

A Randomized Clinical Trial Comparing Fitness and Biofeedback Training versus Basic Treatment in Patients with Fibromyalgia

MARIJKE van SANTEN, PAULIEN BOLWIJN, FRANS VERSTAPPEN, CARLA BAKKER, ALITA HIDDING, HARRY HOUBEN, DESIREE van der HEIJDE, ROBERT LANDEWÉ, and SJEFF van der LINDEN

ABSTRACT. *Objective.* To compare the therapeutic effects of physical fitness training or biofeedback training with the results of usual care in patients with fibromyalgia (FM).

Methods. One hundred forty-three female patients with FM (American College of Rheumatology criteria) were randomized into 3 groups: a fitness program (n = 58), biofeedback training (n = 56), or controls (n = 29). Half the patients in the active treatment groups also received an educational program aimed at improving compliance. Assessments were done at baseline and after 24 weeks. The primary outcome was pain [visual analog scale (VAS)]. Other endpoints were the number of tender points, total myalgic score (dolorimetry), physical fitness, functional ability (Arthritis Impact Measurement Scale and Sickness Impact Profile), psychological distress (Symptom Checklist-90-Revised), patient global assessment (5 point scale), and general fatigue (VAS).

Results. Baseline scores were similar in all 3 groups. Altogether 25 (17.5%) patients dropped out; they were similarly distributed over all groups: 14 patients after randomization and 11 (8%) during the study. A true high impact level for fitness training was not attained by any patient. After treatment, no significant differences in change scores of any outcome were found between the groups (ANOVA, $p > 0.05$). All outcome measures showed large variations intra- and interindividually. The educational program did not result in higher compliance with training sessions (62% vs 71%). Analysis of the subgroup of subjects with a high attendance rate ($> 67\%$) also showed no improvement.

Conclusion. In terms of training intensity and maximal heart rates, the high impact fitness intervention had a low impact benefit. Therefore effectiveness of high impact physical fitness training cannot be demonstrated. Thus compared to usual care, the fitness training (i.e., low impact) and biofeedback training had no clear beneficial effects on objective or subjective patient outcomes in patients with FM. (J Rheumatol 2002;29:575–81)

Key Indexing Terms:

FIBROMYALGIA
AEROBIC EXERCISES

RANDOMIZED CLINICAL TRIAL
BIOFEEDBACK

PHYSICAL FITNESS
ASSESSMENT

From the Department of Internal Medicine, Division of Rheumatology, University Hospital Maastricht; University of Maastricht, Maastricht; Department of Rheumatology, The Atrium Medical Center, Heerlen, The Netherlands; University of Diepenbeek, Diepenbeek, Belgium; Institute for Rehabilitation Research, Hoensbroek, The Netherlands.

Supported by the Dutch Arthritis Association.

M. van Santen-Hoeufft, MD, Rheumatologist; P. Bolwijn, MSc, Department of Internal Medicine, Division of Rheumatology, University Hospital Maastricht; F. Verstappen, MD, PhD, University of Maastricht; C. Bakker, MSc, PhD, Department of Internal Medicine, Division of Rheumatology, University Hospital Maastricht; A. Hidding, MSc, PhD, Senior Researcher Institute for Rehabilitation Research; H. Houben, MD, PhD, Consultant Rheumatologist, Department of Rheumatology, The Atrium Medical Center; D. van der Heijde, MD, PhD, Professor of Rheumatology, University of Diepenbeek; R. Landewé, MD, PhD, Atrium Medical Center; Sj. van der Linden, MD, PhD, Professor of Rheumatology, Institute for Rehabilitation Research.

Address reprint requests to Dr. M. van Santen-Hoeufft, Department of Internal Medicine, Division of Rheumatology, University Hospital Maastricht, PO Box 5800, NL-6202-AZ, Maastricht, The Netherlands. E-mail: msa@sint.azm.nl

Submitted February 26, 2001; revision accepted August 17, 2001.

Fibromyalgia (FM) syndrome is a common condition characterized by widespread pain and tenderness at specific anatomic locations¹. It is associated with a wide variety in somatic and psychological symptoms like irritable bowel syndrome, headaches, sleep disturbances, fatigue, anxiety, and depressive mood. It affects mainly women between 30 and 50 years². Treatment is often unsatisfactory. A critical appraisal of published randomized controlled treatment modalities shows disappointing results: simple analgesics and nonsteroidal antiinflammatory drugs (NSAID) have no convincing positive effects³. Tricyclic antidepressive drugs (amitriptyline) give some relief of sleep disturbances in a subgroup of patients, but pain is not relieved⁴. A short term beneficial effect on some different outcome measures has emerged from trials with fitness programs, biofeedback, hypnotherapy, education and physical therapy, electroacupuncture, and psychomotor therapies³. In general, however, study durations and followups were short, and

there were considerable placebo effects making the reliability of the results questionable. The positive results of 2 controlled studies, namely fitness training by McCain, *et al*⁵ and biofeedback training by Ferracioli, *et al*⁶, led us to compare both interventions in the present study.

