Journal of Back and Musculoskeletal Rehabilitation 18 (2005) 67–78
IOS Press

Feasibility and results of a randomised pilot-study of a work rehabilitation programme

Kathrin Meyer[a,*], Jaap Fransen[b], Hansjörg Huwiler[a], Daniel Uebelhart[a] and Andreas Klipstein[a]
[a]*Department of Rheumatology and Institute of Physical Medicine, University Hospital Zurich, Switzerland*
[b]*Department of Rheumatology, University Medical Centre Nijmegen, The Netherlands*

Abstract. *Objective:* To investigate the feasibility of a randomised controlled trial in a usual clinical setting and the effect of work rehabilitation on improvement to the ability to work in chronic pain patients.
Design: 33 patients with chronic musculoskeletal pain, valid job contracts and absence from work, random allocation. The interdisciplinary work rehabilitation contained work-specific exercises and education and lasted 8 weeks, 3.5 hours each working day. The control treatment took place with the allocating physician, who received recommendations concerning physical therapy and the uptake of work. Assessments took place at 0, 8 and 32 weeks. Measurements: ability to work, actual work status in % of a full time job, functional capacity tasks measured in kilograms.
Results: Feasibility: the recruitment was of a long duration due to the rather restrictive inclusion criteria. Actual work status: the improvement between the groups was not significantly different, but the ability to work improved significantly from an overall median of 0% to 50% at 8 weeks ($p = 0.004$ for the intervention group, $p = 0.026$ for the control group).
Conclusions: The evaluation of a running work rehabilitation programme in a clinical outpatient setting in clinical science is feasible, but a more effective recruitment strategy for a main study is favoured by application of a multi-centric or company based setting.

Keywords: Rehabilitation, vocational, musculoskeletal diseases, pain, evidence-based practice

1. Introduction and objectives

Long-term work incapacity due to musculoskeletal disorders and pain is a burden for patients and society. Among musculoskeletal disorders, low back pain and neck pain are the most common with a point prevalence of 15–30%, and are responsible for large indirect costs due to work incapacity [1–3]. For example, in the German building sector the production loss due to work incapacity from musculoskeletal disorders amounted to EUR 1 billion in the year 2000 [4]. In the Netherlands, the total costs of low back pain were estimated at US$ 5 billion annually [5]. In Switzerland, between 3.5 and 8.8% of the working population is on invalid pensions, depending on large regional differences, mainly because of musculoskeletal and psychiatric diseases [6]. Functional restoration programmes have been established to prevent patients remaining disabled for work. There are a number of such interdisciplinary programmes or work rehabilitation programmes that commonly aim at job reintegration and reducing disability in patients with physically demanding jobs [7–11]. The psychological and physical demands on the patient, as well as the effort on the side of the health professionals are large for a work rehabilitation programme and therefore it should be considered by both sides whether this effort is worth the endeavour. The results of studies carried out in other countries

*Address for correspondence: Kathrin Meyer, MPH, Institute of Physical Medicine, University Hospital Zurich, Gloriastrasse 25, 8091 Zurich, Switzerland. Tel.: +41 1 255 24 84; Fax: +41 1 255 43 88; E-mail: katharina.meyer@usz.ch.

ISSN 1053-8127/05/$17.00 © 2005 – IOS Press and the authors. All rights reserved

are inconclusive or in favour of work rehabilitation [12, 13], but the results may differ between countries due to differences in patient population and the health care system. Therefore we conducted a pilot study with patients with long-term work incapacity due to non-specific chronic pain, in which a functional restoration programme was evaluated. The aim of this study was primarily to test the feasibility of the study in a clinical outpatient setting; an existing clinical treatment procedure was evaluated and the randomisation was implemented in daily business, where health care professionals were doing routine work in a complex interdisciplinary context, in contrast to studies, in which clinical everyday practice is usually changed or amended for the study. Secondarily, the aim was to analyse whether there was an effect on "ability to work", the "actual work status" and "the functional capacity". The experiences and results of this study should help to plan larger studies in order to improve evidence-based practice.

2. Methods

2.1. Subjects

Subjects with an inability to work due to chronic non-specific pain of more than 3 months with musculoskeletal disorders that were referred to our centre from May 2001 to August 2002 were eligible for the study. Inclusion/exclusion criteria were checked at entry by the study physician. Inclusion criteria were:

- Sick-leave for at least 2 months or 50% work incapacity from a full-time job over 3 months
- An existing job contract, a job with light, middle, heavy or very heavy physical demands, according to the dictionary of occupational titles (described by Matheson [14]).
- Failure of physiotherapy or hospitalisation for the same health problem
- Suitability for the work rehabilitation programme, as assessed by three independent assessors (see below)
- Informed written consent

Different aspects of the suitability for the work rehabilitation programme were independently assessed by a physician, a physiotherapist or an occupational therapist and a social worker, by checking physical, psychological, and social factors as shown in Table 1. Suitability testing also included a standardized functional capacity evaluation [15]. Literature indicates that especially a lack of willingness to return to work [16], previous operations in the pain area, high intensity of pain, a positive score of the Waddell signs, lack of willingness to perform exercise even with pain [17], overt pain behaviour [18–20], a low self-estimation of physical performance [21] and the long duration of time off work [22] could be negative predictors for uptake of work. In our experience, also factors such as effort during evaluation of functional capacity, motivation to carry out a fitness programme or consistency of performance could limit the uptake of work and were also taken into consideration. Because the weight of these predictors for the uptake of work is not known enough, every assessor took the factors into account described in Table 1 as a negative or positive condition for rehabilitation and allocated the patient on the basis of the assessor's experience into three different levels of suitability (suitable, fairly suitable or not suitable) for rehabilitation. To minimize the influence of subjective notions we took the information from three different assessors with different professional backgrounds. If the majority of the assessors allocated the patient as suitable, the patient was included, if the majority considered the patient as not suitable, the patient was excluded. If all assessors allocated a patient as fairly suitable for the programme, or every assessor had a different opinion to the other, the pros and cons were discussed between the assessors. In the case of doubt after the discussion, the patient was also excluded.

