

Effectiveness of a graded exercise therapy program for patients with chronic shoulder complaints

Jacques JXR Geraets^{1,2,3}, Mariëlle EJB Goossens^{1,4,5}, Imelda JM de Groot^{1,6},
Camiel PC de Bruijn^{1,2,3}, Rob A de Bie^{1,3,7}, Geert-Jan Dinant^{2,3}, Geert van der Heijden⁸
and Wim JA van den Heuvel^{1,3}

¹Institute for Rehabilitation Research iRv, Hoensbroek, ²Department of General Practice, Maastricht University, ³Care and Public Health Research Institute, Maastricht University, ⁴Department of Medical, Clinical and Experimental Psychology, Maastricht University, ⁵Research School Experimental Psychopathology, Maastricht University, ⁶Department of Rehabilitation, University Hospital, Maastricht University, ⁷Department of Epidemiology, Maastricht University, ⁸Julius Center for Health Sciences and Primary Care, University Medical Centre Utrecht The Netherlands

An operant behavioural and time-contingent graded exercise therapy program was developed to improve functional ability irrespective of pain experience in patients with chronic shoulder complaints. The clinical effectiveness of graded exercise therapy compared to usual care was evaluated in a randomised clinical trial. Assessments were carried out before and after 12 weeks of treatment. Performance of daily activities was measured by two outcome measures: the main complaints instrument and the Shoulder Disability Questionnaire (SDQ). Patients were eligible for participation if they had suffered from shoulder complaints for at least three months. Patients suffering from systemic diseases, referred pain or severe biomedical or psychiatric disorders were excluded. Patients (n = 176) were randomised and allocated either to graded exercise therapy (n = 87) or usual care (n = 89). Graded exercise therapy led to greater improvement in the performance of daily activities than usual care. However, only mean differences between groups in performance of activities related to the main complaints reached statistical significance ($p = 0.049$; 95% CI 0.0 to 15.0). The observed beneficial effects were considered to be small to moderate (calculated effect sizes: 0.30 for the main complaints instrument and 0.07 for the SDQ). Subgroup analysis showed larger improvements on the mean complaints instrument in patients not reporting pain reduction over time. Graded exercise therapy seems to be less effective in restoring performance of daily activities as assessed by the SDQ in patients showing a painful arc during physical examination. Results showed that graded exercise therapy is more effective in restoring the ability to daily activities in patients with chronic shoulder complaints than usual care, although beneficial effects are small. [Geraets JXR, Goossens MEJB, de Groot IJM, de Bruijn CPC, de Bie RA, Dinant G-J, van der Heijden G and van den Heuvel WJA (2005): Effectiveness of a graded exercise therapy program for patients with chronic shoulder complaints. *Australian Journal of Physiotherapy* 51: 87-94]

Key words: Shoulder Complaints; Behavioural Therapy; Randomised Clinical Trial; Physical Therapy

Introduction

Shoulder complaints are a major topic in primary health care: the 1998 point prevalence of self-reported shoulder pain in the general Dutch population was estimated at 21%, while 41% of all patients with shoulder complaints in the past 12 months visited a general practitioner (GP) (Picavet and Schouten 2003). A majority of shoulder complaints last for relatively long periods of time or recur: approximately 50% of all patients who visit their GP with a new episode endure complaints for up to six months (Kuipers et al 2004, van der Windt et al 1996, Winters et al 1997 and 1999) and up to 40% still have complaints after 12 months. Moreover, 46% of all patients with a new episode of complaints recall a previous history of shoulder complaints (van der Windt et al 1996). Eighty-five percent of patients with initial chronic shoulder pain in the general Swedish population still reported pain after 12 years (Andersson 2004).

Patients with shoulder complaints suffer from pain (27%) and disability (29%) (van der Heijden 1999, Picavet and Schouten 2003). Because shoulder complaints can have considerable impact on daily life and often develop into chronic disorders,

it is important to optimise the treatment. It has been hypothesized that the course of shoulder complaints involves biological, psychological, and social factors, as is also the case in other non-specific musculoskeletal pain disorders (Linton 1995). We therefore developed an operant behavioural and time-contingent graded exercise therapy program to improve functional ability in patients with chronic shoulder complaints despite their pain. We chose this approach in view of the proven effectiveness of this type of treatment in patients with low back pain (Fordyce et al 1986, Guzmán et al 2001, Lindstrom et al 1992, Linton 1999, Morley et al 1999, Tulder et al 2001, Vlaeyen and Linton 2000). We have previously described the theory and conceptual model underlying this program, reported on the program's development and content, and presented the design of our study (Geraets et al 2004). The present paper describes the clinical effectiveness of graded exercise therapy in patients with chronic shoulder complaints. Graded exercise therapy is compared with usual care as described in the 1999 guidelines for shoulder complaints issued by the Dutch College of General Practitioners (Winters et al 1999; see also <http://nhg.artsennet.nl/upload/104/guidelines2/E08.htm>). The objective of the study was to assess whether graded exercise

therapy is more effective than usual care after 12 weeks of treatment in terms of restoring the ability to perform daily activities irrespective of pain experience in patients with chronic shoulder complaints.

Methods

Design The short-term clinical effectiveness of graded exercise therapy compared with usual care was assessed in a randomised clinical trial. Short term was defined as the end of the 12-week treatment period. The graded exercise therapy was administered by 20 physiotherapists, whereas usual care was administered by 32 general practitioners (GPs). The study was approved by the iRv/SRL Medical Ethics Committee.

