to determine whether the items and the subitems of the physical functioning item tended to load on a single factor. This analysis also served as a way of testing the assumption that the 10 physical functioning subitems could be added together to form one item.

We hypothesized that the AIMS scales and the FIQ items would be moderately correlated ($r \geq 0.40$) with measures of symptom severity, that the FIQ items would be moderately to highly correlated with comparable AIMS scales, and that the 10 subitems of Item 1 on the FIQ would load on a single factor. We also suspected that the remaining items on the FIQ would load on other factors.

Health status, physical health in particular, has been found to correlate with several demographic variables in studies of patients with rheumatoid arthritis[10-12]. Because fibromyalgia syndrome is not known to be a degenerative process, we did not posit any significant correlations with age or years of the syndrome or with any other demographic variable; however, they were assessed using the questions available on the AIMS. Employment was treated as a dichotomous variable of either employed or unemployed. Occupation was divided into 3 categories: physical labor, clerical, and professional.

RESULTS

Sixty-four patients in the first sample completed the initial testing and 52 finished the 6-week study. The average means and standard deviation of each of the items in the FIQ across 7 data points and the averages for the AIMS across 3 data points are summarized in Table 2.

Reliability. Test-retest reliability correlations (Pearson's r) for each item of the FIQ ranged from an average of 0.56 for pain to 0.95 for physical function over the 6 1-week intervals. Test-retest correlation coefficients for the AIMS scales on which correlations could be obtained ranged between 0.64 and 0.91 for the 2 3-week intervals. A test-retest correlation on the aid to daily living (ADL) scale could not be calculated as it lacked sufficient variance in the item scores.

Content validity. Using the criterion of $\geq 25\%$ impairment for a valid item on the AIMS, we found that none of the 4 activities of ADL scale items met the criterion. One of the 5 dexterity items (opening a new jar of food), 1 of 7 household activity items (do own housework), 2 of 4 mobility items (stay indoors and in bed or chair most of day), and 4 of the 5 physical functioning items (vigorous activity, walking several blocks, bending, and walking one block) qualified. Missing data within the physical functioning subitems of the FIQ were limited to 11% of patients who did not do dishes by hand and 20% who did no yard work. The 2 job items were not relevant for the 38% of patients who were not working outside the home.

Construct validity. Results of the construct validity testing of the AIMS are shown in Table 3. Four of the 9 scales were moderately correlated with the 2 measures of syndrome severity — the AIMS visual analog of syndrome impact and the AIMS syndrome activity question. Sufficient AIMS data to compute correlations for the 13 patients with tender point scores in the first group were available only for the depression, anxiety, pain, and social activity scales. These correlations ranged from 0.11 for anxiety to 0.38 for depression.

Construct validity testing of the FIQ yielded the following results. First, correlation analysis demonstrated that the FIQ physical functioning item had a highly significant correlation of 0.67 with the AIMS lower extremity physical func-

Table 3. *Validity testing of the AIMS*

Scales	Impact Analog* (n = 64)	Syndrome Activity* (n = 64)
Mobility	0.49	0.41
Physical activity	0.49	0.49
Activities of daily living	0.15	0.22
Household activities	0.45	0.29
Dexterity	0.15	0.27
Social activity	0.12	0.01
Pain	0.60	0.78
Depression	0.45	0.46
Anxiety	0.19	0.37

* Correlations above 0.25 are significant at the $p < 0.05$ level.

Table 2. *Average means and SD of the FIQ items and AIMS over the repeated testings*

FIQ Items	Mean	SD	AIMS Scales	Means	SD
Phys. function	3.17	2.48	Mobility	2.61	2.59
Feel good	6.44	3.15	Phys. activity	5.79	2.54
Work missed	0.78	1.71	ADL	0.14	0.50
Job ability	4.08	3.00	Household	0.82	0.86
Pain	5.52	2.54	Dexterity	3.04	3.00
Fatigue	6.61	2.48	Social	4.32	1.73
Morning tired	6.97	2.50	Pain	6.52	2.09
Stiffness	6.11	2.58	Depression	3.16	1.68
Anxiety	4.50	3.01	Anxiety	4.87	2.10
Depression	3.95	3.21			

Note: All items on both scales have been standardized and range from 0–10.

Table 4. *Validity testing of the FIQ*

Items	Impact Analog* (n = 64)	Syndrome Activity* (n = 64)	Tender Points† (n = 13)	Tender Points** (n = 25)
Physical function	0.30	0.49	0.40	0.61
Feel good	0.29	0.57	0.09	0.40
Work missed	0.17	0.47	0.64	0.74
Job ability	0.31	0.63	0.15	0.36
Pain	0.48	0.83	0.37	0.38
Fatigue	0.37	0.48	0.28	0.22
Morning tiredness	0.34	0.48	0.28	0.14
Stiffness	0.31	0.50	0.39	0.36
Anxiety	0.25	0.28	0.43	0.21
Depression	0.29	0.31	0.34	0.23

* All coefficients in these 2 columns are significant at the p = 0.05 level with the exception of the 0.17 in the first column.

