Pain 91 (2001) 65–78

PAIN

www.elsevier.nl/locate/pain

A randomized controlled component analysis of a behavioral medicine rehabilitation program for chronic spinal pain: are the effects dependent on gender?

Irene B. Jensen[a], Gunnar Bergström[a,*], Therese Ljungquist[a], Lennart Bodin[b], Åke L. Nygren[a]

[a]*Karolinska Institute, Section for Personal Injury Prevention, Box 127 18, 112 94 Stockholm, Sweden*
[b]*Örebro University, Department of Statistics, 701 82 Örebro, Sweden*

Received 25 April 2000; received in revised form 8 August 2000; accepted 29 August 2000

Abstract

The aim of the present study was to evaluate the outcome of a behavioral medicine (BM) rehabilitation program and the outcome of its two main components, compared to a 'treatment-as-usual' control group (CG). The study employed a 4×4 repeated-measures design with four groups and four assessment periods (pre-treatment, post-treatment, 6-month follow-up, and 18-month follow-up). The group studied consisted of subjects on sick leave identified in a nationwide health insurance scheme in Sweden. After inclusion, the subjects were randomized to one of four conditions, which were: (1) behavior-oriented physical therapy (PT); (2) cognitive behavioral therapy (CBT); (3) BM rehabilitation consisting of PT + CBT (BM); (4) a 'treatment-as-usual' CG. The treatments were given over a period of 4 weeks, PT and CBT on a part-time basis and BM on a full-time basis. Outcome variables were sick leave, early retirement, and health-related quality of life (measured using the Short Form Health Survey, SF-36). The results showed that the risk of being granted full-time early retirement was significantly lower for females in PT and CBT compared to the CG during the 18-month follow-up period. However, the total absence from work (sick listing plus early retirement) in days over the 18-month follow-up period was not significantly different in the CG compared to the treatments. On the SF-36, women in CBT and BM reported a significantly better health-related quality of life than women in the CG at the 18-month follow-up. No significant differences for men were found on the SF-36 scales. In conclusion, the results revealed gender differences in the outcome of the treatments and that the components of this BM program yielded as good results as the whole program.

Keywords: Randomized controlled trial; Physical therapy; Cognitive behavior therapy; Multidisciplinary rehabilitation; Gender; Vocational rehabilitation

1. Introduction

Chronic pain is a common health problem in economically developed countries and a major cause of work loss and human suffering (Nachemson, 1992; Linton, 1998). In Sweden, the fall in production due to neck and back pain has been approximated to 30 billion Swedish kronor per year (Hansson and Hansson, 1999). Approximately 20% of Swedish patients sick-listed 2 months or longer have been diagnosed with different kinds of spinal pain (Hansson and Hansson, 1999).

A wide array of treatments are available to patients suffering from long-term, non-specific spinal pain, although the scientific evidence for their effectiveness is often weak (Nachemson, 1991; van Tulder et al., 1997). Considering that non-specific spinal pain is a multidimensional phenomenon (Waddell, 1992), one approach has been to combine psychological and physiotherapeutic and/or occupational therapy interventions in so-called multidisciplinary rehabilitation programs. In a review by Flor et al. (1992), it was concluded that chronic back pain patients treated at multidisciplinary centers showed better function and were twice as likely to return to work than controls. In another review by Cutler et al. (1994), it was found that the proportion of patients working increased from 20 to 54% after treatment at non-surgical pain centers.

The evaluations of these multidisciplinary treatment packages usually concern the effect of the intervention as a whole (Linton and Bradley, 1992; Jensen and Bodin, 1998; Johansson et al., 1998), which prevents conclusions regarding the impact of different components of the rehabilitation. A few component studies have been published, however. In

* Corresponding author. Tel.: +46-8-692-2257; fax: +46-8-653-9413.
E-mail address: gunnar.bergstrom@knv.ki.se (G. Bergström).

0304-3959/01/$20.00
PII: S0304-3959(00)00420-6

a study by Heinrich et al. (1985), the outcome of behavior therapy was compared with that of physical therapy and the results showed that patients improved from both interventions and that there were few differences in effect between the treatments. In a later study, Turner et al. (1990) studied the effects of behavioral therapy, physical therapy and a combination of behavioral and physical therapy in comparison with a waiting-list control group. Patients in the combined treatment showed the best short-term improvement but, at 6- and 12-month follow-ups, no significant differences were found between the interventions. Further studies are needed, however, to shed light on the effects of the components of multidisciplinary programs.

An additional issue is the possible impact of gender in the treatment of spinal pain. In a randomized controlled study by Lindström et al. (1992) on subacute low back pain, the results indicated that a graded activity program lowered sick listing for males but not for females. Furthermore, our own research has indicated that the positive effects of a behavioral medicine program for chronic spinal pain seem to be limited to females (Jensen et al., 1994).

The purpose of the present study was to evaluate the long-term outcome of a behavioral medicine rehabilitation program, and the outcome of its two main components, as compared to a 'treatment-as-usual' CG. It was hypothesized that the treatment conditions should be superior to

Table 1
Descriptive information about the conditions

	Behavioral medicine ($N = 63$)	Physical therapy ($N = 54$)	Cognitive behavioral therapy ($N = 49$)	Control group ($N = 48$)	Total ($N = 214$)
Age: mean (SD)	42.5 (11.8)	43.3 (9.4)	43.8 (9.6)	43.9 (10.8)	43.3 (10.4)
Gender: females, *n* (%)	30 (48)	37 (68)	22 (45)	28 (58)	117 (55)
Married: *n* (%)	39 (63)	34 (63)	36 (74)	33 (70)	142 (67)
Swedish origin: *n* (%)	52 (82)	48 (89)	35 (71)	39 (81)	174 (81)
Employed: *n* (%)	52 (84)	42 (78)	42 (86)	44 (94)	180 (85)
Education: n (%)					
Compulsory school	35 (56)	31 (57)	29 (60)	27 (57)	122 (58)
High school	23 (37)	19 (35)	11 (23)	16 (34)	69 (33)
Post-high school	4 (6)	4 (7)	8 (17)	4 (8)	20 (10)
Pending litigation, *n* (%)	17 (31)	11 (20)	21 (43)	16 (36)	65 (32)
Clinically significant depression, *n* (%)[a]	6 (10)	10 (18)	6 (12)	4 (8)	26 (12)
Clinically significant anxiety, *n* (%)[a]	14 (22)	12 (22)	11 (22)	12 (25)	49 (23)
Hazardous alcohol consumption, *n* (%)[b]					
Women	0 (0)	1 (3)	3 (14)	2 (7)	6 (5)
Men	7 (21)	3 (18)	6 (22)	5 (26)	21 (22)
Pain duration (months)					
Mean (SD)	35.6 (68.0)	36.9 (67.1)	22.7 (38.4)	27.3 (39.2)	31.3 (56.9)
Median	8	8.5	9	8	8.5
Total sick leave 6 months prior to inclusion, mean (SD)	127 (40)	120 (50)	129 (44)	119 (43)	124 (44)
Primary pain sites, *n* (%)					
Cervical/thoracic	23 (36)	22 (41)	20 (41)	24 (50)	89 (42)
Lumbar	34 (54)	22 (41)	22 (45)	21 (44)	99 (46)
Mixed pain areas	6 (10)	10 (18)	7 (14)	3 (6)	26 (12)
Spinal mobility, mean (SD)[c]	85.8 (26.3)	84.4 (24.7)	90.4 (26.0)	93.2 (22.4)	88.3 (24.9)
Total neck mobility, mean (SD)[d]	297 (69)	281 (53)	304 (62)	288 (65)	292 (63)
Impaired reflexes, *n* (%)					
Upper extremities	1 (2)	2 (4)	2 (5)	1 (2)	6 (3)
Lower extremities	3 (6)	4 (9)	2 (5)	3 (7)	12 (7)
Positive Laségue, *n* (%)	5 (10)	1 (2)	–	–	6 (3)
Crossed Laségue, *n* (%)	1 (2)	–	–	–	1 (1)

[a] The Hospital Anxiety and Depression Inventory was used to assess depression and anxiety (Zigmond and Snaith, 1983). The cut-off scores were set to 11 points.