Recruitment and allocation of patients Patients with chronic shoulder complaints (lasting for at least three months) living in the Province of Limburg in the Netherlands were invited to participate in the study during the period from January 2002 to July 2003. With 25% and 50% recovered patients respectively in the usual care and graded exercise therapy group at 26 weeks follow-up, we aimed to have 132 patients participate in the study (66 in each intervention group) based on a one-sided alpha of 0.05, a statistical power (1-beta) of 0.90, and a 10% non-differential dropout rate. Patients were recruited either by the GPs during consultation for shoulder complaints, or by advertisement in local newspapers. All potential participants received written information and visited the GP to check eligibility according to the criteria listed below. A research assistant visited the patients at home for the intake procedure within two weeks of the consultation with the GP. Eligible patients were asked to give written informed consent. Patients were included if they were at least 18 years of age and had suffered from shoulder complaints in the shaded area shown in Figure 1 for at least three months.

Exclusion criteria were: medical treatment for shoulder complaints during the previous three months; complete rotator cuff tears; serious prior trauma (i.e. fractures or dislocations); prior surgery of the shoulder, upper limb, neck or thorax; osteoporosis; rheumatoid or bacterial arthritis; tumour; referred pain from internal organs; cervical radicular syndrome; gross shoulder hypermobility; stroke; polyneuropathy; multiple sclerosis; polymyalgia; ankylosing spondylitis; treatment for serious psychiatric disorders; or inability to complete questionnaires in Dutch.

Randomisation and allocation Patients were pre-stratified at physiotherapist practice level. We used block randomisation, with blocks of 10 patients, to allocate patients either to graded exercise therapy or to usual care in each stratum. Concealed randomisation was based on a random number list generated by an experienced researcher who was not involved in the conduct of the study. Patients received prepared opaque sealed and coded randomisation envelopes after the baseline measurements.

Graded exercise therapy Graded exercise therapy is a behavioural treatment program characterised by graded activity, time contingency, and operant conditioning. The primary aim of this exercise program is to increase the patients' ability to perform their own preferred shoulder activities in daily life at home or at work, irrespective of the pain experience. It does not aim to achieve pain relief at all. The program's activities are related to specific shoulder functions such as reaching, supporting, pushing, pulling,

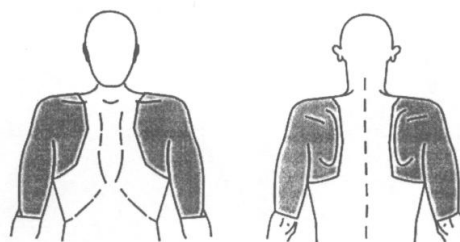


Figure 1. Area of shoulder complaints.

hitting, and stabilising, with work-related activities receiving special attention. The content of the graded exercise therapy program has been described in detail elsewhere (Geraets et al 2004).

An important aim of graded exercise therapy is to increase levels of daily activity by learning from the consequences of behaviours, the so-called operant learning principle (Fordyce 1976, Skinner 1953). Pain behaviour is expected to be consolidated when patients experience the consequences of this behaviour as pleasant and to be extinguished when they experience the consequences as unpleasant. Positive reinforcement of the preferred behavioural change is given to increase and maintain the desired behaviour. In graded activity, levels of activity increase in a step-wise time-contingent fashion, irrespective of pain experience. Goals are preset and agreed upon at the start of the program. The graded exercise therapy program consists of a maximum of 18 group sessions, each lasting approximately 60 minutes, over a period of 12 weeks, administered in groups of three to five persons. During the 'start-up period', the rationale is explained, and pain and pain-related disability are discussed. The 'treatment plus generalisation' period involves the performance of operant conditioning and time-contingent graded exercises. At the start of the program, the frequency of sessions is three times a week, after which it gradually decreases to one session every two weeks. Generalisation to everyday life is strongly emphasised at the end of this phase.

Prior to the start of the study, physiotherapists participating in the graded exercise therapy group (two in each of the 10 local practices) took part in a one-day workshop and two booster sessions under the supervision of experts in the field of cognitive-behavioural treatment.

Usual care Usual care was standardised according to the 1999 version of the guidelines for shoulder complaints issued by the Dutch College of General Practitioners (Winters et al 1999). Usual care consisted of information, recommendations, and pain-contingent medical or pharmaceutical therapy. During the first two weeks following the initial GP consultation, a wait-and-see policy was followed. This policy was supplemented with analgesics or NSAIDs if the GP thought this was necessary. The patient was instructed to re-consult the GP if complaints continued for more than two weeks. The GPs made the specific choice of treatment, like corticosteroid injections or referral to a physiotherapist or orthopaedic surgeon. GPs participating in the usual care group attended a refresher course on the application of the guidelines.

