Pain 128 (2007) 59–68

PAIN

www.elsevier.com/locate/pain

Randomized controlled trial of exercise for chronic whiplash-associated disorders

Mark J. Stewart [a], Chris G. Maher [a,*], Kathryn M. Refshauge [a], Rob D. Herbert [a], Nikolai Bogduk [b], Michael Nicholas [c]

[a] *School of Physiotherapy, University of Sydney, Australia*
[b] *Department of Clinical Research, University of Newcastle, Royal Newcastle Hospital, Newcastle, Australia*
[c] *Pain Management and Research Institute, University of Sydney, Australia*

Received 7 January 2006; received in revised form 10 August 2006; accepted 23 August 2006

Abstract

Whiplash-associated disorders are common and incur considerable expense in social and economic terms. There are no known effective treatments for those people whose pain and disability persist beyond 3 months. We conducted a randomized, assessor-blinded, controlled trial at two centres in Australia. All participants received 3 advice sessions. In addition the experimental group participated in 12 exercise sessions over 6 weeks. Primary outcomes were pain intensity, pain bothersomeness and function measured at 6 weeks and 12 months. Exercise and advice was more effective than advice alone at 6 weeks for all primary outcomes but not at 12 months. The effect of exercise on the 0–10 pain intensity scale was −1.1 (95%CI −1.8 to −0.3, $p = 0.005$) at 6 weeks and −0.2 (0.6 to −1.0, $p = 0.59$) at 12 months; on the bothersomeness scale the effect was −1.0 (−1.9 to −0.2, $p = 0.003$) at 6 weeks and 0.3 (−0.6 to 1.3, $p = 0.48$) at 12 months. The effect on function was 0.9 (0.3 to 1.6, $p = 0.006$) at 6 weeks and 0.6 (−0.1 to 1.4, $p = 0.10$) at 12 months. High levels of baseline pain intensity were associated with greater treatment effects at 6 weeks and high levels of baseline disability were associated with greater treatment effects at 12 months. In the short-term exercise and advice is slightly more effective than advice alone for people with persisting pain and disability following whiplash. Exercise is more effective for subjects with higher baseline pain and disability.

Keywords: Whiplash injuries; Exercise therapy

1. Introduction

Whiplash-associated disorders (whiplash) are a common condition incurring considerable expense in social and economic terms. The most frequent outcome following a whiplash injury is persisting pain and disability. In a cohort study conducted in New South Wales, Australia, only 34% of subjects were free of pain and disability at 3 months (Rebbeck et al., 2006) and this rose to only 39% by 6 months. These figures are consistent with a study conducted in Queensland, Australia (Sterling et al., 2005), which showed that 38% of the cohort were free of pain and disability by 6 months. Recent systematic reviews of the prognosis of acute whiplash have reported similar results in other countries (Cote et al., 2001; Scholten-Peeters et al., 2003).

Current understanding is that by three months, approximately one-third of an inception cohort will have recovered from their initial pain and disability, approximately one-third will have persisting lower levels of pain and disability and approximately one-third of the original cohort will have quite high levels of persisting pain and disability (Rebbeck et al., 2006). The cases

* Corresponding author. Tel.: +61 2 93519192; fax: +61 2 93519601.
E-mail address: C.Maher@fhs.usyd.edu.au (C.G. Maher).

0304-3959/$32.00
doi:10.1016/j.pain.2006.08.030

that do not recover by 3 months are responsible for the majority of whiplash costs (Spitzer et al., 1995). Therefore, treatments that prevent transition to chronicity or are effective for chronic whiplash have the potential to significantly reduce social and economic costs.

Very few trials have evaluated interventions for patients with chronic whiplash i.e. pain and disability that persists beyond three months. In a recent systematic review of conservative treatment (Stewart, 2005) we were only able to locate two trials. One study was of low quality (Fitz-Ritson, 1995) and the other was underpowered (Soderlund and Lindberg, 2001) so firm conclusions about treatment effectiveness were not possible. The only known effective treatment for chronic whiplash is radiofrequency neurotomy (Lord et al., 1996), however this technique is only known to be effective for the subgroup of patients whose symptoms are shown by diagnostic blocks to arise from the lower cervical zygapophyseal joints.

Faced with a lack of known effective treatments for long-term pain and disability following whiplash, clinicians have adopted treatments that are known to work for other forms of chronic spinal pain. As there is evidence that exercise is effective for patients with persisting low back pain (Cost B13, 2004; Hayden et al., 2005), this treatment has been used for patients with chronic whiplash. While there is some logic to this practice, evidence for the lumbar spine cannot be confidently extrapolated to the cervical spine, particularly after a traumatic injury such as whiplash. We therefore directly tested the effectiveness of exercise treatment for people who had sustained a whiplash injury and had ongoing pain and disability that had persisted for greater than three months.

The study aimed to compare the effectiveness of exercise and advice with advice alone for people with persisting pain and disability following a whiplash injury. The study also investigated whether there were characteristics of patients that moderated the effect of exercise treatment.

2. Methods

2.1. Study sites

The trial was conducted at two physiotherapy clinics in Sydney, Australia. The enrolment period was from December 2001 to January 2004. The study protocol was approved by the Human Research Ethics Committee of the University of Sydney and has been published (Stewart et al., 2003).