Other exclusion criteria were:

- Contraindications to the evaluation of functional capacity or physical training, e.g. cardiovascular disease or recent surgical operations.
- Current depression or mental disease
- History of substance abuse
- Spinal nerve root compression or other neurological deficiencies of less than 3 months.
- Active inflammatory disease in an acute phase
- Disability pension claim or lack of cooperation on behalf of the employer to get the patient back to the workplace, or a planned dismissal
- Knowledge of the work rehabilitation programme and therefore not blinded to the programme or the control treatment
- Diverse: immediate uptake of work, night shift, no residency

Patients excluded from the study returned to the referring physician who received information about further treatment. All inclusion and exclusion criteria were checked before randomisation.

Table 1
Factors used as a rationale for indicating the suitability for rehabilitation

	Factor	Positive/Negative Condition for Rehabilitation
Physical Factors	No previous operations of the pain area	Yes/no (information of medical record)
	Low intensity of pain	$< 9/\geqslant 9$ on numeric rating scale
	Effort during the evaluation of the functional capacity	$< 3/\geqslant 3$ signs
	Positive score of the signs of Waddell	$< 3/\geqslant 3$ signs
Psychological Factors	Motivation to carry out a fitness program	Positive/negative answer on question
	Willingness to return to work	Positive/negative answer on question
	No overt pain behaviour	Observation of physiotherapist
	High self-estimation of physical performance	PACT score $> 125/\leqslant 125$ (out of 200)
	Willing to perform exercises even with pain	Completed/not completed step/lifting the arms test
	Consistency of the patients performance during the evaluation of the functional capacity	Score of inconsistency $< 2/\geqslant 2$
Social Factors	No potential dismissal	Yes/no
	Short time off work	$< 6/\geqslant 6$ months
	No intention to get disability pension	Yes/no

2.2. *Random allocation*

An independent person conducted random allocation by using a minimization procedure and a random number table, which is a well-accepted procedure for small sample sizes [23]. The aim of the minimization procedure is to match patients in relation to important variables. Our aim was to achieve a similar distribution in age, gender, duration of work incapacity, comorbidity and physical severity of the job. After the patient's inclusion, a concealed letter concerning the result of the randomisation was given to the therapist to allocate the patient to the respective group.

2.3. *Blinding of patients*

Patients knew about the common aim of the study and control treatment, namely the return to work, but were blinded concerning the two treatments. This meant they were told that they would undergo a fitness programme, but did not know what the exact content of the two treatments was until they started the treatment. The ethic commission agreed with the study and patient information.

2.4. *Intervention*

The intervention, called work rehabilitation programme, lasted 8 weeks, 3.5 hours per day, 5 days per week. The work rehabilitation programme aimed to increase functional capacity and improve the patient's self-efficacy using an operant behavioural therapy approach. The approach was interdisciplinary and involved rehabilitation physicians, a psychologist, a social worker, occupational and physiotherapists. Every patient had a therapist as a case manager to ensure that goals of the rehabilitation are adapted weekly and coordination between all members in the interdisciplinary team were guaranteed.

The programme contained work-specific exercises, progressive exercise therapy with training devices, education in ergonomics, learning strategies to cope with pain and to increase self-efficacy, a group intervention with the psychologist, sports activities for recreation and a workplace visit to develop appropriate workload-related exercises for the programme [24–26]. The uptake of work was designed to be gradual and started 4 weeks after the programme began.

2.5. *Control treatment*

The physician who referred the patient to the hospital administered the control treatment, called progressive exercise therapy. The physician received specific recommendations concerning work reintegration, medication and training. Five physicians specialized in rheumatology had designed the recommendations gathering information about treatment modalities for chronic musculoskeletal pain and rating them with a Delphi procedure [27]. The best-rated therapeutic interventions were exercise therapy such as progressive exercise therapy (with training devices, 3 times per week for 8 weeks) in a physiotherapy practice, or an interdisciplinary pain programme in a clinic for pain patients or sports activities undertaken on one's own initiative, further information about how to cope with pain given by the physician, medication such as antidepressants and/or analgesics, and recommendations for the physician how he should instruct the patient concerning the uptake of work.

2.6. Measurements

Assessments took place at 0 (prerehabilitation), 8 (postrehabilitation) and 32 weeks (follow-up), in both groups. Figure 1 shows the course of the study and the number of patients involved at each stage. The assessor was blinded regarding treatment allocation.

The primary outcome variables were 1) the ability to work in % of a full-time job as assessed by the physician, and 2) the actually performed work status in % of a full-time job. The secondary outcome variables (only at 0 and 8 weeks, since it was a pilot study) were 1) the functional capacity as measured by three standardized lifting tasks, 2) a self-estimation of physical performance using the Performance Assessment of Capacity Testing (PACT) [21], 3) perceived pain using Numerical Rating Scales (NRS), 4) a generic health questionnaire, the SF-36 [28] and, 5) a condition specific questionnaire, the Spinal Function Sort of the North American Spine Society (NASS) [29]. The PACT and SF36 were available in every required language (German, Italian, Portuguese, Spanish, Serbian-Croatian, Albanian and Turkish). The NASS was only available in German, therefore patients were able to fill it out with the assistance of another person, since the questionnaire was not a main outcome.