Table 1. Baseline characteristics.

| | Graded exercise therapy | Usual care | Eligible for analysis | Missing |
|--|-------------------------|-------------|-----------------------|-------------|
| Number | 87 | 89 | 150 | 26 |
| Demographic variables | | | | |
| Age, years (SD) | 51.2 (13.2) | 53.3 (11.7) | 53.1 (12.6) | 47.4 (10.3) |
| Female (%) | 51 | 58 | 54 | 58 |
| Specific disease variables | | | | |
| Present episode of shoulder complaints: | | | | |
| Duration of shoulder pain > 6 months (%) | 79 | 77 | 79 | 73 |
| Sick leave (if paid employment) > 1 week in past 8 weeks (%) | 4 | 12 | 7 | 13 |
| Quick onset (%) | 17 | 27 | 21 | 31 |
| Concomitant neck problems (%) | 54 | 40 | 47 | 50 |
| Cause: sudden movement (%) | 6 | 7 | 6 | 8 |
| Cause: overuse (%) | 25 | 26 | 25 | 27 |
| Cause: slight trauma (%) | 3 | 5 | 3 | 9 |
| Other cause (%) | 12 | 12 | 13 | 8 |
| Unknown cause (%) | 40 | 35 | 37 | 39 |
| Prior episodes of shoulder complaints: | | | | |
| Episodes of at least 1 week in past year (> 5) (%) | 68 | 69 | 70 | 61 |
| Treatment preferences: no/unclear/yes (%) | | | | |
| Behavioural therapy | 6/88/6 | 12/73/15 | 9/81/10 | 12/80/8 |
| Physiotherapy | 6/38/56 | 15/34/51 | 11/36/53 | 8/35/57 |
| Manual therapy | 7/80/13 | 21/59/20 | 12/70/18 | 25/63/12 |
| Analgesics | 67/29/4 | 79/14/7 | 76/19/5 | 56/32/12 |
| NSAIDS | 57/36/7 | 74/20/6 | 69/25/6 | 46/42/12 |
| Corticosteroid injection | 71/26/3 | 75/21/3 | 76/21/3 | 58/38/4 |
| Surgery | 76/22/2 | 84/14/2 | 81/17/2 | 73/23/4 |
| Prognostic variables (%) | | | | |
| Function of shoulder girdle | | | | |
| Passive range of motion exorotation < 60° | 38 | 36 | 33 | 37 |
| Active range of motion abduction/elevation < 150° | 17 | 15 | 17 | 24 |
| Presence of painful arc | 74 | 72 | 76 | 70 |
| Psychosocial (4DSQ) | | | | |
| Distress (medium/high)* | 32 | 18 | 26 | 19 |
| Anxiety (medium/high)† | 2 | 4 | 4 | 0 |
| Depression (medium/high)‡ | 10 | 9 | 10 | 8 |
| Somatisation (medium/high)§ | 43 | 33 | 34 | 54 |
| Outcome variables, mean (SD) | | | | |
| Performance of daily activities | | | | |
| Severity of main complaints (0–100) | 76.2 (19.2) | 71.9 (19.6) | 74.3 (19.2) | 72.3 (21.2) |
| Functional limitations in daily activities, SDQ (0–100) | 66.0 (18.1) | 65.6 (19.9) | 65.7 (17.9) | 66.3 (24.4) |
| Shoulder pain, SPS (7–28) | 17.9 (4.4) | 17.1 (3.8) | 17.4 (4.1) | 17.8 (3.7) |
| Pain intensity, 11-point scale (0–10) | 5.6 (2.2) | 5.4 (1.8) | 5.4 (2.0) | 6.3 (1.8) |
| Quality of life EuroQoL-5D (-1 to 1) | 0.66 (0.23) | 0.69 (0.20) | 0.68 (0.21) | 0.64 (0.23) |
| Process variables, mean (SD) | | | | |
| Kinesophobia, 2-item TSK (1–7) | 2.8 (1.9) | 2.8 (2.0) | 2.9 (1.9) | 2.2 (1.8) |
| Fear-avoidance beliefs, 5-item subscale FABQ (1–7) | 4.3 (1.3) | 4.5 (1.3) | 4.5 (1.3) | 4.2 (1.4) |
| Catastrophising, 12-item subscale PCCL (1–6) | 2.4 (0.8) | 2.3 (0.8) | 2.4 (0.8) | 2.3 (0.9) |
| Coping with pain, 11-item subscale PCCL (1–6) | 3.0 (1.0) | 3.2 (0.9) | 3.1 (0.9) | 3.2 (0.9) |
| Internal locus of control, 11-item subscale PCCL (1–6) | 3.3 (0.9) | 3.7 (0.8) | 3.5 (0.9) | 3.5 (0.8) |
| External locus of control, 7-item subscale PCCL (1–6) | 2.9 (0.8) | 3.0 (0.8) | 2.9 (0.9) | 2.9 (0.8) |

*Distress: low (0–10), medium (11–20), high (21–32)

†Anxiety: low (0–7), medium (8–12), high (13–24)

‡Depression: low (0–2), medium (3–5), high (6–12)

§Somatisation: low (0–10), medium (11–20), high (21–32)

¶PCCL subscale scores: very low (1.0–1.9), low (2.0–3.4), high (3.5–5.0), very high (5.1–6.0)

4DSQ, Four Dimensions of Psychological Symptomatology Questionnaire. FABQ, Fear-Avoidance Beliefs Questionnaire.

NSAIDS, nonsteroidal anti-inflammatory drugs. PCCL, Pain Coping and Cognition List. SDQ, Shoulder Disability Questionnaire.

SPS, Shoulder Pain Score. TSK, Tampa Scale for Kinesophobia.

