Abstract

Objectives: To examine the construct validity of the International FItness Scale (IFIS) (ie, self-reported fitness) against objectively measured physical fitness in women with fibromyalgia and in healthy women; and to study the test-retest reliability of the IFIS in women with fibromyalgia.

Design: Cross-sectional study.

Setting: Fibromyalgia patient support groups.

Participants: Women with fibromyalgia (n = 413) and healthy women (controls) (n = 195) for validity purposes and women with fibromyalgia (n = 101) for the reliability study. The total sample was N=709.

Interventions: Not applicable.

Main Outcome Measures: Fitness level was both self-reported (IFIS) and measured using performance-based fitness tests. For the reliability study the IFIS was completed on 2 occasions, 1 week apart.

Results: Women with fibromyalgia who reported average fitness had better measured fitness than those reporting very poor fitness (all \( P < .001 \), except 6-minute walk test where \( P < .05 \)), with similar trends observed in healthy control women. The test-retest reliability of the IFIS, as measured by the average weighted \( \kappa \), was .45.

Conclusions: The IFIS was able to identify women with fibromyalgia who had very low fitness and distinguish them from those with higher fitness levels. Furthermore, the IFIS was moderately reliable in women with fibromyalgia.

Archives of Physical Medicine and Rehabilitation 2016;97:395-404

© 2016 by the American Congress of Rehabilitation Medicine

Fibromyalgia is a complex disorder characterized by widespread musculoskeletal pain and other symptoms (eg, fatigue, sleep disorder, cognitive problems, depression, anxiety). Fibromyalgia is markedly more prevalent in women, and quality of life in patients is poor. People with fibromyalgia have lower physical fitness than age-matched healthy peers, and their fitness has been reported to be similar to that of healthy older adults. Higher levels of physical fitness have been related to lower levels of pain, lower levels of fibromyalgia severity, and...
better quality of life in patients with fibromyalgia. Therefore, rehabilitation programs focused on improving physical fitness are commonly recommended in fibromyalgia.

Busch et al emphasized the importance of characterizing physical fitness in people with fibromyalgia. Mannerkorpi et al suggested that fitness testing could complement clinical examination when planning treatment for patients with fibromyalgia. Indeed, our group has shown that fitness testing is a powerful tool to discriminate between patients with and without fibromyalgia. The aforementioned studies suggest that physical fitness is an important clinical indicator of health in people with fibromyalgia. Additionally, the assessment of physical fitness is key to individualize the optimal dose of exercise for rehabilitation purposes.

Laboratory- or field-based fitness tests are not always practical or possible to conduct in clinical settings or in large surveys and epidemiologic studies because of limited time, equipment, or qualified personnel. It seems that a cheap, quick tool to estimate physical fitness would be helpful in fibromyalgia. Additionally, a discrepancy between self-reported disability and clinically observed disability has been reported in people with fibromyalgia. Therefore, even when performance-based fitness testing is feasible, a complementary tool to assess self-reported fitness levels might provide useful information. In this context, many researchers have recommended that both self-report and performance-based functional measures are used to assess physical fitness.

The International Fitness Scale (IFIS) is a self-reported measure of physical fitness that could easily be implemented in patients with fibromyalgia. This scale has been validated in children, adolescents, and young adults; however, it is unknown whether the IFIS represents a valid and reliable estimate of physical fitness in women with fibromyalgia. The symptoms and comorbidities found in people with fibromyalgia (eg, pain, chronic fatigue, nonrestorative sleep, depression, cognitive impairment) could influence the accuracy of self-reported physical fitness. Therefore, the present study aimed to examine the construct validity of the IFIS (ie, self-reported fitness) against objectively measured physical fitness in women with fibromyalgia and in healthy women (substudy 1) and examine the test-retest reliability of the IFIS in women with fibromyalgia (substudy 2).

Methods

Study design and sample

Women with fibromyalgia and healthy women volunteers (controls) were recruited from the 8 provinces of the Andalusia Region of Southern Spain between 2011 and 2013 (and were part of a larger study, the al-Ándalus project). Local patient associations were contacted by e-mail, letter, or telephone. Controls were participants’ acquaintances and members of the general public recruited via e-mail and Internet advertisements. Data from these participants were used to examine the construct validity of the IFIS (substudy 1).

A pilot study was performed before the al-Ándalus project, with women with fibromyalgia from a local patient association in Granada, Spain. Data from the pilot study were used to address the test-retest reliability of the IFIS (substudy 2).