† None of the coefficients in this column are statistically significant.

** First 3 coefficients in this column are significant at the p = 0.05 level.

tioning component. The pain, depression, and anxiety analog scales also showed highly significant correlations of 0.69, 0.73, and 0.76 with their respective AIMS scales. Second, the FIQ item correlations with the AIMS visual analog of syndrome impact and the AIMS syndrome activity question are shown in Table 4. The correlations were highest between the AIMS analog scale and 3 items, pain, fatigue and morning tiredness and lowest with days of work missed. Item correlations with the AIMS syndrome activity question tended to be higher, ranging from 0.28 to 0.83. Most correlations between the FIQ item scores and the number of tender points were in the 0.30 to 0.40 range in the first sample. However, they were not significant because of the small number of patients. The last column in Table 4 shows the correlations between the number of tender points and the FIQ items for the second sample of patients. These correlations ranged from 0.14 for morning tiredness to 0.74 for work missed.

Third, the principal components procedure yielded 5 factors. The first 10 subitems of the FIQ loaded on the first factor with component loadings ranging from 0.50 to 0.95. No other items loaded on the first factor. Factor 2 consisted of work difficulty, feeling good, pain, fatigue, rest, and stiffness. Anxiety, depression, and days of work missed all loaded on separate factors.

There were no significant correlations between any items on the FIQ and age, time since diagnosis, education or income. In addition, there were no significant differences in item means between those who were employed and those unemployed or among any of the 3 occupational categories.

DISCUSSION

The results of our study focus on 2 separate but related issues. First, does the AIMS, which is known to be reliable and valid in several rheumatic disease subgroups, have the same positive properties in another subgroup, particularly one with fewer objectives diagnostic features? Second, does a new instrument geared more specifically to fibromyalgia have psychometric properties that make it a potentially valuable addition to the measurement of outcomes in rheumatic disease care?

The 3-week test-retest correlation coefficients on the AIMS scales were moderate to high and similar, although not quite as high, to those reported by Meenan and colleagues for a 2-week interval[8]. These differences are most easily attributed to the longer time between retests. However, it is also possible that they reflect instability of the fibromyalgia syndrome or error in measurement.

Overall, the AIMS appears to have construct validity in this fibromyalgia sample. Most scales correlated adequately with the 2 symptom severity measures that we obtained. However, we were unable to evaluate adequately the AIMS in relations to tender point scores. This needs to be done since at this time tender points are the key to the diagnosis and evaluation of patients with fibromyalgia. The social functioning scale is somewhat problematic because in our study it had little correlation with the measures of syndrome severity and, therefore, may not be a valid indicator of health status in this population.

The content validity of the physical functioning component was somewhat problematic. For example, items on the household activities scale, such as taking medicine, handling money, using the telephone, as well as all the ADL items (bathing, dressing, toileting, moving about) were not problems for the patients with fibromyalgia. On the other hand, some of the impairments in mobility and physical functioning marked by over 25% of the patients, such as the perception of having to stay indoors most of the day, being confined to bed or a chair, and being unable to walk more than one block represent a level of impairment beyond what is currently appreciated. It is possible that some changes in the AIMS could enhance its validity in the population with fibromyalgia. However, any changes should not be made

without a careful study of the psychometric ramifications.

Results of the psychometric testing of the FIQ indicate that it has some positive properties. The instrument has sufficient test-retest reliability to be used in further research. Physical functioning and amount of work missed were highly reliable on test-retest, showing little random error in their repeated measurement. Ability to do job, anxiety, depression, and days felt good all exceeded 0.70 test-retest correlations. Pain, stiffness, fatigue, and morning tiredness had lower but still respectable test-retest correlations. Since the visual analog approach is generally accepted as a reliable and valid way to obtain patient perception data, these moderate test-retest correlation coefficients may be reflecting the changing, unpredictable nature of fibromyalgia symptoms. Fatigue and feeling tired on awakening may be especially vulnerable to these fluctuations.

Validity testing of an instrument is always a more theoretical process than reliability testing. In the case of the FIQ it was especially difficult to test validity because the fibromyalgia syndrome lacks the type of objective criteria (gold standard) available for some other rheumatic diseases such as rheumatoid arthritis and also because until now no health status instrument has been tested for its reliability and validity in this population. Thus, we were in the position of first testing the psychometric properties of the AIMS in our fibromyalgia sample and then cautiously comparing the 2 scales.

As hypothesized, the FIQ pain, depression, and anxiety visual analogs and FIQ physical functioning item were all moderately to highly correlated with the comparable AIMS scales. These significant correlations between the items on the FIQ and the comparable scale of the AIMS are indicative of some degree of convergent construct validity of both instruments.

The correlations between the FIQ items and the 2 self-rated measures of syndrome severity were within the boundaries usually seen for subjective indicators; and in the case of the item measuring syndrome severity during the past month, the correlations were appreciably higher than the correlations between the same question and the AIMS scales. The tender point-FIQ correlations in the second sample were consistent with the first sample and on 4 items somewhat higher.