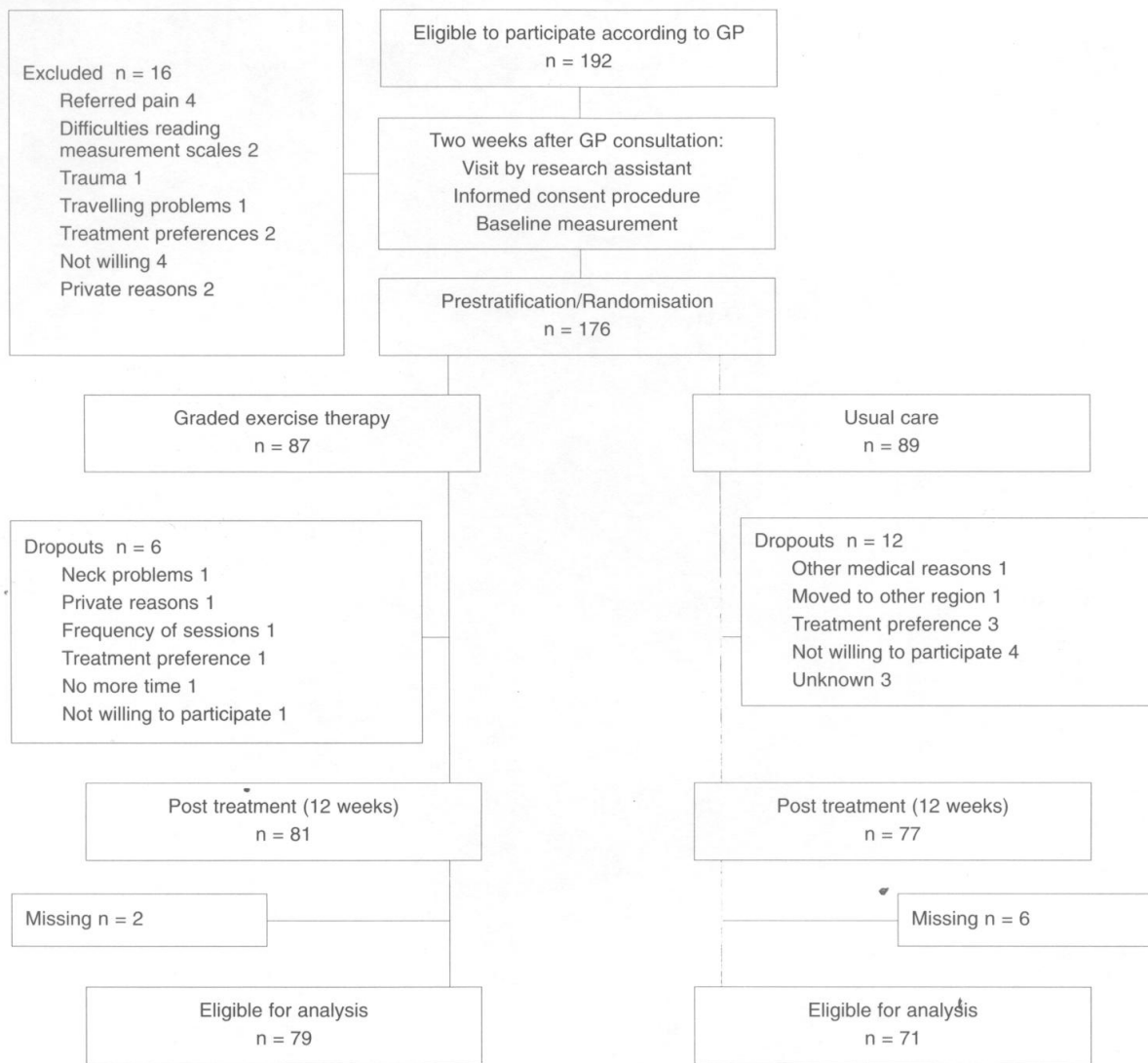


Figure 2. Flow-chart showing progression of the study.

Blinding Two research assistants who were not involved in the randomisation procedure carried out the data collection of baseline characteristics and outcome measurements. Blinding of patients and health care providers (physiotherapists and GPs) was not possible. Data entry and statistical analysis were carried out according to a pre-specified analysis plan and blinded for treatment allocation.

Baseline and outcome assessment At baseline, demographic variables, disease characteristics and treatment preferences were documented (Table 1). Before randomisation and after 12 weeks of treatment continuous outcome variables were documented.

The primary outcome variable, the performance of the level of daily activities, was assessed using two different assessment instruments: the main complaints instrument (a measure for the performance of activities related to patient-specific complaints) and the Shoulder Disability Questionnaire (SDQ) (van der Heijden et al 2000). In the first instrument, patients selected at baseline three daily activities they regarded as most important in relation to their shoulder complaints, but not necessarily related specifically to pain. They rated their ability to perform these activities during the past week on an ordinal (11-point) scale at baseline and during follow-up. The second instrument, the SDQ, is a functional status measure consisting of 16 statements

Table 2. Mean improvement in outcome measures after 12 weeks of treatment (after imputation of the overall mean).