2.2. Participants

To be eligible for inclusion, participants must have presented for medical care of a whiplash-associated disorder grades I–III within one month of the accident. Participants must also have reported that they were at least "mildly" disabled with respect to pre-injury status and have had significant pain or disability as indicated by a score of at least 20% on any one of the primary outcome measures. Participants were recruited by sending letters to claimants who had a whiplash injury at least 3 but not greater than 12 months earlier based upon lists of claimants provided to the authors by the Motor Accident Authority of New South Wales. Exclusion criteria were previous neck surgery, known or suspected serious pathology, nerve root compromise (i.e., whiplash is currently grade III), contraindication to exercise (ACSM, 1995) severe or greater depressive symptoms as measured by the Depression Anxiety Stress Scale (DASS), no neck radiograph obtained since the accident and current physiotherapy neck treatment. Because of the importance of clear understanding of the messages contained within the advice sessions, subjects with poor English comprehension were excluded from the study. Prior to entry to the trial, consent was obtained from all participants.

2.3. Randomization

After completing the baseline assessment, participants were randomly allocated to one of two intervention groups: exercise and advice or advice alone. The allocation schedule was generated by a person not involved in the trial and placed in sequentially numbered, sealed, opaque envelopes. The trial coordinator handed the next-numbered envelope to the treating physical therapist who then opened the envelope out of sight of the trial coordinator. Participants were considered to have been randomized and to have entered the trial at the time the envelope was opened. This process ensured allocation was concealed from participants, referring medical practitioners, trial staff who determined eligibility, and the assessor of outcomes.

2.4. Treatments

After baseline measures were taken by the trial coordinator, subjects were comprehensively assessed by the treating physiotherapist who took a complete history and performed a physical examination. During this assessment, current and pre-injury ability to perform work and home activities were established and measures of upper and lower body mobility, coordination, strength and aerobic fitness were taken. On the basis of this assessment and using the baseline measures, each subject was given advice (as below). Randomization to treatment group occurred after this advice had been given.

2.4.1. Advice

Subjects in the advice group received standardised education, reassurance and encouragement to resume light activity alone. The advice was given in one consultation and two follow-up phone contacts. The treating physiotherapist emphasised the favourable prognosis of whiplash and addressed common inaccurate beliefs about whiplash. Subjects were told that physical activity, even that which evokes moderate pain, is unlikely to further damage the neck and back, and that excessive voluntary limitation of activity may cause muscle weakness and impede recovery. The physiotherapist explored and discussed the patient's understanding of whiplash and attitudes and beliefs about activity resumption (e.g. fear of increased pain or re-injury) that may have imposed barriers to recovery.

Subjects were given a written report covering the main points of the advice session. Two and four weeks later subjects were contacted by telephone and the standardised advice was reinforced.

2.4.2. Exercise

The six-week graded exercise program was carried out under supervision from a physiotherapist (3 sessions in the first and second weeks; 2 sessions in the third and fourth weeks; and 1 session in the fifth and sixth weeks). Each session consisted of 1 h of exercises, 30 min of which was supervised by the therapist. The exercise program was an individualized, progressive, submaximal program designed to improve participants' ability to complete functional activities specified by the participant as being difficult because of whiplash. Each participant carried out a form of aerobic exercise (e.g., a walking or cycling program), stretches, functional activities, activities to build speed, endurance and coordination, and trunk and limb strengthening exercises. The therapist used principles of cognitive behavioral therapy including setting goals of progressively increasing difficulty, shaping, encouraging self-monitoring of progress and self-reinforcement (Nicholas and Tonkin, 2005). Participants were also given an individualized home exercise program that was regularly evaluated by the physiotherapist, and they were encouraged to continue the home program after the intervention period finished.

To ensure consistent administration of treatments, a treatment manual was developed and each physiotherapist was trained in the study protocol and interventions. Physiotherapists were educated by an experienced clinical psychologist about the principles of the cognitive behavioral approach. In addition, one of the investigators regularly visited each treatment site to monitor delivery of treatment.

Treatment compliance was established by recording the number of appointments attended, session duration and the amount of time spent with each participant. Participants were discouraged from seeking other health care during the 6-week treatment period. Participants who discontinued treatment were encouraged to return for follow-up assessments.

2.5. Outcome measures

Immediately prior to randomization, demographic, historical and social data were collected as well as treatment outcome measures. Six weeks and 12 months after randomization, additional data were collected on medication use, side-effects, adverse events and number and type of co-interventions.

Primary and secondary outcome measures were nominated *a priori* (Stewart et al., 2003). The primary outcomes were pain intensity, pain bothersomeness and functional ability at 6 weeks and 12 months. Secondary outcomes were disability, global perceived effect, health-related quality of life and work status at 6 weeks and 12 months.

Pain intensity and bothersomeness were both rated as the average over the last 24 h rated on a 0–10 box scale. Functional ability was measured using the Patient-Specific Functional Scale, range 0–10. Global perceived effect was measured on an 11-point scale, ranging from −5 (vastly worse) to 5 (completely recovered), with 0 being unchanged. Disability was measured with the Neck Disability Index, score range 0–50. Health-related quality of life was measured by the physical and mental summary scores of the SF36.

At both 6 weeks and 12 months, information about adverse effects of treatment was sought from all subjects using open-ended questioning. In addition, subjects' perceptions of the credibility of intervention were assessed at 6 weeks by questionnaire. Compliance with the activity program was also assessed at 6 weeks by examining attendance records and exercise diaries. Use of treatment (other than that provided as part of the trial) was also assessed at 6 weeks and 12 months by questionnaire.