We report the results of a randomized clinical trial comparing fitness and biofeedback training versus usual care in female patients with FM. Half the patients in the active intervention groups additionally received an educational program, aimed at promoting compliance.

MATERIALS AND METHODS

Recruitment of patients. Patients with FM were recruited from the central registry for the diagnosis of rheumatic diseases. Within 2 years the departments of rheumatology in Heerlen and Maastricht had registered 453 patients with a clinical diagnosis of FM (about 15% of all newly referred outpatients). For practical reasons only female patients aged 18 to 60 years and living within 30 km of either center were invited to participate. By prescreening medical records, we excluded subjects with known comorbidity and those with more localized myalgia (MvS). Of the remaining 268 patients, a total of 168 gave written informed consent. Main reasons for nonparticipation were lack of interest ($n = 64$) and the refusal to reply even after a second letter ($n = 33$). Before the study, one of us (MvS) assessed all patients to confirm the diagnosis of FM according to the American College of Rheumatology (ACR) criteria¹ and to exclude patients with ischemic heart disease, arrhythmia, exercise induced asthma, unsettled disability compensation disputes, or incapacitating psychological distress. A total of 10 patients did not meet the ACR criteria, 8 were excluded because of emotional disorders, and 7 were not considered suitable for other reasons (pregnancy, on waiting list for elective surgery, vacation during trial). Thus 143 patients were enrolled.

Interventions. Fitness training. A professional female instructor applied and closely supervised the fitness training program to groups of 15 to 17 patients at a time. The groups exercised in one of 2 private fitness centers offering similar facilities. Patients trained 60 min, twice weekly, for 24 consecutive weeks. Subjects were encouraged to attend a third, unsupervised weekly session, also 60 min, and to use the sauna or swimming pool after all the sessions. Each session comprised a warmup period with aerobic exercises and postural muscle stretching (10 min), intensive aerobic exercises alternated with general flexibility and balance exercises (30 min), and isometric muscle strengthening (10 min). The session ended with a cooling down period comprising aerobic, stretching, and relaxation exercises (10 min). This protocol is in accordance with the guidelines of the American College of Sports Medicine⁷. It should be noted, however, that the training intensity was left up to each patient. That is, an individual patient was allowed to stop or interrupt this process of exercising whenever she felt too tired or experienced too much pain.

Biofeedback training. Biofeedback training comprised individual 30 min sessions in the hospital twice weekly during 8 weeks. In the first session general suggestions were given to each subject to accomplish muscle relaxation. To have feedback on the level of relaxation a biofeedback apparatus was used as tonometer. Measures were recorded with a Myotron 22 or a Myomad 222 using standard electrode placements on the forehead, giving feedback with a visual signal. A change in tension of the musculus frontalis is shown by a change in digital numbers on the screen of the biofeedback apparatus. In the subsequent 15 sessions patients were taught the progressive relaxation technique, consisting of alternately tightening and relaxing of separate groups of muscles. The aim was to bring down the number recorded on the biofeedback apparatus. Each patient was trained by one regular supervisor (psychologist or physiotherapist). They additionally encouraged each subject to practice the progressive relaxation technique twice daily at home by using an audio tape with instructions. After the 8

weeks of supervised sessions all patients were instructed to continue their twice daily relaxation exercises during the next 16 weeks.

Educational program. Per center, subjects within each intervention group were randomized to additionally receive an educational program aimed at compliance with fitness or biofeedback training. The educational program consisted of 6 health promotion sessions of 90 min each, spread over the 24 weeks, supervised by a health educator. The sessions included information on FM, general health education, self-management, and relapse prevention principles. The importance of physical conditioning was stressed in the education classes with subjects randomized to fitness training, and the importance of relaxation strategies in the education classes with subjects randomized to biofeedback training.

Control treatment. Control patients received the usual care at the outpatient department and by their general physician. That is, they were treated with analgesics, NSAID, or tricyclic antidepressive agents, if considered appropriate. Physiotherapy and medical counseling were allowed. However, their general physicians were informed that aerobic exercises and relaxation therapy were neither to be prescribed nor encouraged. Control patients were offered the best study treatment free of charge after completion of the study.

Assessment. Pain was the primary endpoint. The other endpoints were physical fitness, health status, psychological distress, and patient global assessment. All patients were assessed at baseline and after 12 and 24 weeks by the same observers, blinded for the treatment allocation. Patients were asked not to discuss the allocation with the assessors. Whenever possible the followup measurement was done at the same time of the day as the baseline assessment to control for diurnal fluctuation.