| | Graded exercise therapy n = 87 mean (SD) | Usual care n = 89 mean (SD) | Mean difference | Differences between groups | | p value | Effect size |
|-----------------------------|---|--------------------------------|-----------------|----------------------------|-------|---------|-------------|
| | | | | 95% CI of diff* Lower | Upper | | |
| Primary outcome measures | | | | | | | |
| Main complaints (0–100) | 32.8 (25.7) | 25.3 (24.5) | 7.5 | 0.0 | 15.0 | 0.05† | 0.30 |
| SDQ (0–100) | 17.0 (26.0) | 15.3 (21.6) | 1.7 | -5.4 | 8.8 | 0.64 | 0.07 |
| Secondary outcome measures | | | | | | | |
| Shoulder pain (7–28) | 3.0 (3.3) | 2.3 (3.4) | 0.7 | -0.3 | 1.7 | 0.17 | 0.21 |
| Pain intensity (0–10) | 1.7 (2.2) | 1.5 (2.2) | 0.3 | -0.4 | 1.0 | 0.38 | 0.13 |
| Quality of life (-1–1) | 0.06 (0.22) | 0.06 (0.18) | 0.00 | -0.06 | 0.06 | 0.93 | 0.00 |
| Fear avoidance (1–7) | 0.4 (1.4) | 0.2 (1.0) | 0.1 | -0.2 | 0.5 | 0.45 | 0.12 |
| Kinesophobia (1–7) | 0.2 (1.6) | 0.0 (1.8) | 0.2 | -0.3 | 0.7 | 0.41 | 0.13 |
| Catastrophising (1–6) | 0.4 (0.7) | 0.2 (0.5) | 0.2 | 0.0 | 0.4 | 0.03 | 0.33 |
| Coping (1–6) | 0.1 (0.8) | 0.0 (0.7) | 0.1 | -0.2 | 0.3 | 0.50 | 0.11 |
| | n (%) | n (%) | Risk difference | Chi-squared | | p value | |
| Perceived recovery (yes/no) | 23 (29.1) | 13 (18.3) | 0.11 | 0.12 | | | |

*95% confidence interval of difference

†Statistical significance, $p < 0.05$

regarding pain and limitations while performing common daily activities during the past 24 hours.

The secondary outcome measures we assessed were perceived recovery, shoulder pain, generic health-related quality of life, catastrophising, coping with pain, kinesophobia, and fear-avoidance beliefs. Perceived recovery was measured on an eight-point ordinal scale (0, fully recovered, to 7, very much deteriorated). Patients were regarded as recovered if they felt either fully recovered (0) or very much improved (1). Shoulder pain was scored on the Shoulder Pain Score (SPS) and pain intensity was measured on an 11-point ordinal scale (Winters et al 1996). Generic health-related quality of life was rated on the EuroQol-5D (Brooks 1996). Catastrophising and coping with pain were scored on subscales of the Pain Coping and Cognition List (PCCL) (Berg et al 2001). Kinesophobia was assessed by two items (items 1 and 9) of the Tampa Scale for Kinesophobia (TSK-DV) (Vlaeyen et al 1995). Fear-avoidance beliefs were scored on the subscale of fear-avoidance beliefs about physical activity, adjusted for shoulder complaints, of the Fear-Avoidance Beliefs Questionnaire (FABQ-DV) (Vendrig et al 1998).

Statistical analyses All data were analysed according to the intention-to-treat principle. Analyses were carried out with SPSS statistical software (version 11.5). A p value of < 0.05 was considered to be statistically significant (two-tailed) for all comparisons.

The primary endpoints, i.e. the change in the performance of daily activities at 12 weeks as assessed by the main complaints instrument and the SDQ, were compared using the average change over time in both groups. For all cases, we

calculated the highest score of three selected main complaints at baseline and differences at follow-up. If two or three of the main complaints were rated equally highly at baseline, we calculated mean differences over time.

Differences between groups were calculated and analysed by means of Student's t -tests for all outcome variables measured on continuous scales and having a Gaussian distribution, and by Mann-Whitney tests for non-Gaussian distributions. Average changes for groups, mean differences between groups, and 95% confidence intervals of the differences were calculated. Chi-squared tests were performed for ordinal and dichotomous outcome variables. Risk differences were calculated and presented for dichotomous outcome measures. Effect sizes were calculated by taking the mean differences between groups and dividing them by the standard deviation of the average change of the total population. We considered effect sizes of 0.2, 0.5, and 0.8 as small, medium and large beneficial effects respectively (Cohen 1977).

We used imputation of the overall mean for missing data if patients for whom data were missing and patients eligible for analyses were comparable at baseline, i.e. if all baseline means for missing cases were no more than half of a standard deviation away from the overall means for continuous outcome measures and if baseline proportions differed by no more than 5% for dichotomous outcome measures.

We considered the following variables as putative prognostic factors: function of the shoulder girdle as assessed by physical examination (passive range of external rotation, active range of abduction/elevation, and presence of painful arc), psychosocial variables (anxiety, depression, somatisation, distress) (Four Dimensions of Psychological