2.7. Statistics

Because the data were not parametric, they are shown with the median and interquartile range (Q_1, Q_3) and non-parametric statistics were used. Differences between demographic and baseline variables between the two groups were examined with the Mann-Whitney test, Chi-square test or Fishers exact-test, where appropriate. Differences in working capacity within groups were tested using the Wilcoxon test for paired samples and differences between groups were tested using the Mann-Whitney test. To investigate changes in the questionnaires, repeated measures ANOVA was conducted. P-values < 0.05 were regarded as statistically significant. To check whether there were confounding variables we also looked for associations between baseline variables, the assignment to the group and the outcome. This was done with the Chi-square test, Fishers exact-test, Spearman Rank correlation and Mann-Whitney test where appropriate. The analysis was conducted using SPSS.

3. Results

3.1. Recruitment and sample

We were able to include 3 to 4 patients per month from the hospital and a network of general practitioners. Only 33 of 270 patients referred could be included in the study (Table 2). The 33 patients had chronic pain with a median (Q_1, Q_3) duration of 36.0 (8.3, 84.0) months at several locations, as is shown in Fig. 2, and a duration of absence from work with a median (Q_1, Q_3) of 6 (4.0, 9.3) months. Of these 33 patients 20 did not have a professional education, 27 were immigrants. In Table 3 the baseline variables are shown. There were no significant differences between the groups, although the scores of the PACT appeared to be better in the control group, but this was not significant. Two patients of the intervention group did not finish the intervention and one patient of the control group did not want to take part at the functional capacity testing after the study period (Fig. 1). The data of the main outcome of these 3 patients who did not finish the treatment could be collected at all points of time, using patient records.

3.2. Blinding

The blinding of patients to the alternative treatment was successful. We evaluated the status of blinding at the baseline interview by indirectly questioning the patients as to their perceptions regarding the treatment they were about to receive and excluded patients if they had any knowledge of the work rehabilitation programme through the referring physician. As the control group was treated in diverse external facilities, contact between patients in the different groups was avoided. Assessors were blinded, but treatment personnel were not.

3.3. Adherence

None of the intervention group received co-interventions and all patients adhered to the work rehabilitation, except two dropouts that were unhappy with the programme. Recommendations for the control group consisted of progressive exercise therapy in a physiotherapy practice (2 times 45 minutes per week), a suggestion for medication and information for the physician how he/she should augment the step-by-step uptake of work. Neither an interdisciplinary pain programme nor sports activities on a patient's own initiative were recommended for any patient. The physicians carried out

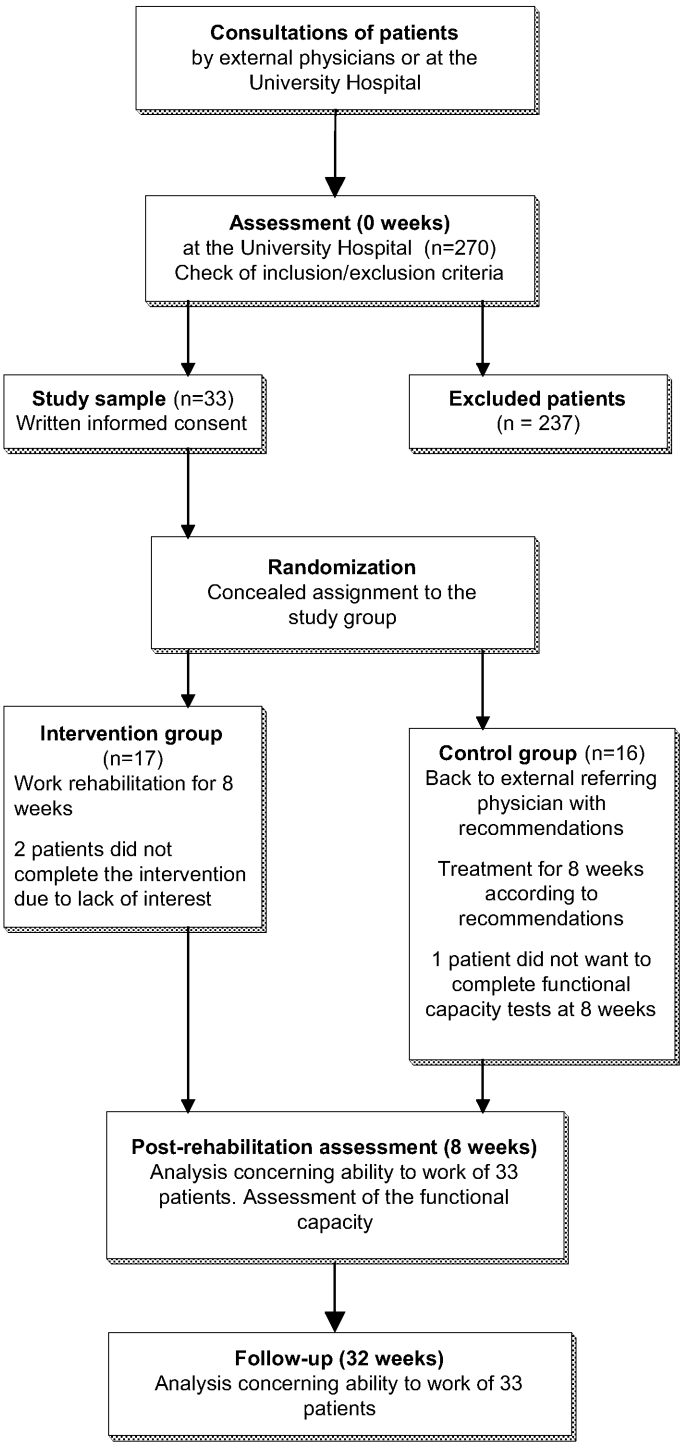

Fig. 1. Flow diagram of the study.