Instruments. Pain. Patients marked the estimated extent of pain they had experienced during the previous week on a horizontal 100 mm visual analog scale (VAS: 0 = no pain, 100 = the worst imaginable). In addition, the assessor palpated tender and control points and applied dolorimetry. The 18 tender points according to the ACR criteria for the classification of FM were scored as being positive when the subject noticed pain at thumb pressure of about 4 kg. A dolorimeter with a 9 kg scale with 1.0 kg marks and a pressure area of 1.54 cm² (Chatillon, New York, NY, USA) measured pain threshold. The assessor applied vertical force with the instrument at a rate of about 1 kg/s and instructed the subject as follows: "Tell me when this becomes painful." As described by McCain, dolorimetry was performed at 5 tender point pairs. The sum of these individual scores in kg/cm² was named "the total myalgic score." The control points for the dolorimetry were the same as those used by the ACR Multicenter Criteria Committee¹.

Physical fitness. An electronically braked bicycle ergometer measured peak workload (Jaeger, ER 800, Breda, The Netherlands). The saddle height was adjusted to the subject's height. The internal power delivered by the bicycle remained constant over a pedaling range of 45 to 75 rpm. The protocol started at 50 watts (W) for 5 min as warmup. The resistance level was increased with 10 W every minute until subjective maximum workload (Wmax). Heart rate was monitored continuously by a pulse watch recorder (Sports tester PE 3000, Finland) and registered at a steady state level at 50 W and at Wmax. The fitness test was interrupted if the heart rate exceeded 200 beats/min or the patient developed chest pain. Perceived exertion was scored after 5 min and at the end of the test by applying the Borg scale (range 6–20; 6 = extremely easy, 20 extremely heavy; the Borg score should be equal to the heart rate divided by 10).

Health status. The Arthritis Impact Measurement Scales (Dutch-AIMS)⁸ and the Sickness Impact Profile (SIP)⁹ are reliable generic assessments of self-reported health status. The AIMS comprises 53 items, which can be categorized in 11 subscales covering physical, psychological, and social well being. As the internal consistency of the scales was good (Cronbach's alpha 0.66–0.89 for patients with rheumatoid arthritis^{8,10}) and we speculated this would count for FM patients as well, we computed a total AIMS score [weighted mean score of all 11 subscales (range 0–10), in which 0 is no problem/impact at all] for the purpose of this article. The SIP comprises

136 statements on health related dysfunction; they probe actual performance, not potential ability or capacity. Although 12 category scores can be computed, only the total score and the physical and psychological subscore are reported here (all ranges from 0 to 100).

Psychological distress. The Symptom Checklist-90-Revised (SCL-90)¹⁰ indicates the presence or absence of psychological symptoms. This self-report inventory evaluates 9 primary symptom clusters (phobic anxiety, anxiety, depression, somatization, obsession/convulsion, interpersonal sensitivity, hostility, sleep, and psychoticism) that can be added up to a global severity index of psychological distress. Adequate test-retest reliability and internal consistency have been documented¹⁰.

Patient global assessment. Subjects indicated the global assessment of their general sense of well being on a marked 5 point scale (1 = very bad, 5 = very good).

General fatigue. Extent of fatigue during the last week was marked on a 100 mm VAS (0 = no fatigue, 100 = worst imaginable).

Rating of the interventions. Participants marked their active intervention at the end of it on a 10 point scale (1 = worthless, 10 = optimal).

The amount of additional therapy. All subjects were asked to keep a weekly diary in which they noted the number of physician consultations, physiotherapy sessions, pharmacotherapy use, and activities like swimming and gymnastics.

Measurements. Divided over 2 days within one week the following measurements were completed: self-reported questionnaires (requiring up to 2 h to complete), number of tender points and dolorimetry (10 min), and fitness tests (30 min).

Reliability assessment. In a random sample of 30 subjects test-retest reliability was assessed 5 days after baseline for the following instruments: number of tender points, total myalgic score, pain (VAS), patient global assessment, fitness (Wmax), and functional ability (AIMS and SIP).

Statistics. Sample size calculation was based on pain as primary outcome. Given an expected mean baseline pain score of 60 with the VAS, we set the minimum relevant decrease at 30%, or 20 mm VAS less than baseline. Standard deviation values from a relevant population were used⁵ and error rates set at alpha (2 tailed) = 0.05 and beta = 0.10. A total of 25 patients per group was calculated as necessary to detect the minimum clinically important difference of 20 mm (VAS for pain). To determine whether adding an educational program to 50% of the intervention participants would improve compliance, we doubled the sample sizes for the active interventions to 50 participants.