Substudy 1: construct validity

To be included in the study, women with fibromyalgia had to be previously diagnosed with fibromyalgia by a rheumatologist and meet the 1990 American College of Rheumatology fibromyalgia classification criteria. People were excluded from the study if they had an acute or terminal illness or cognitive impairment (defined as a Mini-Mental State Examination score ≤10). The participants from the control group had to meet the same criteria except those that were disease-specific (ie, fibromyalgia diagnosis and to meet the 1990 American College of Rheumatology criteria). We set 3 additional inclusion criteria for the present study. First, to allow us to obtain age-matched groups, participants had to be aged between 37 and 65 years. Second, for the validity analyses, participants should not have participated in any other assessments involving physical fitness within the last 3 months to avoid influence on self-reported fitness. Finally, participants had to have data available for at least 1 fitness test. The final sample was comprised of 413 women with fibromyalgia and 195 controls. The flowchart of participants is presented in figure 1.

Substudy 2: test-retest reliability of the IFIS

After applying the same inclusion criteria previously described, 101 women with fibromyalgia were included in the pilot study for the al-Ándalus project and completed the IFIS on 2 occasions, 1 week apart. To avoid any influence that the results of the physical fitness tests could have on the participants’ perception of their own fitness level (ie, IFIS), the physical fitness testing took place 1 week after the retest of the IFIS. The reliability substudy did not include a control group.

Ethical issues

The study was performed following the ethical guidelines of the Declaration of Helsinki, which was last modified in 2000. Ethics approval was obtained by the Ethics Committee of the Virgen de las Nieves Hospital (Granada, Spain), and all patients gave written informed consent.

Outcomes

Self-reported physical fitness

Self-reported fitness was assessed by means of the IFIS, previously validated in European adolescents and in young Spanish adults. The IFIS is comprised of 5 Likert-scale questions about self-reported fitness. Each question has 5 response options (very poor, poor, average, good, very good) about overall fitness and specific fitness components: cardiorespiratory fitness, muscular strength, speed-agility, and flexibility.

Physical fitness

Physical fitness was assessed by means of the standardized Functional Senior Fitness Test Battery, and the handgrip strength test. These tests are commonly used in fibromyalgia.

List of abbreviations:

ANOVA analysis of variance
IFIS International Fitness Scale
and have demonstrated to be feasible and reliable in women with this condition. This battery of tests assesses cardiorespiratory fitness, muscular strength, speed-agility, and flexibility. To prevent fatigue, the tests were carried out alternating upper- and lower-body tests, with 1-minute rest between tests, in the following order: chair sit-and-reach, back scratch, handgrip, 30-second chair stand, arm curl, 8-foot up-and-go, and 6-minute walk tests.

### Physical fitness: Cardiorespiratory fitness

The 6-minute walk test measures the maximum distance (m) that participants are able to walk in 6 minutes along a 45.7-m rectangular course. This test has been shown to be a valid and reliable measure of cardiovascular fitness in women with fibromyalgia.

### Physical fitness: Lower-body muscular strength

The 30-second chair stand test involves counting the number of times, in 30 seconds, that an individual can rise to a full stand from a seated position with back straight and feet flat on the floor, without pushing off with the arms. Participants performed 1 trial after becoming familiar with the procedure.

### Physical fitness: Upper-body muscular strength

The arm curl test involves determining the number of times a hand weight (2.3kg for women) can be curled through a full range of elbow motion in 30 seconds. The test was performed once with each arm and the average number of repetitions was recorded. The handgrip test was performed with a digital dynamometer as described by Ruiz et al. Participants performed the test twice for each hand (alternating between hands), with 1-minute rest between trials. The average of the best score for each hand was used.
Physical fitness: Speed-agility
The 8-foot up-and-go test involves standing up from a chair, walking 2.44m (8ft) to a cone, and returning to the chair in the shortest period of time.43 The best time of 2 trials was recorded.

Physical fitness: Lower-body flexibility
In the chair sit-and-reach test, the participant was seated with 1 leg extended, slowly bending forward and sliding the hands down the extended leg in an attempt to touch (or pass) the toes. The number of centimeters short of reaching the toes (negative score) or reaching beyond them (positive score) was recorded.43 The test was performed twice for each leg (alternating between legs), and the average of the best score for each leg was selected.