the treatment following the recommendations with two exceptions: one of the patients received physiotherapy and another acupuncture. Eleven patients of the control group adhered to the progressive exercise therapy. In contrast to the intervention group, the treatment of the control group was not interdisciplinary. The adherence

Table 2
Reasons why patients did not satisfy the inclusion criteria

Reason	n
Lack of suitability (see Table 1)	102
Lack of valid job contract	69
Too short time off work or working more than 50% of a full-time job	21
Medical contraindications	11
Disability pension claim	10
Did not want to take part	10
Were aware of the rehabilitation and were not blinded to alternative treatment	7
Diverse (e.g.: goes immediately back to work ($n = 5$), night shift ($n = 1$), no residency ($n = 1$))	7

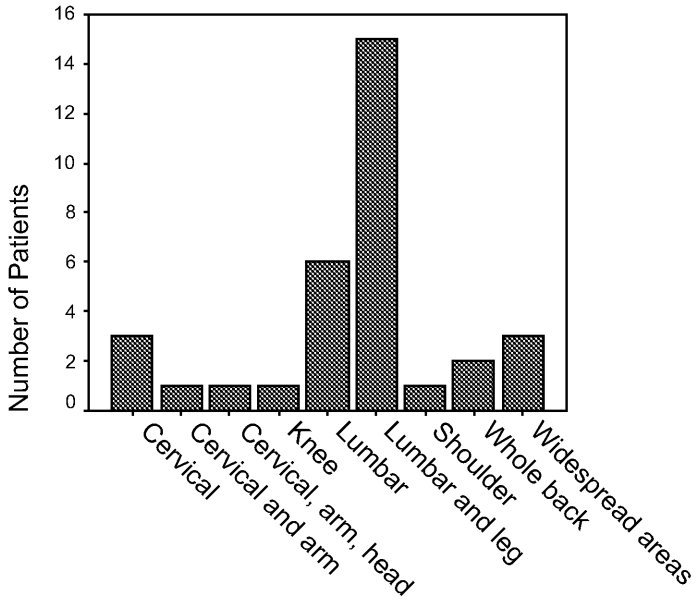

Fig. 2. Location of the main pain area.

to the treatment was checked by a telephone interview with the patient and with the treating physician.

3.4. Outcomes

After 8 weeks, the ability to work improved in both groups from an overall median of 0% to 50%. The improvement after the rehabilitation was significant for the ability to work for both groups and for the actual work status of the intervention group. However, there were no significant differences between the work rehabilitation and the control treatment (Table 4, Figs 3 and 4). The improvement could be maintained from 8 to 32 weeks in the intervention group and increased slightly but not significantly in the control group.

The prerehabilitation value of the functional capacity evaluation of the test carrying a weight over a short distance was with 2.5 kg somewhat higher in the work rehabilitation group (Fig. 5), but the difference was not significant ($p = 0.26$). The improvement in this test was generally small and had the tendency to be a little bit higher in the control group, but the change over time did not differ significantly between the groups ($p = 0.60$) (Fig. 5). The other tests "lifting a weight from floor to waist" and "lifting a weight to the height of the head" showed similar patterns, which were also not significant and therefore not shown in detail. Some patients consciously limited themselves for reasons of pain and did not reach the functional limit. The functional limit is the maximum weight that can be lifted in a safe ergonomic way. The tester evaluates the functional limit according to given observational criteria, such as the effort of muscles, the inability to stabilize joints, the heart rate and the change of movement patterns related to pain [15]. If a patient limits him/herself because of reasons of pain or of low effort, the objective criteria cannot be observed and therefore the result of the test cannot be analysed, which leads to a different number

Table 3
Baseline variables

		Work rehabilitation group	Control group	p-value
Gender	Male	13	13	1.0
	Female	4	3	
Age	Years, Median (Q_1, Q_3)	42 (35.0, 49.5)	44 (36.0, 50.0)	0.73
Duration of incapacity to work	Months, Median (Q_1, Q_3)	7 (4.0, 9.3)	5.6 (3.7, 9.7)	0.68
Physical demands at workplace	Light and middle physical job burden	7	8	0.73
	Heavy and very heavy physical job burden	10	8	
Comorbidity	No other diagnosis	11	9	0.73
	1 or more other diagnoses	6	7	
Radiating pain	Yes	13	13	1.0
	No	4	3	
Self-estimation of physical performance (PACT)	Score between 0–200, Median (Q_1, Q_3)	82 (59, 125)	136 (75, 161)	0.15

Table 4
Ability to work and actual work status, prerehabilitation, postrehabilitation and follow-up ($n = 33$)

Outcome	Group	0 weeks Median (Q_1, Q_3)	8 weeks Median (Q_1, Q_3)	0–8 weeks p-value	32 weeks Median (Q_1, Q_3)	8–32 weeks p-value
Ability to work in % of a full-time job	I	0.0 (0.0, 50.0)	50.0 (0.0, 70.0)	0.004	50.0 (25, 100)	0.547
	C	0.0 (0.0, 50.0)	50.0 (0.0, 71.3)	0.026	62.5 (0, 100)	0.167
Actual work status in % of a full-time job	I	0.0 (0.0, 50.0)	50.0 (0.0, 60.0)	0.009	50.0 (0, 50)	0.472
	C	0.0 (0.0, 50.0)	25.0 (0.0, 50.0)	0.068	50.0 (0, 100)	0.167

I = Intervention group.
C = Control group.
Differences between groups (I versus C) $p > 0.46$ in each case.