Table 3. Crude and adjusted analysis of primary outcome measures.*

| Variables | B [†] | B [‡] | B [§] | 95% CI of difference | | p value | |
|--|--------------------------|----------------|----------------|----------------------|-------|---------|---------|
| | | | | lower | upper | | |
| Main complaints; mean differences between groups (0–100) | | | | | | | |
| Model 1.1 | Graded exercise therapy | 7.5 | – | – | 0.0 | 15.0 | 0.05 |
| Model 1.2 | Graded exercise therapy | – | 7.6 | – | 0.9 | 14.3 | 0.03 |
| | Change in pain intensity | – | 26.8 | – | 19.3 | 34.4 | < 0.001 |
| | Depression | – | 8.3 | – | 0.1 | 16.6 | 0.048 |
| Shoulder Disability Questionnaire; mean differences between groups (0–100) | | | | | | | |
| Model 2.1 | Graded exercise therapy | 1.7 | – | – | -5.4 | 8.8 | 0.64 |
| Model 2.2 | Graded exercise therapy | – | 0.6 | – | -5.6 | 6.9 | 0.84 |
| | Change in pain intensity | – | 23.6 | – | 16.3 | 30.8 | < 0.001 |
| | Painful arc | – | -10.3 | – | -17.4 | -3.3 | 0.004 |
| Model 2.3 | Graded exercise therapy | – | – | 10.8 | -1.2 | 22.7 | 0.08 |
| | Change in pain intensity | – | – | 24.6 | 17.3 | 31.8 | < 0.001 |
| | Painful arc | – | – | -3.5 | -13.3 | 6.3 | 0.49 |
| | GET ≠ painful arc | – | – | -14.0 | -28.1 | 0.1 | 0.05 |

*After imputation of the overall mean. For the graded exercise therapy factor, graded exercise therapy was coded 1 and usual care was coded 0. For the change in pain intensity factor, a change was coded 1 and no change was coded 0. For the painful arc factor, presence of a painful arc was coded 1 and absence was coded 0.

†Regression coefficient model adjusted only for graded exercise therapy.

‡Regression coefficient model adjusted for depression and presence of painful arc at baseline and for change in pain intensity.

§Regression coefficient model adjusted for presence of painful arc at baseline, interaction between graded exercise therapy and presence of painful arc at baseline and for change in pain intensity.

GET = graded exercise therapy.

Symptomatology Questionnaire (4DSQ)) (Terluin 1998), and treatment preferences. Using a stepwise forward procedure ($p < 0.10$), the main complaint instrument and SDQ as outcome measures, the influence of these putative prognostic factors and post-randomisation differences between groups was evaluated by means of multiple linear regression analyses.

A change in shoulder pain intensity was considered a priori to be a modifier. We assumed that pain reduction over time might naturally result in increased performance of daily activities. Since the aim of graded exercise therapy is to increase performance of daily activities, irrespective of pain experience, we assumed that the program might be more effective in patients not reporting pain reduction over time. Therefore, the influence of changes in pain intensity on performance of daily activities as a modifying factor was analysed. We defined subgroups based on pain reduction by more than one standard deviation away from baseline. We studied the effect of graded exercise therapy in these subgroups using multivariate linear regression for the main complaint instrument and for the SDQ. We also evaluated the interaction with treatment. Crude and adjusted regression coefficients and 95% confidence intervals are presented.

Results

The GPs considered 192 patients to be eligible to participate in the study. Sixteen patients were excluded before randomisation (for reasons, see Figure 2), so 176 patients were randomised and allocated to either the graded exercise therapy group ($n = 87$) or the usual care group ($n = 89$)

(Figure 2). Eighteen patients withdrew from the study during the treatment period. Six withdrawals had been allocated to graded exercise therapy (for reasons, see Figure 2), while 12 had been allocated to usual care (for reasons, see Figure 2). Data for eight patients were missing at 12 weeks (two graded exercise therapy and six usual care).

Table 1 shows that at baseline treatment groups were comparable on all continuous outcome measures. Patterns of treatment preferences were similar for both groups, although patients allocated to usual care more frequently favoured behavioural therapy and less frequently favoured manual therapy and drugs therapy (analgesics and NSAIDs). Patients allocated to usual care scored slightly lower on the distress subscale of the 4DSQ and less frequently reported concomitant neck problems but more frequently reported quick onset of their current complaints. Otherwise, prognostic status at baseline was comparable for both groups.

Patients who withdrew from the study during treatment period ($n = 18$) or had missing data at 12 weeks ($n = 8$) were more likely to report a quick onset and shorter duration of complaints, and had fewer prior episodes of shoulder complaints during the past year. In terms of treatment preferences, they less frequently favoured manual therapy and more frequently favoured drug therapy and injection therapy. Patients who withdrew scored slightly higher on the somatisation subscale of the 4DSQ. No difference on the baseline values of outcome measures was found.

Table 2 shows that patients allocated to graded exercise

therapy improved more than patients allocated to usual care as measured on both primary outcome measures. Differences between groups for improvement of the main complaints reached statistical significance ($p = 0.05$). Effect sizes for the main complaints instrument and the SDQ were 0.30 and 0.07 respectively. Improvements on all secondary outcome measures favour graded exercise therapy, but only differences between groups for improvement of catastrophising thoughts reached statistical significance. Imputation did not alter these results.

Crude and adjusted analyses are presented in Table 3. The between-group difference for reduction of severity of the main complaint at 12 weeks post randomisation for patients receiving graded exercise therapy was on average 7.5 points larger and statistically significant compared with those receiving usual care (model 1.1). The between-group difference for reduction in SDQ score at 12 weeks post randomisation favoured graded exercise therapy, but failed to reach statistical significance (model 2.1).

Pain intensity was unchanged in 132 patients, while pain reduction was seen in 44 patients. There was a significant relationship between reduction of the severity of the main complaint at 12 weeks on the one hand and pain reduction and depression scores at baseline on the other hand (model 1.2). After adjustments for pain reduction and depression score at baseline, the difference in reduction of severity of the main complaint between treatment groups was retained at the conventional level (7.5 versus 7.6). When analysed according to strata, mean differences between treatment groups for reduction of severity of the main complaint were larger in patients reporting no change in pain over time (7.6) than in patients reporting pain reduction (1.6).