Analysis. An intention-to-treat and a completers analysis were performed. It was previously decided that only patients who had participated in the allocated program at least once should be analyzed by intention-to-treat. Dropouts during the study were not included in the completers analysis. The number of missing data can be found in Table 2. In addition, an explorative analysis was done for those patients who participated in at least 34 (67%) of the supervised fitness sessions or in at least 11 (67%) of 16 biofeedback sessions. To determine whether the educational program improved attendance rates, subjects should have been present in at least 3 of 6 health promotion sessions.

Data were summarized by means, standard deviations, and 95% confidence intervals of change scores, or by medians and ranges when appropriate. At 24 weeks the groups were compared for improvement by analysis of variance. Reliability was assessed by intraclass correlation, but for patient global assessment the kappa statistic was used.

Ethics. The study was approved by the ethical committee of the University Hospital of Maastricht and the Atrium Medical Center Heerlen.

RESULTS

Demographic data. At baseline no important differences were present among the groups regarding age, years of education, or socioeconomic status (Table 1). The control

Table 1. Demographic data of the 129 participating patients with fibromyalgia.

	Fitness, n = 50	Group Biofeedback, n = 50	Control, n = 29
Age, yrs	46.2 (26–59)	44.4 (26–60)	42.8 (26–59)
Mean (range)			
Duration of complaints, yrs			
Mean (range)	9.7 (1–37)	10.1 (1–38)	15.4 (3–40)
Marital status			
Married, %	84	86	83
Divorced, %	4	10	7
Education level			
< 6, %	8	27	10
7–12, %	78	57	66
> 12, %	14	16	24
Net family income			
Low, %	58	62	70
Middle, %	42	23	30
High, %	0	15	0

group had a longer disease duration and a slightly better physical condition at baseline (Table 2). The regular use of medication and physiotherapy preceding the study was comparable.

Reliability. Intraclass correlation coefficients were moderately high (about 0.70) for pain (VAS and total myalgic score), for fitness (Wmax), and for health status (AIMS and SIP). Intraclass correlation coefficient for the number of tender points was poor at 0.51 and the unweighted kappa statistic for patient global assessment was marginally satisfactory at 0.66.

Dropouts. Figure 1 is the flowchart of this study. Altogether 143 patients with FM were randomized (fitness 58, biofeed-

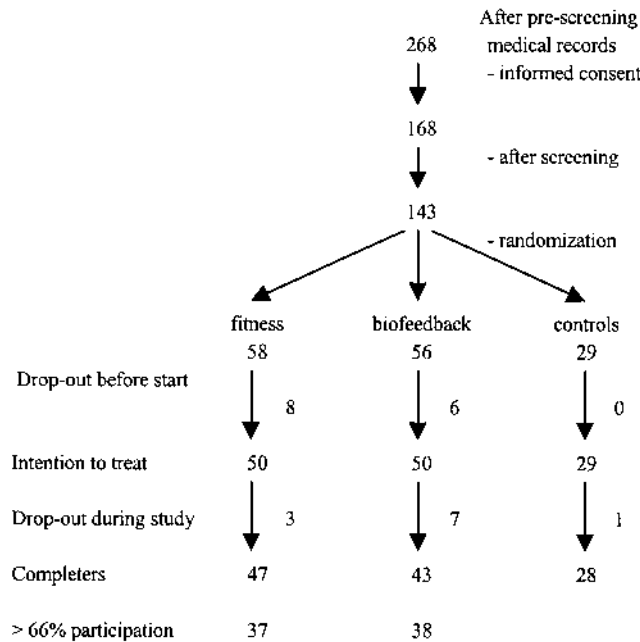


Figure 1. Flow chart of patient numbers.

Table 2. Assessments at baseline and change scores at 24 weeks for the 3 study groups. Completers analysis. Values in italics indicate deterioration.