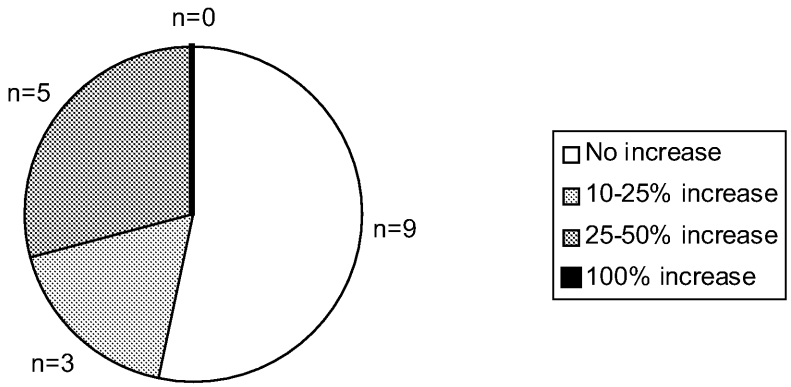

Fig. 3. Improvement of the intervention group. Improvement as shown in current work status (% of a full-time job; pre-/postrehabilitation).

of patients in each test. The number of patients who limited themselves by discontinuing a test varied from two to three from test to test, but were approximately even in both groups.

We did not find an association between improvement in functional capacity and an increase in the ability to work. The intensity of the "actual" and the "worst"/"lowest pain" during the previous seven days did not change significantly in either group (Table 5). After therapy, pain in new localizations appeared in 2 patients of the intervention group and in 5 patients of the control group. Other adverse reactions were not reported. The PACT and SF-36 scores did not show significant differences at baseline or between timepoints (Table 5). We did not observe any differences for the NASS. In the beginning the scores for all patients were 5.4 for the dimension of low back pain, 3.5 for the dimension of neurogenic complaints and 3.7 for the dimension of function ($p > 0.6$ for every comparison). There was no change between timepoints concerning the NASS.

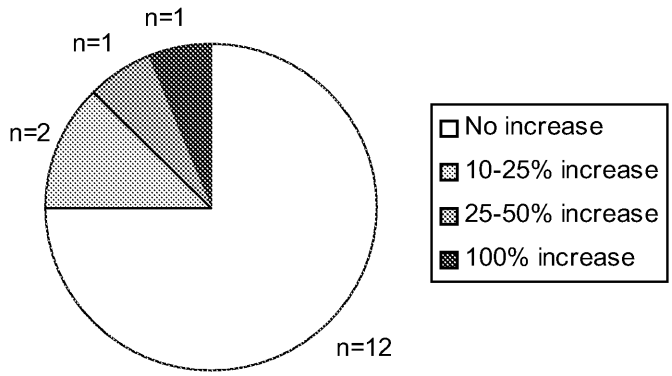

Fig. 4. Improvement of the control group. Improvement as shown in current work status (% of a full-time job; pre-/postrehabilitation).

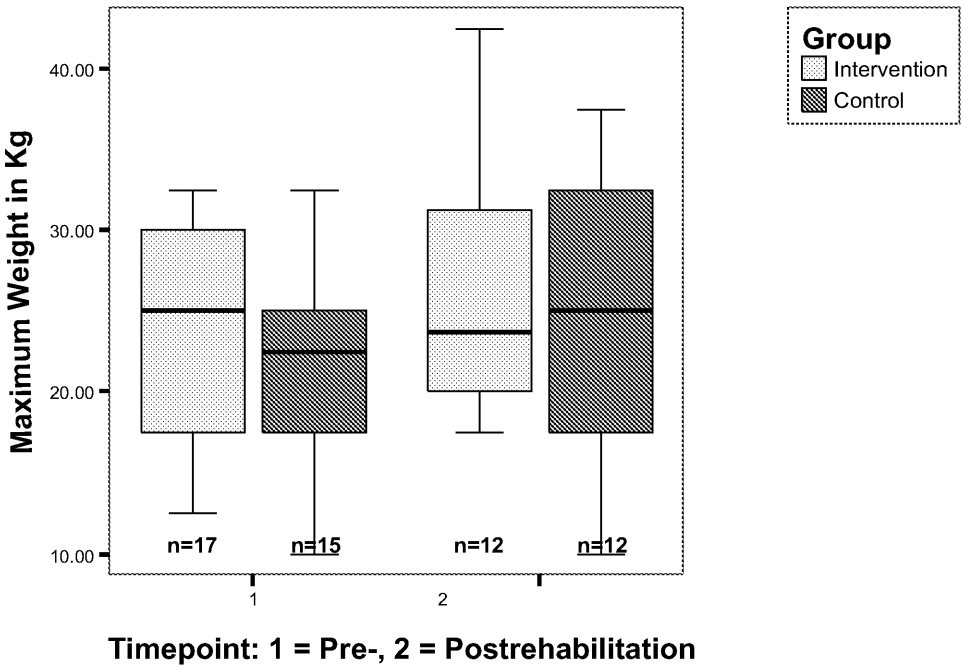

Fig. 5. Test of the functional capacity: carrying a weight over a short distance.

Female gender may be a positive predictor for work uptake; more women appeared to improve than men ($p = 0.07$). No patients received invalid pensions after the rehabilitation.

4. Discussion

4.1. Discussion concerning feasibility

The main issue of the study concerned the feasibility of the study. We believe this pilot study to be of interest because of the presumption that study costs could eventually be saved with studies undertaken in a daily clinical setting. Answering this question of feasibility, randomisation, blinding, intervention, data collection and analysis were carried out without problems, despite the usual daily clinical setting. But questions arise concerning the recruitment.