Physical fitness: Upper-body flexibility
The back scratch test provides a measure of the overall shoulder range of motion, as the distance between (or overlap of) the middle fingers behind the back with a ruler.43 Participants performed the test twice (alternating between arms), and the average of the best score for each hand was used.

Statistical analysis
Construct validity of the IFIS in fibromyalgia
The ability of the IFIS to correctly rank patients with fibromyalgia and healthy women into appropriate physical fitness levels was determined by means of a 1-way analysis of variance (ANOVA). Post hoc group comparisons with Tukey correction were applied to assess the differences in measured fitness across categories of self-reported fitness (the IFIS). Performance-based fitness variables were entered as dependent variables, and self-reported fitness variables were used as fixed factors. Differences in performance-based fitness z scores among self-reported fitness categories were analyzed by means of a 1-way ANOVA, and post hoc group comparisons with Tukey correction were applied to assess the differences in performance-based fitness across categories on self-reported fitness (the IFIS). Before the analysis, a clustered score of physical fitness (general fitness) was computed as the average of the standardized scores ([value−mean]/SD) from the 7 fitness tests, weighted for the number of tests assessing the same component (ie, handgrip, arm curl, and chair stand tests assessed muscle strength; back scratch and chair sit-and-reach assessed flexibility). The standardized score from each of the physical fitness tests and general fitness were entered as dependent variables in separate models, and the corresponding self-reported fitness component and overall fitness were entered into each model as fixed factors.

Reliability of the IFIS in fibromyalgia
The t test was used to analyze the differences between women with fibromyalgia and controls in body mass index and tender points. The test-retest reliability of the IFIS was examined by means of weighted κ coefficients. Weighted κ coefficients were calculated, which is appropriate when dealing with ordered categorical data,47 because weighted κ accounts for strict agreement and provides weighting to adjacent categories. Linear (instead of quadratic) weights were chosen because the distance between adjacent categories was equally important.35 Statistical analysis was performed with SPSS,b and the significance of the tests was interpreted as suggested by Sterne and Smith.48 Cohen weighted κ is not available in the standard SPSS package, but command syntax is available from the Knowledgebase at www.SPSS.com.b Data for imputation into the syntax were generated from cross-tabulation. The level of significance was set at P<.05.

Results
The sociodemographic characteristics of the study groups are shown in table 1. As a group, women with fibromyalgia had higher body mass index (P<.001) than controls. Moreover, educational status (P=.015) and occupational status (P<.001) were statistically different between groups. No differences were found for age or marital status (P>.05). The distribution of the answers of the IFIS was positively skewed, with only a small number of women with fibromyalgia reporting good and very good fitness (fig 2). More than half of the healthy women, however, reported their overall fitness to be average, good, or very good (see fig 2). Controls also had a more normative distribution of responses to
For fibromyalgia, there were very few subjects (range, 0.5%–1%) in category 4 (good–very good); consequently, this group is underrepresented, and results should be interpreted carefully.

The comparison of performance-based fitness across self-reported fitness categories in women with fibromyalgia and controls is shown in table 2. Overall, a linear relation between self-reported and performance-based fitness was observed in women with fibromyalgia and controls. The post hoc analysis revealed that for all the fitness variables studied, those women with fibromyalgia reporting an average fitness had better performance-based fitness than those reporting very poor fitness levels (all P<.001, except 6-minute walk test P<.05). Moreover, control women reporting good/very good fitness had better performance-based fitness than those reporting poor or very poor fitness levels (all P<.001, except back scratch test P<.05). The association between self-reported and performance-based physical fitness is presented in figures 3 and 4. There was a linear relation between item categories of the IFIS in both groups, in all fitness components, and in overall fitness. Overall, the 6-minute walk test showed the weakest relation between self-reported fitness and performance-based fitness in women with fibromyalgia, whereas flexibility tests showed the strongest relation.

Construct validity of the IFIS in women with fibromyalgia

The comparison of performance-based fitness across self-reported fitness categories in women with fibromyalgia and controls is shown in table 2. Overall, a linear relation between self-reported and performance-based fitness was observed in women with fibromyalgia and controls. The post hoc analysis revealed that for all the fitness variables studied, those women with fibromyalgia reporting an average fitness had better performance-based fitness than those reporting very poor fitness levels (all P<.001, except 6-minute walk test P<.05). Moreover, control women reporting good/very good fitness had better performance-based fitness than those reporting poor or very poor fitness levels (all P<.001, except back scratch test P<.05). The association between self-reported and performance-based physical fitness is presented in figures 3 and 4. There was a linear relation between item categories of the IFIS in both groups, in all fitness components, and in overall fitness. Overall, the 6-minute walk test showed the weakest relation between self-reported fitness and performance-based fitness in women with fibromyalgia, whereas flexibility tests showed the strongest relation.

