Cite this article as: BMJ, doi:10.1136/bmj.38282.669225.AE (published 29 November 2004)

# Primary care

# United Kingdom back pain exercise and manipulation (UK BEAM) randomised trial: effectiveness of physical treatments for back pain in primary care

UK BEAM Trial Team


## Abstract

**Objective** To estimate the effect of adding exercise classes, spinal manipulation delivered in NHS or private premises, or manipulation followed by exercise to "best care" in general practice for patients consulting with back pain.
**Design** Pragmatic randomised trial with factorial design.
**Setting** 181 general practices in Medical Research Council General Practice Research Framework; 63 community settings around 14 centres across the United Kingdom.
**Participants** 1334 patients consulting their general practices about low back pain.
**Main outcome measures** Scores on the Roland Morris disability questionnaire at three and 12 months, adjusted for centre and baseline scores.
**Results** All groups improved over time. Exercise improved mean disability questionnaire scores at three months by 1.4 (95% confidence interval 0.6 to 2.1) more than "best care." For manipulation the additional improvement was 1.6 (0.8 to 2.3) at three months and 1.0 (0.2 to 1.8) at 12 months. For manipulation followed by exercise the additional improvement was 1.9 (1.2 to 2.6) at three months and 1.3 (0.5 to 2.1) at 12 months. No significant differences in outcome occurred between manipulation in NHS premises and in private premises. No serious adverse events occurred.
**Conclusions** Relative to "best care" in general practice, manipulation followed by exercise achieved a moderate benefit at three months and a small benefit at 12 months; spinal manipulation achieved a small to moderate benefit at three months and a small benefit at 12 months; and exercise achieved a small benefit at three months but not 12 months.


## Introduction

Back pain is a common and costly problem.[1] The role of different physical treatments is not clear. Evidence suggests that encouraging patients to keep active is effective,[2] but evidence for the effectiveness of spinal manipulation is conflicting.[3 4] Although specific exercises seem to be ineffective,[3] weak evidence exists for general programmes that encourage physical activity as a treatment for back pain.[2 5]

This trial compared a class based general exercise programme and a spinal manipulation package with "best care" in general practice, based on "active management." A previous UK trial reported that treatment by private chiropractors was superior to routine outpatient care,[6] but the trial received criticism for not considering the potentially biasing effect of treatment location.[7] Therefore, we also compared the effect of manipulation delivered in private premises with that of manipulation in premises owned by the NHS.

Our main aim was to estimate, for patients consulting their general practitioner with back pain, the effectiveness of adding the following to best care in general practice[8]: a class based exercise programme ("back to fitness"),[9] a package of treatment by a spinal manipulator (chiropractor, osteopath, or physiotherapist),[10] or manipulation followed by exercise. We also aimed to test whether the manipulation package was more or less effective in manipulators' private premises than in NHS premises.

## Methods

### Protocol

*Study design*

We randomised participants between spinal manipulation delivered in NHS premises, the same in private premises, and "best care" in general practice. We also randomised them between the exercise programme and best care. Of six groups of participants, one received only best care in general practice. The other five received best care plus an intervention—exercise, manipulation in private or NHS premises, or manipulation in private or NHS premises followed by exercise (fig 1). Statistically this is a three by two factorial design.[11]

We selected 14 centres, including two for the feasibility study. All centres had general practices from the Medical Research Council (MRC) General Practice Research Framework (mrc-gprf.ac.uk), with a total of at least 40 000 registered patients within travelling distance of treatment locations for manipulation and exercise; two manipulators (chiropractors, osteopaths, or physiotherapists) with private premises, willing to work in NHS premises, with colleagues to cover absences; a physiotherapist to deliver the exercise programme, with a colleague to cover absences; NHS premises in the community suitable for spinal manipulation; and premises in the community suitable for exercise classes.

*Recruitment of participants*

In participating practices, research nurses identified patients consulting with back pain, both directly from general practitioners and their staff and by searching computerised records. They assessed potential participants' eligibility and interest by brief