Outcomes were administered and scored by staff blinded to treatment group, but treatment outcomes were self-assessed by participants so complete assessor-blinding was not possible.

2.6. Statistical analysis

Sample size was determined *a priori* with analyses comprehensively described in the trial protocol (Stewart et al., 2003). Data were double-entered and then analyzed by a blinded statistician, (author RH) on an intention-to-treat basis. Analysis followed a published protocol specified *a priori*. The only variation to the published protocol (Stewart et al., 2003) was to use logistic regression rather than Cox regression to analyze work status data. A few participants did not complete every item in a questionnaire. When this occurred missing items were substituted by the average item score of the questionnaire or subscale (for that participant). Participants were omitted from the analysis if all follow-up data were missing.

Effects of treatment were analyzed using a linear regression approach to analysis of covariance. We specified *a priori* that baseline scores of each outcome would be entered as a covariate in the analysis of that outcome. In secondary analyses we evaluated whether patient characteristics, all specified *a priori*, moderated the effect of treatment. This was accomplished by adding a term to each regression model for the interaction between the putative moderator variable and group membership. The emphasis was on point and interval estimation, but we also conducted hypothesis tests. The test-wise Type I error rate was set at 0.05.

3. Results

3.1. Recruitment and follow-up of participants

Letters were sent to 4889 potential participants. There were 297 respondents of whom 163 were ineligible (Fig. 1). Of the 134 participants randomized, 132 (99%) attended the 6-week follow-up and 125 (93%) attended the 12 month follow-up.

3.2. Baseline characteristics

The groups were similar at baseline (Table 1). Participants had moderate levels of pain and fear avoidance, were mildly disabled and distressed and were generally confident that the two treatments would be helpful.

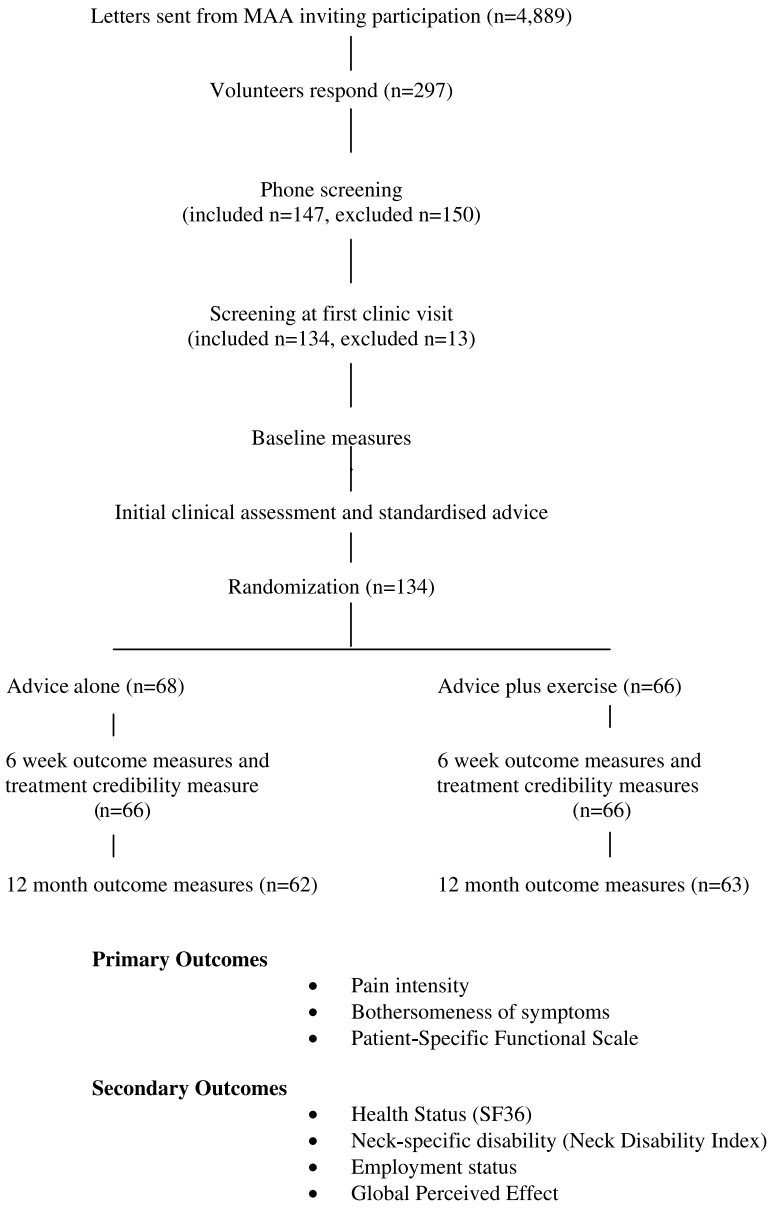

Fig. 1. Flow chart of whiplash trial.

3.3. Compliance with treatment

The mean number of exercise sessions attended was 9.9 (range 0–12 sessions) of the maximum 12 sessions. The mean number of advice sessions was 2.9 (range 1–3 sessions).