[b] The Alcohol Use Disorders Identification Test, AUDIT (Saunders et al., 1993) was used as a measure of alcohol consumption. The cut-off score for hazardous consumption was set to 8 points.

[c] Spinal mobility = thoracic + lumbar mobility measured with a kyphometer (Debrunner, 1972).

[d] Total neck mobility measured with a myrin goniometer (American Academy of Orthopaedic Surgeons, 1966).

the control condition and that the full-time BM program should be superior to its main components in the primary endpoints. Subgroup comparisons were planned with regard to gender.

2. Methods

2.1. Design

The study was a randomized, controlled multicenter trial employing a 4×4 repeated-measures design with four groups and four assessment periods (pre-treatment, post-treatment, 6-month follow-up, and 18-month follow-up). Due to the fact that the treatments were given in groups (see below) with a minimum of four individuals per group and that the number of participants available on each screening occasion were not sufficient to fill more than one treatment group, the randomization was also carried out on a group level. This meant that individuals included in the study on the same occasion were randomized to the same treatment. A block randomization procedure was employed to ensure an even distribution of the treatment conditions. However, the sizes of the groups were not expected to be exactly the same because different numbers of participants were included on different screening occasions. The randomization was carried out with the use of opaque envelopes, each envelope containing a slip of paper with one of the conditions written on it. The envelopes were then shuffled and drawn according to the sealed envelope technique. The screening personnel were blinded to the results of the randomization.

2.2. Subjects

The subjects were 214 individuals (97 men, 117 women) suffering from long-term, non-specific spinal pain. Data was gathered between May 1995 and October 1999. The selected rehabilitation clinics were situated in Stockholm, Gothenburg, Helsingborg, and Malmö. The inclusion criteria were as follows: non-specific spinal pain, currently and continuously sick-listed for at least 1 month and a maximum of 6 months before inclusion (because of spinal pain), fluency in Swedish and 18–60 years of age. Exclusion criteria were: serious spinal pathology (e.g. tumors or spinal fractures), exposure to physical trauma within 6 months of examination (e.g. a whiplash associated disorder), objective neurological signs indicating a need for surgery, comorbidities (e.g. alcohol abuse, acute psychosis), ongoing rehabilitation, and verified pregnancy.

The source population were persons covered by the AGS insurance scheme, which is a health insurance plan covering 2.4 million employees in Sweden. (AGS is a sick pay and disability pension insurance scheme which is part of the agreement between the Swedish Employers' Confederation (SAF) and the Swedish Trade Union Confederation (LO)). The recruitment procedure can be divided into four steps. (1) On a monthly basis, all new reported cases with a musculoskeletal diagnosis and residing in the described geographical areas were sent a short questionnaire regarding their symptoms. Two reminders were sent to non-respondents. (2) Individuals meeting the inclusion criteria according to the questionnaire were then interviewed by telephone. (3) Individuals meeting the inclusion criteria at the time of the telephone interview were offered a medical and functional examination by personnel from the research team. (4) At that examination, the final decision regarding admission to the study was made by a licensed physician under the supervision of an orthopedic specialist. In the AGS database, 2104 individuals were identified (step 1), 1330 were telephone-interviewed (step 2), 405 were offered the described examination (step 3), 235 individuals met the inclusion criteria (step 4), and 21 of these declined participation. Consequently, 214 individuals gave their written consent to participate and were randomized to the different conditions (BM, $N = 63$; PT, $N = 54$; cognitive behavioral therapy (CBT), $N = 49$; CG, $N = 48$).

The entire recruitment procedure (steps 1–4) took a maximum of 8 weeks and the reasons for exclusion were as follows: 50% did not meet the criteria regarding sick leave, 14% were excluded due to medical findings, 12% declined participation in the study, 9% underwent other rehabilitation interventions, 8% were not fluent in the Swedish language, and 7% were excluded for other reasons, such as personal or family reasons. Descriptive information for the subjects in each treatment condition and for the controls is presented in Table 1.

All ethical considerations regarding the study were approved by the Committee on Ethics of the Karolinska Institutet, Stockholm.

2.3. Treatment conditions

Common features of the three treatment conditions were as follows. (1) The interventions lasted for 4 weeks and were conducted in groups of 4–8 participants. (2) All treatments included a physician who examined the patients and was available throughout the intervention for consultations regarding the patients medical concerns. (3) All treatments included two didactic sessions on psychological aspects of chronic pain (conducted by the psychologist), two didactic sessions on ergonomics (conducted by the physical therapist), and two sessions on medical aspects of chronic spinal pain (conducted by the physician). (4) All treatments included scheduled times for visits to the workplace, and work managers and rehabilitation officials were invited to participate in the discharge session at which a rehabilitation plan was agreed upon. This plan covered issues concerning work, exercise and social and everyday activities and was developed during the intervention. (5) Six booster sessions (90 min per session) were held over a period of 1 year after the treatment.

2.3.1. Behavior-oriented physical therapy

The PT intervention was carried out on a part-time basis (approximately 20 scheduled hours per week) and was aimed at enhancing the physical functioning and facilitate a lasting behavior change of the individual. Each participant was assigned to an individually tailored training program. The PT intervention involved a pedagogical approach in that the participants were given practical examples of how to perform different activities in their everyday lives which they had reported to be problematic. The program included individual goal setting, gradually increased exercises to improve muscular endurance, aerobic training (e.g. cycling on a test bicycle), water exercise (pool training), relaxation techniques according to Jacobson (1938) and Westin (1985) and body awareness therapy. Practical sessions in ergonomics were also included as well as didactic presentations of anatomy and exercise physiology. Home work assignments for physical activities were given according to the individual's interests and problem areas.

2.3.2. Cognitive behavioral therapy (CBT)

The CBT intervention comprised, on average, 13–14 scheduled hours per week and was aimed at improving the subjects ability to manage their pain and to resume a normal level of activity. If no other literature references are given, the techniques referred to in the following have been described in Philips (1988) and/or Turk et al. (1983). The CBT program included activity planning and goal setting, problem solving, applied relaxation (Lisspers and Hallgren, unpublished data, 1994), cognitive coping techniques (e.g. distracting imagery, external focusing, coping self-statements), activity pacing, the role of vicious circles and how to break them, the role of significant others and assertion training. Individually tailored homework assignments were given at the end of each session pertaining to what had been dealt with during the session. The general structure of the sessions was as follows: discussion of the homework assignments given at the last session, introduction of a new topic or a new step in the treatment, a practical example of the topic, and, finally, new homework assignments.

2.3.3. Full-time behavioral medicine rehabilitation

This condition included both the PT and CBT programs.

2.3.4. 'Treatment-as-usual' control group

The CG was not offered any types of intervention in the research project. Consequently, they were subjected to the normal routines in health care. Our own research indicates that only a minority of individuals with long-term spinal pain are offered more comprehensive rehabilitation programs in Sweden (Jensen et al., 1998).

2.3.5. Therapists' compliance and participants' attendance rates

Licensed physiotherapists, psychologists, and physicians were involved in the rehabilitation programs. They were all experienced in the treatment of long-term non-specific spinal pain and they had to adhere to a specified treatment manual. The therapists involved in the components were also involved in the full-time BM program. After each session, the therapist (physiotherapist or psychologist) had to fill out a detailed checklist on what had been included in the session (or if any scheduled elements had not been carried out). Furthermore, structured telephone interviews were held directly after rehabilitation with a number of randomly chosen participants ($N = 34$) regarding the content of the intervention (the therapists were 'blinded' as to which patients in which groups were to be interviewed). Both the checklists filled out by the therapists and the telephone interviews with the participants indicated that approximately 95% of the scheduled activities were carried out in accordance with the specifications. This was regarded as acceptable therapist compliance.