	Fitness, n = 44		Biofeedback, n = 38		Control, n = 27		p*
	Baseline	Change at 24 Weeks	Baseline	Change at 24 Weeks	Baseline	Change at 24 Weeks	
Pain							
VAS	66.8 (15.3)	-5.5 (-10.9 to -0.1)	59.1 (18.5)	-0.6 (-6.5 to 5.3)	62.4 (20.5)	<i>1.3 (-4.5 to 7.1)</i>	0.3
TP	9.9 (3.6)	-0.6 (-2.0 to 0.8)	9.8 (3.6)	-1.4 (-2.4 to 0.4)	9.7 (2.9)	-1.9 (-0.5 to -3.3)	0.4
TMS	146.4 (53.8)	12.8 (-1.0 to 26.6)	157.1 (52.3)	15.5 (0.0 to 31.0)	156.6 (54.8)	25.3 (-5.5 to 45.1)	0.6
Physical fitness							
Wmax	122.3 (28.2)	-9.9 (-16.6 to -3.2)	131.2 (37.9)	<i>-13.0 (-7.2 to -18.8)</i>	136.3 (30.5)	<i>-27.1 (-34.8 to -19.4)</i>	0.01
Borgmax	17.8 (1.2)	-0.5 (-0.9 to -0.1)	17.4 (1.5)	-0.3 (-0.8 to 2.0)	17.6 (1.5)	<i>0.1 (-0.7 to 0.9)</i>	0.7
Patient global	2.8 (0.7)	0.5 (0.2 to 0.8)	2.9 (0.8)	0.3 (0.6 to 0.0)	3.0 (0.8)	0.5 (0.8 to 0.2)	0.5
General fatigue	69.1 (12.3)	-5.1 (-9.8 to 0.6)	60.1 (19.4)	-3.1 (1.4 to 7.6)	61.9 (20.1)	-1.9 (-7.6 to 3.8)	0.3
Psychological distress	182.4 (48.0)	-6.8 (-20.1 to 6.5)	176.5 (40.5)	-9.4 (-22.9 to 4.1)	183.9 (51.3)	-8.1 (-19.8 to 3.6)	0.9
Health status							
SIP total	14.4 (7.8)	-1.9 (-3.9 to 0.1)	14.0 (9.4)	-2.3 (-4.3 to -0.3)	11.4 (9.4)	-1.4 (-3.4 to 0.6)	0.8
SIP physical	11.3 (7.7)	-1.7 (-3.7 to 0.3)	11.4 (11.2)	-1.6 (-3.4 to 0.2)	9.8 (9.3)	-0.6 (-2.9 to 1.7)	0.7
SIP psychological	16.3 (11.8)	-3.2 (-6.2 to 0.2)	15.8 (11.8)	-3.7 (-4.9 to -2.5)	18.1 (13.9)	-3.5 (-7.0 to 0.0)	0.09
AIMS	1.9 (2.1)	0.1 (-0.6 to 0.8)	3.1 (2.1)	0.4 (-0.1 to 0.9)	5.4 (2.0)	0.8 (-1.8 to -0.2)	0.2

Baseline values: mean (SD); change scores: mean (95% confidence interval). TP: tenderpoints (0–18); TMS: total myalgic score (kg/cm²); Wmax: maximal Watt bicycle ergometer; Borgscale: rating of perceived exertion (6–20); global assessment (0–5); SIP: Sickness Impact Profile; AIMS: Arthritis Impact Measurement Scales. * Between-group differences (ANOVA).

back 56, controls 29), but 14 (10%) of them withdrew within a few days after randomization, before the interventions started: 8 had been randomized to fitness training and 6 to biofeedback training. The reasons stated were: too much stress (n = 8), severe disease of husband (n = 2), hospitalization (n = 1), full time job (n = 2), and not satisfied with allocated fitness intervention (n = 1). The subjects who withdrew immediately after randomization, before the interventions started, were comparable in demographics with the others (results not shown). Altogether 57 patients started the sessions aimed at promoting compliance with the intervention. During the study 11 patients dropped out: 3 from fitness training, 7 from biofeedback (of which 2 and 4, respectively, followed the educational program), and one from the control group. The reasons were: stress at home (n = 6), death of father (n = 1), no benefit from biofeedback (n = 1), and stress due to biofeedback (n = 2).

In total 118 patients (82% of those randomized; 92% of actual participants) completed the study: 47 had fitness training, 43 biofeedback training, and 28 were controls. A total of 37 (79%) from the 47 patients who completed the fitness program attended more than 67% of the supervised training sessions, whereas 38 (88%) attended more than 67% of the biofeedback sessions. Nobody in the fitness group made use of the opportunity to train a third time per week individually. The additional educational program was followed by 53 patients (87%) in more than 50% of the sessions.

Outcomes. The intention-to-treat analysis and the completers analysis showed similar results. Because we were primarily interested in whether fitness and/or biofeedback training improved outcomes compared to controls, we

present the results of only the 118 completers who had actually followed therapy (Table 3). As shown in Table 2 no intervention led to statistically significant or clinically relevant improvement of pain or any other outcome measure. Indeed, physical fitness worsened significantly during the trial in all groups, although the decrease was significantly less in the fitness group compared to the control group (-9.9 W vs -27 W; $p < 0.01$).

Monitoring heart rates during the exercise program revealed that no single participant showed rates exceeding 130 beats/min for longer than a few minutes. In terms of training intensity and maximal heart rates achieved, the fitness intervention was actually a low impact training program, despite efforts to encourage patients to follow a high impact version.