*Authorship details and a list of collaborators are on bmj.com*

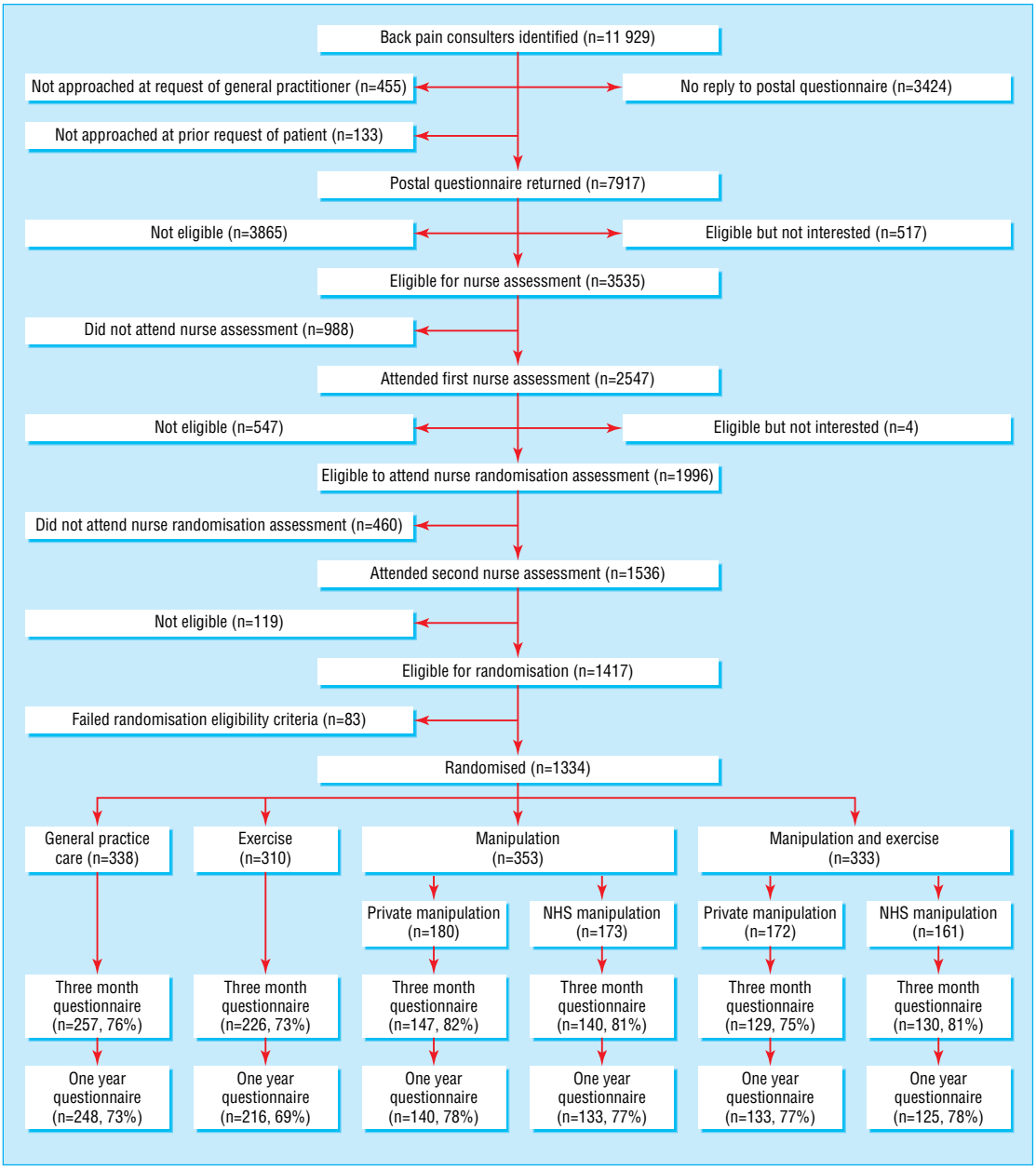


**Fig 1** Progress of the UK BEAM trial

postal questionnaires. They saw interested patients on two occasions. The first was to explain the trial and assess eligibility (boxes). The second was to confirm eligibility and general practitioners' consent, seek participants' consent, collect baseline data, and randomise participants. To exclude patients whose back pain resolved rapidly, randomisation occurred at least four weeks after the initiating consultation.[12]

*"Best care" in general practice— the "comparator" treatment*

Like other evidence based guidelines,[13] the UK national acute back pain guidelines advise continuing normal activities and avoiding rest.[2] To maximise recruitment and base the comparator treatment on these guidelines,[11] we invited clinical and support staff from all participating practices to training sessions on the "active management" of back pain.[8] We also provided copies of *The Back Book*,[14] the corresponding patient booklet, for practice reception areas and for patients with back pain.

*Interventions*

We defined "basic minimum treatment" as initial assessment plus one class for exercise and as two sessions, including assessment, for manipulation.

*Exercise programme*—We developed the exercise programme ("back to fitness"[9]) from previous trials.[15 16] It comprises initial individual assessment followed by group classes incorporating cognitive behavioural principles. We trained physiotherapists with at least two years' experience since qualification to deliver this programme. Classes ran in local community facilities. Up to 10 people took part in each session. We invited participants to attend up to eight 60 minute sessions over four to eight weeks and a "refresher" class 12 weeks after randomisation.

**Exclusion criteria**

Patients were not eligible if:

- They were aged 65 or over, because the spinal manipulation package could be more hazardous in older people with osteoporosis
- There was a possibility of serious spinal disorder, including malignancy, osteoporosis, ankylosing spondylitis, cauda equina compression, and infection
- They complained mainly of pain below the knee, as clinical outcome was likely to be different
- They had previously had spinal surgery, as clinical outcome was likely to be very different
- They had another musculoskeletal disorder that was more troublesome than their back pain
- They had previously attended, or been referred to, a specialised pain management clinic
- They had a severe psychiatric or psychological disorder
- They had another medical condition, such as cardiovascular disease, that could interfere with therapy
- They had moderate to severe hypertension (systolic blood pressure $>180$ mm Hg or diastolic blood pressure $>105$ mm Hg, on at least two separate occasions)
- They were taking anticoagulant treatment
- They were taking long term steroids, which might lead to osteoporosis
- They could not walk 100 m when free of back pain, because exercise would be difficult
- They could not get up from and down to the floor unaided
- They had received physical therapy (including acupuncture) in the previous three months
- They had a Roland disability questionnaire score of three or less on the day of randomisation
- They could not read and write fluently in English

*Spinal manipulation package*—A multidisciplinary group developed a package of techniques representative of those used by the UK chiropractic, osteopathic, and physiotherapy professions.[10] The three professional associations agreed to the use of this package in this trial. Similar numbers of qualified manipulators from each of these professions treated participants. They all had a minimum of two years' clinical experience and were skilled in a range of manipulative techniques, including high velocity thrusts.[10] Participants randomised to private manipulation received treatment in manipulators' own consultation rooms. Those randomised to NHS manipulation saw the same manipulators in NHS premises. Following initial assessment, manipulators chose from the agreed manual and non-manual treatment options. They agreed to do high velocity thrusts on most patients at least once. We invited participants to attend up to eight 20 minute sessions, if necessary, over 12 weeks.

**Inclusion criteria**

Patients were eligible if:

- Their ages were between 18 and 65 years
- They were registered for medical care with a participating practice
- They had consulted with simple low back pain—pain of musculoskeletal origin in the area bounded by the lowest palpable ribs, the gluteal folds, and the posterior axillary lines, including pain referred into the legs provided it was mainly above the knee
- They had a score of four or more on the Roland disability questionnaire at randomisation
- They had experienced pain every day for the 28 days before randomisation or for 21 out of the 28 days before randomisation and 21 out of the 28 days before that
- They agreed to avoid physical treatments, other than trial treatments, for three months

*Combined treatment*—We invited participants to attend eight sessions of manipulation over six weeks, eight sessions of exercise in the next six weeks, and a refresher class at 12 weeks. Other aspects of treatment were identical to those in the manipulation only or exercise only groups.

*Follow up and monitoring adverse events*

Participants completed questionnaires on general health, back pain, beliefs, and psychological wellbeing before randomisation and at one, three, and 12 months thereafter. The research team monitored serious adverse events, defined as treatment related events leading to hospital admission or death within one week.