Reliability of the IFIS in women with fibromyalgia

The test-retest reliability statistics for the 5 items comprising the IFIS are displayed in table 3. Weighted κ ranged from .40 (muscular fitness) to .63 (flexibility), and the averaged weighted κ was .45.
Discussion

This study demonstrated that the IFIS has moderate validity and test-retest reliability for ranking women with fibromyalgia according to their objectively measured physical fitness level. The IFIS was able to identify women with fibromyalgia who had very low fitness and distinguish them from those with average fitness levels. Although objective measures of physical fitness are generally recommended, these findings suggest that the IFIS could be useful when performance-based testing is not feasible. The IFIS may also be useful as a screening test in epidemiologic studies to detect women with fibromyalgia who have poor fitness because low fitness is a consistent indicator of more severe symptomatology in this population. Therefore, both self-reported questionnaires and performance-based tests are essential components of a comprehensive clinical assessment.

Consistent with previous work, participants with fibromyalgia in the current study had lower educational status than healthy women, and many were retired, were on sick leave, or were unemployed. The reduced employment status of those with fibromyalgia could be the result of the symptomatology and physical disability associated with the condition, emphasizing the importance of a physical fitness assessment in fibromyalgia.

Validity

Women with fibromyalgia reported worse physical fitness levels than age- and sex-matched controls. Approximately 65% to 85% of women with fibromyalgia reported very poor or poor physical fitness, whereas only 8% to 50% of control women reported very poor or poor physical fitness. The results from performance-based fitness testing confirmed that women with fibromyalgia had lower physical fitness levels than age- and sex-matched controls, which is also supported by previous studies. The IFIS was able to detect these differences in fitness, supporting the validity of the IFIS to discriminate between levels of fitness in women with fibromyalgia and healthy controls.

Fig 3  Comparison between self-reported fitness and measured fitness for cardiorespiratory fitness, speed-agility, and flexibility. Data represented means and 95% confidence intervals. One-way ANOVA was carried out to test whether participants reporting better fitness (IFIS) had significantly better fitness performance. Each measured fitness variable was compared with the corresponding item in the IFIS. Common superscripts indicate significant differences ($P<.001$) between groups. Results for the extreme categories (ie, very poor and good/very good) should be interpreted cautiously because of the small sample size (range, 4–27). $\dagger$The lower the score, the better the performance. $\#$Confidence interval for speed-agility was −.51 to .51, and for the flexibility chair sit-and-reach it was −.99 to .99.
Our results indicate that the IFIS can identify women with fibromyalgia who have very poor fitness, and these women are significantly different from those with an average fitness level. Paradoxically, no differences were found for many of the tests studied between patients reporting a very poor fitness and those reporting good/very good fitness. We believe that this lack of significance is explained by the low statistical power caused by the low number of participants reporting good/very good fitness. The opposite situation was found in the healthy control women, in that there were few reporting very poor fitness. Consequently, the results involving participants reporting good/very good fitness in the fibromyalgia group and very poor fitness in the control group must be interpreted with caution. In this regard, it is also possible that the clinical utility of the IFIS is limited in women with fibromyalgia with high fitness levels. However, the IFIS could have still an important role in clinical settings to identify patients with very low fitness. In the general population, individuals with a very low fitness (ie, first quintile) have more than double risk of cardiovascular disease and death from any cause than individuals with higher fitness levels. The health risks are much smaller among those groups with higher fitness levels (ie, second to fifth quintiles).13,56-59 This evidence strongly supports that the key group to be screened and targeted is the one belonging to the first quintile or equivalent (ie, individuals with a very low fitness level). From a clinical point of view, it is also important to identify those individuals with fibromyalgia with very low physical fitness because

![Fig 4](image-url)

Fig 4  Comparison between self-reported fitness and measured fitness for muscular strength and overall fitness. Data represented means and 95% confidence intervals. One-way ANOVA was carried out to test whether participants reporting better fitness (IFIS) had significantly better fitness performance. Each measured fitness variable was compared with the corresponding item in the IFIS. Common superscripts indicate significant differences (P<.001) between groups. *Results for the extreme categories (ie, very poor and good/very good) should be interpreted cautiously because of the small sample size. #Confidence interval for 30-s chair stand test was -2.41 to 2.41, for arm curl test it was -2.01 to 2.01, for handgrip test it was -1.05 to 1.05, for overall fitness (very poor) it was -1.57 to 1.57, and for overall fitness (poor) it was -0.78 to 0.78.