There was a significant relationship between reduction in SDQ scores at 12 weeks on the one hand and pain reduction and presence of a painful arc at baseline on the other hand (model 2.2), and in addition for the interaction between graded exercise therapy and painful arc (model 2.3). After adjustment for pain reduction, for the presence of painful arc at baseline, and for the interaction between graded exercise therapy and painful arc, the difference in reduction of severity of disability (SDQ) between treatment groups changed substantially (from 1.7 to 10.8). When analysed according to strata, mean differences between treatment groups for reduction of severity disability (SDQ) were smaller and favoured usual care in patients with a painful arc (-0.2) as compared with patients with no painful arc (7.3). These findings suggest that in terms of improvement of the SDQ scores graded exercise therapy was less effective in patients showing a painful arc at baseline.

Discussion

Our findings suggest that the operant behavioural graded exercise therapy program is more effective than usual care in restoring daily activities in patients with chronic shoulder complaints. However, only differences between groups for the main complaint instrument reached statistical significance and we consider the observed beneficial effects to be small.

Performance of daily activities was measured by two outcome measures (main complaints instrument and SDQ). Patients allocated to graded exercise therapy improved more on both measures than those allocated to usual care, although differences between the groups were only statistically

significant for changes in performance of activities related to patients' main complaints. The SDQ is an instrument to assess activities in terms of functional limitations related to pain. Given the aim of the graded exercise therapy program, which is not primarily to reduce pain intensity, SDQ changes might be expected to be less obvious. Nevertheless, patients allocated to graded exercise therapy showed greater changes in shoulder pain than those allocated to usual care. Shoulder pain reduction was considered a priori to be a modifying factor. Mean differences between groups as assessed by the main complaints instrument were higher in patients not reporting pain reduction. Graded exercise therapy was more effective in restoring ability to perform daily activities as assessed by the SDQ in patients not reporting a painful arc.

For clinical practitioners like GPs and physiotherapists, these results mean that graded exercise therapy seems to be particularly effective in improving the performance of daily activities among patients who eventually show little reduction of pain intensity. Graded exercise therapy seems less effective in patients who report a painful arc during physical examination. However, since only 44 participants (26%) reported pain reduction post treatment, and since a painful arc was absent in only 48 patients (28%), these outcomes need to be interpreted with caution. We suggest that these factors should be evaluated in greater detail in further studies.

Graded exercise therapy is a behavioural treatment and not a cognitive behavioural treatment. Nevertheless, data show that graded exercise therapy had positive effects on all psychological outcome measures. The positive effect on catastrophising thoughts was even statistically significant ($p = 0.03$).

Thirty-six patients (23 graded exercise therapy versus 13 usual care) reported recovery post treatment (full recovery or very much improved). Differences between groups in perceived recovery rates were 11% and not statistically significant ($p = 0.12$). But even though patients did not perceive themselves as recovered, performance of daily activities improved. These results indicate that patients allocated to graded exercise therapy consider their ability to perform daily activities at home or at work less frequently dependent of perceived recovery or pain experience.

As patients were only excluded from the trial on the basis of systemic diseases, referred pain or severe biomedical or psychiatric disorders, we regard our study population as fairly representative of persons with shoulder complaints in the general practice population. Treatment groups were comparable with respect to outcome measures at baseline. Blinding of patients and health care providers (physiotherapists and GPs) for the nature of allocated treatment was not possible. However, although treatment preferences could therefore introduce bias these were found to have no influence on outcome measures after 12 weeks of treatment.

Blinding could explain the differential rates for dropouts and loss to follow-up. Eighteen patients (10.2%) withdrew from the study, and data for eight patients (4.5%) were missing for post-treatment analysis. Although more withdrawals had been allocated to usual care (12 usual care versus six graded exercise therapy), the numbers of patients who withdrew because of the content of the interventions (frequency of treatment sessions and treatment preferences) were small and similar for both groups (three usual care versus two graded exercise therapy).

Since patients who withdrew from the study during treatment period (n = 18) or had missing data at 12 weeks (n = 8), had a shorter duration of complaints and had fewer prior episodes of shoulder complaints during the past year, the prognostic status at baseline of these patients was probably slightly better compared with patients available for analysis (Windt van der et al 1996). Because the numbers of withdrawals per group were not equally balanced, we applied imputation of the overall mean for missing data. However, imputation of missing outcome data did not alter the results of this study.

Conclusion

This study confirms the hypothesis that behavioural factors play a role in the course of shoulder complaints, which is comparable to that in other non-specific musculoskeletal pain disorders (Linton 1995, Vlaeyen and Linton 2000). Further evaluation of long-term effectiveness and cost-effectiveness is needed before implementation of graded exercise therapy can be recommended.

Our overall conclusion is that graded exercise therapy seems to be more effective than usual care in restoring the performance of daily activities in patients with chronic shoulder complaints, although beneficial effects are small. Mean improvements for the main complaints instrument were larger in patients not reporting pain reduction. Graded exercise therapy has a favourable effect on catastrophising thoughts among patients with chronic shoulder complaints. It seems to be less effective in patients showing a painful arc during physical examination.