The recruitment was difficult, as has also been reported in previous studies. The time delay in carrying out our study was mainly due to the rather restrictive inclusion criteria, where 102 patients did not fit the suitability criteria and 69 did not have a valid job contract. We chose these inclusion criteria due to the three following reasons: firstly, the inclusion criteria

Table 5
Pain intensity and scores of the questionnaires PACT, SF-36; prerehabilitation, postrehabilitation

Outcome	Group	0 weeks Median (Q_1, Q_3)	Change[a] Median (Q_1, Q_3)	p-value[b]
Intensity of actual pain (NRS)	I	6.0 (3.5, 7.0)	1.0 (−1.0, 2.0)	0.34
	C	5.5 (5.0, 6.0)	1.0 (0.0, 2.0)	
Intensity of worst pain during the last 7 days (NRS)	I	9.0 (7.5, 10)	0.0 (−1.0, 1.0)	0.39
	C	8.0 (5.3, 8.0)	0.0 (−0.8, 2.0)	
Intensity of lowest pain during the last seven days (NRS)	I	4.0 (2.5, 5.0)	1.0 (−1.0, 2.0)	0.83
	C	4.5 (2.0, 5.7)	1.0 (−1.8, 3.0)	
Self-estimation of physical performance (PACT)	I	82.0 (58.5, 124,5)	10.0 (−9.0, 37.0)	0.60
	C	136.5 (74.8, 160.8)	3.0 (−10.0, 36.0)	
SF-36 (dimension of physical health)	I	30.0 (25.5, 34.5)	−0.4 (−3.8, 2.6)	0.17
	C	32.8 (28.4, 37.4)	−4.4 (−9.0, 0.4)	
SF-36 (dimension of mental health)	I	32.0 (26.7, 44.0)	6.0 (−5.4, 12.9)	0.9
	C	29.6 (25.2, 37.4)	2.5 (−3.6, 11.9)	

I = Intervention group.
C = Control group.
[a]Value for change between 0 and 8 weeks.
[b]p-value for change from prerehabilitation to postrehabilitation between groups.
NRS: numeric analogue scale (rating from 0–10).
PACT: perceived assessment of capacity testing (score between 0–200).
SF-36: generic health questionnaire (score between 0–100).

"suitability for rehabilitation" takes into account the motivation of a patient to take part in a treatment as stated in the study from Keel et al. [30]. Patients with little motivation mostly terminated their involvement too early or did not comply with the treatment. Furthermore, it does not make sense to conduct an expensive programme if the patient does not agree with the aims of the treatment. Therefore, the restricted inclusion criteria of the study mainly represent the restrictions applied in daily practice. Secondly, to include patients without an existing job contract would produce bias, because the situation on the labour market is a potential economic confounder. In the review of Elders [31] and in the study of Halderson [32] a valid job contract was also an inclusion criterion. This also represents the legal situation in Switzerland where a very weak protection against dismissal during sick leave exists.

The third reason is that it is better to study a homogenous group of patients in the first step, as the evidence of work rehabilitation is only moderate. In a further step, it would be appropriate to study the effect in a more heterogeneous sample to obtain a representative picture of patients with chronic pain.

The design of the main study shall be improved according to our experiences with the long lasting recruitment as follows:

The recruitment in the main study will take place with the help of companies with a worker's absence control system, to assess a greater number of patients who have a shorter absence from work and therefore are suitable for rehabilitation. Workers with absences from work will be identified by the companies, directly referred to the study assessment and be checked for the study inclusion criteria. This guarantees a valid job contract, a quicker referral to the study and a shorter duration of absence from work, contrary to the pilot study. A quicker referral augments the chance for more motivation to return to work and to be suitable for the study. The 6 months absence from work in the pilot study was due to the time delay that arose between the time that patients first contacted their general practitioner and a referral to a rheumatologist. The long absence from work diminishes the chance for the uptake of work. The importance of including patients with a shorter absence from work is also stated in another study [16].

For the main study a further important action to ensure a quicker enrolment is to conduct the study in a multi-centric manner to profit from a larger region.

4.2. Advantages of the design

The study was conducted in a daily clinical setting in terms of evidence-based practice, which could become a very important issue to augment cost-benefit of treatment in the future.

The rather restrictive inclusion criteria only allowed selection of patients who were expected to benefit from a work rehabilitation programme, and this resulted in a small rate of patients (9.1%) who did not continue the treatment. This is in contrast with the dropout rate of 24% in the study from Lindh [33]. A small dropout rate

is important to minimize bias, as patients who quit the treatment could have different characteristics to those who finish the treatment.

Selection bias was further avoided with the more representative sample obtained from the network of physicians. This led to a larger case mix than usual at any University Hospital.

4.3. Limitations

A selection bias due to inclusion restrictions cannot be excluded and may limit generalization of results. The objective of generalizing results has to be weighed out against evaluating specific work rehabilitation programmes in order to improve evidence based practice and cost-effectiveness. A possible overestimation of the ability to work directly after the intervention in both groups may have occurred due to the desire of the treating physicians to demonstrate a high return-to-work rate after treatment. But any potential overestimation should disappear by the 6-month follow-up due to natural progression and should show up within this period of time. Some studies only dealt with low back pain to reduce the variability of the patients, but we included every type of non-specific musculoskeletal disorder, because the aim of the treatment was more important than the diagnosis and we individually adapted exercises for each patient, so this should not contribute to an adverse effect. The idea of treating patients with various diagnoses is also supported by other studies [32,33], which included patients with many types of soft tissue damage or musculoskeletal disorders.