| Table 3 Test-retest reliability of the IFIS in women with fibromyalgia |
|-----------------|-----------------|-----------------|
| IFIS Items      | Weighted κ Coefficients | 95% Confidence Interval |
| Overall fitness | .45              | .07             | .31—.59          |
| Cardiorespiratory fitness | .46              | .08             | .30—.61          |
| Muscular fitness | .40              | .07             | .25—.55          |
| Speed-agility   | .46              | .07             | .33—.59          |
| Flexibility     | .63              | .07             | .52—.74          |

www.archives-pmr.org
they are at higher risk of general future diseases and perhaps also at a higher risk of worsening specific fibromyalgia symptoms.

The findings of this study are consistent with previous work examining the validity of the IFIS in children, adolescents, and young adults, supporting a linear association between self-reported and objectively measured physical fitness in both fibromyalgia and control women. Our results also concur with previous studies using other questionnaires or scales to estimate fitness, in which a moderate to good agreement with measured fitness was observed in middle-age and older adults.

Reliability

The test retest reliability of the IFIS was .45, indicating moderate reliability of the IFIS in women with fibromyalgia. The most reliable fitness component was flexibility (κ = .63), whereas muscular strength was the least reliable (κ = .40). The large variability of the fibromyalgia-related symptoms, with intermittent periods of exacerbation and remission, may explain these differences. For example, symptoms (e.g., fatigue, quality of rest on the night before testing, weather) could vary from one day to another, which could influence self-reported fitness responses. Even if the physical performance of women with fibromyalgia remains steady from day-to-day, daily variations in symptoms severity might influence their perceptions about physical capability. The average test-retest weighted κ for the IFIS when used in healthy young adults was .59, which is slightly higher than observed in the present study, but leads to a similar conclusion (i.e., the IFIS is a moderately reliable tool in these 2 populations). It must also be taken into account that the biologic variability (in addition of the instrument’s variability) might have an effect on the reliability of the IFIS. Previous research reported significant test-retest differences in the arm curl, 30-second chair stand, and 8-foot up-and-go tests in women with fibromyalgia; however, the mean intratrial difference for these measurements was low, and the effect size of the mean differences was small (Cohen d < .25). It appears therefore that from a practical point of view, these tests can be used to evaluate physical fitness in women with fibromyalgia. Our findings support this previous study in that muscular strength was the least reliable of all fitness components in women with fibromyalgia.

Study limitations

This study had some limitations (e.g., there was no control group for the reliability study). The IFIS was designed to be used in large populations; however, the number of participants in some categories in our study was small, and the results regarding categories with small sample sizes should be interpreted with caution. For this reason, 2 categories of the IFIS were merged (good and very good), which could have influenced the results. Although the physical fitness tests used have been shown to be feasible and reliable in women with fibromyalgia, the validity of some of the fitness tests used has not been studied; therefore, their use as a comparative standard should be taken with caution. Finally, the design of this study did not allow for examination of the possible variability of physical capabilities at different time points.

Strengths

On the other hand, this is one of the largest studies examining the construct validity of a self-reported fitness tool against performance-based physical fitness in women with fibromyalgia. Moreover, the fact that sampling was designed to be representative of the fibromyalgia population in Southern Spain should also be acknowledged.

Conclusions

The results of this study demonstrate that the IFIS can be a useful tool to identify women with fibromyalgia who had very low fitness and distinguish them from those with higher fitness levels, which is important from a clinical point of view. On the other hand, because of the low frequency of participants reporting good/very good fitness levels, the utility of the IFIS to detect women with FM with high fitness is uncertain. Furthermore, the IFIS demonstrated moderate test-retest reliability in women with fibromyalgia.

Keywords

Chronic pain; Fibromyalgia; Muscle strength; Physical fitness; Rehabilitation; Self-assessment

Acknowledgments

We thank the AGRAFIM (Association of Fibromyalgia from Granada, Southern Spain) and the Andalusian Federation of People with Fibromyalgia (Federación Alba Andaluca) members involved in the field work for their effort.

References


47. Cohen J. Weighted kappa: nominal scale agreement provision for scaled disagreement or partial credit. Psychol Bull 1968;70:213-20.