3.4. Other treatments

3.4.1. Additional treatment during intervention period (0–6 weeks)

Fifteen subjects (23%) in the advice group reported seeking additional treatment during the intervention period despite encouragement to refrain from extra therapy during the period. Of these, nine subjects attended physiotherapy, two attended massage therapy, one began a gym program, one had chiropractic treatment and one subject had both physiotherapy and chiropractic treatment and one subject did not specify the treatment. In the advice plus exercise group, 10 subjects (15%) sought additional treatment during the treatment period. Three subjects sought chiropractic treatment and massage, one sought other physical therapy treatment and massage, one sought other physical therapy treatment, one attended hydrotherapy, one attended acupuncture treatment, one attended Pilates classes, one attended work conditioning and one attended massage therapy.

Table 1
Baseline characteristics of study population

Variable	Advice Group ($N = 68$)[a]	Advice plus exercise group ($N = 66$)	All subjects ($N = 134$)[a]
Age (years)	42.7 (14.4)	43.9 (15.1)	43.3 (14.7)
Gender (count and %)	41F (62%)	48F (73%)	89F (66%)
Duration of Symptoms (months)	8.6 (2.5)	9.5 (2.1)	9.0 (2.4)
Pain (0–10 scale)	5.3 (2.0)	5.2 (2.0)	5.3 (2.0)
Bothersomeness (0–10 scale)	7.1 (2.3)	6.8 (2.4)	7.0 (2.3)
PSFS Average[b] (0–10 score)	4.1 (1.6)	3.9 (1.7)	4.0 (1.6)
GPE[c] (−5 to 5)	0.3 (2.4)	0.6 (2.4)	0.5 (2.4)
NDI[d] (0–50)	19.7 (6.9)	18.2 (6.3)	19.0 (6.6)
DASS21[e]	8.4 (9.3)	8.1 (9.0)	8.3 (9.1)
Advice expectation[f] (0–10 scale)	6.9 (2.8)	6.9 (2.5)	6.9 (2.7)
Exercise expectation[g] (0–10 scale)	8.2 (2.2)	8.0 (1.9)	8.1 (2.0)
Tampa[h] (0–68 scale)	41.0 (7.7)	38.3 (7.8)	39.7 (7.8)
SF36 – Physical summary score	36.8 (8.6)	36.4 (9.9)	36.7 (9.3)
SF36 – Mental summary score	48.0 (11.4)	49.0 (11.0)	48.7 (11.0)

Data are means and SDs except where indicated.

[a] Data include 2 subjects in advice group who dropped out of study after randomization (6 week data not obtained).

[b] Patient-Specific Functional Scale score (average of 3 scores, all out of 10, 0 representing unable to perform activity, 10 representing able to perform activity at pre-injury level).

[c] Global Perceived Effect Score (−5 representing vastly worse, 0 representing unchanged, 5 representing completely recovered).

[d] Neck Disability Index Score (0–50 scale, 10 items, higher scores represent greater disability).

[e] DASS 21 Score (21 items, depression score only shown. Scale range 0–42 score).

[f] Patient perception of advice treatment. Question asked "How helpful do you believe advice alone will be for your whiplash problems?" (0–10 scale, 0 representing not at all helpful, 10 representing extremely helpful).

[g] Patient perception of exercise treatment. Question asked "How helpful do you believe advice plus an exercise program will be for your whiplash problems?" (0–10 scale, 0 representing not at all helpful, 10 representing extremely helpful).

[h] Tampa score (17–68 scale, 17 items, higher scores represent greater fear avoidance).

3.4.2. Additional treatment during follow-up year (6 weeks–12 months)

Thirty-five subjects (56%) in the advice group sought additional treatment during the 12 month follow-up period. Of these, 17 had further physiotherapy, five had massage, four had chiropractic treatment, two had osteopathic treatment, two attended hydrotherapy (one of these in conjunction with a prescribed gym program), two had acupuncture (one of these in conjunction with massage therapy), one attended a work conditioning program, one consulted multiple practitioners including a physiotherapist, a massage therapist and a chiropractor, and one subject made use of magnetic therapy. Eighteen subjects (29%) in the exercise group sought additional treatment during the 12 month follow-up period. Six had further physiotherapy, seven had massage therapy (one in conjunction with hydrotherapy and one in conjunction with facet joint injections), two had chiropractic treatment, one had acupuncture, one attended a work conditioning program and one subject attended hydrotherapy.

3.5. Effectiveness of treatment

All groups improved on average on the primary and secondary outcomes from baseline to 6 weeks but the improvements from 6 weeks to one year were more modest (Table 2).

The primary analysis compared effects of exercise and advice with advice alone on pain intensity, pain bothersomeness and patient-specific function at the 6-week and 12-month follow-ups (shaded cells in Table 2). The effects of exercise (i.e. the adjusted difference in outcomes between exercise and advice and advice on a 0–10 scale) at 6 weeks were −1.1 (95%CI −1.8 to −0.3, $p = 0.005$) for pain intensity, −1.0 (−1.9 to −0.2, $p = 0.003$) for pain bothersomeness, and 0.9 (0.3–1.6, $p = 0.006$) for patient-specific function. These effects were smaller and no longer statistically significant at 12 months.

The secondary analysis showed that exercise and advice was more effective than advice alone in reducing disability, improving health-related quality of life (SF36 physical and mental summary scores), and had a greater global perceived effect. With the possible exception of global perceived effect the effects were very small at 6 weeks; and at 12 months all were smaller and no longer statistically significant. There was no effect of exercise and advice treatment on employment status at 6 weeks or 12 months (Table 3).