Improvements in global assessment were neither statistically nor clinically significant among the groups. The scores on the AIMS-total and SIP-physical, -psychological, and -total of our FM population were comparable to Steinbrocker functional class 2 to 3¹¹ in patients with rheumatoid arthritis and worse than in patients with

Table 3. Numbers of patients included in the analysis without missing data at baseline or followup.

	Feedback	Group Biofeedback	Control
Clinical variables*	44	38	27
SCL-90R**	45	40	28
SIP/AIMS†	46	42	28

Missing data: * n = 9 (baseline 6; followup 3). ** n = 5 (baseline 3; followup 2). † n = 2 (baseline 0; followup 2).

spondylitis¹², and did not change irrespective of the intervention. Data about sensitivity to change of these outcomes in FM patients are not available. Most measures showed high variability between and within the patients as indicated by change score SD that were similar to or greater than SD at baseline.

Participants of biofeedback training, aimed at relaxation, showed no improvement on the psychological outcomes in general or the more specific anxiety subscales of the SCL-90-R and the AIMS (data not shown). The control group used less or equal amounts of NSAID (44%), analgesics (33%), antidepressive drugs (13%), tranquilizers (21%), and physiotherapy (35%) compared to the active intervention groups. Educational sessions aimed at promoting compliance failed to do so: more than 2/3 of fitness sessions were attended by 62% who did not have and 67% who did have special compliance promotion sessions. Unexpectedly, for the biofeedback population this was 80% versus 57%. Patients with a high participation rate (> 67%) in the fitness or biofeedback sessions had similar outcomes (data not shown) compared to those with a low participation rate.

Based on a scale in which 10 is the best rating possible, fitness patients rated their intervention 7.9 (SD 1.0), biofeedback patients 6.3 (SD 1.9), and the educational program 7.8 (SD 1.0). Almost all fitness patients continued their program once weekly under supervision of the same study instructor, but at their own expense. Less than half the biofeedback patients stated they continued their relaxation practices at home individually.

DISCUSSION

The main conclusion of our study is that patients with FM did not improve in pain or in any other outcome measure after fitness training or biofeedback training, compared to controls. Our negative results are in contrast with those of a clinical trial described by McCain, *et al*⁵, in which patients were randomized to high impact cardiovascular fitness training or to control treatment. Physical fitness improved with more than 25% in this study in the intervention group and there was a significant improvement in pain and global assessments. Patients included in the McCain study, however, had been preselected on the basis of a successful treadmill test and had a better compliance during the intervention; they also achieved higher intensity compared to our fitness program. Differences in study design and patient motivation or capability to reach a cardiovascular training effect may therefore be responsible for the opposite results of both studies. To explore this issue, we plan to compare high versus low impact fitness in patients with FM.

Ferracioli, *et al*⁶ studied the efficacy of biofeedback training in patients with FM. They found important improvements in pain, fatigue, morning stiffness, and other measures. The compliance in that study was very low (only 6 patients had attended two-thirds of the sessions or more)

so that it is unclear whether perceived improvements can be attributed to a specific treatment effect. Evidence to endorse the muscle tension hypothesis and the positive role of relaxation treatment supported by electromyographic biofeedback can be derived from studies in chronic back pain¹³. Our study was the first that followed Ferracioli's protocol. Although we had only 14% dropouts during the supervised biofeedback training period of 2 months, almost all participants stated that they could not sustain the effort of twice daily relaxation practices at home for the rest of the 6 months, as was advised. A total of 38 (88%) patients were able to continue unsupervised relaxation practices twice weekly. We could not measure any beneficial effect. Our patient population scored rather high on all subscores of the SCL-90-R at baseline, which may be an indication that the treatment protocol (biofeedback) is too restricted and/or not intensive enough to induce improvement.

Buckelew, *et al*¹⁴ conducted a clinical trial in which they performed biofeedback/relaxation training, exercise training, a combination of both, or an educational program (as a control). They have shown modest within-group improvements concerning pain, pain behavior, and self-efficacy in all 3 intervention groups, but between-group differences were small and not statistically significant. The improvement in physical activities was based on changes in the AIMS-physical only: the global severity index of the SCL-90-R remained unchanged. An important difference with our study was that patients were treated on an individual basis once weekly during 6 weeks and continued in groups during a longer period of time (2 yrs). These differences may partly explain why patients in our study performed worse. Recently, Mannerkorpi, *et al*¹⁵ investigated the effects of a 6 month once weekly pool exercise program, with intensity matched to the individual's threshold for pain and fatigue, combined with education focused on strategies of coping with symptoms. They found significant improvements on the Fibromyalgia Impact Questionnaire and the 6 minute walk test compared with the control group.