The therapists recorded each participant's attendance rate at the treatments and the booster-sessions. The attendance rates in percent per treatment were: BM (94%), PT (94%), and CBT (94%). The corresponding figures for the booster sessions were: BM (65%), PT (64%), CBT (65%).

2.4. Measures

Absence from work and health-related quality of life were chosen as primary endpoints. Due to the fact that 85% of the sample was employed, absence from work was defined as sick listing plus early retirement. Information on absence from work was obtained from the National Social Insurance Board (NSIB) for a period of 18 months before inclusion in the study and 18 months after rehabilitation. The NSIB register covers all employees in Sweden. To receive sick leave compensation for longer than 2 weeks, it is necessary to notify the NSIB. As a result, virtually all sick-absenteeism exceeding 14 days is on record.

The Short Form 36 (SF-36) was used to assess health-related quality of life (McHorney et al., 1994). The SF-36 is a 36-item instrument developed for measuring eight physical and mental health constructs: physical functioning, limitations on usual role activities because of physical or emotional problems, bodily pain, general health perceptions, vitality, social functioning, and mental health. In a recent study the responsiveness of a global SF-36 scale has been shown to be satisfactory (Bronfort and Bouter, 1999). The global score was calculated by taking the mean of the eight constructs mentioned above. This global score was also used in this study. A Swedish translation of the SF-36 has shown an acceptable internal consistency (alpha-coefficient ranged from 0.79 to 0.93) and construct validity (Sullivan et al., 1995). Each scale ranges from 0 to 100 with higher scores indicating a better perceived health status (e.g. less pain, better physical functioning etc).

2.4.1. Perceived relevance of rehabilitation

At the end of the first day of rehabilitation, the partici-

pants were asked one question worded as follows: 'How certain are you that this type of treatment is relevant to your pain?' A numerical rating scale of 0–10 with the anchors 'not certain at all' and 'absolutely certain' was used. At the post-treatment assessment the participants were asked the following questions: 'Do you consider this rehabilitation program to have been relevant to your neck/back pain?' and 'If you had a relative with problems similar to your own, to what degree would you recommend this type of rehabilitation?' Numerical rating scales of 0–10 with the anchors 'Not relevant at all' and 'Yes, absolutely relevant' and 'Would not recommend it at all' and 'Yes, would absolutely recommend it' were employed.

2.4.2. Adherence to treatment plan

At the 6-months follow-up the treated participants were asked about their adherence to the treatment plan agreed upon at discharge. Three areas were covered: (1) plans concerning their paid jobs; (2) plans concerning domestic activities; (3) plans concerning lifestyle (e.g. exercise or smoking cessation). The response categories were: 'No, not at all', 'Yes, to a certain degree', 'Yes, to a high degree', 'Yes, completely', and 'Didn't have any plan'.

Several other psychosocial and physical functioning variables were measured at the pre-treatment assessment and at one or more of the follow-ups (Jensen et al., 1999). Since absence from work and health-related quality of life were postulated as primary endpoints during the planning of the study, we will confine the reporting in this article to these variables.

2.5. Procedure

Data for the pre-treatment assessment were gathered during the medical and functional examination described above (1–2 weeks prior to rehabilitation). Subjects were presented with a number of self-report questionnaires to be filled out under the supervision of a psychologist. These data were gathered by personnel from the research team to ensure that the same procedure was used at all four clinics. Directly after treatment and at 6-month and 18-month follow-ups, the questionnaires were mailed to the participants.

2.6. Attrition rates

2.6.1. Intention to treat

The data obtained from the NSIB on sick listing and early retirement were complete for all individuals. Directly after treatment the overall non-response rate for the SF-36 was 8.9 and 13.1%, respectively, at the 6-month follow-up, and 12.1% at the 18-month follow-up. At the 18-month follow-up, the non-response rates per condition were 12.7% (BM), 1.9% (PT), 14.3% (CBT), and 20.8% (CG), and these figures were significantly different between the conditions ($X^2 = 9.0$, d.f. $= 3$, $P = 0.030$).

2.6.2. Per protocol

Twenty-eight participants dropped out of treatment (BM, 14; PT, 6; CBT, 8). These proportions did not differ significantly between treatments ($X^2 = 2.6$, d.f. $= 2$, $P = 0.28$). Consequently, 49 participants completed the BM program, 48 the PT intervention, and 41 the CBT program. The overall non-response rate for the SF-36 was 7.0% directly after treatment, 10.8% at the 6-month follow-up, and 9.7% at the 18-month follow-up. At the 18-month follow-up, the non-response rates per condition were 8.2% (BM), 0% (PT), 9.8% (CBT), and 20.8% (CG), and these figures were significantly different between the conditions ($X^2 = 12.1$, d.f. $= 3$, $P = 0.007$).

An analysis of the non-respondents intention to treat (ITT) at the 18-month follow-up showed no statistically significant differences in age, pre-treatment values on the SF-36 variables, in absence from work the quarter before rehabilitation, or in total absence from work during the follow-up period. The 28 participants who did not complete their treatment consisted of 17 men and 11 women, the mean age was 43.8 years (SD 10.7), the median duration of pain was 8 months, 21% were depressed, and 41% of the males reported a hazardous alcohol consumption.

2.7. Statistical analyses

Analysis of variance, Cox regression, and logistic regression were employed to evaluate treatment effects on absence from work, full-time early retirement, and the SF-36 variables at the 18-month follow-up. Absence from work was analyzed in four ways, which were: (1) total absence from work over the whole follow-up period; (2) the duration of absence from work during the 18-month follow-up period; (3) the risk of having any registered absence from work during the month prior to the 18-month follow-up; (4) the risk of being granted full-time early retirement during the follow-up period. Data on sick leave the quarter before randomization to the treatment conditions were used as a covariate (controlled for) in all the analyses of absence from work.

For the SF-36, comparisons between pre-treatment figures and data from the 18-month follow-up were preferred so as to minimize the percentage of missing cases in the analyses. Initially, a MANCOVA was applied to test for multivariate effects followed by ANCOVAs for each of the SF-36 scales. Furthermore, as a complementary analysis, data from pre-treatment, post-treatment and the 18-month follow-up were also analyzed applying MANCOVA with a repeated-measures design to the global score on the SF-36. The pre-treatment values of the outcome variables were used as covariates in the analyses of health-related quality of life.

Data on adherence to the treatment plan was analyzed by chi-square analysis.

2.7.1. Statistical power

The goal set during the planning of the study was to ensure the detection of a 25% reduction in the primary endpoint (sick leave) with 80% power, starting from a baseline value of approximately 100 days. This figure was derived from a consideration of a 35% decrease in the BM program, 25% decrease in the PT and CBT conditions respectively, and no decrease in the CG. With a common standard deviation of 50 these figures gave an effect size for a one-way ANOVA of 0.067. A power calculation revealed that, in order to detect that effect size with a power of 80% (5% significance level) we needed approximately 40 subjects in each of the four study groups ($n = 160$). In the power calculation, we used information on subjects from earlier studies using the same inclusion criteria as in the present study.

The initial design of the study called for a gender-differentiated analysis. This doubled the size of the study population ($n = 320$) needed to obtain a power of 80%. After the recruitment of subjects had started, we found that it would not be possible to reach our goal of 320 participants within the planned study period. Thus, our aim to conduct a gender-differentiated analysis with a power of 80% could not be achieved in this study.

3. Results

Due to the fact that the results differed between men and women with regard to several of the primary endpoints, and in order to avoid being too lengthy, the analyses presented below are gender-differentiated. Figures from ITT analyses are presented. However, as per protocol (PP) analyses were also done comments are made if the ITT and PP analyses yielded different results in terms of statistical significance.

3.1. Perceived relevance of rehabilitation

As can be seen from Table 2, none of the differences revealed between the programs were statistically significant. However, in general the BM and PT program have a higher face validity than the CBT intervention on scrutinizing the absolute values.