Acknowledgements This study was funded by the Netherlands Organization for Scientific Research (NWO-MW, grant number 904-65-901) and by the 'De Drie Lichten' Foundation, Hilversum, The Netherlands.

Correspondence Jacques Geraets, Institute for Rehabilitation Research iRv, PO Box 192, 6430 AD Hoensbroek, The Netherlands. Email: <jacques.geraets@planet.nl>

References

- Andersson HI (2004): The course of non-malignant chronic pain: A 12 year follow-up of a cohort from the general population. *European Journal of Pain* 8: 47–53.
- van den Berg SGM, Vlaeyen JWS, ter Kuil MM, Spinhoven PH, van Breukelen G and Kole-Snijders AMJ (2001): *Meetinstrumenten chronische pijn deel 2. Pijn Coping en Cognitie Lijst*. Maastricht: Pijn Kennis Centrum.
- Brooks R (1996): EuroQol: The current state of play. *Health Policy* 37: 53–72.
- Cohen J (1977): *Statistical Power Analysis for the Behavioral Sciences*. London: Academic Press.
- Fordyce WE (1976): *Behavioral Methods for Chronic Pain and Illness*. Saint Louis: Mosby.
- Fordyce WE, Brockway J, Bergman J and Spengler D (1986): A control group comparison of behavioural versus traditional management methods in acute low back pain. *Journal of Behavioral Medicine* 2: 127–140.
- Geraets JJXR, Goossens EJB, de Bruijn CPC, Köke AJA, de Bie RA, Pelt RAGB, van den Heuvel WJA and van der Heijden GJMG (2004): A behavioural treatment for chronic shoulder complaints: Concepts, development, and study design. *Australian Journal of Physiotherapy* 50: 33–38.
- Guzmán J, Esmail R, Karjalainen K, Irvin E, and Bombadier C (2001): Multidisciplinary rehabilitation for chronic low back pain: Systematic review. *BMJ* 322: 511–516.
- van der Heijden GJMG (1999): Shoulder disorders: A state-of-the-art review. *Bailliere's Clinical Rheumatology* 13: 287–309.
- van der Heijden GJMG, Leffers P and Bouter LM (2000): Shoulder Disability Questionnaire design and responsiveness of a functional status measure. *Journal of Clinical Epidemiology* 53: 29–38.
- Kuipers T, van der Windt DA, van der Heijden GJMG and Bouter LM (2004): Systematic review of prognostic cohort studies on shoulder disorders. *Pain* 109: 420–431.
- Lindstrom I, Ohlund C, Eek C, Wallin L, Peterson LE, Fordyce WE and Nachemson AL (1992): The effect of graded activity on patients with subacute low back pain: A randomised prospective clinical study with an operant conditioning behavioural approach. *Physical Therapy* 72: 279–290.
- Linton SJ (1995): An overview of psychosocial and behavioral factors in neck-and-shoulder pain. *Scandinavian Journal of Rehabilitation Medicine* 32: 67–77.
- Linton SJ (1999): Prevention with special reference to chronic musculoskeletal disorders. In Gatchel RJ and Turk DC (Eds): *Psychosocial Factors in Pain*. New York: Guilford Press, pp. 374–389.
- Morley S, Eccleston C, and Williams A (1999): Systematic review and meta-analysis of randomised controlled trials of cognitive behaviour therapy and behaviour therapy for chronic pain in adults, excluding headache. *Pain* 88: 1–13.
- Picavet HSJ and Schouten JSAG (2003): Musculoskeletal pain in the Netherlands: Prevalences, consequences and risk groups, the DMC (3)-study. *Pain* 102: 167–178.
- Skinner B (1953): *Science and Human Behavior*. New York: Macmillan.
- Terluin B (1998): De Vierdimensionale Klachtenlijst (4DKL) in de huisartspraktijk. *De Psycholoog*: 18–24.
- Tulder MW van, Ostelo RWJG, Vlaeyen JWS, Linton SJ, Morley SJ and Assendelft WJJ (2001): Behavioral treatment for chronic low back pain. A systematic review within the framework of the Cochrane Back Review Group. *Spine* 26: 270–281.
- Vendrig A, Deutz P et al (1998): Nederlandse vertaling en bewerking van de Fear-Avoidance Beliefs Questionnaire. *Nederlands Tijdschrift voor Pijn en Pijnbestrijding* 18: 11–14.
- Vlaeyen JWS and Linton SJ (2000): Fear-avoidance and its consequences in chronic musculoskeletal pain: A state of the art. *Pain* 85: 317–332.
- van der Windt DA, Koes BW, Boeke AJ, Deville W, de Jong BA and Bouter LM (1996): Shoulder disorders in general practice: Prognostic indicators of outcome. *British Journal of General Practice* 46: 519–523.
- Winters JC, Sobel JS, Groenier KH, Arendzen JH and Meyboom-de Jong B (1996): A Shoulder Pain Score: A comprehensive questionnaire for assessing pain in patients with shoulder complaints. *Scandinavian Journal of Rehabilitation Medicine* 28: 163–167.
- Winters JC, de Jongh AC, van der Windt DAWM, Jonquiere M, de Winter AF, van der Heijden GJMG, Sobel JS and Goudswaard AN (1999): NHG-Standaard Schouderklachten. *Huisarts en Wetenschap* 42: 222–231.