The results of all other randomized controlled trials¹⁶⁻²² evaluating the effect of moderately intense exercise therapy cannot be compared completely because of high dropout rates^{17,18} and different measurement techniques. In all trials at least some increased feeling of well being was found, but the results concerning pain, fatigue, and fitness were inconsistent.

Instead of improvement, our patients showed statistically significant deterioration in physical fitness (measured by Wmax) and no effect on pain, psychological distress, health status, or general fatigue with the applied measures. This observation is in striking contrast with the high therapy ratings given by most patients and with the broad willingness to continue therapy at their own expense for at least a few years after the study stopped. The best explanation for this discrepancy is that we failed to measure the obvious

positive effect of fitness training with fellow sufferers. Unlike the biofeedback participants who were treated individually, those who followed group fitness training indicated the importance of sharing experiences of symptoms and problems with each other, receiving confirmation and finding some self-confidence. In contrast with the Mannerkorpi, *et al* study in which the sense of pleasure was ascribed to a mode of physical training that did not increase pain, our patients mentioned substantial post-exercise pain, fatigue, and stiffness. This delayed onset generalized muscle soreness occurred one to 3 days after the fitness session (in which eccentric exercises were avoided) and was worst during the first 3 months. Nevertheless, only 6% dropped out in our fitness arm compared to 24% in the Mannerkorpi study. We cannot exclude that the stimulating instructor played a role in treatment adherence: the groups continued for a considerable time even after the study ended. In summary, it seems that patients with FM use group fitness training for contacting fellow patients and sharing experiences rather than improve physical fitness. In their opinion physical fitness training is a legitimate means to show society they are positive about their own health; willingness to pay for this intervention may therefore be a better indicator of improvement than increased physical function or decreased pain.

Another explanation for not measuring improvement is that measurements evaluating change in FM are purely subjective. Our observation that patients with FM experienced and reported far more disability than assessors could measure when observing them performing tasks²³ is an example of a disturbed perception with respect to their own physical condition. Disturbed perception, however, does not adequately explain deterioration of physical fitness, as in this study. It is not conceivable that physical function really worsened within 6 months. A better explanation is that patients performed worse during the study because they wanted to avoid the direct postexercise pain, which may have been due to an insufficient warmup period before the ergometer test, and to delayed onset generalized muscle soreness²⁴⁻²⁶. This theory was supported by the observation that maximum workload measured one week after first measurement was significantly lower (Wmax 134 and heart rate 163 vs Wmax 124 and heart rate 156; $p = 0.005$), while patients reported similar levels of exertion.

A last explanation for not measuring improvements despite patients' willingness to continue treatment relates to the high variability of scores, both between patients and within patients. The high variability of measurement results seriously interferes with sensitivity to detect change. Unambiguous data on sensitivity to change in outcome measures such as AIMS, SIP, and SCL-90 in FM trials are not available. It is obvious from our results that the outcome measurements used in this study are, at least in patients with FM, not appropriate to evaluate improvement.

In conclusion, 6 month low impact fitness training and biofeedback did not improve pain, patient global assessment of well being, physical fitness, functional ability, or psychological distress in patients with FM. The high variability of measures between and within subjects was remarkable. There was no correlation between efficacy and participation rate of fitness or biofeedback sessions, which in itself was not influenced by participation in meetings promoting compliance. Usual care provided the same results as fitness or biofeedback. Nevertheless patients did continue low impact fitness training at their own expense, suggesting some form of unmeasured benefit. Again, our results seem to reflect the heterogeneity of patients with fibromyalgia and the disputable suitability of measures evaluating change in this condition.

ACKNOWLEDGMENT

We express our gratitude to all participating patients and to Rob de Bie, Rinie Geenen, Xandra Gielen, Jos Kleynen, Erik de Klerk, Nico Groenman, Ton Lenssen, Sander van Sloun, Carlo Theunissen, Vicky Verstappen, and Johan Vlaeyen for their contributions at various stages of the study.