3.2. Adherence to treatment plan

The percentages of females reporting that they had adhered fully or to a high degree to the treatment plan with regard to their paid jobs were 56, 58 and 60%, respectively, for the BM, PT and CBT conditions, and the corresponding figures for males were 65% (BM), 70% (PT) and 57% (CBT). With regard to domestic activities, the figures for females were 45% (BM), 46% (PT) and, 54% (CBT) and, for males, 36% (BM), 71% (PT), and 54% (CBT). The figures for the treatment plans involving lifestyle were 50% (BM), 41% (PT), and 54% (CBT) for females and 25% (BM), 54% (PT), and 83% (CBT) for males. The percentages for the whole group (males and females) reporting that they did not have any plans were 2% for paid jobs, 18% for domestic activities, and 11% for lifestyle factors. The only statistically significant difference between treatment conditions was for lifestyle, where the males in the CBT program had the highest percentage for adherence and the males in the BM program the lowest percentage (chi-square test, $P = 0.016$).

3.3. Absence from work

Figs. 1 and 2 show the course of absence from work (sick listing plus early retirement) per quarter for males and females for the period 18 months before inclusion in the study and 18 months after rehabilitation. Visual examination of the figures reveals that, from quarter 2 onward, the decrease in absence from work for females (Fig. 1) is more pronounced in the treatment groups than in the controls. Among females, the total absence from work expressed in days is 299 in the BM group, 290 in the PT group, 355 in the CBT group and 316 in the CG.

Among males (Fig. 2), from quarter 2 to quarter 5, there is

Table 2
Perceived relevance of rehabilitation

	Behavioral medicine (BM)	Behavior-oriented physical therapy (PT)	Cognitive behavioral therapy (CBT)	P[a]
Before treatment: how certain program is relevant, mean (SD)				
Females	6.0 (1.6)	5.0 (3.0)	3.8 (3.1)	n.s.
Males	4.9 (2.2)	4.7 (2.7)	3.9 (2.7)	n.s.
After treatment: regards program as appropriate, mean (SD)				
Females	6.4 (3.1)	7.1 (3.4)	4.8 (3.7)	n.s.
Males	6.0 (3.6)	6.3 (3.2)	4.5 (3.2)	n.s.
After treatment: would recommend program, mean (SD)				
Females	7.0 (3.7)	7.9 (3.2)	6.0 (3.8)	n.s.
Males	6.1 (4.3)	7.6 (3.8)	5.1 (4.2)	n.s.

[a] P-values from Kruskal–Wallis test.

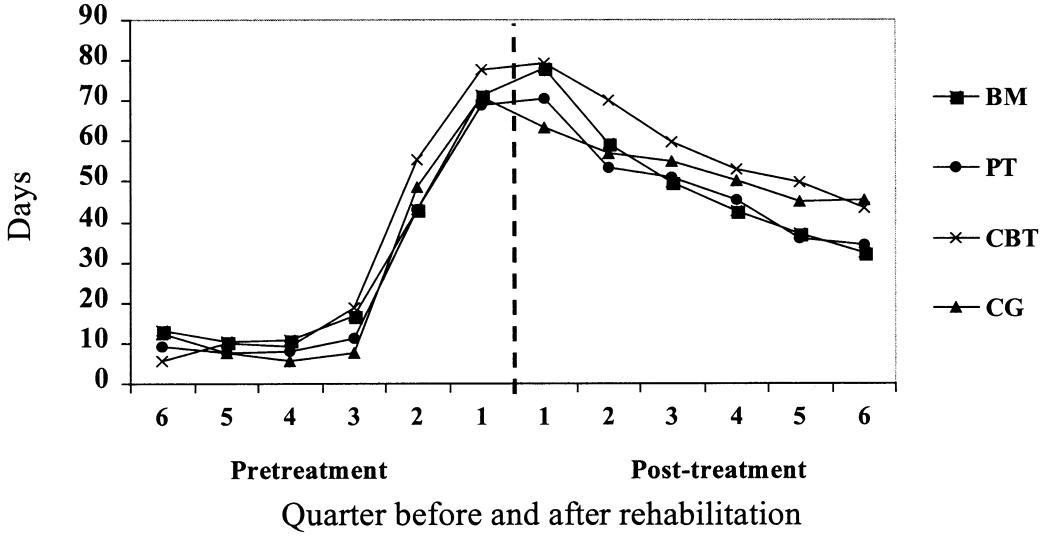

Fig. 1. Absence from work: intention to treat (females).

a gap between the CBT group and the other conditions in that men in the CBT program display a higher rate of absence from work. It should also to be noted that males in the CG group display increased absence from work after quarter 4. Among males, the total absence from work expressed in days is 295 in the BM group, 301 in the PT group, 397 in the CBT group and 309 in the CG.

Total absence from work was evaluated by ANCOVA. The parameter estimates for differences in total absence from work over 18 months are given in Table 3. No significant differences were found, although the estimates for males were relatively high. However, it should be noted that the CBT group showed an increased absence from work in relation to the CG, whereas the other two treatment groups showed a decreased absence from work (non-significant differences).

Cox regression was used to analyze the duration of absence from work during the 18-month follow-up period. The hazard ratios for women with 95% confidence intervals in parenthesis were as follows (figures above 1 indicate a shorter duration and below 1 a longer duration compared to the CG): BM 1.2 (0.7–2.2); PT 1.1 (0.6–1.9); CBT 1.0 (0.5–1.8). The corresponding figures for men were: BM 1.1 (0.6–2.0); PT 1.3 (0.6–2.7); CBT 0.5 (0.3–1.1). As can be seen from the confidence intervals, no differences were statistically significant.

The risk of having any recorded absence from work during the month prior to the 18-month follow-up and the risk of being granted full-time early retirement during the follow-up period were evaluated by logistic regression. The CG constituted the reference category. Descriptive data can be found in Table 4. The odds ratios with 95% confidence

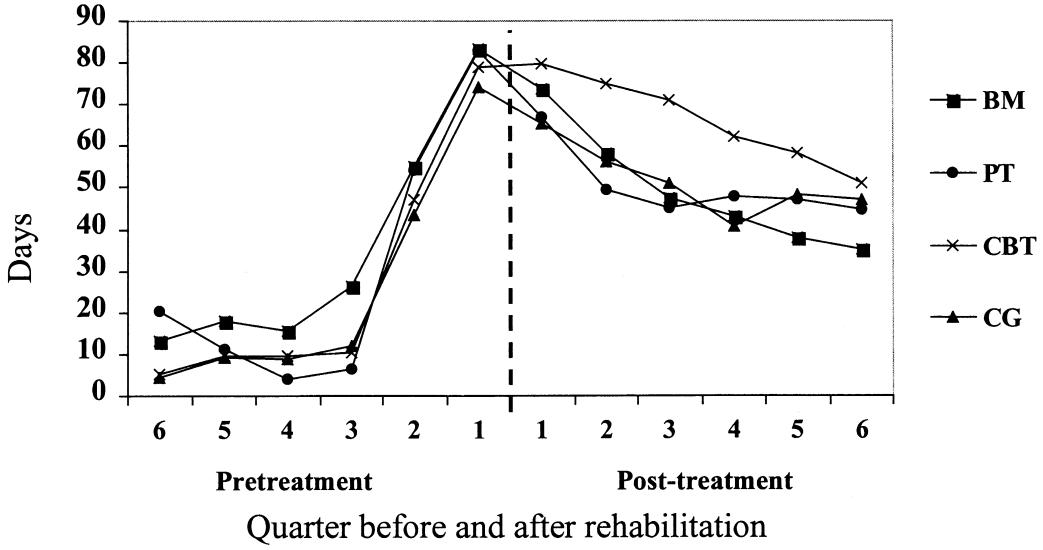

Fig. 2. Absence from work: intention to treat (males).