3.6. Adverse events

There were no instances in this trial where subjects required referral to a medical practitioner for adverse effects due to participation in the trial. At 6 weeks

Table 2
Mean (SD) for outcome measures at baseline, 6 weeks, and 12 months and effects of exercise vs advice

Outcome measures	Group means (SD)		Effects (95% CI)	
	Advice	Exercise and advice	Exercise and advice vs advice	*p* value
Pain (0–10)				
Baseline	5.3 (2.0)	5.2 (2.0)		
6 weeks	4.3 (2.5)	3.2 (2.2)	−1.1 (−1.8, −0.3)	0.005
12 months	3.8 (2.7)	3.5 (2.3)	−0.2 (−1.0, 0.6)	0.590
Bothersomeness (0 to 10)				
Baseline	7.1 (2.3)	6.8 (2.4)		
6 weeks	4.8 (2.9)	3.6 (2.6)	−1.0 (−1.9, −0.2)	0.019
12 months	4.0 (3.0)	4.1 (2.5)	0.3 (−0.6, 1.3)	0.480
PSFS (0–10)[a]				
Baseline	4.1 (1.6)	3.9 (1.7)		
6 weeks	5.6 (2.0)	6.4 (2.1)	0.9 (0.3, 1.6)	0.006
12 months	6.0 (2.4)	6.6 (1.9)	0.6 (−0.1, 1.4)	0.100
Neck Disability Index[c] (0–50 Scale)				
Baseline	19.7 (6.9)	18.2 (6.3)		
6 weeks	15.7 (7.9)	12.0 (6.8)	−2.7 (−4.5, −0.9)	0.004
12 months	15.5 (9.9)	12.1 (7.5)	−2.3 (−4.9, 0.3)	0.080
SF36 Physical Summary Score (0–100 scale)				
Baseline	36.8 (8.6)	36.4 (9.9)		
6 weeks	38.9 (9.3)	42.1 (8.9)	3.6 (1.3, 6.0)	0.003
12 months	40.7 (11.3)	42.3 (9.8)	1.9 (−1.4, 5.1)	0.260
SF36 Mental Summary Scale (0–100 scale)				
Baseline	48.0 (11.4)	49.0 (11.0)		
6 weeks	46.4 (12.9)	51.4 (9.7)	4.6 (1.4, 7.9)	0.005
12 months	46.1 (12.4)	48.4 (11.4)	1.8 (−1.8, 5.4)	0.330
Global Perceived Effect[b] (−5 to 5 scale)				
Baseline	0.3 (2.4)	0.6 (2.4)		
6 weeks	1.5 (2.5)	2.5 (1.8)	0.9 (0.3,1.6)	0.006
12 months	1.9 (2.3)	2.3 (2.0)	0.3 (−0.5, 1.0)	0.480

Primary outcomes are shaded.
[a] Patient-Specific Functional Scale score (Average of 3 scores, all out of 10, 0 representing unable to perform activity, 10 representing able to perform activity at pre-injury level).
[b] Global Perceived Effect Score (−5 representing vastly worse, 0 representing unchanged, 5 representing completely recovered).
[c] Neck Disability Index Score (0–50 scale, 10 items, higher scores represent greater disability).

subjects were asked if they had suffered any adverse effects arising from participation. Twelve subjects (18%) in the advice group answered 'yes' to this question. The main complaint was muscle pain (4) followed by an increase in headaches (2) and ongoing pain (2). Thirteen subjects (20%) in the advice plus exercise group

Table 3
Employment status

	Time of accident		Baseline		6 weeks		12 months	
	Advice ($n = 68$)	Exs ($n = 66$)	Advice ($n = 68$)	Exs ($n = 66$)	Advice ($n = 66$)	Exs ($n = 66$)	Advice ($n = 66$)	Exs ($n = 66$)
Full time/full duties	46 (68%)	37 (56%)	30 (44%)	25 (38%)	34 (52%)	27 (41%)	31 (50%)	24 (38%)
Full time/some duties	1 (1%)	0 (0%)	7 (10%)	0 (0%)	4 (6%)	0 (0%)	2 (3%)	1 (1.5%)
Part time/full duties	8 (12%)	12 (18%)	5 (7%)	8 (12%)	7 (11%)	7 (11%)	8 (13%)	9 (14%)
Part time/some duties	2 (3%)	0 (0%)	8 (12%)	5 (8%)	7 (11%)	5 (8%)	5 (8%)	6 (9.5%)
Not working/employed	0 (0%)	1 (2%)	2 (3%)	2 (3%)	2 (3%)	0 (0%)	2 (3%)	0 (0%)
Not working/employed/retraining	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Not working/unemployed/retraining	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	2 (3%)
Not working/unemployed	11 (16%)	16 (24%)	16 (24%)	26 (39%)	12 (18%)	27 (41%)	14 (23%)	21 (33%)

Number (percentage) in each employment category at time of accident, at trial entry, at 6 week follow up and at 12 month follow up.