REFERENCES

1. Wolfe F, Smythe HA, Yunus MB, et al. The American College of Rheumatology 1990 criteria for the classification of fibromyalgia: report of the multicenter criteria committee. *Arthritis Rheum* 1990;33:160-72.
2. Wolfe F, Ross K, Anderson J, Russell IJ, Hebert L. The prevalence and characteristics of fibromyalgia in the general population. *Arthritis Rheum* 1995;38:19-28.
3. Blecourt de ACE, Knipping AA. Treatment of fibromyalgia. In: *Fibromyalgia: towards an integration of somatic and psychological aspects (thesis)*. Groningen: Netherlands; Rijksuniversiteit Groningen; 1995:237-67.
4. Carette S, Bell MJ, Reynolds J, et al. Comparison of amitriptyline, cyclobenzaprine, and placebo in the treatment of fibromyalgia. *Arthritis Rheum* 1994;8:245-50.
5. McCain GA, Bell DA, Mai FM, Halliday PD. A controlled study of the effects of a supervised cardiovascular fitness training program on the manifestations of the primary fibromyalgia syndrome. *Arthritis Rheum* 1988;31:1135-41.
6. Ferracioli GF, Ghirelli L, Scita F, et al. EMG-biofeedback training in fibromyalgia syndrome. *J Rheumatol* 1987;14:820-5.
7. American College of Sports Medicine. Resource manual for guidelines for graded exercise testing and prescription. Philadelphia: Lea and Febiger; 1988.
8. Taal E, Jacobs JW, Seydel ER, Wiegman O, Rasker JJ. Evaluation of the Dutch Arthritis Impact Measurement Scales (Dutch-AIMS) in patients with rheumatoid arthritis. *Br J Rheumatol* 1989;28:487-91.
9. Bergner M, Babbitt RA, Pollard WE. The Sickness Impact Profile: validation of a health status measure. *Med Care* 1976;14:57-67.
10. Derogatis LR. SCL-90-R administration, scoring and procedures manual. 2nd ed. Towson, MD: Clinical Psychometric Research; 1977.
11. Riemsma RP, Taal E, Rasker JJ, Houtman PM, van Paassen HC, Wiegman O. Evaluation of a Dutch version of the AIMS2 for patients with rheumatoid arthritis. *Br J Rheumatol* 1996;35:755-60.
12. Hidding A, de Witte L, van der Linden S. Determinants of self-reported health status in ankylosing spondylitis. *J Rheumatol* 1994;21:275-9.

13. Vlaeyen JW, Haazen IW, Schuerman JA, Kole-Snijders AM, van Eek H. Behavioural rehabilitation of chronic low back pain: comparison of an operant treatment, an operant-cognitive treatment and an operant respondent treatment. *Br J Clin Psychol* 1995;34:95-118.
14. Buckelew SP, Conway R, Parker J, et al. Biofeedback/relaxation training and exercise interventions for fibromyalgia: a prospective trial. *Arthritis Care Res* 1998;11:196-209.
15. Mannerkorpi K, Nyberg B, Ahlmen M, Ekdahl C. Pool exercise combined with an education program for patients with fibromyalgia syndrome. A prospective, randomized study. *J Rheumatol* 2000;27:2473-81.
16. Burckhardt CS, Mannerkorpi K, Hedenberg L, Bjelle A. A randomized, controlled clinical trial of education and physical training for women with fibromyalgia. *J Rheumatol* 1994;21:714-20.
17. Martin L, Nutting A, MacIntosh BR, et al. An exercise program in the treatment of fibromyalgia. *J Rheumatol* 1996;23:1050-3.
18. Mengshoel AM, Komnaes HB, Forre O. The effects of 20 weeks of physical fitness training in female patients with fibromyalgia. *Clin Exp Rheumatol* 1992;10:345-9.
19. Nichols DS, Glenn TM. Effects of aerobic exercise on pain perception, affect, and level of disability in individuals with fibromyalgia. *Phys Ther* 1994;74:327-32.
20. Wigers SH, Stiles TC, Vogel PA. Effects of aerobic exercise versus stress management treatment in fibromyalgia: a 4.5 year prospective study. *Scand J Rheumatol* 1996;25:77-86.
21. Gowans SE, deHueck A, Voss S, Richardson M. A randomized, controlled trial of exercise and education for individuals with fibromyalgia. *Arthritis Care Res* 1999;12:120-9.
22. Meyer BB, Lemley KL. Utilizing exercise to affect the symptomatology of fibromyalgia: a pilot study. *Med Sci Sports Exerc* 2000;32:1691-7.
23. Hidding A, van Santen M, de Klerk E, et al. Comparison between self-report measures and clinical observations of functional disability in ankylosing spondylitis, rheumatoid arthritis and fibromyalgia. *J Rheumatol* 1994;21:818-23.
24. Verstappen FTJ, van Santen-Hoeufft HMS, van Sloun S, et al. Fitness characteristics of female patients with fibromyalgia. *J Musculoskel Pain* 1995;3:45-59.
25. Bennet RM, Clark SR, Goldberg L, et al. Aerobic fitness in patients with fibrositis. A controlled study of respiratory gas exchange and 133 xenon clearance from exercising muscle. *Arthritis Rheum* 1989;32:454-60.
26. Mengshoel AM, Vollestad NK, Forre O. Pain and fatigue induced by exercise in fibromyalgia patients and sedentary healthy subjects. *Clin Exp Rheumatol* 1995;13:477-82.