Table 3
Parameter estimates (B) from covariance analyses for differences between treatment groups versus the control group at the 18-month follow-up[a]

	Work absence (days)	Bodily pain	Physical functioning	General health	Mental health	Vitality	Social functioning
Behavioral medicine							
Women, B (CI 95%)	−18 (−106–70)	11.7 (−0.6–24.1)	7.1 (−3.3–17.5)	13.9 (3.8–23.9)**	15.3 (3.8–26.9)*	11.5 (−0.8–23.8)	18.8 (4.1–33.5)*
Men, B (CI 95%)	−58 (−159–43)	12.1 (−2.3–26.4)	7.3 (−4.5–19.2)	8.2 (−5.0–21.4)	8.6 (−4.0–21.1)	14.9 (0.1–29.7)*	6.4 (−9.3–22.0)
Behavior-oriented physical therapy							
Women, B (CI 95%)	−20 (−104–64)	7.0 (−4.8–18.7)	3.1 (−6.8–13.0)	7.4 (−2.3–17.0)	13.5 (2.4–24.6)*	10.0 (−1.9–21.8)	12.5 (−2.0–27.0)
Men, B (CI 95%)	−50 (−167–67)	7.7 (−8.1–23.6)	2.3 (−10.8–15.3)	8.0 (−6.7–22.6)	9.8 (−4.0–23.5)	14.9 (−1.3–31.2)	−0.5 (−17.8–16.8)
Cognitive behavioural therapy							
Women, B (CI 95%)	20 (−76–116)	11.5 (−1.9–24.8)	13.3 (2.1–24.6)*	13.7 (2.8–24.7)*	20.9 (8.4–33.4)**	15.6 (2.1–29.1)*	29.8 (14.0–45.6)***
Men, B (CI 95%)	65 (−39–169)	7.6 (−7.4–22.6)	2.9 (−9.3–15.1)	5.1 (−8.6–18.8)	5.1 (−7.9–18.1)	9.7 (−5.6–25.0)	−5.6 (−21.9–10.7)

[a] ***$P < 0.001$; **$P < 0.01$; *$P < 0.05$.

Table 4
Data on early retirement and absence from work

	Behavioral medicine (BM)	Behavior-oriented physical therapy (PT)	Cognitive behavioral therapy (CBT)	Control group (CG)
Percentage granted full-time early retirement				
Females	20	8	9	36
Males	18	24	18	25
Percentage with any registered work absence during the month prior to the 18-month follow-up				
Females	50	54	54	54
Males	52	65	59	70

intervals for females were as follows (figures above 1 indicate higher risk and below 1 a lower risk of having any recorded absences from work during the month prior to the 18-month follow-up): BM, 0.9 (0.3–2.4); PT, 1.0 (0.4–2.8); CBT, 1.0 (0.3–3.1); for early retirement (figures above 1 indicates higher risk and below 1 lower risk of being granted early retirement): BM, 0.4 (0.1–1.4)); PT, 0.1 (0.0–0.6); CBT 0.1 (0.0–0.8). The corresponding data for males were (for having any recorded absence from work): BM 0.4 (0.1–1.4); PT 0.7 (0.2–3.0); CBT 0.6 (0.2–2.1); for early retirement: BM 0.4 (0.1–1.9); PT, 0.6 (0.1–2.9); CBT, 0.5 (0.1–2.3). Consequently, the risk of being granted full-time early retirement was significantly lower for females in the PT and CBT groups compared to the CG.

In summary, it was found that females in the PT and CBT groups had a lower risk of being granted early full-time retirement during the follow-up period. The PP analyses of absence from work yielded similar results with regard to statistical significance as the ITT analyses.

3.4. Health-related quality of life

Data on health-related quality of life are presented for all four measurements in Tables 5 and 6. Six of the SF-36 variables at the 18-month follow-up were found to be normally distributed and were, as intended, analyzed by analysis of variance. The MANCOVA which considered these 6 SF-36 variables simultaneously, was statistically significant for the women: Wilks's Lambda = 0.72, $F(18, 255) = 1.7$, $P = 0.036$. The corresponding figures for men were: Wilks's Lambda = 0.80, $F(18, 187) = 0.9$, $P =$ n.s. The results of the ANCOVAs are presented in Table 3. As can be seen from the table females in the CBT group report significant improvement in five of the six SF-36 variables compared to the CG. Three of the six SF-36 variables were reported to have improved for the BM condition and, for the PT condition, the mental health variable improved significantly. Among males, the only significant difference was an improvement in vitality in the BM program compared to the CG, although a non-significant general trend to improvement could be seen in all the three treatment programs (especially in the BM program).

Among females, the MANCOVA (4[group] × 2[time]) with repeated measures on the global score yielded a non-significant multivariate main effect for time (Wilk's lambda = 0.98, $F(1, 97) = 2.08$, $P =$ n.s.), a significant group effect ($F(3, 97) = 3.3$, $P = 0.023$), and a non-significant time × group effect (Wilk's lambda = 0.95, $F(3, 97) = 1.6$, $P =$ n.s.). The figures were significantly higher for the BM program directly after treatment ($P = 0.032$) and for the BM and CBT conditions at the 18-month follow-up ($P = 0.016$ and $P = 0.004$, respectively) than for the CG. The improvement approached significance for the PT program at the 18-month follow-up ($P = 0.057$).

Among males, the analysis produced a non-significant multivariate main effect for time (Wilk's lambda = 0.99, $F(1, 71) = 0.62$, $P =$ n.s.), a non-significant group effect ($F(3, 71) = 1.4$, $P =$ n.s.), and a non-significant time X group effect (Wilk's lambda = 0.97, $F(3, 71) = 0.74$, $P =$ n.s. No significant differences between conditions were revealed at any of the measuring points, but the score for the BM program approached significance ($P = 0.06$) at the 18-month follow-up.

The role physical (RP) and role emotional (RE) scales were not normally distributed. Change scores (pre-treatment minus post-treatment values) were computed for these variables and dichotomized into one category reporting no change or worsened health (not improved health) and one category reporting improved health. Logistic regression was then employed with the variable improved/not improved health as the dependent variable and the treatment condition as the independent variable. The CG served as the reference category. For females, the odds ratios with 95% confidence intervals for RP and RE, respectively, were as follows (values above 1 indicate a higher chance (risk) and those below 1 a lower chance (risk) of reporting improved health): BM, 2.9 (0.8–10.2), 3.4 (0.9–12.6); PT, 3.2 (0.9–10.6), 2.5 (0.7–9.1); CBT, 6.3 (1.6–24.5), 3.0 (0.7–12.2). The corresponding figures for men were: BM, 2.0 (0.6–7.5), 2.8 (0.7–10.9); PT 1.3 (0.3–5.7), 3.9 (0.9–17.3); CBT 1.5 (0.4–5.9), 2.1 (0.5–8.6). As can be seen from these figures, the only significant difference was that females in the CBT group reported improved physical role function compared to the CG.