Two different assessors tested the functional capacity. Because the prerehabilitation tester was rotated with the work plan of the clinic, he/she did not know whether a potential study patient was being assessed. The postrehabilitation assessor was an external and blinded assessor, who was always the same person. For the main study a further recommendation is to conduct pre and post tests with the same test person, because intrarater reliability of the functional capacity testing is better than interrater reliability. The intrarater reliability has a kappa-value between 0.73 and 0.81, whereas the interrater reliability has values between 0.6 and 0.7 [34]. With the same tester, the results of the functional capacity testing would be more reliable.

4.4. Discussion concerning the results

The secondary aim of the study was to test whether there was an effect on the ability to work, the actual work status, and functional capacity. The differences in the main and secondary outcomes between the intervention and control group were only small.

We assume that these small differences in the main outcome may be due to the high quality of the treatment of the control group, because the recommendations for the control group treatment were developed by means of a Delphi procedure. So the unusual situation occurred that the control treatment included the best current modalities if the treatment is not an interdisciplinary work rehabilitation programme. This complex control treatment could have diluted the active treatment effects of the work rehabilitation programme. Another reason for the small differences may also be due to the "extra effort" from the treating physicians, knowing that their results were being monitored as part of the study.

This extra effort with a high quality control treatment was a result of negative experiences from a previous national research programme: in the previous study the randomisation had to be stopped because physicians were disappointed with the treatment of the control group, which only received usual care.

The improvement of functional capacity after the intensive training was smaller in the intervention group than in the control group (not significant). This was not expected after intensive training. But the higher values of the intervention group at prerehabilitation, even when not significant, may explain the smaller improvement of functional capacity, as it is difficult to improve the initial values if they are too high.

For the main study, the contrast between the two treatments should be larger. Patients will be referred directly and quickly to the study assessment by companies, consequently having little chance to receive physiotherapy before referral. In this case, physiotherapy without other treatments would be a rational and ethical treatment for the control group, as also shown in another study [12]. This is in contrast to the pilot study, where long illness history generated the chance for physiotherapy before the study began and therefore complex recommendations with several treatment modalities had to be made to provide an ethical treatment.

5. Conclusion

The study to evaluate an existing work rehabilitation programme in an outpatient setting was feasible. The pilot study could be carried out successfully in a daily

clinical setting with the limitation of a long recruitment time. There was a considerable overall improvement on the ability to work and actual work status, but the differences of the improvement were not significant between the groups. However, because of the small sample size, one cannot conclude that the more active and interdisciplinary work rehabilitation did not work. Possibly both treatments were effective, which could possibly be in contrast with a solely passive control treatment.

Recommendations for a following main study: for more rapid recruitment in the main study, only patients who potentially meet the inclusion criteria will be assigned to the assessment, therefore the recruitment will take place with the help of diverse companies, which guarantee a quicker referral of the patient to the study. Further, the study must be multi-centric. An evaluation of cost-effectiveness should be added. To increase the contrast between the groups, the recommendations for the control group should be a minimal intervention. Nevertheless, the control group should receive some kind of attention from the research team.

Acknowledgements

The authors would like to thank Anne Mannion for her useful and precise critique and Leanne Pobjoy for her help in the preparation of the manuscript. The financial support of the foundation "Hermann Klaus-Stiftung", "Kurt & Senta Herrmann Stiftung" and the "Forschungsfonds des Kantonalverbandes Zürich und Glarus des SPV" enabled us to conduct the study.

References

[1] A. Nachemson, G. Waddell and A.I. Norlund, Epidemiology of Neck and Low Back Pain, in: *Neck and Back Pain,* A. Nachemson, ed., Jonsson, Lippincott Williams & Wilkins, Philadelphia, 2000, pp. 166, 171, 172.

[2] J.W. Frymoyer and C.L. Durett, The Economics of Spinal Disorders, in: *The Adult Spine,* J.W. Frymoyer, T.B. Ducker, N.M. Hadler, J.P. Kostuik, J.N.Weinstein and T.S. Whitecloud, eds, Lippincott-Raven, Philadelphia, 1997, pp. 143–150.

[3] J.K. Moffett, G. Richardson, T.A. Sheldon and A. Maynard, Back pain: its management and cost to society, *York CHE Discussion Paper* **2** (1995), 5–67.

[4] B. Badura, M. Litsch and C.H. Vette, *Fehlzeiten-Report 2001,* Springer, Berlin, 2002.

[5] M.W. Van Tulder, B.W. Koes and L.M. Bouter, A cost-of-illness study of back pain in the Netherlands, *Pain* **62** (1995), 233–240.

[6] S. Spycher, J. Baillod, J. Guggisberg and M. Schär Moser, Analyse der interkantonalen Unterschiede innerhalb der Invalidenversicherung. National Research Programme 45 "Probleme des Sozialstaates", Project No. 4045-064749. SNF, Bern, 2004.

[7] P. Loisel, L. Abenhaim, P. Durand, J.M. Esdaile, S. Suissa, L. Gosselin, R. Simard, J. Turcotte and J. Lemaire, A population-based, randomised clinical trial on back pain management, *Spine* **22** (1997), 2911–2918.

[8] R.G. Hazard, J.W. Fenwick, S.M. Kalisch, J. Redmond, V. Reeves, S. Reid and J.W. Frymoyer, Functional restoration with behavioral support: A one-year prospective study of patients with chronic low-back pain, *Spine* **14** (1989), 157–161.

[9] A.F. Bendix, T. Bendix, M. Labriola and P. Boekgaard, Functional restoration for chronic low back pain, *Spine* **23** (1998), 717–125.