Table 4
Treatment credibility

	Advice ($n = 66$)	Advice plus exercise ($n = 66$)
How confident do you feel that this treatment can relieve your pain?	2.84 (2.11)	4.54 (1.35)
How confident do you feel that this treatment will help you manage your pain?	3.09 (2.12)	4.79 (1.27)
How confident would you be in recommending this treatment to a friend who suffered from similar complaints?	3.32 (2.36)	5.35 (1.14)
How logical does this type of treatment seem to you?	3.85 (2.25)	5.36 (1.03)

Mean (SD) treatment credibility ratings, each item is rated on a scale from 0 to 6. For all items unpaired t tests were statistically significant.

answered 'yes' to this question. The main complaint in this group was muscle pain with exercise (3) followed by knee pain (2) and lumbar spine pain (2).

3.7. Treatment credibility

There was a significant difference between groups on all 4 measures of treatment credibility with subjects in the exercise group perceiving treatment to be more credible than those in the advice group ($p < 0.0001$ on all 4 questions) (Table 4).

3.8. Prediction of response to treatment

Predictors of response to treatment are shown in Table 5. High levels of baseline pain intensity were associated with greater effects of exercise on the intensity and bothersomeness of pain at 6 weeks and high levels of baseline disability were associated with greater treatment effects at 12 months.

3.9. Post hoc analyses

A journal reviewer requested that we assess the impact of the treatment credibility ratings and use of additional health care on outcomes as these differed between groups. To assess the influence of additional health care use we repeated our analysis of the 6-week and 12-month outcomes adding the 'use of additional health care variable' to the treatment effectiveness regression equations (and so controlling for this in the analysis). This did not change the original results: there remained an effect of treatment group at 6 weeks, but not 12 months, for all 7 outcomes after controlling for use of additional health care. To assess the influence of treatment credibility the four individual treatment credibility items were summed to create a single treatment credibility variable that was entered into the treatment effectiveness regression equations. This substantially changed the original results: there was now no longer an effect of treatment group at 6 weeks for any of the 7 outcomes and at 12 months there was a marginally significant effect ($p = 0.033$) for one outcome.

At the request of the same reviewer we also evaluated whether duration of symptoms was a predictor of response to treatment. To do this we added the interaction term, duration of whiplash × treatment group, to the regression equation used to evaluate effectiveness of treatment. For all primary outcomes at 6 weeks and 12 months the duration of whiplash was not a predictor of response to treatment.

Table 5
Magnitude of treatment response

	Factors influencing magnitude of treatment response		
	Baseline pain	Baseline disability	Baseline fear of movement
6 weeks			
Pain intensity	−0.38 (−0.74, −0.02), $p = 0.037$	−0.26 (−0.71, 0.18), $p = 0.24$	−0.03 (−0.13, 0.06), $p = 0.51$
Pain bothersomeness	−0.50 (−0.90, −0.10), $p = 0.016$	−0.27 (−0.77, 0.25), $p = 0.31$	−0.08 (−0.19, 0.04), $p = 0.18$
Function	0.17 (−0.14, 0.49), $p = 0.28$	0.11 (−0.29, 0.51), $p = 0.58$	0.08 (−0.01, 0.17), $p = 0.07$
12 months			
Pain intensity	−0.22 (−0.63, 0.19), $p = 0.29$	−0.59 (−1.11, −0.07), $p = 0.028$	−0.04 (−0.16, 0.08), $p = 0.492$
Pain bothersomeness	−0.42 (−0.87, 0.04), $p = 0.073$	−0.46 (−1.05, 0.12), $p = 0.118$	−0.06 (−0.19, 0.07), $p = 0.360$
Function	0.29 (−0.08, 0.67), $p = 0.127$	−0.15 (−0.60, 0.30), $p = 0.518$	0.08 (−0.02, 0.19), $p = 0.108$

Pain intensity and bothersomeness were both rated as the average over the last 24 h on a 0–10 box scale. Function measured using Patient-Specific Functional Scale score (Average of 3 scores, all out of 10, 0 representing unable to perform activity, 10 representing able to perform activity at pre-injury level). Fear of movement using the Tampa score (Vlaeyen et al., 1995).
Effect of adding putative effect modifiers to regression equation. Data are regression coefficients with 95%CI.

4. Discussion

This study provides the first rigorous test of exercise and advice as treatments for people who have sustained a whiplash injury and have ongoing pain and disability that persists beyond three months. Compared with advice alone, exercise produced small reductions in pain intensity and bothersomeness, and small improvements in function, disability, quality of life and global perceived effect for people with mildly disabling chronic whiplash at 6-weeks follow-up. These effects were not apparent at 12 months. Individuals with high initial levels of pain and disability improved to a greater extent with exercise than those with low levels.

A critical issue is whether the magnitude of the effects of exercise is clinically worthwhile. Short-term effects of exercise on pain, bothersomeness of symptoms and disability at 6 week follow-up were of a similar magnitude. Farrar et al. (2001) reported that a difference of 1 point on a 0–10 pain scale was the threshold for patients to consider themselves to be at least ''minimally improved''. Similar results have been reported by Kelly (2001) and Bijur et al. (2003). Furthermore Farrar et al. (2001) estimated that an improvement of 1.7 points was the threshold for ''much improved''. In the current study, subjects in the advice group improved by 1.0 point from baseline and subjects in the exercise group improved by 2.0 points from baseline on a 0–10 pain scale. The effect of exercise was therefore equivalent to a change from ''minimally improved'' to ''much improved'' and is likely to be considered worthwhile by many patients with chronic whiplash, particularly because there are no other evidence-based treatments available for this traditionally treatment-resistant patient group.