*Patient assessed outcomes*

Questionnaires (table 1) included two back specific instruments—the Roland Morris disability questionnaire and the modified Von Korff scales,[17 18] two measures of belief—the back beliefs questionnaire and the fear avoidance beliefs questionnaire,[19 20] and two generic measures—the SF-36 and the EuroQol,[21 22] reported in the accompanying economic paper. [23].

*Assignment*

After consenting participants had completed baseline assessments, nurses contacted the remote randomisation service. This stratified participants by practice and allocated them between the six groups in figure 1 by randomly permuted blocks. In the feasibility study participants were allocated between the six groups in equal proportions. In the main trial one quarter of participants were allocated to each of four groups—"best care" in general practice, best care plus exercise, best care plus manipulation, and best care plus manipulation and exercise; participants due to receive manipulation were then subdivided equally between NHS and private settings (fig 1). Nurses informed participants of their allocations immediately. Alerted by letter, manipulators and exercise physiotherapists contacted participants to arrange their first appointments.

*Sample size*

Many experts regard a difference of 2.5 in Roland disability questionnaire scores (which have a population standard deviation of about 4 points[11]) as being clinically important in back pain trials.[24] This translates into a difference of 1.67 between manipulation in NHS and private premises.[11] To yield 80% power of detecting this difference between private and NHS treatment at a 1% significance level requires data on about 130 participants per group. To allow for clustering effects by manipulator, we sought data from 450 participants randomised to manipulation. Allowing for 67% follow up at 12 months, we sought to recruit 1350 participants in all.[11] To minimise cost, we counted participants from practices in the feasibility study who followed the main protocol by training in active management.[11]

*Masking*

As UK BEAM was a pragmatic trial to estimate the effectiveness of manipulation and exercise in routine clinical practice, blinding of participants and professionals was neither desirable nor possible.

**Table 1** Patient characteristics at randomisation by group. Values are numbers* (percentages) for binary characteristics and mean (SD) for quantitative characteristics

| Characteristics | Best care in general practice | Best care plus exercise alone | Best care plus private manipulation alone | Best care plus NHS manipulation alone | Best care plus private manipulation plus exercise | Best care plus NHS manipulation plus exercise | Total |
|---|---|---|---|---|---|---|---|
| **Binary** | | | | | | | |
| Female sex | 178/338 (53) | 170/310 (55) | 114/180 (63) | 98/173 (57) | 100/172 (58) | 89/161 (55) | 749/1334 (56.1) |
| Ethnic status white | 307/322 (95) | 287/300 (96) | 166/172 (97) | 154/167 (92) | 158/162 (98) | 153/157 (97) | 1225/1280 (95.7) |
| Left full time education aged ≤16 | 188/329 (57) | 176/305 (58) | 101/178 (57) | 97/168 (58) | 94/169 (56) | 92/159 (58) | 748/1308 (57.2) |
| Not doing work or normal activities because of poor health | 22/327 ( 7) | 35/298 (12) | 12/176 ( 7) | 14/167 ( 8) | 11/166 ( 7) | 16/154 (10) | 110/1288 (8.5) |
| Off work in past four weeks because of back or leg pain | 76/277 (27) | 75/255 (29) | 42/143 (29) | 45/148 (30) | 42/145 (29) | 43/131 (33) | 323/1099 (29.4) |
| Current episode of back pain has lasted more than 90 days | 190/323 (59) | 173/301 (57) | 93/172 (54) | 109/166 (66) | 94/158 (59) | 90/155 (58) | 749/1275 (58.7) |
| **Quantitative (maximum n=1334)** | | | | | | | |
| Age (years) | 42.5 (10.6) | 44.0 (11.0) | 42.5 (10.6) | 43.2 (12.1) | 42.7 (12.0) | 43.5 (11.7) | 43.1 (11.2) |
| Roland disability questionnaire (0-24, 0=best) | 9.0 ( 3.9) | 9.2 (4.3) | 8.9 (4.0) | 8.9 (4.0) | 8.9 (3.6) | 9.1 (4.3) | 9.0 (4.0) |
| Modified Von Korff scale: | | | | | | | |
| Disability (0-100, 0=best) | 44.9 (21.0) | 47.7 (22.6) | 46.9 (22.0) | 46.6 (22.7) | 45.4 (21.7) | 44.5 (22.2) | 46.1 (22.0) |
| Pain (0-100, 0=best) | 60.5 (17.6) | 60.8 (17.6) | 61.4 (19.0) | 61.6 (19.0) | 59.5 (17.4) | 60.6 (18.5) | 60.7 (18.0) |
| "Troublesome" back pain (1-5, 1=best) | 3.3 (0.8) | 3.5 (0.8) | 3.5 (0.9) | 3.5 (0.9) | 3.4 (0.8) | 3.4 (0.8) | 3.4 (0.8) |
| "Troublesome" leg pain (1-5, 1=best) | 2.1 (1.1) | 2.1 (1.0) | 2.1 (1.1) | 2.1 (1.0) | 2.0 (1.0) | 2.0 (1.0) | 2.1 (1.0) |
| Change in back pain in past 4 weeks (1-7, 1=best, 4=no change) | 4.0 (0.8) | 4.1 (0.9) | 4.1 (0.9) | 4.0 (1.0) | 4.0 (0.8) | 4.1(0.9) | 4.0 (0.9) |
| Back beliefs questionnaire (9-45, 45=best) | 28.4 (7.1) | 27.2 (7.1) | 28.4 (7.4) | 27.9 (7.0) | 28.7 (6.7) | 27.4 (7.4) | 28.0 (7.1) |
| Fear avoidance beliefs questionnaire—physical scale (0-24, 0=best) | 14.9 (5.1) | 15.0 (5.3) | 14.7 (5.4) | 15.0 (5.1) | 14.9 (4.8) | 15.0 (5.7) | 14.9 (5.2) |
| SF-36 (mean=50, SD=10, 100=best): | | | | | | | |
| Physical component score | 41.0 (6.4) | 40.5 (6.7) | 41.1 (6.4) | 40.8 (6.6) | 40.9 (6.6) | 40.6 (7.2) | 40.8 (6.6) |
| Mental component score | 46.6 (10.4) | 45.4 (10.8) | 45.0 (10.0) | 45.8 (9.7) | 45.8 (10.1) | 45.7 (10.0) | 45.8 (10.3) |

*Denominator varies according to number of valid responses.