The results of the PP analyses differed for the BM program in that women also reported significant improve-

Table 5
Females: descriptive information for the SF-36[a]

Measure and group	Pre-treatment	Post-treatment	6-month follow-up	18-month follow-up
Bodily pain				
BM	23.6 (11.8)	33.0 (15.9)	38.2 (22.8)	41.6 (27.0)
PT	23.6 (14.9)	27.7 (12.7)	31.8 (21.9)	36.7 (24.4)
CBT	26.9 (8.2)	30.9 (11.9)	34.8 (16.7)	41.9 (13.2)
CG	24.9 (11.9)	26.5 (15.8)	27.2 (16.2)	30.1 (14.8)
Physical functioning				
BM	49.8 (20.2)	56.7 (20.8)	58.5 (25.2)	57.3 (26.8)
PT	47.2 (20.6)	50.4 (22.8)	49.0 (23.5)	51.6 (25.9)
CBT	51.0 (17.1)	54.3 (14.0)	59.0 (16.3)	63.3 (15.7)
CG	54.5 (19.1)	57.1 (18.8)	55.6 (22.3)	52.6 (16.6)
Role physical				
BM	7.5 (16.3)	17.5 (29.3)	30.4 (41.6)	34.8 (42.7)
PT	2.7 (7.9)	5.6 (15.9)	20.8 (31.9)	37.9 (43.5)
CBT	10.2 (24.0)	22.5 (30.2)	32.9 (41.7)	51.3 (40.9)
CG	10.7 (27.6)	13.0 (23.7)	22.7 (29.8)	14.8 (27.5)
General health				
BM	47.1 (24.3)	50.1 (25.6)	49.6 (23.1)	53.8 (25.7)
PT	45.1 (19.0)	43.7 (23.4)	48.8 (23.9)	45.9 (23.2)
CBT	43.1 (20.3)	45.8 (20.8)	47.6 (18.8)	50.7 (20.9)
CG	51.0 (22.8)	48.4 (22.0)	47.0 (22.8)	44.7 (24.2)
Vitality				
BM	33.2 (18.8)	41.6 (24.2)	43.0 (26 3)	42.1 (25.6)
PT	29.3 (14.0)	35.4 (17.3)	35.2 (25.6)	38.2 (25.7)
CBT	26.4 (14.4)	39.3 (18.7)	43.4 (25.0)	43.0 (20.4)
CG	34.6 (22.0)	38.8 (21.9)	33.0 (22.9)	33.0 (23.0)
Social functioning				
BM	53.3 (23.2)	61.2 (24.9)	67.4 (25.5)	68.8 (27.3)
PT	47.0 (22.9)	54.1 (23.6)	52.1 (30.3)	58.0 (29.3)
CBT	57.4 (26.6)	75.0 (23.3)	74.3 (29.9)	80.6 (24.8)
CG	67.0 (28.1)	63.5 (27.6)	67.6 (31.3)	59.1 (31.4)
Role emotional				
BM	34.4 (43.3)	54.0 (46.6)	67.9 (44.9)	61.9 (46.9)
PT	32.4 (43.4)	50.5 (42.8)	51.9 (45.4)	48.1 (45.4)
CBT	40.9 (47.1)	55.0 (43.6)	70.2 (39.9)	65.0 (42.5)
CG	52.4 (47.5)	59.4 (42.6)	51.5 (43.3)	54.5 (47.8)
Mental health				
BM	55.7 (20.6)	66.8 (19.6)	64.0 (20.0)	64.9 (22.9)
PT	53.9 (19.5)	60.4 (21.4)	59.4 (23.5)	61.8 (21.6)
CBT	55.6 (21.3)	67.4 (20.3)	71.0 (21.3)	70.2 (18.6)
CG	69.9 (21.3)	68.2 (19.5)	65.5 (24.8)	58.2 (28.7)
Global health				
BM	38.1 (14.5)	47.6 (18.0)	52.4 (21.6)	53.1 (24.5)
PT	35.1 (11.4)	41.0 (15.1)	43.6 (22.7)	47.2 (24.7)
CBT	38.9 (13.7)	48.8 (16.8)	54.2 (19.3)	58.2 (18.4)
CG	45.6 (16.5)	47.0 (15.2)	46.3 (19.3)	43.4 (20.1)

[a] BM, behavioral medicine; PT, behavior-oriented physical therapy; CBT, cognitive behavioral therapy; CG, control group.

ment in the bodily pain, physical functioning, and vitality variables compared to the CG.

Altogether, females in the BM and CBT groups reported improvement in several of the SF-36 variables compared to

Table 6
Males: descriptive information for the SF-36[a]

Measure and group	Pre-treatment	Post-treatment	6-month follow-up	18-month follow-up
Bodily pain				
BM	27.4 (17.4)	31.2 (19.2)	40.2 (22.1)	44.0 (25.2)
PT	32.5 (19.4)	37.4 (20.7)	35.3 (17.4)	41.4 (22.8)
CBT	22.0 (13.9)	28.9 (11.6)	31.7(15.6)	38.0 (26.8)
CG	27.0 (16.5)	30.7 (15.6)	38.5 (23.9)	32.1 (13.0)
Physical functioning				
BM	55.3 (20.2)	58.5 (20.6)	62.6 (19.7)	63.3 (20.1)
PT	60.0 (14.9)	69.1 (15.6)	62.0 (18.0)	64.4 (17.1)
CBT	59.8 (21.6)	57.1 (18.2)	57.4 (21.0)	62.3 (27.5)
CG	60.6 (19.8)	59.7 (20.6)	62.5 (26.3)	62.8 (24.4)
Role physical				
BM	6.1 (15.4)	21.4 (34.5)	25.9 (35.7)	37.7 (42.1)
PT	6.3 (19.4)	12.5 (27.4)	31.7 (36.0)	23.5 (32.4)
CBT	14.8 (32.7)	14.4 (22.6)	23.8 (36.6)	35.2 (44.1)
CG	8.8 (24.7)	8.3 (22.5)	35.9 (42.8)	21.9 (34.0)
General health				
BM	48.0 (22.3)	49.8 (20.1)	53.8 (20.6)	53.4 (22.4)
PT	48.0 (20.4)	49.4 (18.2)	46.5 (21.0)	51.3 (19.9)
CBT	54.0 (17.8)	51.9 (21.4)	44.8 (21.8)	53.0 (29.6)
CG	58.6 (20.0)	58.5 (17.0)	54.6 (24.0)	49.3 (19.9)
Vitality				
BM	32.2 (19.8)	42.0 (19.7)	45.4 (23.5)	46.7 (21.3)
PT	35.0 (14.5)	40.0 (20.5)	47.0 (19.8)	47.1 (21.5)
CBT	37.8 (18.9)	40.6 (21.9)	34.3 (20.7)	42.8 (29.6)
CG	38.8 (24.5)	40.0 (19.5)	41.3 (30.7)	34.1 (25.1)
Social functioning				
BM	64.8 (24.1)	67.4 (23.9)	72.3 (23.2)	73.2 (26.8)
PT	63.2 (20.0)	66.4 (21.3)	65.0 (20.2)	66.2 (21.1)
CBT	66.7 (24.0)	62.0 (27.3)	58.5 (27.4)	61.9 (32.6)
CG	63.1 (23.1)	57.5 (23.1)	65.6 (18.0)	68.0 (26.6)
Role emotional				
BM	42.4 (45.1)	54.8 (44.6)	69.1 (40.5)	72.8 (39.3)
PT	29.2 (38.3)	33.3 (43.9)	51.1 (43.4)	56.9 (45.3)
CBT	34.6 (43.8)	43.6 (46.9)	45.0 (46.2)	50.0 (46.9)
CG	40.0 (39.9)	44.4 (43.0)	56.3 (39.9)	39.6 (42.6)
Mental health				
BM	56.9 (22.1)	62.9 (21.0)	64.9 (19.2)	66.4 (18.7)
PT	59.8 (20 6)	62 5 (15.9)	64.0 (20.5)	68.2 (19.5)
CBT	60.7 (17.7)	60.0 (23.0)	52.5 (23.3)	63.2 (25.3)
CG	63.0 (19.6)	61.3 (17.6)	57.8 (27.8)	59.8 (18.8)
Global health				
BM	41.6 (14.6)	48.5 (17.2)	54.3 (18.3)	57.2 (21.8)
PT	42.6 (13.3)	46.3 (14.4)	50.3 (16.5)	52.4 (17.9)
CBT	43.8 (16.0)	44.8 (18.0)	43.5 (19.1)	50.8 (27.9)
CG	45.0 (14.7)	45.1 (13.2)	51.5 (24.2)	45.9 (21.2)

[a] BM, behavioral medicine; PT, behavior-oriented physical therapy; CBT, cognitive behavioral therapy, CG = control group.

the CG at the 18-month follow-up. Among males, the only significant difference was an improvement in vitality in the BM program compared to the CG, although a non-significant general trend to improvement could be seen in all the three treatment programs (most pronounced in the BM program).