Only two physical treatments have been investigated in randomized controlled trials for patients with chronic whiplash, exercise and cognitive behavioral therapy. Prior to this study, there were only one randomized controlled trial that specifically investigated the use of exercise in patients with chronic whiplash (Fitz-Ritson, 1995). However the findings of that study should be interpreted with caution because of the poor methods (non-blinded, lack of baseline comparability, no intention-to-treat analysis) and the lack of description of the exercise. The authors concluded that 'phasic' exercises in conjunction with chiropractic treatment were more effective than 'rehabilitation' exercises and chiropractic treatment. A further randomized controlled trial investigated a cognitive behavioral therapy approach in the treatment of chronic whiplash (Soderlund and Lindberg, 2001). In that study, Soderlund and Lindberg (2001) compared cognitive behavioral therapy and standard physiotherapy program with standard physical therapy alone and found that there was no statistically or clinically significant differences in primary pain or disability outcomes between the treatment groups (Soderlund and Lindberg, 2001). Unfortunately only 33 subjects participated in this study and while the authors noted trends in favor of the cognitive behavioral therapy group, the study lacked the statistical power necessary to demonstrate a clinically significant difference between the two groups. It is also worth noting that both groups received supervised active treatments in the physiotherapy program involving neck and shoulder strength and stability exercises, range of motion, stretching and muscular coordination exercises.

The only therapy for chronic whiplash currently known to be effective appears to be radiofrequency neurotomy (Lord et al., 1996) however this treatment is technically difficult to administer proficiently and is not appropriate for all chronic whiplash sufferers. In light of the lack of evidence-based therapy the results of the current study should be viewed with some optimism. The challenge is how to adjust the dose of the exercise program in order to increase both the size of the treatment effect and the durability of the effect. Changing the intensity and/or duration of the initial exercise program or adding follow-up reminder phone calls and/or face-to-face consultations are options we are currently considering.

A secondary aim of this trial was to determine if pain severity, level of disability or fear of movement predicted response to the physical activity program. There was a statistically significant effect of both initial pain and disability on treatment outcomes. Subjects with higher initial levels of pain and disability experienced a greater treatment effect than subjects with lower levels. This is an important finding for patients, clinicians and insurers alike. Chronic whiplash patients with high levels of pain and disability potentially have a worse prognosis than those with lower levels of pain (Cote et al., 2001; Scholten-Peeters et al., 2003). Subjects in the exercise group clearly thought the treatment they received was more credible and logical than did patients in the advice group (Table 4). This finding may be due either to the type or amount of treatment received. This finding is similar to that of Bronfort et al. (2001), who found that subjects were more satisfied with spinal manipulative therapy supplemented by exercise than with spinal manipulative therapy alone. Similarly Taimela et al. (2000) found active intervention was more effective in terms of self-reported benefits than education about neck pain and advice to exercise. These findings might suggest that patients who take an ''active'' role in treatment may be more satisfied than subjects who remain ''passive'', or alternately that subjects who receive a greater amount of treatment may be more satisfied with that treatment.

A further indicator that subjects in the exercise group were more satisfied with treatment was the difference in numbers of subjects in each group seeking additional treatment during the course of the trial. Subjects in the

advice group were more likely to seek other treatment during the treatment period and the subsequent 12 months. This indicates that simple advice may not be as economical as it would first seem. The number of subjects in the advice group who sought additional treatment indicates that either these individuals did not feel that the treatment was helping them adequately or they did not feel that they were receiving sufficient treatment. The cost of the additional treatment (which was extensive in some cases) should be considered before recommending chronic whiplash be treated with advice alone.

The first of the post hoc analyses requested by the reviewer demonstrated that after controlling for the use of other treatments, exercise and advice was still not more effective than advice alone at 12 months. It is perhaps not surprising that controlling for the use of other treatments did not affect the results because there are no known effective treatments for whiplash beyond radiofrequency neurotomy. If this post hoc analysis did substantially change the effectiveness analysis it would be of interest to researchers attempting to understand the results of this trial but of little or no relevance to clinicians wishing to apply these treatments in clinical practice. The reason for this is that it is not possible or ethical to prevent patients seeking other treatments.

The second post hoc analysis revealed that after controlling for treatment credibility in the treatment effectiveness analyses there was no longer an effect of treatment group at 6 weeks. This result does not change our conclusion that combined exercise and advice treatment is more effective than advice alone however it does offer an insight into why combined exercise and advice is more effective. While it is tempting to view this result as evidence that patients' perceptions of treatment credibility influence outcome it is also possible to argue that outcome influences patients' perceptions of treatment credibility. The latter position is tenable because treatment credibility ratings were taken at the end of the 6 week treatment program. Further research is required to clarify the direction of this relationship. We would also caution that the treatment credibility result has limited application to clinicians attempting to decide on appropriate individualized treatment options for their patients. The first limitation is that in our trial treatment credibility ratings were taken at the end of the 6 week treatment program and so unavailable to the clinician at the time of selecting therapy. Secondly, we analyzed predictors of outcome, and this is a separate issue to predicting response to treatment. It is the latter which is important when selecting treatment.