### Analysis

We used two sided significance tests to analyse the primary outcome—Roland disability questionnaire score after three or 12 months—by intention to treat. We used analysis of covariance to adjust this score for centre and baseline score. We analysed the data in steps. Firstly, we used multilevel modelling to make allowance for the innate clustering of participants by centre, exercise class, manipulator, and practice. Secondly, we tested the effect of exercise without manipulation by comparing participants allocated to best care with those allocated to best care plus exercise (table 2). Thirdly, we tested the effect of manipulation without exercise by comparing participants allocated to best care with those allocated to best care plus manipulation (table 3). If this was significant, we tested for differences between manipulation in NHS and private premises. Finally, if either exercise or manipulation gave significant results, we tested for interactions between exercise and manipulation—that is, whether the estimated improvement in participants allocated to best care, manipulation, and exercise (table 4) differed significantly from the sum of the estimated improvement due to manipulation (table 3) and that due to exercise (table 2).

Of the 14 effects of combined treatment estimated in table 4, most were larger but not much larger than the corresponding estimates for exercise only (table 2) or manipulation only (table 3). Indeed, only that relating to back beliefs at 12 months was larger than the sum of the corresponding estimates for the individual treatments in tables 2 and 3. So participants who have already received manipulation apparently get less benefit from exercise than do those who have not received manipulation. Rather than ascribe a significance level to these 14 correlated findings, we report that at three months the interaction between manipulation and exercise in their effect on disability scores was just significant at the 5% level; the standardised regression coefficient was 1.0 (95% confidence interval 0.0 to 2.1). At 12 months, however, the corresponding interaction was not significant; the standardised regression coefficient was 0.1 ( – 1.0 to 1.2). Nevertheless, to avoid underestimating the effect of treatment, we estimated that of exercise as in the second step and that of manipulation as in the third step.

Because correlation within clusters proved smaller than projected,[11] multilevel modelling generated estimates and confidence intervals very similar to those generated by the simpler analysis of covariance (tables 2 and 3). We present the simpler estimates here.

## Results

### Participant flow and follow up

We recruited 1334 participants from 181 general practices around 14 centres across the United Kingdom (fig 1). These practices were broadly typical of UK practices in size and deprivation. The feasibility study recruited 164 participants between

**Table 2** Changes in outcome attributable to exercise.† Values are mean (SE) scores unless stated otherwise

| Outcome measure (description) | At three months | | | At 12 months | | |
|---|---|---|---|---|---|---|
| | Best care (max n=256/338; 76%) | Best care plus exercise (max n=225/310; 73%) | Net benefit from exercise (95% CI) | Best care (max n=248/338; 73%) | Best care plus exercise (max n=216/310; 70%) | Net benefit from exercise (95% CI) |
| Roland disability questionnaire (0-24, 0 =best) | 6.83 (0.28) (n=256) | 5.47 (0.29) (n=225) | 1.36** (0.63 to 2.10) | 6.13 (0.30) (n=248) | 5.74 (0.31) (n=216) | 0.39 (–0.41 to 1.19)‡ |
| Modified Von Korff scale: | | | | | | |
| Disability (0-100, 0=best) | 34.73 (1.53) (n=239) | 29.73 (1.64) (n=205) | 5.03* (1.02 to 9.05) | 34.29 (1.57) (n=235) | 29.73 (1.68) (n=202) | 4.56* (0.34 to 8.78) |
| Pain (0-100, 0=best) | 49.32 (1.58) (n=239) | 44.73 (1.71) (n=204) | 4.59* (0.43 to 8.75) | 48.44 (1.71) (n=235) | 41.54 (1.84) (n=200) | 4.90* (0.30 to 9.50) |
| Back beliefs questionnaire (9-45, 45=best) | 27.62 (0.41) (n=239) | 29.79 (0.44) (n=201) | 2.16*** (1.09 to 3.24) | 27.09 (0.42) (n=233) | 28.54 (0.45) (n=197) | 1.46* (0.33 to 2.58) |
| Fear avoidance beliefs questionnaire—physical scale (0-24, 0=best) | 13.23 (0.40) (n=236) | 10.78 (0.43) (n=200) | 2.46*** (1.41 to 3.50) | 12.72 (0.42) (n=208) | 11.64 (0.46) (n=187) | 1.08 (–0.05 to 2.22) |
| SF-36 (mean=50, SD=10, 100=best): | | | | | | |
| Physical component score | 43.94 (0.44) (n=227) | 46.35 (0.48) (n=191) | 2.41*** (1.23 to 3.58) | 42.84 (0.59) (n=221) | 44.39 (0.63) (n=194) | 1.55 (–0.02 to 3.11) |
| Mental component score | 46.49 (0.62) (n=227) | 47.24 (0.67) (n=191) | 0.74 (–0.90 to 2.39) | 46.44 (0.77) (n=221) | 46.77 (0.81) (n=194) | 0.34 (–1.69 to 2.37) |

†Estimated by analysis of covariance with adjustment for centre and baseline score. Because correlation coefficients within clusters (centres, exercise classes, manipulators, and practices) are small, this analysis generates estimates similar to the corresponding, but more complex, multilevel model.
‡For example the multilevel model estimates the effect of exercise on the Roland disability questionnaire at 12 months as 0.284 (95% CI –0.38 to 0.94).
*Significant at 5% level.
**Significant at 1% level.
***Significant at 0.1% level.

March 1998 and April 1999.[11] The main trial recruited 1170 participants between August 1999 and April 2001. These participants attended exercise classes in 18 community settings and received manipulation in 45 premises, 27 private and 18 owned by the NHS. At three months, 1029 (77%) returned questionnaires; at 12 months, 995 (75%) returned questionnaires. Responders were much more likely than non-responders to be female, above average age, and educated beyond age 16 and to have had severe back pain at randomisation. As these trends were consistent across randomised groups, however, little risk of bias exists.