[10] D. Hart, Development of Clinical Standards in Industrial Rehabilitation, *JOSPT* **19** (1994), 232–241.

[11] T.G. Mayer and R.J. Gatchel, *Functional Restoration for Spinal Disorders: The Sports Medicine Approach,* Lea & Febiger, Philadelphia, 1988.

[12] E. Schonstein, D.T. Denny, J. Keating and B.W. Koes, Work Conditioning, Work Hardening and Functional Restoration for Workers with Back and Neck Pain, http://www.update-software.com/abstracts/AB001822.htm (Cochrane Review). The Cochrane Library, Issue 1, Accessed on October 30, 2003, Oxford: Update Software.

[13] J. Guzman, R. Esmail, K. Karjalainen, A. Malmivaara, E. Irvin and C. Bombardier, Multidisciplinary rehabilitation for chronic low back pain: systematic review, *BMJ* **322** (2001), 1511–1516.

[14] L.N. Matheson, Getting a Handle on Motivation: Self-efficacy in Rehabilitation, in: *The Comprehensive Guide to Work Injury Management,* S.J. Isernhagen, ed., Aspen Publishers, Gaithersburg, 1995, pp. 514–539.

[15] M. Oliveri, F. Denier-Bont and M.L. Hallmark, Evaluation der funktionellen Leistungsfähigkeit EFL nach Susan Isernhagen, *SUVA-Medizinische Mitteilungen* **69** (1996), 15–27.

[16] A.M. Carosella, J.M. Lackner and M. Feuerstein, Factors associated with early discharge from a multidisciplinary work rehabilitation program for chronic low back pain, *Pain* **57** (1994), 69–76.

[17] Ösch P, Möglichkeiten und Grenzen der Physiotherapie in der Rehabilitation chronischer Schmerzpatienten, *Physiotherapie* **11** (1998), 11–16.

[18] H.G. Kopp, M. Oliveri and A. Thali, Im Grenzbereich zwischen Körper, Psyche und sozialen Schwierigkeiten. Teil l: Neue Entwicklungen in der Diagnose und Therapie von somatoformen Störungen (am Beispiel von chronischen Schmerzpatienten), *Schweiz Med Wochenschr* **127** (1997), 1380–1390.

[19] H.G. Kopp, M. Oliveri and A. Im Grenzbereich zwischen Körper, Psyche und sozialen Schwierigkeiten. Teil ll: Die psychiatrische Begutachtung von somatoformen Störungen (am Beispiel von chronischen Schmerzpatienten), *Schweiz Med Wochenschr* **127** (1997), 1420–1439.

[20] L.N. Matheson, Symptom Magnification Syndrome, in: *Work Injury,* S.J. Isernhagen, ed., Aspen Publishers, Gaithersburg, 1988, pp. 257–282.

[21] L.N. Matheson, M.L. Matheson and J. Grant, Development of a measure of perceived functional ability, *J Occup Rehab* **3** (1993), 15–30.

[22] K. Beissner, R.L. Saunders and B.G. McManis, Factors related to successful work hardening outcomes, *Phys Ther* **76** (1996), 1188–1201.
[23] D.G. Altmann, *Practical Statistics for Medical Research,* Chapman & Hall/CRC, London, 1999, 443–445.
[24] F. Denier-Bont, M.L. Hallmark and H.J. Huwiler, Ergonomie in der Rehabilitation, 1. Teil, *Physiotherapie* **4** (1999), 13–16.
[25] F Denier-Bont, M.L. Hallmark and H.J. Huwiler, Ergonomie in der Rehabilitation, 2. Teil, *Physiotherapie* **6** (1999), 22–27.
[26] J. Weineck, *Optimals Training,* Perimed-spitta Med Publisher, Erlangen, 1994.
[27] J. Jones and D. Hunter, Consensus Methods for Medical and Health Services Research, in: *Qualitative Research in Health Care,* N. Mays and C. Pope, eds, BMJ Publishing Group, London, 1996, pp. 46–58.
[28] J.E. Ware, K.K. Snow, M. Kosinski and B. Gandek, *SF-36 Health Survey, Manual & Interpretation Guide,* The Health Institute, Boston, 1997.
[29] T. Schochat, W. Rehberg, J. von Kempis, G. Stucki and W.H. Jäckel, The North American Spine Society Lumbar Spine Outcome Assessment Instrument, *Z Rheumatol* **59** (2000), 303–313.
[30] P. Keel, C. Perini and D. Schütz-Petitjean, Chronifizierung von Rückenschmerzen: Hintergründe, Auswege, Final Report National Research Programme Nr. 26B, EULAR, Basel, 1996, 11–30.
[31] L.A.M. Elders, A.J. van der Beek and A. Burdorf, Return to work after sickness absence due to back disorders – a systematic review on intervention strategies, *Int Arch Occup Environ Health* **73** (2000), 339–348.
[32] E.M.H. Haldorsen, K. Kronholm, J.S. Skouen and H. Ursin, Multimodal cognitive behavioral treatment of patients sicklisted for musculoskeletal pain, *Scand J Rheumatol* **27** (1998), 16–25.
[33] M. Lindh, M. Lurie and H. Sanne, A randomized prospective study of vocational outcome in rehabilitation of patients with non-specific musculoskeletal pain: a mulitdisciplinary approach to patients identified after 90 days of sick-leave, *Scand J Rehab Med* **29** (1997), 103–112.
[34] S.J. Isernhagen, D.L. Hart and L.M. Matheson, Reliability of independent observer judgments of level of lift effort in a kinesiophysical Functional Capacity Evaluation, *Work* **12** (1999), 145–150.