4. Discussion

Overall, the components of this BM program yielded as good results as the entire program when compared to the CG. No significant differences were revealed over the 18-month follow-up period in total absence from work between the treatment groups and the CG. However, a number of differences were found regarding early retirement and health-related quality of life to the advantage of the interventions in relation to the CG. The positive effects of the treatments were confined to females.

Contrary to our hypothesis, the analyses regarding different aspects of absence from work (except early retirement) did not show any differences between the treatment groups compared to the CG. The hypotheses regarding full-time early retirement were partially confirmed in that women in the PT and CBT groups had a lower risk of being granted early retirement. From a clinical point of view it is of vital importance to ask what was lacking or what can be improved in the rehabilitation programs in question with regard to their ability to facilitate return to work among the participants. Although the interventions evaluated in the present study did comprise contacts with supervisors and the drawing up of a written rehabilitation plan for return to work, this was not sufficient to lower absence from work. In a recent study by Loisel et al. (1997) in sick-listed back pain sufferers, it was shown that a combined intervention program (clinical intervention plus occupational intervention) did reduce the duration of absence from regular work, and that the occupational intervention program did account for most of this reduction in absence from work. Therefore, more active collaboration with the workplace may be needed.

The point in time for the intervention may also be crucial and, as reported by Linton (1999) the longer one postpones interventions, the more comprehensive the treatment probably must be. On the average, the participants in this study had been sick-listed for more than 4 months before inclusion and the interventions may not have been able to address the specific needs of these individuals in the light of this long period of absence from work.

The use of a population-based sample instead of referred patients could have also influenced the results. For instance, due to the fact that the participants were not considered for rehabilitation in the 'ordinary' way (referral by physicians or rehabilitation officials at local social insurance offices, etc.) they may not have received the same recognition and support after rehabilitation, neither at the workplace nor from a referring agent for resuming their paid work. The fact that these subjects had not been referred to any type of more comprehensive rehabilitation may also in itself indicate that they have been considered by referral agents to have a relatively low potential for rehabilitation. Furthermore, absence from work is thought to be affected by several factors not directly targeted in a rehabilitation program (e.g. employer policies or a disability compensation system) (Fordyce, 1995).

The hypothesis that the BM program would be consistently superior to the PT and CBT conditions in a comparison with the CG was not supported. Actually, with regard to early retirement for females, the PT and CBT conditions showed better results than the BM program. This is puzzling since the BM program consisted of the PT program plus the CBT program and was carried out by the same therapists as in the component treatments. It may be suggested that this difference could reflect that women in the PT or CBT conditions were equally likely to seek, but less likely to be awarded, early retirement than participants in the other conditions. However, it appears troublesome to explain why such a bias caused by administrative or other social/legal considerations would selectively affect the results for women in those conditions (PT and CBT). Moreover, although it may further be argued that the randomization procedure was not successful in allocating females with the same characteristics between conditions, or that the statistical efficiency of the analysis was insufficient, we do not have any satisfactory explanations of this result. It may be wise to await an attempt to replicate the results before drawing any further conclusions concerning this issue.

With regard to health-related quality of life of the females, the BM condition was superior to the PT program but not to the CBT condition, as compared to the CG. Consequently, the results suggest that the CBT component is necessary for improving health-related quality of life among females.

Furthermore, one unexpected finding was that the participants in the CBT program showed a higher total absence from work than the CG (especially among the males). Although the difference was non-significant doubt may be raised with regard to the efficacy of the CBT program to facilitate return to work among males. This finding needs further attention in upcoming research on CBT as a treatment for non-specific spinal pain.

A number of questions can be raised concerning the fact that the significant improvements in outcome with regard to early retirement and health-related quality of life were confined to females (even though a general, non-significant trend toward improvement in health-related quality of life was reported by the males). Do females partly have a different constellation of coping strategies than males which results in their benefiting more from the pain management techniques practiced in the programs in question? In a review by Unruh (1996) of gender differences in clinical pain experience, a number of differences in coping patterns were reported for males and females. It may either be that females have a greater need for the specific coping strategies presented in the programs or that they are able to apply these strategies in a more efficient manner. For instance, due to the fact that women have been found to have a higher total workload (paid plus unpaid work) than men (because of more time spent on household work and child care) (Lundberg et al., 1994) they may profit more from applying coping strategies in their home environment.

Another possibility might be that females accept a psychological view of chronic pain more willingly and subsequently adhere better to the treatment regimen, which results in a better outcome. However, this argument is not supported by the data, neither when examining expectations prior to treatment or adherence to the treatment plan. Admittedly, a host of other factors of psychological, social, cultural, or biological origin may influence the gender differences reported in this study (Unruh, 1996; Berkley, 1997; Riley et al., 1998).

No significant differences were found before treatment between the programs with regard to the participants' ratings of the perceived relevance of the rehabilitation. However, in absolute values, the CBT condition was rated to have somewhat lower relevance than the BM and PT conditions and, furthermore, this tendency still existed after treatment. It may not be surprising from a clinical point of view that a purely psychological treatment is perceived by the participants as somewhat less credible than the PT and BM programs for chronic spinal pain. Nevertheless, the figures point to the need to further develop ways of introducing and implementing treatments such as CBT to heighten their relevance from the point of view of the patients.

The finding that the males in the CBT group were more compliant than males in the BM group may be a consequence of the fact that males in the CBT group were sick-listed more often and therefore had more time to carry out various activities such as, for instance, exercise.

The present study was randomized and controlled and this design has been suggested to be the best way to compare the outcome of different interventions (Altman, 1996). As primary endpoints, it used both register data and self-reports to give a more comprehensive picture of the outcome. The register data were gathered from the NSIB and this means that we have complete data on absences from work for the whole follow-up period. The self-report outcome data were gathered by employing a widely used instrument (the SF-36) which has displayed satisfactory psychometric qualities. Furthermore, the participants in this study were gathered from a well-defined source population, the AGS insurance scheme (see the Subjects section for a definition of AGS). This means that the results can be generalized to a majority of employed individuals in Sweden suffering from similar problems involving long-term, non-specific spinal pain.

A number of limitations of the study should also be acknowledged. Firstly, the wide confidence intervals in the analyses of sick leave and health-related quality of life indicates a low statistical power, which increases the risk of type II errors. However, to perform the analyses without respect to gender was considered to be a less satisfactory alternative since a lot of information could be missed in that case as the results differed between males and females. Secondly, multiple tests or comparisons have been made in the data analyses which increases the risk of type I errors (e.g. in the comparisons between treatments with regard to adherence to the treatment plan). Consequently, the results must be interpreted with caution. Thirdly, the gender of the therapist may have influenced the results. Approximately 85% of the participants had female therapists exclusively. On the other hand, it is not clear whether the gender of the therapist interacts with the gender of the patient nor, furthermore, how this presumed interaction affects treatment outcome (Flaskerud, 1990).

In addition, as is usual in applied research, a large number of conditions may have affected the results, e.g. the 'charisma' of different therapists, the readiness and resources of the participants' family or social network to support work resumption and life-style changes, the existence of subgroups of patients who may benefit to different degrees from treatment (Rudy et al., 1995), or the fact that the readiness of patients to adopt self-management strategies for coping with pain may differ (Kerns et al., 1997), etc. However, since a randomized design was used, measures were taken to distribute these factors evenly among the conditions. Furthermore, it should also be noted that the non-response rates for the SF-36 were significantly different between conditions at the 18-month follow-up with the highest rate in the CG and the lowest non-response rate in the PT condition. It seems plausible that the participants in the CG were less motivated to fill in the questionnaires since they were not offered any treatment and therefore might have experienced less involvement in the study. However, as mentioned, the data on absence from work was complete for all participants.