The last post hoc analysis requested by the reviewer examined whether subjects with a greater duration of symptoms were less likely to respond to exercise treatment than those with a shorter duration. The analysis demonstrated that the duration of whiplash symptoms was not a predictor of response to exercise treatment. Because we restricted enrolment to subjects with duration of symptoms of greater than 3 months and less than 12 months it is not possible to exclude the possibility that people with whiplash symptoms that have persisted beyond 12 months may be less likely to respond to exercise treatment.

5. Conclusions

The results of this randomized controlled trial indicate that exercise and advice produces better outcomes than advice alone for people who have sustained a whiplash injury and have ongoing pain and disability that persist beyond three months. The effect of exercise is greater in people with high levels of pain and disability than those with low levels. The superior results with exercise and advice compared to advice alone are however small and only apparent in the short term. The small, short-term effect of exercise should be interpreted in light of the fact that there are very few treatments known to be effective for this treatment-resistant patient group. Whether the superior short-term outcomes in the exercise and advice group, compared to the advice alone group, are justifiable in terms of treatment and labor costs is a decision for individual patients, medical practitioners and insurers.

Acknowledgements

This trial was funded by the NSW Motor Accidents Authority. We thank staff at RehabOne Physiotherapy for providing the treatment in the trial. A/Prof Chris Maher's research fellowship is funded by the National Health and Medical Research Council of Australia.

References

ACSM. ACSM's Guidelines for exercise testing and prescription. 5th ed. Baltimore: Williams and Wilkins; 1995.

Bijur P, Latimer C, Gallagher E. Validation of a verbally administered numerical rating scale of acute pain for use in the emergency department. Acad Emerg Med 2003;10:390–2.

Bronfort G, Evans R, Nelson B, Aker P, Goldsmith C, Vernon H. A randomized clinical trial of exercise and chiropractic spinal manipulation for patients with chronic neck pain. Spine 2001;26:788–97.

Cost B13 working group. European Guidelines for the Management of Chronic Non-Specific Low Back Pain, Cost B13 working group on guidelines for chronic low back pain; 2004.

Cote P, Cassidy D, Carroll L, Frank J, Bombardier C. A systematic review of the prognosis of acute whiplash and a new conceptual framework to synthesize the literature. Spine 2001;26:E445–58.

Farrar J, Young J, LaMoreaux L, Werth J, Poole M. Clinical importance of changes in chronic pain intensity measured on a 11-point numerical pain rating scale. Pain 2001;94:149–58.

Fitz-Ritson D. Phasic exercises for cervical rehabilitation after "whiplash" trauma. J Manipulative Physiol Ther 1995;18:21–4.

Hayden JA, van Tulder MW, Tomlinson G. Systematic review: strategies for using exercise therapy to improve outcomes in chronic low back pain. Ann Intern Med 2005;142:776–85.

Kelly A. The minimum clinically significant difference in visual analogue scale pain score does not differ with severity of pain. Emerg Med J 2001;18:205–7.

Lord SM, Barnsley L, Wallis BJ, McDonald GJ, Bogduk N. Percutaneous radio-frequency neurotomy for chronic cervical zygapophyseal-joint pain. N Engl J Med 1996;335:1721–6.

Motor Accidents Authority of New South Wales. Whiplash and the NSW motor accidents scheme. Statistical Information Paper no 7. Motor Accidents Authority of NSW, Sydney, 2001.

Nicholas MK, Tonkin L. Persisting pain: using cognitive-behavioural principles for activity-based pain management. Chapter 8. In: Refshauge K, Gass E, editors. Musculoskeletal physiotherapy. Clinical science and evidence-based practice. Oxford: Butterworth Heinemann; 2005.

Scholten-Peeters G, Verhagen A, Bekkering G, van der Windt A, Barnsley L, Oostendorp R, Hendriks E. Prognostic factors of whiplash associated disorders: a systematic review of prospective cohort studies. Pain 2003;104:303–22.

Soderlund A, Lindberg P. Cognitive behavioural components in physiotherapy management of chronic whiplash associates disorders (WAD) – a randomized group study. Physiother Theory Pract 2001;17:229–38.

Spitzer W, Skovron M, Salmi LR, Cassidy J, Duranceau J, Suissa S, Zeiss E. Scientific monograph of the Quebec Task Force on Whiplash-Associated Disorders: Redefining 'whiplash'and its management. Spine 1995;20:1S–73S.

Sterling M, Jull G, Vicenzino B, Kenardy J, Darnell R. Physical and psychological factors predict outcome following whiplash injury. Pain 2005;114:141–8.

Stewart M. Treatment of chronic whiplash. PhD Thesis University of Sydney, Australia; 2005.

Stewart M, Maher C, Refshauge K, Herbert R, Bogduk N, Nicholas M. Advice or exercise for chronic whiplash disorders? Design of a randomized controlled trial. BMC Musculoskelet Disord 2003;4:18.

Taimela S, Takala E, Asklof T, Seppala K, Parviainen S. Active treatment of chronic neck pain: a prospective randomised intervention. Spine 2000;25:1021–7.

Rebbeck T, Sindhusake D, Cameron I, et al. A prospective cohort study of health outcomes following whiplash associated disorders in an Australian population. Injury Prevent 2006;12:86–93.

Vlaeyen J, Kole-Snijders A, Boersen R, van Eek H. Fear of movement/(re)injury in chronic low back pain and its relation to behavioural performance. Pain 1995;62:363–72.