### Baseline data

The mean (SD) age of participants at randomisation was 43 (11) years; 56% were female, and 9% were not working because of poor health. More than half had had pain for more than 90 days. Mean (SD) Roland disability score at randomisation was 9.0. The six randomised groups had similar characteristics (table 1).

### Process

The message about active management reached most participants: when asked at randomisation, 1160 (87%) recalled seeing *The Back Book*.[14] Of 686 participants allocated to manipulation, 633 (92%) received "basic minimum treatment." Of 643 participants allocated to exercise, 408 (63%) received basic minimum treatment. No serious adverse events occurred.

### Analysis

Roland disability questionnaire scores improved by a mean (SD) of 3.3 (4.5) points at three months and 3.5 (4.7) points at 12 months. Figure 2 shows progress in disability scores following randomisation between the four basic interventions.

#### *Exercise programme*

Exercise produced statistically significant improvements in mean Roland disability score at three months only (difference = 1.4; 95% confidence interval 0.6 to 2.1), in mean Von Korff disability and pain scores and back beliefs score at both three and 12 months, and in mean SF-36 physical score and fear avoidance beliefs physical score at three months only (table 2). Mean SF-36 mental score did not differ.

**Table 3** Changes in outcome attributable to manipulation.† Values are mean (SE) scores unless stated otherwise

| Outcome measure (description) | At three months | | | At 12 months | | |
|---|---|---|---|---|---|---|
| | Best care (max n=256/338; 76%) | Best care plus manipulation (max n=287/353; 81%) | Net benefit from manipulation (95% CI) | Excluding manipulation (max n=248/338; 73%) | Best care plus manipulation (max n=273/353; 77%) | Net benefit from manipulation (95% CI) |
| Roland disability questionnaire (0-24, 0=best) | 6.66 (0.30) (n=256) | 5.09 (0.28) (n=287) | 1.57*** (0.82 to 2.32) | 6.16 (0.31) (n=248) | 5.15 (0.29) (n=273) | 1.01* (0.22 to 1.81)‡ |
| Modified Von Korff scale: | | | | | | |
| Disability (0-100, 0=best) | 35.11 (1.61) (n=239) | 31.14 (1.48) (n=275) | 3.97 (–0.050 to 7.98) | 35.50 (1.60) (n=235) | 29.85 (1.50) (n=262) | 5.65** (1.57 to 9.72) |
| Pain (0-100, 0=best) | 49.59 (1.62) (n=239) | 40.90 (1.50) (n=275) | 8.90*** (4.84 to 12.95) | 47.56 (1.69) (n=235) | 41.68 (1.58) (n=264) | 5.87** (1.58 to 10.17) |
| Back beliefs questionnaire (9-45, 45=best) | 28.32 (0.42) (n=239) | 31.27 (0.38) (n=274) | 2.94*** (1.91 to 3.98) | 27.78 (0.43) (n=233) | 29.22 (0.41) (n=264) | 1.43* (0.33 to 2.54) |
| Fear avoidance beliefs questionnaire—physical scale (0-24, 0=best) | 12.86 (0.37) (n=236) | 11.95 (0.35) (n=271) | 0.91 (–0.03 to 1.84) | 12.78 (0.39) (n=208) | 12.88 (0.37) (n=260) | –0.10 (–1.09 to 0.89) |
| SF-36 (mean=50, SD=10, 100=best): | | | | | | |
| Physical component score | 44.04 (0.48) (n=227) | 46.56 (0.45) (n=259) | 2.52*** (1.30 to 3.74) | 42.50 (0.60) (n=221) | 44.18 (0.55) (n=252) | 1.68* (0.18 to 3.19) |
| Mental component score | 46.77 (0.60) (n=227) | 49.64 (0.56) (n=259) | 2.88*** (1.36 to 4.40) | 46.41 (0.75) (n=221) | 48.09 (0.69) (n=252) | 1.68 (–0.21 to 3.57) |

†Estimated by analysis of covariance with adjustment for centre and baseline score. Because correlation coefficients within clusters (centres, exercise classes, manipulators, and practices) are small, this analysis generates estimates very similar to the corresponding, but more complex, multilevel model.
‡For example, the multilevel model estimates the effect of manipulation on the Roland disability questionnaire at 12 months as 0.95 (95% CI 0.33 to 1.56).
*Significant at 5% level.
**Significant at 1% level.
***Significant at 0.1% level.

**Table 4** Changes in outcome attributable to manipulation followed by exercise.† Values are mean (SE) scores unless stated otherwise

| Outcome measure (description) | At three months | | | At 12 months | | |
|---|---|---|---|---|---|---|
| | Best care (max n=256/338: 76%) | Best care plus manipulation plus exercise (max n=258/333: 77%) | Net benefit from manipulation plus exercise (95% CI) | Best care (max n=248/338: 73%) | Best care plus manipulation plus exercise (max n=257/333: 77%) | Net benefit from manipulation plus exercise (95% CI) |
| Roland disability questionnaire (0-24, 0=best) | 6.71 (0.28) (n=256) | 4.84 (0.28) (n=258) | 1.87*** (1.15 to 2.60) | 6.02 (0.30) (n=248) | 4.72 (0.29) (n=257) | 1.30** (0.54 to 2.07) |
| Modified Von Korff scale: | | | | | | |
| Disability (0-100, 0=best) | 34.56 (1.50) (n=239) | 29.05 (1.49) (n=246) | 5.51** (1.75 to 9.28) | 34.80 (1.60) (n=235) | 28.09 (1.59) (n=246) | 6.71** (2.62 to 10.80) |
| Pain (0-100, 0=best) | 48.96 (1.60) (n=239) | 40.76 (1.59) (n=246) | 8.21*** (4.20 to 12.21) | 46.39 (1.66) (n=235) | 39.68 (1.65) (n=245) | 6.71** (2.47 to 10.95) |
| Back beliefs questionnaire (9-45, 45=best) | 27.97 (0.42) (n=239) | 31.25 (0.42) (n=245) | 3.28*** (2.23 to 4.33) | 27.61 (0.44) (n=233) | 30.56 (0.43) (n=244) | 2.96*** (1.84 to 4.07) |
| Fear avoidance beliefs questionnaire—physical scale (0-24, 0=best) | 13.08 (0.40) (n=236) | 10.68 (0.39) (n=245) | 2.40*** (1.41 to 3.39) | 12.81 (0.45) (n=208) | 11.58 (0.48) (n=204) | 1.24* (0.07 to 2.41) |
| SF-36 (mean=50, SD=10, 100=best): | | | | | | |
| Physical component score | 43.91 (0.48) (n=227) | 46.46 (0.48) (n=231) | 2.55*** (1.34 to 3.75) | 42.58 (0.62) (n=221) | 45.11 (0.64) (n=221) | 2.53** (0.96 to 4.09) |
| Mental component score | 46.59 (0.58) (n=227) | 48.89 (0.59) (n=231) | 2.30** (0.82 to 3.78) | 46.71 (0.73) (n=221) | 48.01 (0.75) (n=221) | 1.30 (–0.55 to 3.14) |