In conclusion, we found that the treatment conditions were superior to the control condition in one or two of the three primary endpoints at the 18-month follow-up and that the positive effects were restricted to women. Females in the CBT and PT groups had a lower risk of being granted full-time early retirement and females in the CBT and BM groups reported a better health-related quality of life than females in the CG. However, the total absence from work (sick leave plus early retirement) was not significantly different between the control and the treatment groups. Furthermore, on the whole, the full-time BM program was not more effective than the PT and CBT programs compared to results in the CG.

Finally, as recognized in the Introduction, the economic costs associated with musculoskeletal pain are enormous in the western world. This makes it necessary to evaluate vocational rehabilitation programs in terms of their cost-effectiveness as well. Therefore, a 3-year follow-up of the participants in this study, which will address health care utilization and long-term health-economic aspects of the described programs, is ongoing.

Acknowledgements

We are grateful to Christian Garheden, Anders Grahn, Lena Hurtig, Anders Hägg, Walis Ludvigsson, Gull-Britt

Norrgård, Eva Nilsson, and Mathias Nilsson, who have all in different capacities helped us with the data. We also want to express our gratitude to the participants in the study and the personnel at the involved clinics. Furthermore, we wish to thank Mr Isaac Austin for his revision of the English in the manuscript. This study was supported by AMF-sjukförsäkring, Stockholm, Sweden.

References

American Academy of Orthopaedic Surgeons. Joint motion: method of measuring and recording, Churchill Livingstone, Edinburgh, 1966, pp. 87.

Altman DG. Better reporting of randomised controlled trials: the CONSORT statement [editorial, see comments]. Br Med J 1996;313:570–571.

Berkley, K.J. Sex differences in pain. Behav Brain Sci 1997;20:371–380; discussion 435–513.

Bronfort G, Bouter LM. Responsiveness of general health status in chronic low back pain: a comparison of the COOP charts and the SF-36. Pain 1999;83:201–209.

Cutler RB, Fishbain DA, Rosomoff HL, Abdel-Moty E, Khalil TM, Rosomoff RS. Does nonsurgical pain center treatment of chronic pain return patients to work. Spine 1994;19:643–652.

Debrunner HV. Das Kyphometer. Zeitschrift fur Orthopädie und ihre Grenzgebiete 1972;110:389–392.

Flaskerud JH. Matching client and therapist ethnicity, language, and gender: a review of research. Issues Ment Health Nurs 1990;11:321–336.

Flor H, Fydrich T, Turk DC. Efficacy of multidisciplinary pain treatment centers: a meta-analytic review. Pain 1992;49:221–230.

Fordyce, W.E. Back pain in the workplace. Seattle, WA: IASP Press, 1995, pp. 75.

Hansson, E., Hansson, T. Medicinska åtgärder för sjukskrivna med rygg- och nackbesvär. Rygg och nacke 3. Riksförsäkringsverket och Sahlgrenska universititssjukhuset, Stockholm, 1999, pp. 74.

Heinrich RL, Cohen MJ, Naliboff BD, Collins GA, Bonebakker AD. Comparing physical and behavior therapy for chronic low back pain on physical abilities, psychological distress, and patient perceptions. J Behav Med 1985;8:61–78.

Jacobson E, editor. Progressive relaxation. Chicago, IL: University of Chicago Press, 1938.

Jensen IB, Bodin L. Multimodal cognitive-behavioural treatment for workers with chronic spinal pain: a matched cohort study with an 18-month follow-up. Pain 1998;76:35–44.

Jensen IB, Nygren Å, Lundin A. Cognitive-behavioural treatment for workers with chronic spinal pain: a matched and controlled cohort study in Sweden. Occup Environ Med 1994;51:145–151.

Jensen, I., Bergström, G., Nygren, Å., Ljungquist, T. Kartläggning av rehabiliteringsinsatser för långtidssjukskrivna/förtidspensionerade arbetare och tjänstemän med besvär från ryggkotpelaren, 3, Sektionen för Personskadeprevention, Karolinska Institutet, Stockholm, 1998.

Jensen, I., Bergström, G., Ljungquist, T. Rehabilitering av patienter med smärttillstånd i ryggkotpelaren. Sektionen för Personskadeprevention, Karolinska Institutet, Stockholm, 1999.

Johansson C, Dahl J, Jannert M, Melin L, Andersson G. Effects of a cognitive-behavioral pain-management program. Behav Res Ther 1998;36:915–930.

Kerns RD, Rosenberg R, Jamison RN, Caudill MA, Haythornthwaite J. Readiness to adopt a self-management approach to chronic pain: the Pain Stages of Change Questionnaire (PSOCQ). Pain 1997;72:227–234.

Lindström I, Ohlund C, Eek C, Wallin L, Peterson LE, Fordyce WE, Nachemson AL. The effect of graded activity on patients with subacute low back pain: a randomized prospective clinical study with an operant-conditioning behavioral approach. Phys Ther 1992;72(279-290):291–293.

Linton SJ. The socioeconomic impact of chronic back pain: is anyone benefiting? [editorial]. Pain 1998;75:163–168.

Linton SJ. Prevention with special reference to chronic musculoskeletal disorders. In: Gatchel RJ, Turk TC, editors. Psychological factors in pain, New York: Guilford Press, 1999, pp. 374–389.

Linton SJ, Bradley LA. An 18-month follow-up of a secondary prevention program for back pain: help and hindrance factors related to outcome maintenance. Clin J Pain 1992;8:227–236.

Loisel P, Abenhaim L, Durand P, Esdaile JM, Suissa S, Gosselin L, Simard R, Turcotte J, Lemaire J. A population-based, randomized clinical trial on back pain management. Spine 1997;22:2911–2918.

Lundberg U, Mårdberg B, Frankenhaeuser M. The total workload of male and female white collar workers as related to age, occupational level, and number of children. Scand J Psychol 1994;35:315–327.

McHorney CA, Ware JE, Lu JFR, Sherbourne CD. The MOS 36-item Short Form Health Survey (SF-36). III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. Med Care 1994;32:40–66.

Nachemson, A. Back pain (Ont i ryggen. Orsaker, diagnostik och behandling). The Swedish Council on Technology Assessment in Health Care (SBU), Stockholm, 1991.

Nachemson AL. Newest knowledge of low back pain. Clin Orthop Rel Res 1992;279:8–20.

Philips HC, editor. The psychological management of chronic pain, New York: Springer, 1988. p. 226.

Riley III JL, Robinson ME, Wise EA, Myers CD, Fillingim RB. Sex differences in the perception of noxious experimental stimuli: a meta-analysis. Pain 1998;74:181–187.

Rudy TE, Turk DC, Kubinski JA, Zaki HS. Differential treatment responses of TMD patients as a function of psychological characteristics. Pain 1995;61:103–112.

Saunders JB, Aasland OG, Babor TF, De La Fuente JR, Grant M. Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption (II). Addiction 1993;88:791–804.

Sullivan M, Karlsson J, Ware Jr JE. The Swedish SF-36 health survey. I. Evaluation of data quality, scaling assumptions, reliability, and construct validity across general populations in Sweden. Soc Sci Med 1995;41:1349–1358.

Turk DC, Meichenbaum D, Genest M, editors. Pain and behavioral medicine: a cognitive-behavioral perspective, New York: Guilford Press, 1983.

Turner JA, Clancy S, McQuade KJ, Cardenas DD. Effectiveness of behavioral therapy for chronic low back pain: a component analysis. J Consult Clin Psychol 1990;58:573–579.

Unruh AM. Gender variations in clinical pain experience. Pain 1996;65:123–167.

van Tulder MW, Koes BW, Bouter LM. Conservative treatment of acute and chronic nonspecific low back pain. Spine 1997;22:2128–2156.

Waddell G. Biopsychosocial analysis of low back pain. Bailliére's Clin Rheumatol 1992;6:523–553.

Westin, B. 4-punktsppaus ad modum Bibbi Westin. Sjukgymnasten 1985:20–22

Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. Acta Psychiatr Scand 1983;67:361–370.