The factor analysis demonstrated that, much like the AIMS, the FIQ is also multidimensional. The first 10 subitems form a strong physical functioning item. Although these items are heavily weighted toward women, so is the syndrome. It is possible that the addition of more items in the physical functioning section could increase its comprehensiveness; however, the current subitems do measure areas of concern to most women patients with fibromyalgia. A possible addition to the questionnaire could be a perceived physical discomfort factor; this would be different from the psychological discomfort of anxiety and depression or the more objective factor of actual days of work missed. Several patients commented that the number of days of work they missed had little to do with how they felt on any particular day.

The lack of correlations or significant mean differences between the FIQ items and the demographic variables are interesting clinically. Basically, what we found here was that age and years of having fibromyalgia were not related to functioning. Older women or those who had had the syndrome longer did not have higher impact from their fibromyalgia than younger women. Unlike other reports that have found a consistent inverse relationship between lower educational level and higher functional impairment, we found no relationship between education and the individual items on the FIQ. Also, our arbitrary categorization of occupation yielded no significant results. We had few women who were in heavy factory worker-type positions. It may be more productive to categorize occupations in terms of their repetitive or static types of activities (long periods of standing in one place or sitting at a computer terminal) or level of perceived job stress.

At this point in its development, the FIQ appears to be a potentially valuable addition to outcome measurement in fibromyalgia research and clinical care. It has evidence of construct validity, sufficient test-retest reliability, and content relevance. Also, it is short and includes items on job performance, fatigue, and morning tiredness. We recognize that with brevity comes some compromise in depth of information. For example, only one item relates to depression. Nonetheless, our data show a high correlation between the visual analog item of depression and the 6-question AIMS depression scale. Thus, for purposes of outcome measurement when one score is being used as the measure of depression, we would argue at this point, that the visual analog item tells us as much as the aggregated AIMS depression scale.

Further work is needed. Longitudinal testing in clinical trials could provide information on the reliability of the instrument versus the stability or instability of the syndrome. Sensitivity to change is a critical feature of any health status instrument and the robustness of the FIQ in this dimension needs to be established. Although the correlations obtained between the FIQ items and number of tender points were in the same range as most correlations of health status with objective measures, more data is needed to establish the degree of relationship between the 2 measures. Lastly more evidence of construct validity needs to be gathered. The FIQ should be compared to other validated instruments. This is not an easy task as most outcome studies of instruments used in rheumatology have not included patients with fibromyalgia. Nevertheless, the physical functioning item in our instrument could be compared to the Beck Depression and Anxiety Inventories or the McGill Pain Questionnaire, for example. We are currently gathering additional data to support the sensitivity of the FIQ to clinical change resulting from participation in the comprehensive treatment program.

REFERENCES

1. Meenan RF, Gertman PM, Mason JH: Measuring health status in arthritis: The Arthritis Impact Measurement Scales. *Arthritis Rheum 1980;*23:146–52.
2. Fries JF, Spitz P, Kraines RG, Holman HR: Measurement of patient outcome in arthritis. *Arthritis Rheum 1980;*23:137–45.
3. Jette AH: The Function Status Index: Reliability and validity of a self-report functional disability measure. *J Rheumatol 1987;*(suppl 15)14:15–9.
4. Cathey MA, Wolfe F, Kleinheksel SM, Hawley DJ: Socioeconomic impact of fibrositis. *Am J Med 1986;*(suppl 3A)81:78–84.
5. Cathey MA, Wolfe F, Kleinheksel SM, Miller S, Pitetti KH: Functional ability and work status in patients with fibromyalgia. *Health Care Res 1988;*1:85–98.
6. Hawley DH, Wolfe F, Cathey MA: Pain, functional disability and psychological status: A 12-month study of severity in fibromyalgia. *J Rheumatol 1988;*15:1551–6.
7. Wolfe F, Smythe HA, Yunus MB: The American College of Rheumatology 1990 criteria for the classification of fibromyalgia. *Arthritis Rheum 1990;*33:160–72.
8. Meenan RF, Gertman PM, Mason JH, Dunaif R: The Arthritis Impact Measurement Scales: Further investigations of a health status measure. *Arthritis Rheum 1982;*25:1048–53.
9. Mason JH, Anderson JJ, Meenan RF: A model of health status for rheumatoid arthritis. *Arthritis Rheum 1988;*31:714–20.
10. Pincus T, Callahan LF, Sale MG, Brooks AL, Payne LE, Vaughn WK: Severe functional declines, work disability, and increased mortality in seventy-five rheumatoid arthritis patients studied over nine years. *Arthritis Rheum 1984;*27:864–72.
11. Pincus T, Callahan LF: Formal education as a marker for increased mortality and morbidity in rheumatoid arthritis. *J Chron Dis 1985;*38:973–84.
12. Meenan RF, Kazis LE, Anderson JJ: The stability of health status in rheumatoid arthritis: A five-year study of patients with established disease. *Am J Public Health 1988;*78:1484–7.