†Estimated by analysis of covariance adjusting for centre and baseline score. Because correlation coefficients within clusters (centres, exercise classes, manipulators, and practices) are small, this analysis generates estimates very similar to the corresponding, but more complex, multilevel model.
*Significant at 5% level.
**Significant at 1% level.
***Significant at 0.1% level.

*Manipulation package*

Manipulation produced statistically significant improvements in Roland disability scores at three months (1.6; 0.8 to 2.3) and at one year (1.0; 0.2 to 1.8); in mean Von Korff pain score, back beliefs score, and SF-36 physical score at both three and 12 months; in mean Von Korff disability score at 12 months only; and in mean SF-36 mental score at three months only (table 3). Mean fear avoidance beliefs physical score did not differ.

We found no significant differences between the outcome of manipulation delivered in NHS or private premises. The adjusted difference in disability scores was 0.2 (–0.6 to 0.9) in favour of private premises at three months and 0.1 (–0.7 to 0.9) in favour of NHS premises at 12 months.

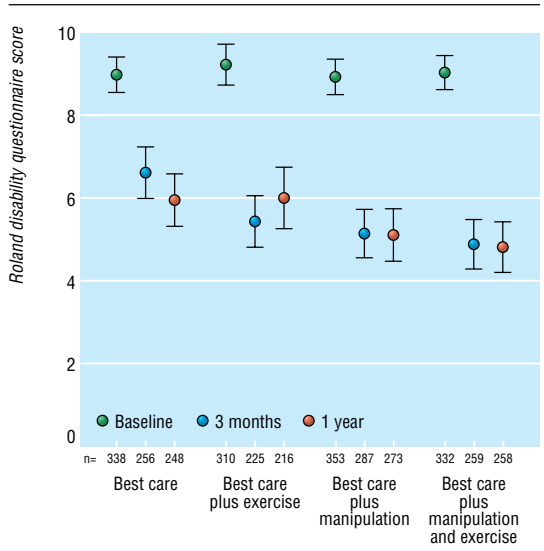


**Fig 2** Mean Roland disability questionnaire scores (with 95% confidence intervals) over 12 months by group: "best care" in general practice, best care plus exercise alone, best care plus manipulation alone, and best care plus manipulation and exercise

*Manipulation followed by exercise*

Manipulation followed by exercise produced significant improvements in Roland disability scores at three months (1.9; 1.2 to 2.6) and at one year (1.3; 0.5 to 2.1); in mean Von Korff disability and pain scores and back beliefs, fear avoidance beliefs, and SF-36 physical scores at both three and 12 months; but in mean SF-36 mental score only at three months (table 4). Three of these 13 significant improvements were significantly greater than the corresponding improvements from manipulation without exercise—in fear avoidance beliefs scores at three and 12 months and back beliefs scores at 12 months.

## Discussion

### Principal findings

The "back to fitness" programme led by physiotherapists encourages participants to increase their physical activity in a socially supportive milieu. At three months, participants randomised to this programme reported significant improvements in the primary functional outcome measure (Roland disability score) and several secondary outcomes—disability and pain, back beliefs, fear avoidance, and general physical health. Their mean improvement in disability score was equal to 35% of the population standard deviation—a "standardised difference" of 0.35. At 12 months, they maintained their reductions in disability and pain in full, and their improved beliefs about back pain in part, but not their other improvements, notably in Roland disability scores.

At three and 12 months, participants randomised to the spinal manipulation package delivered by chiropractors, osteopaths, and physiotherapists reported significant improvements in Roland disability scores and several secondary outcomes—pain, back beliefs, and general physical health. Their disability scores improved by a standardised difference of 0.39 at three months and 0.25 at 12 months. They also reported improved mental health at three months, and improved disability at 12 months. These benefits did not differ between NHS and private premises.

At three and 12 months, participants randomised to combined treatment reported significant improvements in all reported outcomes except general mental health, which was significant only at three months. Their Roland disability scores improved by a standardised difference of 0.47 at three months and 0.32 at 12 months. However only two outcomes—back beliefs and fear avoidance—achieved significant improvements over manipulation alone. Although combined treatment offers little more than manipulation alone, firm recommendations depend on detailed economic analysis (reported in accompanying paper).[23]

**Strengths and weaknesses of the study**

This trial examined the pragmatic question of how general practitioners should manage patients with back pain that does not resolve spontaneously. The patients of nearly 200 participating general practices were broadly typical of the United Kingdom. As we randomly allocated manipulators delivering therapy between their own and NHS premises, the absence of any difference in outcome answers some of the criticisms of the previous MRC trial of chiropractic.[6 7] Thus the effectiveness of manipulation may not depend on location.

The nature of the "comparator" treatment may have limited the size of our positive findings. By training volunteer general practices in the active management of back pain,[8] and providing trial participants with a copy of *The Back Book*,[14] we were using "best care" in general practice as the comparator, thus reducing the opportunities for additional improvement. That only 63% of participants allocated to the exercise programme received "basic minimum treatment" may have reduced its effectiveness. We cannot be sure whether limiting the treatments available to manipulators reduced or enhanced their effectiveness.

**Meaning of the study**

The Roland disability questionnaire comprises 24 items designed to measure functional disability due to back pain, including walking, bending, sitting, lying down, sleeping, dressing, self care, and other daily activities.[17] Each item contributes one point to the total score. We found that exercise enabled participants to perform an average of 1.4 additional personal functions at three months, manipulation generated 1.6 additional personal functions at three months and 1.0 at 12 months, and combined treatment generated 1.9 additional personal functions at three months and 1.3 at 12 months.

Thus exercise improves back function by a small, but statistically significant, margin at three months; it also achieves sustained reductions in disability and pain, and in adverse beliefs about back pain. Manipulation improves back function by a small to moderate margin at three months and a small but significant margin at 12 months; it also achieves sustained improvements in disability and pain, adverse back beliefs, and general physical health. Combined treatment improves back function by a moderate margin at three months and a small but significant margin at 12 months; generally it achieves little more than manipulation, except for much greater improvements in beliefs about back pain and fear avoidance.

**Unanswered questions**

Are these small to moderate clinical benefits worth the cost of therapy? The large cost of back pain means that small differences in clinical outcomes may have large economic effects. We report the costs and benefits in quality of life of manipulation, exercise, and combined treatment in the accompanying economic paper.[23]

We thank all participants—patients, primary care staff, and the collaborators listed on bmj.com—for their contributions. Members of the UK BEAM Trial Team: Ian Russell, Martin Underwood, Stephen Brealey, Kim Burton, Simon Coulton, Amanda Farrin, Andrew Garratt, Emma Harvey, Louise Letley, Andrea Manca, Jeannett Martin, Jennifer Klaber Moffett, Veronica Morton, David Torgerson, Madge Vickers, Ken Whyte, Melanie Williams.

Contributors: SB has been assistant trial manager at the York coordinating centre since August 2000. He contributed to collecting, validating, analysing and interpreting data, and drafting this paper. KB was a member of the MRC Back Pain Working Party, participated in the design and implementation of the trial, and was a member of the group that developed the spinal manipulation package. As joint national clinical coordinator, he was responsible for gaining the collaboration of the osteopathic profession in delivering this package and for implementing the resulting programme across the United Kingdom. He contributed to drafting this paper. SC was trial data manager, responsible for designing and implementing the data management plan and verifying and validating data. He contributed to implementing the trial design, developing and implementing randomisation procedures, developing questionnaires and other data collection instruments, analysing and interpreting data, and drafting this paper. AF was trial statistician, responsible for designing and implementing the analysis plan and validating data statistically. She contributed to implementing the trial design, developing and implementing randomisation procedures, developing questionnaires and other data collection instruments, analysing and interpreting data, and writing this paper. AG was a member of the MRC Back Pain Working Party, participated in the design and implementation of the trial, and was responsible for the evaluation and selection of the patient assessed health instruments. He contributed to developing questionnaires, analysing and interpreting data, and drafting this paper. EH was trial manager at the York coordinating centre, responsible for establishing and maintaining the 14 treatment centres and recruiting clinical advisers and local coordinators. She contributed to development and delivery of training in active management, implementing the "back to fitness" exercise programme, the spinal manipulation package, and the trial design; developing questionnaires and other materials; collecting, analysing, and interpreting data; and writing this paper. LL supervised the activities of practice based research nurses and coordinated nurse training. She contributed to implementing the trial design, selecting and recruiting practices for the trial, developing trial documentation including standard operating procedures, and drafting this paper. AM has been assistant trial economist since April 2002, responsible for designing and implementing the economic analysis plan. He estimated unit costs, analysed and interpreted economic data, and contributed to drafting this paper. JM had overall responsibility for all nursing activity within the trial including training and quality control through the GPRF regional training nurses. She contributed to implementing the trial design, selecting and recruiting prac-

## What is already known on this topic

The role of different physical treatments for back pain is not clear

Evidence for the effectiveness of spinal manipulation is conflicting; one trial reported that treatment by private chiropractors was superior to routine outpatient treatment in the NHS but did not consider the effect of treatment location

Weak evidence exists for general programmes that encourage physical activity as a treatment for back pain

## What this study adds

The spinal manipulation package improves back function by a small to moderate margin at three months and by a smaller but still statistically significant margin at one year, irrespective of location

The exercise programme improves back function by a small but significant margin at three months but not at one year

Manipulation followed by exercise improves back function by a moderate margin at three months and by a smaller but still significant margin at one year

tices, trial documentation including standard operating procedures, and drafting this paper. JKM was a member of the MRC Back Pain Working Party, participated in the design and implementation of the trial, contributed to the original grant application, and acted as grant holder responsible for physiotherapy within the trial. She developed the "back to fitness" exercise programme and was a member of the group that developed the spinal manipulation package. As joint national clinical coordinator, she was responsible for gaining the collaboration of the physiotherapy profession in both the exercise programme and the manipulation package and for implementing the exercise programme across the United Kingdom. She contributed to drafting this paper. VM has been assistant trial statistician since April 2002. She contributed to validating, analysing, and interpreting data and drafting this paper. IR was a member of the MRC Back Pain Working Party, participated in the design and implementation of the trial, was primarily responsible for the original grant application, and acted as grant holder and principal investigator in York. He chaired the trial management group and contributed to analysing and interpreting data and to writing this paper. He guarantees its statistical content. DT was trial economist, responsible for designing and implementing the economic evaluation and validating economic data. He contributed to implementing the trial design, developing questionnaires and other data collection instruments, analysing and interpreting data, and drafting this paper. MU was a member of the MRC Back Pain Working Party, participated in the design and implementation of the trial, contributed to the original grant application, and acted as grant holder responsible for medical aspects of the trial. He was trial manager at the London coordinating centre until October 1999, then joint national clinical coordinator. He contributed to developing questionnaires and developing other data collection instruments, the development and delivery of the training in active management, recruiting practices and participants, following up participants, managing, analysing and interpreting data, and writing this paper. He guarantees its clinical content. MV was a member of the MRC Back Pain Working Party, participated in the design and implementation of the trial, contributed to the original grant application, and acted as grant holder and principal investigator in London, responsible for recruiting practices and participants, and following up participants. She contributed to analysing and interpreting data and drafting this paper. KW has been trial manager at the London coordinating centre since October 1999. He was responsible for monitoring participant recruitment and for patient and practice based follow up procedures. He contributed to drafting this paper. MW was assistant trial manager at the York coordinating centre until August 2000 and acted as trial manager from then until April 2001. She was responsible for maintaining the 14 treatment centres, recruiting and training local coordinators, managing NHS treatment costs, and collecting and validating data. She contributed to drafting this paper.

Funding: Medical Research Council (research costs); National Health Service in England, Northern Ireland, Scotland, and Wales (excess treatment and service support costs). The MRC established a trial steering committee to advise the grant holders and trial team on trial design; the collection, analysis, interpretation; and writing up of data; and publication policy.

Competing interests**:** LL, JM, MU, MV, and KW have received salaries from the MRC. MU has received fees for speaking from Menarini Pharmaceuticals, the manufacturers of dexketoprofen and ketoprofen, and Pfizer, the manufacturers of celecoxib and valdecoxib. The other 12 authors have nothing to declare.

Ethical approval: The Northern and Yorkshire multicentre research ethics committee and 41 local research ethics committees approved the trial protocol.

1 Maniadakis N, Gray A. The economic burden of back pain in the UK. *Pain* 2000;84:95-103.
2 Waddell G, Feder G, McIntosh G, Lewis M, Hutchinson A. *Low back pain evidence review.* London: Royal College of General Practitioners, 1999.
3 Van Tulder MW, Koes BW, Bouter LM. Conservative treatment of acute and chronic nonspecific low back pain: a systematic review of randomized controlled trials of the most common interventions. *Spine* 1997;22:2128-56.
4 Assendelft WJ, Morton SC, Yu EI, Suttorp MJ, Shekelle PG. Spinal manipulative therapy for low back pain: a meta-analysis of effectiveness relative to other therapies. *Ann Intern Med* 2003;138:871-81.
5 Van Tulder M, Malmivaara A, Esmail R, Koes B. Exercise therapy for low back pain: a systematic review within the framework of the Cochrane Collaboration Back Review Group. *Spine* 2000;25:2784-96.
6 Meade TW, Dyer S, Browne W, Townsend J, Frank AO. Low back pain of mechanical origin: randomised comparison of chiropractic and hospital outpatient treatment. *BMJ* 1990;300:1431-7.
7 Assendelft WJ, Bouter LM, Kessels AG. Effectiveness of chiropractic and physiotherapy in the treatment of low back pain: a critical discussion of the British randomized clinical trial. *J Manipulative Physiol Ther* 1991;14:281-6.
8 Underwood M, O'Meara S, Harvey E, UK BEAM Trial Team. The acceptability to primary care staff of a multidisciplinary training package on acute back pain guidelines. *Fam Pract* 2002;19:511-5.
9 Klaber Moffett J, Frost H, UK BEAM Trial Team. Back to fitness programme: the manual for physiotherapists to set up the classes. *Physiotherapy* 2000;85:295-305.
10 Harvey E, Burton AK, Moffett JK, Breen A, UK BEAM Trial Team. Spinal manipulation for low back pain: a treatment package agreed by the UK chiropractic, osteopathy and physiotherapy professional associations. *Man Ther* 2003;8:46-51.
11 UK BEAM Trial Team. UK back pain exercise and manipulation (UK BEAM) trial—national randomised trial of physical treatments for back pain in primary care: objectives, design and intervention. *BMC Health Serv Res* 2003;3:16.
12 Knipschild P, Leffers P, Feinstein AR. The qualification period. *J Clin Epidemiol* 1991;44:461-4.
13 Koes BW, van Tulder MW, Ostelo R, Burton AK, Waddell G. Clinical guidelines for the management of low back pain in primary care. *Spine* 2001;26:2504-14.
14 Roland M, Waddell G, Klaber Moffett J, Burton AK, Main CJ, Cantrell T. *The back book.* Norwich: Stationery Office, 1996.
15 Frost H, Klaber Moffett J, Moser JS, Fairbank JC. Randomised controlled trial for evaluation of fitness programme for patients with chronic low back pain. *BMJ* 1995;310:151-4.
16 Klaber Moffett J, Torgerson D, Bell-Syer S, Jackson D, Llewlyn-Phillips H, Farrin A, et al. Randomised controlled trial of exercise for low back pain: clinical outcomes, costs, and preferences. *BMJ* 1999;319:279-83.
17 Roland M, Morris R. A study of the natural history of back pain. I: development of a reliable and sensitive measure of disability in low-back pain. *Spine* 1983;8:141-4.
18 Underwood MR, Barnett AG, Vickers MR. Evaluation of two time-specific back pain outcome measures. *Spine* 1999;24:1104-12.
19 Symonds TL, Burton AK, Tillotson KM, Main CJ. Do attitudes and beliefs influence work loss due to low back trouble? *Occup Med* 1996;46:25-32.
20 Waddell G, Newton M, Henderson I, Somerville D, Main CJ. A fear-avoidance beliefs questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. *Pain* 1993;52:157-68.
21 Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I: conceptual framework and item selection. *Med Care* 1992;30:473-83.
22 The EuroQol Group. EuroQol—a new facility for the measurement of health-related quality of life. *Health Policy* 1990;16:199-208.
23 UK BEAM Trial Team. United Kingdom back pain exercise and manipulation (UK BEAM) randomised trial: cost effectiveness of physical treatments for back pain in primary care. *BMJ* 2004;329:doi = 10.1136/bmj.38282.607859.AE.
24 Bombardier C, Hayden J, Beaton DE. Minimal clinically important difference in low back pain outcome measures. *J Rheumatol* 2001;28:431-8.

*(Accepted 13 October 2004)*

doi 10.1136/bmj.38282.669225.AE

Correspondence to: Martin Underwood, professor of general practice, Centre for General Practice and Primary Care, Institute of Community Health Sciences, Barts and the London, Queen Mary's School of Medicine and Dentistry, Queen Mary University of London, London E1 4NS m.underwood@qmul.ac.uk

**Amendment**

This is Version 2 of the paper. In this version, fig 1 has been amended so that the "manipulation" groups are correctly divided into "private manipulation" and "NHS manipulation" [in the previous version, all subgroups were labelled "private manipulation].